THE UNIVERSITY OF TEXAS AT AUSTIN
CENTER FOR TRANSPORTATION RESEARCH

TECHNICAL REPORT 0-6967-1
TxDOT PROJECT NUMBER 0-6967

Evaluate the Use of PWL as Payment Adjustment Factor for Asphalt Mixtures

Dr. Syeda Rahman
Dr. Amit Bhasin
Dr. Zhanmin Zhang
Dr. Andre Smit

August 2019; Published March 2020

http://library.ctr.utexas.edu/ctr-publications/0-6967-1.pdf

Technical Report Documentation Page

1. Report No.
FHWA/TX-20/0-6967-1

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Evaluate the Use of PWL as Payment Adjustment Factor for Asphalt Mixtures

5. Report Date
August 2019; Published March 2020

6. Performing Organization Code

7. Author(s)
Dr. Syeda Rahman, Dr. Amit Bhasin, Dr. Zhanmin Zhang, Dr. Andre Smit

8. Performing Organization Report No.
0-6967-1

9. Performing Organization Name and Address
Center for Transportation Research
The University of Texas at Austin
3925 W. Braker Lane, 4th Floor
Austin, TX 78759

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
0-6967

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 2017–July 2019

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 payment awarded to a Contractor for production and placement of asphalt mixes under Specification Items 341, 344, and 346 involves a payment adjustment factor that either awards a bonus or assigns a penalty to the base payment depending on the test results for laboratory-molded density (production) and in-place air voids (placement or construction). A characteristic feature of the current payment adjustment factors for production and placement is that these factors are based on average values of in-place air voids and absolute deviation from the target laboratory-molded density. In some cases, this practice results in a scenario where the Contractor delivers inconsistent quality but still receives a bonus because of the average performance. In order to overcome this weakness, several different quality measures including the Percent Within Limits (PWL) were reviewed as a part of this study. Data from TxDOT's SiteManager database for 2004 and 2014 specification years were collected for all projects under the specification Items 341, 344, and 346. Data sets from 2014 specification years, covering approximately the last five years of construction across the entire state, were then used with hypothetical models for PWL and payment adjustment factor schemes to demonstrate: (i) the feasibility and financial impact of using the PWL based approach for bonuses and penalties, and (ii) the step-by-step process of using this approach for future implementation. An analysis tool was developed to evaluate the implications of adjusting various parameters in the PWL model using past data. This tool is intended to facilitate implementation of the PWL based approach. Implementation of PWL is certainly an improvement over the existing specification that is based only on average values. This approach also ensures that the quality of material produced or placed is consistent and rewards Contractors for meeting the specification limits consistently.

17. Key Words
Hot mix asphalt, Quality measures, Percent within limits, Variability, Risk analysis

18. Distribution Statement
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) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 163	22. Price

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

THE UNIVERSITY OF TEXAS AT AUSTIN
CENTER FOR TRANSPORTATION RESEARCH

Evaluate the Use of PWL as Payment Adjustment Factor for Asphalt Mixtures

Dr. Syeda Rahman
Dr. Amit Bhasin
Dr. Zhanmin Zhang
Dr. Andre Smit

CTR Technical Report:	0-6967-1
Report Date:	August 2019; Published March 2020
Project:	0-6967
Project Title:	Evaluate the Use of Percent Within Limits as Payment Adjustment Factor for Placement of Asphalt Mixes
Sponsoring Agency:	Texas Department of Transportation
Performing Agency:	Center for Transportation Research at 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
3925 W. Braker Lane, 4th Floor
Austin, TX 78759

http://ctr.utexas.edu/

Disclaimers

Engineering Disclaimer

Research Supervisor: Amit Bhasin

Acknowledgments

The authors acknowledge the financial support of the Texas Department of Transportation for project 0-6967. The authors would like to thank Darrin Jensen and Chris Glancy, who served as project managers. The authors are also thankful to Ryan Barborak and Travis Patton for their assistance and suggestions in conducting the project. Special thanks to Andre Smit for his support and assistance with data collection and management and Aaron Coenen for sharing valuable information about implementation of PWL-based quality assurance plan in Wisconsin.

EXECUTIVE SUMMARY

The payment for production and placement of asphalt mixes typically requires that the Contractor meets requirements for several metrics including gradation, binder content, laboratory-molded density, segregation, longitudinal joint density, thermal profile and in-place air voids. Of these, payment adjustment factors are computed to adjust the payment awarded to the Contractor based on test results for laboratory-molded density (production) and in-place air voids (placement or construction). The payment adjustment factors are computed separately for production and placement and averaged for the final payment. The payment adjustment factors can increase the Contractor payment by as much as 5%, 7.5%, and 10% as a bonus for specification Items 341, 344, and 346, respectively or decrease by 28% for Item 341 and 30% for Items 344 and 346 as a penalty. As an extreme case, the Contractor may be required to remove or replace or forfeit the payment if production and placement specifications are not met.

A characteristic feature of the current payment adjustment factors for production and placement is that these factors are based on average values of in-place air voids and absolute deviation from the target laboratory-molded density. These average values are obtained using measurements made over typically four sublots for each lot of production. In some cases, this practice results in a scenario where the Contractor delivers inconsistent quality but still receives a bonus because of the average performance. The goal of this project was to evaluate the use of a payment adjustment factor that is not based solely on the average test results for production and placement but is based on a measure of quality that reflects both the average and the variability in production and placement.

A review of the existing literature and practice in other states and recommended practice by the Federal Highway Administration (FHWA) revealed that this can be achieved using a metric that is referred to as the Percent Within Limits (PWL). In simple terms, PWL provides an estimate for the probability that a test result from a lot falls within the upper and lower specification limits. The PWL is calculated using the upper specification limit, the lower specification limit, the sample mean, the sample standard deviation, and an appropriate probability distribution table. A minimum value of this probability or PWL, referred to as the acceptable quality level or AQL, is typically prescribed as a part of the specification for the Contractor to receive full payment. A payment adjustment factor can be developed to award a bonus for exceeding this AQL and a penalty for falling below this AQL value. A rejectable quality level or RQL is also prescribed as the lower threshold below which the Contractor will not receive any payment or may be required to remove or replace the work. The main difference between the current and the PWL based payment adjustment factor model is that the latter accounts for variability and consistency in quality, whereas the former does not.

As a part of this study, several different quality measures including PWL were reviewed. Data for all projects covering all districts from TxDOT's SiteManager database were collected for both the 2004 and 2014 specification years for Items 341, 344, and 346. These data sets were analyzed to determine the variability in production and placement metrics and highlight the current gap in payment adjustment factor, i.e. the possibility of awarding a bonus to a Contractor based on the average test results even when the variability is high. These data were also then used with a hypothetical model for PWL and payment adjustment factor scheme to demonstrate: (i) the feasibility and financial impact of using the PWL based approach for bonuses and penalties, and (ii) the step-by-step process of using this approach for future implementation. Implementation of the PWL based payment adjustment scheme may require input from stake holders. An analysis tool has been provided that will allow the Agency (TxDOT) to create any hypothetical model and evaluate its implications using past data.

Implementation of PWL is certainly an improvement over the existing specification that is based only on average values. This approach also ensures that the quality of material produced or placed is consistent and rewards Contractors for meeting the specification limits consistently.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW

The payment for production and placement of asphalt mixes typically requires that the Contractor meets requirements for several metrics including gradation, asphalt content, laboratory-molded density, segregation, longitudinal joint density, thermal profile and in-place air voids. Of these, payment adjustment factors are computed to adjust the payment awarded to the Contractor based on test results for laboratory-molded density (production) and in-place air voids (placement or construction). The payment adjustment factors are computed separately for production and placement for each sublot and averaged for the final payment for a lot. The payment adjustment factors can increase the Contractor payment by as much as 5%, 7.5%, and 10% as a bonus for specification Items 341, 344, and 346, respectively or decrease by 28% for Item 341 and 30% for Items 344 and 346 as a penalty. As an extreme case, the Contractor may be required to remove or replace or forfeit the payment if production and placement specifications are not met.

A characteristic feature of the current payment adjustment factors for production and placement is that these factors are based on average values of in-place air voids or absolute deviation from the target laboratory-molded density. These average values are obtained using measurements made over typically four sublots for each lot of production. In some cases, this practice results in a scenario where the Contractor delivers inconsistent quality but still receives a bonus because of the average performance. The goal of this project was to evaluate the use of a payment adjustment factor that is not based solely on the average test results for production and placement but is based on a measure of quality that reflects both the average and the variability in production and placement.

This chapter presents a brief background in terms of the current payment method practiced by the state of Texas. Several available systems, including percent within limits (PWL), percent defective (PD), average absolute deviation (AAD), conformal index (CI), and moving average, are then reviewed to identify the best approach for determining the payment adjustment factor. Among these systems, PWL is the most recommended method by other agencies and used commonly outside of the state of Texas. The PWL approach has potential to be adopted as the basis for a payment adjustment factor for production and placement of HMA. Finally, practices followed by other states that adopted PWL as payment adjustment method, were studied to propose the best scheme for the state of Texas. Challenges that some of the states faced in introducing a new system like PWL and strategies to manage those challenges have also been summarized to provide guidance for future implementation.

1.1 PROJECT BACKGROUND

The current TxDOT specification [11] determines bonuses to Contractors for placement payment based on the average of the in-place air voids and for production payment based on the average of the absolute deviation from the target laboratory-molded density for a number of sublots. The average placement and production payment adjustment factors (PAF) for the entire lot is then calculated by averaging PAFs for all sublots. This practice can potentially create a scenario where Contractors can deliver inconsistent quality but still receive a bonus because of the average performance. Therefore, the primary goals of this project are

- to evaluate the use of percent within limits (PWL) that incorporates variability as the payment adjustment factor for asphalt mixes,
- to assess practices for PWL adopted by other states,
- to collect and analyze data from various databases hosted by TxDOT (SiteManager or SMGR and Design and Construction Information System or DCIS) to assess the overall relationship between payment adjustment factors and mix properties and performance predicting test data,
- to develop several alternative schemes for payment adjustment factors, and
- to identify the cost benefit and associated risk to the agency using each scheme.

1.2 REPORT STRUCTURE

In order to achieve the objectives of this study, it was divided into five major tasks:

- a review of the literature,
- data mining,
- global correlations between payment adjustment factor, consistency in quality and performance,
- developing potential alternative schemes to current payment adjustment factor, and
- computing hypothetical cost to TxDOT for alternative schemes.

The structure of this report follows the above tasks and objectives.

1.3 LITERATURE REVIEW

The remainder of this chapter begins with a brief background in terms of the current payment method practiced by the state of Texas. Several available systems, including percent within limits (PWL), percent defective (PD), average absolute deviation (AAD), conformal index (CI), and moving average, are then reviewed to identify the best approach for determining the payment adjustment factor. Among these systems, PWL is perceived as the best method that warrants consistency in the production and placement of HMA. Finally, practices followed by other states that adopted PWL as payment adjustment method, were studied to propose alternative payment adjustment factor schemes for the state of Texas. Challenges that some of the states faced in introducing a new system such as the PWL and strategies to manage those challenges have also been summarized in the last section of this chapter to provide guidance for future implementation.

1.4 CURRENT PRACTICE

The current TxDOT specification [11] determines the total adjustment pay as the average of the production payment adjustment and the placement payment adjustment for each lot. Each lot consists of upto four equal sublots. The first lot has a default size of 1000 tons. The remaining lots can vary in size from the default size to a maximum of 4000 tons. The production payment adjustment factor for each sublot is determined based on the absolute value of the deviation from the target laboratory-molded density obtained from Engineer's test results. A graphical representation of the production PAF for specification Item 341, for instance, is provided in Figure 1.1. The production payment adjustment factor for completed lots is the average of the payment adjustment factors for the four sublots sampled within that lot.

The placement payment adjustment is determined from tabulated data (as presented graphically in Figure 1.2 for Item 341) for each sublot that requires in-place air void measurement. If any sublot is not subject to in-place air-void determination due to random sampling, a payment adjustment factor of 1 is assigned to the sublot. The placement payment adjustment factor for completed lots is the average of the placement payment adjustment factors for up to four sublots within that lot. Payment for each sublot, including applicable payment adjustment bonuses, is only paid for sublots for which the Contractor supplies the Engineer with the required documentation for production and placement quality control (QC)/ quality assurance (QA), thermal profiles, segregation density profiles, and longitudinal joint densities. No placement penalty is assessed for any sublot placed in Lot 1 when the in-place air voids are greater than or equal to 2.7% and less than or equal to 9.9% for Item 341. The maximum allowable in-place air void values differ for Items 344 and 346 and cannot exceed 9.0% for Item 344 and 8.0% for Item 346.

In the event that the Engineer's test results require suspension of production or remove or replace condition for placement, whereas the Contractor's test results are within the specification limits, the Contractor may appeal for referee testing. If after referee testing, the laboratory-molded density for any production sublot results in a "remove and replace" condition, the Engineer may require removal and replacement or may allow the sublot to be left in-place without payment. If the new payment adjustment factor is 0.720 or greater for Item 341 and does not drop below 0.700 for Items 344 and 346, the new payment adjustment factor applies to that sublot. If the new payment adjustment factor is less than these values, no payment is made for the sublot. The sublot is then subject to removal and replacement, or the Engineer may allow the sublot to be left in-place without payment.

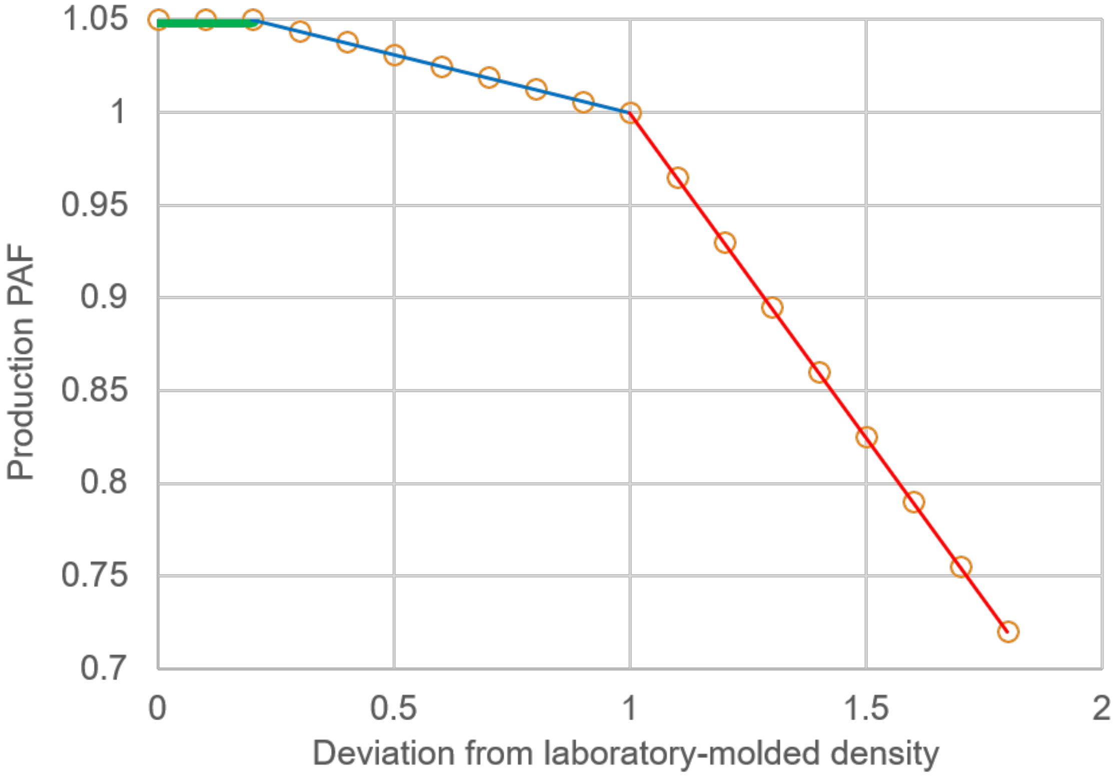


Figure 1.1. Current payment adjustment factors (PAF) for absolute deviations from the target laboratory-molded density for specification Item 341.

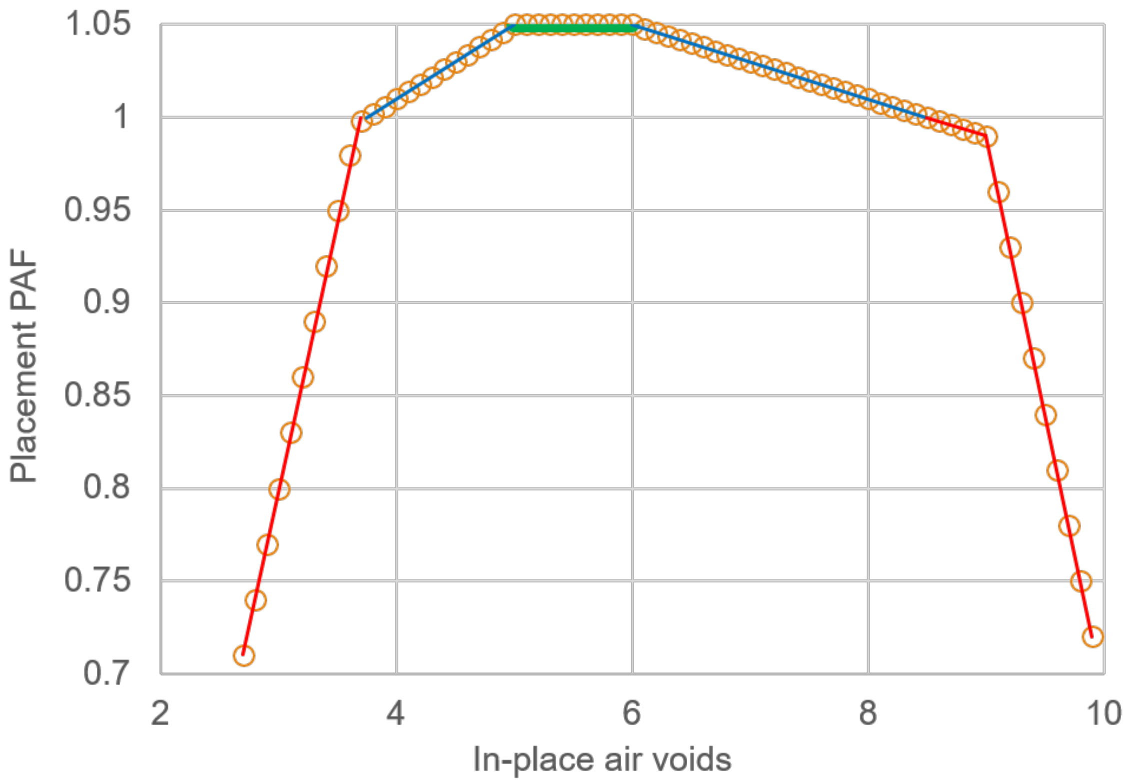
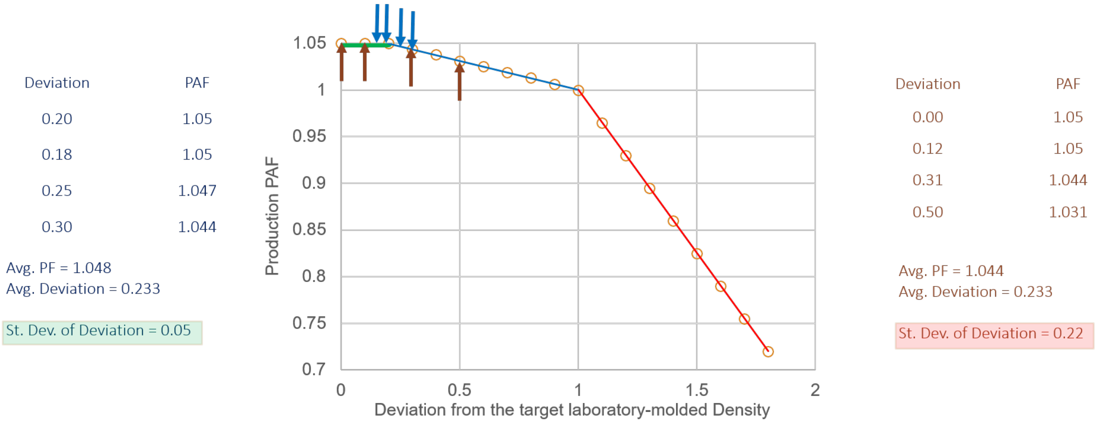


Figure 1.2. Current placement payment adjustment factors (PAF) for average in-place air voids for specification Item 341.

The problem with this averaging technique lies in the fact that such QC/QA approach does not take variability into account and therefore is not a good indicator of consistency. In other words, this approach cannot differentiate between construction practices that may meet specification limits on an average but may have substantial differences in terms of the consistency with which this quality metric is achieved. This could be simply visualized with the following two examples as shown in Figures 1.3 and 1.4. In the first example, Figure 1.3, one lot (blue down arrows) has four sublots with absolute deviations of 0.20, 0.18, 0.25, and 0.30 from the target laboratory-molded density with the corresponding payment adjustment factors of 1.05, 1.05, 1.047, and 1.044, respectively, receiving an average PAF of 1.048. On the other hand, a second lot (brown up arrows) receives a PAF of 1.044 from four sublots with payment adjustment factors of 1.05, 1.05, 1.044, and 1.031 for deviations of 0.00, 0.12, 0.31, and 0.50, respectively. The latter has a much higher standard deviation (0.22) than the previous lot (0.05) but is receiving almost the same bonus of 4.4% as the first lot that receives a bonus of 4.8%. Similarly, from Figure 1.4, it can be seen that one construction practice wins a bonus of 4.5% with a coefficient of variation of 4.5% whereas another construction wins almost the same bonus of 4.3% with a much higher coefficient of variation of 14%. The current practice does not account for this variation and hence cannot differentiate between consistent and inconsistent performance. Therefore, a robust PAF scheme is necessary that not only controls the quality characteristics within the specification limits but also ensures consistency in the quality of production and placement of hot mix asphalt.



Deviation	PAF
0.20	1.05
0.18	1.05
0.25	1.047
0.30	1.044

Avg. PF = 1.048
Avg. Deviation = 0.233

St. Dev. of Deviation = 0.05

Deviation	PAF
0.00	1.05
0.12	1.05
0.31	1.044
0.50	1.031

Avg. PF = 1.044
Avg. Deviation = 0.233

St. Dev. of Deviation = 0.22

Figure 1.3. Example showing consistent and inconsistent performance receiving the same bonus for production. Consistent performance (blue down arrows) has low deviations that are very close to the target, whereas inconsistent performance (brown up arrows) has high variability.

1.5 DIFFERENT QUALITY MEASURES

Burati et al., in an FHWA report [3], summarized several quality measures for payment adjustment factor including percent within limits (PWL) or percent defective (PD), average absolute deviation (AAD), conformal index (CI), and the moving average.

Percent within limits or PWL, also known as percent conforming refers to the percentage of the lot that falls above the lower specification limit and below the upper specification limit. In this quality measure, both the sample mean and the sample standard deviation are used to estimate the percentage of the population (lot) that is within the specification limits. It is analogous to determining the area under the normal curve where the sampled population is bell shaped and not highly skewed or bimodal. PD is related to the PWL by the simple relationship PWL = 100 - PD.

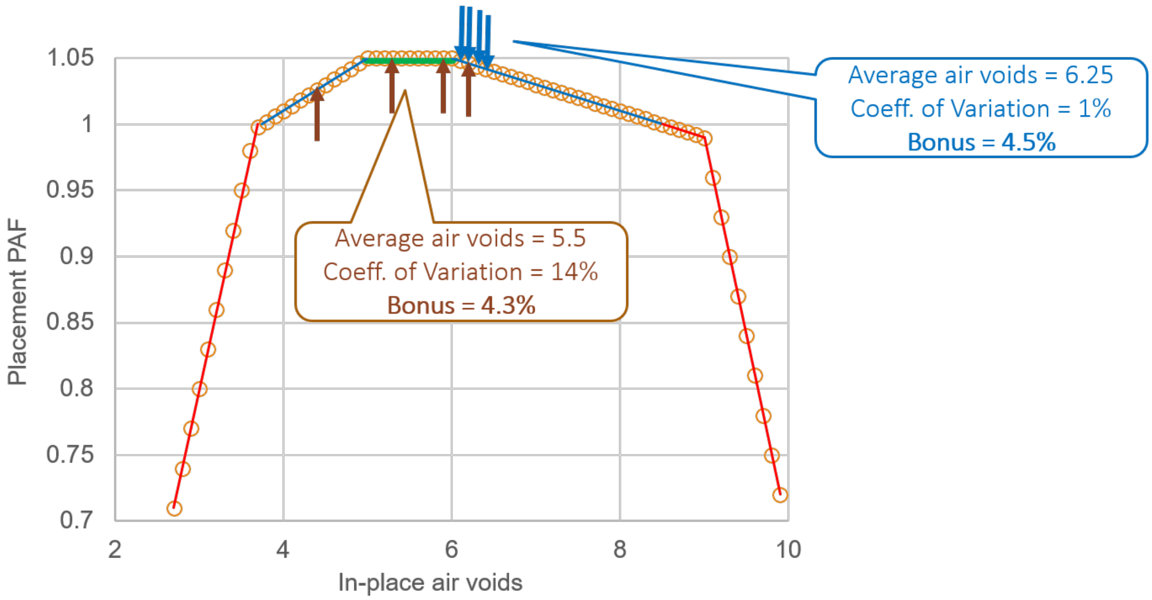


Figure 1.4. Example showing consistent and inconsistent performance receiving almost the same bonus for placement. Inconsistent construction practice, indicated with brown up arrows, receives a bonus of 4.3% with a coefficient of variation as high as 14%.

Average absolute deviation or AAD is defined as the mean of absolute deviations from a target or specified value for a series of test results. The purpose of choosing the absolute value of the deviation from the target is to prohibit Contractor from benefiting by offsetting negative deviations by positive deviations. The existing TxDOT production payment adjustment factor falls under this system.

Conformal index or CI is the root mean square of the quantity obtained by summing the squares of the deviations from the target value and dividing by the population number. The CI is conceptually similar to the standard deviation: the standard deviation measures dispersion of a series of results around a mean, whereas the CI measures the dispersion around a target or specified value. Another acceptance procedure is developed based on the moving or running average of the quality characteristic. For moving averages, a default sample size is first determined. The first average is calculated from the first set of population. For the second moving average, the latest value replaces the first value in the population, and so on.

Among these methods, PWL and PD have been widely accepted as the most robust method as both of these utilize the sample mean and sample standard deviation and are applicable to both one-sided and two-sided acceptance properties. One might assume that AAD and CI may be good measures of quality since the lower the dispersion, the closer the process is to the target value. However, these methods do not determine the variability adequately as the same AAD or CI can be obtained from very different sets of test results with different means and standard deviations. FHWA has recommended PWL for many years, and AASHTO Quality Assurance Specification also uses PWL for determining HMA payment adjustment factor.

At this point before proceeding into discussion about PWL schemes adopted by other states, it is necessary to introduce the common terminology associated with the PWL based QC/QA approach. This terminology will be used throughout the remainder of the report to explain the step by step process involved in development of the PWL based payment adjustment factor scheme.

1.6 TERMINOLOGY

The following terminology is used in the remainder of this chapter

- Sample mean ($\bar{X}$) - Average of the test results.
- Sample standard deviation (s) - Measures the spread of a given data set with respect to its mean.
- Quality characteristic - A material or construction characteristic that is measured to determine acceptability of that material or construction. For example, laboratory-molded density is a quality characteristic that is measured to determine acceptability of HMA production. Similarly, field density is a quality characteristic that is measured to determine acceptability of placement.
- Quality measure - A mathematical or statistical tool that quantifies the quality of material production or construction practice. Examples of a quality measure are the mean, average absolute deviation from the target, or PWL.
- Specification limit(s) - The limiting value(s) placed on a quality characteristic, established preferably by statistical and sensitivity analysis, for evaluating whether material production or construction meets the design requirements. The term can refer to either an individual upper or lower specification limit, USL or LSL, called a single specification limit, or to USL and LSL together, called double specification limits.
- Percent within limits (PWL) - The estimated percentage of the material produced or placed on site that falls above the lower specification limit and below the upper specification limit. In other words, this is the estimated percentage of material that

falls between the upper and lower specification limits.

- Percent defective (PD) - The percentage of the material produced or placed on site that falls outside the specification limits. Based on this definition, $PD = 100 - PWL$.
- Acceptable quality level (AQL) - The minimum level of established quality for a quality characteristic that is fully acceptable. AQL is not estimated, rather it is established by the agency as the value of the PWL at or above which the material is considered to be fully acceptable by the agency.
- Rejectable quality level (RQL) - The maximum level of established quality for a quality characteristic that is fully unacceptable. In other words, level of quality characteristic below which the Contractor may be required to reject and replace (or in some cases forfeit the payment). This is the value of the PWL (not estimated) that is established by the agency, typically at or below which the work is rejected and the Contractor does not receive any payment.
- Payment adjustment factor (PAF) - A multiplication factor used to determine the Contractor's payment for a unit of work.
- Operating characteristic (OC) curve - A graphical representation of an acceptance plan that shows the relationship between the actual quality of the material produced or placed on site and (1) the probability of its acceptance or (2) the probability of its acceptance at various payment levels.
- Seller's risk (α) or Type I risk or α error - The risk to the Contractor of having acceptable quality material or construction rejected. This is the probability that an acceptance plan will erroneously reject material or construction that is otherwise acceptable.
- Buyer's risk (β) or Type II risk or β error - The risk to the agency of accepting rejectable quality level material or construction. This is the probability that an acceptance plan will erroneously accept material or construction that should have been rejected.

1.7 PWL ADOPTED BY OTHER STATES

The FHWA QA assessment report published in 2014, documented that around twenty nine state agencies utilize some form of PWL specifications. Since then six more states have adopted PWL as highlighted in Figure 1.5. The following subsections summarize the practices followed by some of the states that apply PWL to calculate payment adjustment factor for the production and placement of HMA. A summary of this comparison is provided in

Figure 1.7.

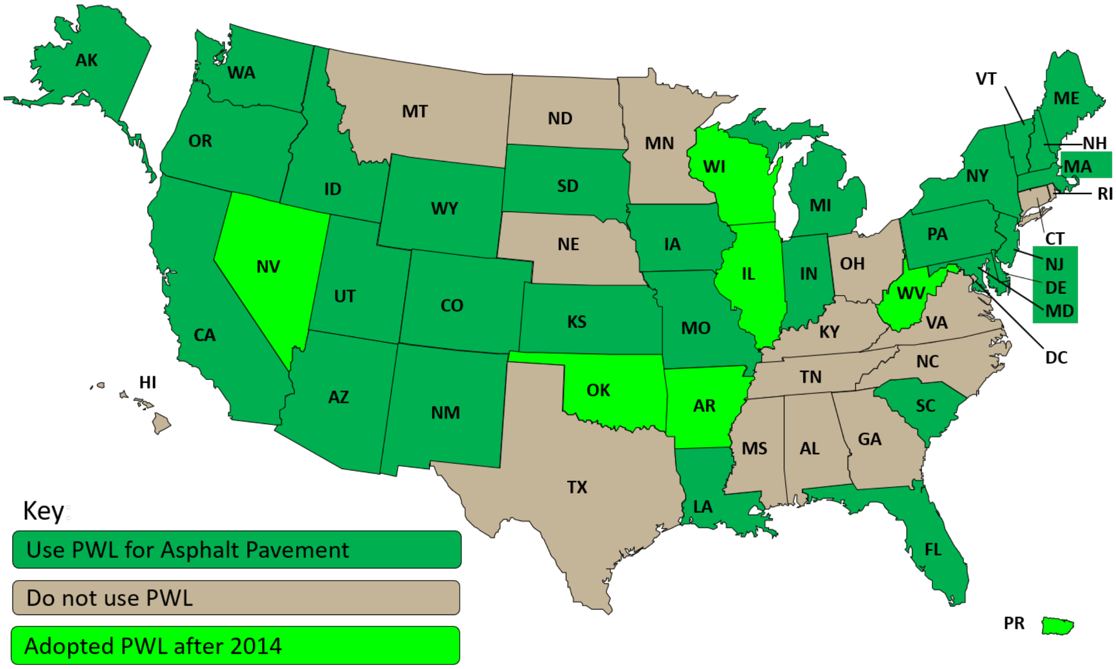


Figure 1.5. States using PWL for HMA production and placement. Source: FHWA WI Division, and 2014 FHWA QA Assessment.

1.7.1 Oklahoma

Oklahoma Department of Transportation [8] uses in-place density, air voids, asphalt cement content, and voids in mineral aggregate (VMA) as payment characteristics for a lot. A standard lot consists of four equal sublots of 1000 tons each. The lot size can vary from a minimum of 2750 tons to a maximum of 6250 tons. PWL is determined based on the amount of materials and construction that falls within the upper and lower specified limits with respect to the JMF for each characteristic. When the PWL for any characteristic is greater than 50, pay factor (PF), is calculated applying a quadratic equation as a function of PWL as

$$PF = 0.024 \times PWL - 0.0001 \times PWL^2 - 0.35. \tag{1.1}$$

A composite pay factor (CPF) is then determined for the entire lot using a weighted average method, whereby the highest importance is given to the in-place density,

$$CPF = \frac{4 \times PF_D + 3 \times PF_V + 2 \times PF_A + PF_{VMA}}{10}, \tag{1.2}$$

where: PF_D = pay factor for in-place density, PF_V = pay factor for air voids, PF_A = pay factor for asphalt content, and PF_{VMA} = fay factor for voids in mineral aggregates.

A lot is considered acceptable with positive payment adjustment with respect to a particular characteristic if the PWL is greater than or equal to 90. A lot is considered rejectable with respect to a particular characteristic if the PWL is less than 50. Lots exceeding the rejectable quality level, but falling below the acceptable quality level, is considered for negative payment adjustments. If a lot fails to exceed the rejectable quality level in one or more characteristics, the Engineer may require its removal and replacement at the Contractor's expense. If the PWL is less than 50 and the Engineer does not require removal and replacement of the lot, the lot may be left in-place subject to a pay factor of zero for the respective quality characteristic. The Engineer performs all tests for acceptance for all quality control test characteristics on a lot-to-lot basis. Outliers are marked based on a mathematical calculation performed according to ASTM E 178 (the upper 2.5% significance level). Engineer may discard the outlier and supplement the remaining test results if necessary. In an event that large differences exist between the Engineer and the Contractor test results, Contractor may appeal for referee testing.

1.7.2 Indiana

The Indiana Department of Transportation [6] performs quality acceptance testing on samples obtained during pavement construction for three properties. These properties include asphalt content, gyratory-compacted specimen air voids (V_a) and voids in mineral aggregate (VMA). The department also measures density from core samples taken from the pavement after construction. The department uses pseudo-random number system to direct the Contractor on where to obtain QA mixes and density samples. Pay factors, PFs, are calculated for asphalt content, air voids at N_{des}, void in mineral aggregates, VMA at N_{des}, and in-place density, G_{mm} separately. The appropriate pay factor is calculated as a continuous function (equation form) of PWL, as

$$PF = (105.00 - 0.50 \times (100 - PWL))/100 \tag{1.3}$$

for estimated PWL greater than 90, and

$$PF = (105.00 - 0.000020072 \times (100 - PWL)^{3.5877})/100 \tag{1.4}$$

for estimated PWL greater than or equal to 50 and less than or equal to 90.

If the lot PWL for any one of the properties is less than 50 or a sublot has an air void content less than 1.0% or greater than 7.0%, the lot is adjudicated as a failed material. A composite pay factor for each lot based on test results for mix properties and density is determined by a weighted formula as

$$PF = 0.35 \times PF_D + 0.35 \times PF_V + 0.20 \times PF_A + 0.10 \times PF_{VMA}. \tag{1.5}$$

where: PF_D = pay factor for in-place density, % G_{mm} PF_V = pay factor for air voids at N_{des}, PF_A = pay factor for asphalt content, and PF_{VMA} = fay factor for voids in mineral aggregate at N_{des}.

The lot quality assurance adjustment for mix properties and density is also quantified in terms of lot pay factor and mix adjustment factor. Only the lower quality index is specified for in-place density whereas both upper and lower quality indices are specified for mix properties. Test results exceeding the tolerance limits are considered as a failed material and need to be adjudicated. If the Contractor conducted QC test results do not agree with the acceptance test results, the Contractor may request for additional testing. The appeal results replace all previous test results for acceptance of mix and are used to determine the pay factor.

1.7.3 Louisiana

Louisiana Department of Transportation and Development (DOTD) [9] specifies tabulated percentage payment depending on the PWL for mainline pavement density. The state also uses PWL as a means to validate JMF. The Contractor are responsible to perform the minimum quality control tests that include theoretical maximum specific gravity, G_{mm}, asphalt content, gradation, temperature of mix for loose mix and G_{mm} at $N_{initial}$, G_{mm} at N_{des}, V_a, VMA, voids filled with asphalt (VFA) for compacted mix for each plant lot (P-lot) of 1000 tons. The quality of plant is monitored through the five sublots rolling average and standard deviation for aggregate gradation, asphalt content, air voids, and G_{mm}. When rolling five test results of air voids or G_{mm} fall below 71 PWL, average VFA and gradation for no. 8 and no. 200 are outside the specification limits, and the asphalt content is ±0.2% of the JMF, corrective actions need to be taken or the production needs to be ceased.

The mainline lot size is 37,500 linear lane feet consisting of five sublots, each with 7,500 linear lane feet. The sublot is divided into three segments of 2,500 linear feet, and one core is randomly collected from each segment by the Engineer. The District Laboratory conducts the density testing of each acceptance in-place core. Contractors, with proven plant production consistency, may be allowed to sample and test in-place cores for acceptance in lieu of District Laboratory acceptance testing, and when recommended by the District Laboratory Engineer (DLE) and approved by the Materials Engineer. The plant production consistency is determined by continuously monitoring plant data and in-place data by JMF, by plant, and by contractor. The DLE conducts statistical analysis for means and variances (F and t) tests for a set of a minimum of 45 Contractor acceptance tests and 15 DOTD verification tests results. If the Contractor data fail the F and t test analysis, DOTD acceptance testing of in-place cores will resume.

The plant pay quantity is determined by multiplying the measured quantity of asphalt mixes with the adjustment factors shown in Table 1.1.

Table 1.1. Pay quantity adjustment factors for asphalt mixes for Louisiana DOTD.

Theoretical maximum specific gravity, (G_{mm})(DOTD TR 327)	Adjustment factor
2.340 - 2.360	1.02
2.361 - 2.399	1.01
2.400 - 2.540	1.00
2.541 - 2.570	0.99
2.571 - 2.590	0.98

The adjustment factor for mixes with theoretical maximum specific gravities less than 2.340 or more than 2.590 is determined by the following formulas:

$$F = \frac{2.400}{S} \tag{1.6}$$

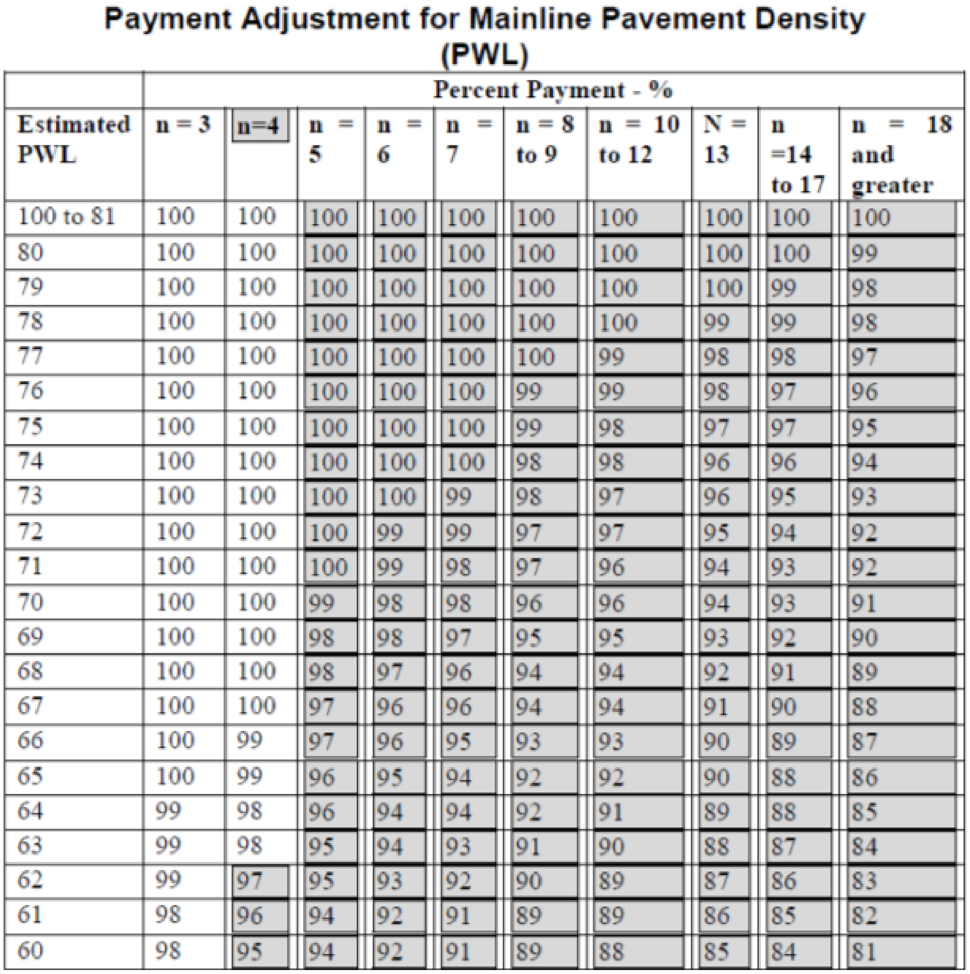
for theoretical maximum gravity less than 2.340, and

$$F = \frac{2.540}{S} \tag{1.7}$$

for theoretical maximum gravity more than 2.590,

where: F = quantity adjustment factor, and S = theoretical maximum specific gravity of mix from approved job mix formula.

For all mainline mixes, adjustments in contract unit price for in-place density are based on PWL using the tabulated payment adjustment factors and are applied to the theoretical mainline lane quantity and contract unit price. A portion of the Louisiana DOTD payment adjustment factors for mainline pavement density is presented in Figure 1.6.



Payment Adjustment for Mainline Pavement Density (PWL)

	Percent Payment - %									
Estimated PWL	**n = 3**	**n=4**	**n = 5**	**n = 6**	**n = 7**	**n = 8 to 9**	**n = 10 to 12**	**N = 13**	**n =14 to 17**	**n = 18 and greater**
100 to 81	100	100	100	100	100	100	100	100	100	100
80	100	100	100	100	100	100	100	100	100	99
79	100	100	100	100	100	100	100	100	99	98
78	100	100	100	100	100	100	100	99	99	98
77	100	100	100	100	100	100	99	98	98	97
76	100	100	100	100	100	99	99	98	97	96
75	100	100	100	100	100	99	98	97	97	95
74	100	100	100	100	100	98	98	96	96	94
73	100	100	100	100	99	98	97	96	95	93
72	100	100	100	99	99	97	97	95	94	92
71	100	100	100	99	98	97	96	94	93	92
70	100	100	99	98	98	96	96	94	93	91
69	100	100	98	98	97	95	95	93	92	90
68	100	100	98	97	96	94	94	92	91	89
67	100	100	97	96	96	94	94	91	90	88
66	100	99	97	96	95	93	93	90	89	87
65	100	99	96	95	94	92	92	90	88	86
64	99	98	96	94	94	92	91	89	88	85
63	99	98	95	94	93	91	90	88	87	84
62	99	97	95	93	92	90	89	87	86	83
61	98	96	94	92	91	89	89	86	85	82
60	98	95	94	92	91	89	88	85	84	81

Figure 1.6. Payment adjustment for mainline pavement density for Louisiana DOTD.

1.7.4 Wisconsin

The Wisconsin Department of Transportation (WisDOT) [13] required QC properties include mix bulk specific gravity (G_{mb}), G_{mm}, V_a, VMA, aggregate gradation, and percent asphalt content. Wisconsin also uses running average values to regulate production quality. If two consecutive running average values exceed the warning limits, the production needs to be ceased and requires adjustment. The approximate location of each sample within the prescribed sublots is determined by selecting random numbers using ASTM D3665 or by using a calculator or computerized spreadsheet that has a random number generator. To verify product quality, bulk specific gravity of the mix, maximum specific gravity of the mix, air voids, and voids in the mineral aggregate are measured by the department. When quality verification (QV) test results deviate from the QC test results, the bureau's AASHTO accredited laboratory referee tests the retained portion of the QV samples and the retained portion of the nearest available previous QC sample. If after referee testing the material is deemed unacceptable, it has to be removed and replaced. After verifying the material characteristics, the department reduces pay for the tonnage of nonconforming mix, if the Engineer allows that mix to remain in-place. A minimum of 75 PWL for pavement density is required to qualify for payment. 100% pay requires V_a to be above 3.2 and below 5.8 and VMA to be greater than 0.5. Whereas, 50% pay corresponds to mixes with 5%< V_a<1.5% and VMA >1% that is allowed to remain in-place. For materials that fall above 50% pay, different sets of continuous functions in the form of linear equations are used to determine the payment adjustment for high and low air voids and low VMA [12]. If the lot density is greater than the minimum required and all individual air voids test results for the same day mix fall between 2.5% - 4.0%, the department specifies bonuses in dollar amount for that lot. Pay incentive and disincentive for HMA pavement density are provided in Table 1.2.

1.7.5 Kansas

Kansas Department of Transportation (KDOT) [7] uses PWL specification to pay the Contractor for two quality characteristics, V_a at N_{des} and in-place density ($\%G_{mm}$). Payment is given for each lot, where lot size corresponds to a day's production or 3,000 tons consisting of four equal sublots of 750 tons. Ten Contractor QC tests and five agency verification tests are performed for each lot. The QC test results are used for payment as long as the variances, determined by the F test, and the means, determined by the t test, are in compliance with the KDOT quality verification test results. If the F and t test show that the QC test results do not comply with the verification test results within the significance level of 0.01, the KDOT results are used for material acceptance, material rejection, and pay determination for the air void and road way density. The upper and lower specification limits (USL and LSL) for V_a are 5.00% and 3.00%, respectively. For density, KDOT uses one-sided specification, LSL of 91% for thickness $\leq 2''$ and 92% for thickness $> 2''$. Pay factors for the air void and in-place density are given by

Table 1.2. Disincentive and incentive payment adjustment for HMA pavement density for Wisconsin DOT.

Disincentive pay reduction for HMA pavement density	
Percent lot density below specified minimum	Payment factor (percent of contract price)
From 0.5 to 1.0 inclusive	98
From 1.1 to 1.5 inclusive	95
From 1.6 to 2.0 inclusive	91
From 2.1 to 2.5 inclusive	85
From 2.6 to 3.0 inclusive	70
More than 3.0	Remove and replace or remain in-place with a 50 percent payment factor
Incentive payment adjustment for HMA pavement density	
Percent lot density above specified minimum	Pay adjustment per ton
From -0.4 to 1.0 inclusive	$0
From 1.1 to 1.8 inclusive	$0.40
More than 1.8	$0.80

$$P_{V_a} = 0.003 \times PWL_{V_a} - 0.270 \tag{1.8}$$

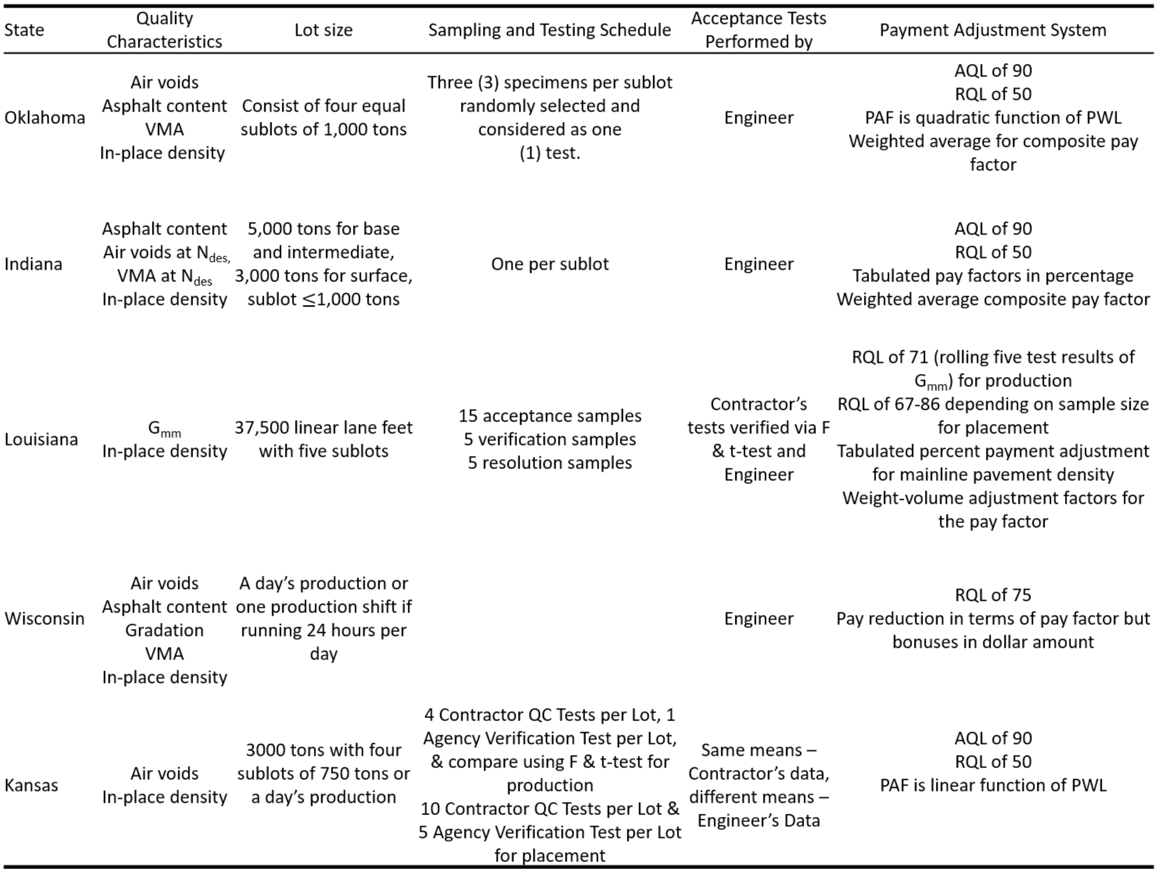
and

$$P_{compaction} = 0.004 \times PWL_{compaction} - 0.360, \tag{1.9}$$

where: P_{V_a} = payment adjustment factor for air void, PWL_{V_a} = estimated PWL for air voids, $P_{compaction}$ = payment adjustment factor for in-place density, and $PWL_{compaction}$ = estimated PWL for in-place density. A 90 PWL is required for a full pay, whereas a 50 PWL corresponds to “remove and replace” condition.

1.8 STRATEGIES AND SUGGESTIONS BY OTHER STATES

This section summarizes challenges, strategies, and suggestions by different state DOTs in implementing PWL as payment adjustment method for production and placement of HMA.



State	Quality Characteristics	Lot size	Sampling and Testing Schedule	Acceptance Tests Performed by	Payment Adjustment System
Oklahoma	Air voids Asphalt content VMA In-place density	Consist of four equal sublots of 1,000 tons	Three (3) specimens per sublot randomly selected and considered as one (1) test.	Engineer	AQL of 90 RQL of 50 PAF is quadratic function of PWL Weighted average for composite pay factor
Indiana	Asphalt content Air voids at N_{des}, VMA at N_{des} In-place density	5,000 tons for base and intermediate, 3,000 tons for surface, sublot ≤1,000 tons	One per sublot	Engineer	AQL of 90 RQL of 50 Tabulated pay factors in percentage Weighted average composite pay factor
Louisiana	G_{mm} In-place density	37,500 linear lane feet with five sublots	15 acceptance samples 5 verification samples 5 resolution samples	Contractor's tests verified via F & t-test and Engineer	RQL of 71 (rolling five test results of G_{mm}) for production RQL of 67-86 depending on sample size for placement Tabulated percent payment adjustment for mainline pavement density Weight-volume adjustment factors for the pay factor
Wisconsin	Air voids Asphalt content Gradation VMA In-place density	A day's production or one production shift if running 24 hours per day		Engineer	RQL of 75 Pay reduction in terms of pay factor but bonuses in dollar amount
Kansas	Air voids In-place density	3000 tons with four sublots of 750 tons or a day's production	4 Contractor QC Tests per Lot, 1 Agency Verification Test per Lot, & compare using F & t-test for production 10 Contractor QC Tests per Lot & 5 Agency Verification Test per Lot for placement	Same means – Contractor's data, different means – Engineer's Data	AQL of 90 RQL of 50 PAF is linear function of PWL

Figure 1.7. Summary of PWL approaches adopted by different states.

1.8.1 Wisconsin

Wisconsin DOT, one of the more recent state agencies to implement the PWL, was contacted to learn about some of the challenges the department faced and the strategies the department embraced to overcome those challenges. The department followed a step wise slow process to implement the method. A pilot program was first initiated where the neutral zone was broadened to provide leeway to the Contractors so that they do not get penalized easily. The department aimed to lead two projects in each of the five regions of Wisconsin. However, there were only two projects in the first year. By the second year, there were around ten or eleven projects in all five regions. Since the concept was new, there was no pour for three to five years. Multiple initiatives were taken to encourage Contractors to pour. The department set up a test strip with 1000 tons to help Contractors establish compaction pattern. Also, the department made QC test data a form of payment adjustment basis as long as the QC test results pass the F test and t-test verification with respect to the QA test results. This is done because it is believed that the QC data are more inclusive and representative of the entire project than the QA test results as QC sampling and testing are done more frequently covering wide numbers of sublots. It has been reported that PWL resulted in more pavements with "replace and remove" condition than the previous method. The reasons behind for obtaining more "replace and remove" condition were that the Contractors did not take the new change seriously and did not make any adjustment to cope with the new modifications. It has also been reported that the variability in the production characteristics has been dramatically reduced. For example, the variation in Gmb has been reduced to seven to eight thousandth compared to that of 0.07 to 0.08 from the previous system. One of the claimed challenges from Contractor was that high-ESAL mixes were difficult to achieve density. This claim was mostly driven by the previously low target (without correlating gauges to cores, as gauges have a tendency to read relatively high on dense-graded mixes). In order to handle this problem, the department advised them to improve their construction practices that would help them achieve the required density. Such improvements were to decrease spacing between rollers, include additional rollers, and set initial/breakdown rollers closer to the paver, etc. Also to reduce variability, a valuable lesson that has been learned in the process, is to improve the test requirements, reduce sampling variability, and utilize better test technique and sound test equipment. The department also acknowledged that "PWL requires Contractors to gravitate toward best practices, at which point, reasonable targets are achievable! This was another instance of the steep learning curve here in Wisconsin, but also a positive outcome of enforcing PWL."

1.8.2 Oklahoma

To evaluate the consistency and suitability of the PWL specifications, to reward the Contractors properly, and to enhance the relationship between the ODOT and its Contractors, the department sponsored an evaluation study [4], where two HMA paving jobs were led for construction. In addition to the normal quality control, acceptance, and assurance sampling and testing, the ODOT Materials Division performed more extensive sampling and testing on randomly selected lots from each pavement which were designated "super lots." In each HMA super lot, each of five sublots were sampled and tested three times so that there were a total of 15 additional tests. The Contractor performed his regular specified sampling and testing for each super lot. Sampling and testing consisted of the same tests prescribed by the proposed PWL specifications. The study showed that the proposed PWL specification was sound and required very little additional work to transform from the existing method. A partial comparison of the ODOT PWL to that adopted by Kansas and Missouri showed that ODOT PWL resulted in a pay reduction than the other two states. The study also identified the necessity of establishing the acceptance limits for specifications: too tight specification may deprive the Contractor of a reasonable opportunity to meet the specifications whereas too lose may prove ineffective in controlling quality. This study also noted that aggregate gradation had the least effect in dictating the pavement performance and suggested using mix volumetric properties, such as VMA instead of the aggregate gradation as payment adjustment criteria. The report recommended that the risk analysis should be performed using advanced simulation schemes capable of fully evaluating the risk associated with the complex pay factor equation and multiple acceptance properties.

1.9 SUMMARY

The current payment adjustment factors for production and placement are based on average of the absolute deviation from the laboratory-molded density or the average of in-place air voids. These averages are obtained using measurements made upto four sublots for each lot of production. Due to this averaging technique, often Contractors receive bonuses even with high variability in the quality characteristics. Therefore, a robust payment adjustment factor scheme is required where the payment adjustment is not solely based on the average test results for production and placement but is based on a measure of quality that accounts for both the average and the variability in production and placement.

A review of the existing literature and practice in other states and recommended practice by the Federal Highway Administration (FHWA) revealed that this can be achieved using a metric that is referred to as the Percent Within Limits or PWL. In simple terms, PWL provides an estimate for the probability that a test result from a lot falls within the upper and lower specification limits. The PWL is calculated using the upper specification limit, the lower specification limit, the sample mean, the sample standard deviation, and an appropriate probability distribution table. A minimum value of this probability or PWL, referred to as the acceptable quality level or AQL, is typically prescribed as a part of the specification for the Contractor to receive full payment. A payment adjustment factor can be developed to award a bonus for exceeding this AQL and a penalty for falling below this AQL value. A rejectable quality level or RQL can also be prescribed as the lower threshold below which the Contractor will not receive any payment or may be required to remove or replace the work.

CHAPTER 2. DATA EXTRACTION AND PRELIMINARY ANALYSIS

This chapter summarizes the work done to extract cost, quality control, and quality assurance data from the available TxDOT databases. A comprehensive database comprising the necessary fields was generated that covered all projects from all districts across the state. This database was then used to develop a robust payment adjustment model and to ensure consistency in both the HMA [1] production and placement. The work related to the payment adjustment model is described in more detail in a subsequent chapter. The following sections briefly describe the existing TxDOT databases that have been used to extract the data used in this study. Finally, several preliminary results are presented to demonstrate the nature of the problem that was being addressed in this study.

2.1 OVERVIEW

TxDOT stores information regarding the design, quality of production and placement of materials, and pavement performance in three different database systems:

- Sitemanager (SMGR)
- Design and Construction Information System (DCIS)
- Pavement Management Information System (PMIS)

Brief descriptions of these database systems are provided below.

2.1.1 SiteManager (SMGR)

The TxDOT SiteManager database is a comprehensive record of material, design, and construction quality control and quality assurance (QC/QA) information collected on all contracted pavement projects constructed by TxDOT. This database is composed primarily of the HMA information collected from projects that include specification Items 340/1 dense-graded, 342 permeable friction course (PFC), 344 performance designed (e.g., Superpave), 346 stone matrix asphalt (SMA), 347 thin overlay mixes (TOM), and 348 thin bonded wearing courses, and those constructed under special specifications including crack attenuating mixes (CAM). The database traditionally comprised HMA items per the TxDOT 2004 book [10] and was recently updated to include 2014 specification items [11]. A distinction

---

[1] Note: The term HMA is used in this report for brevity but it also includes asphalt mixes produced as Warm Mix Asphalt (WMA).

is made between mix design information and QC/QA data collected during construction. These information types are collected from separate templates and, indeed, it is possible that the same mix design could be used on multiple different projects. The database includes a referencing system that allows linking of different database tables within SiteManager. It is this referencing system that can be used to extract material and design information for an HMA project and ultimately link it to performance information.

2.1.1.1 Mix design

Pertinent HMA material and mix design information in SiteManager is updated from mix design templates, previously for 2004 (TX2MIXDE4), now for 2014 (TX2MIXDE14) specification items. The mix design template is essentially an Excel-based spreadsheet application that is completed by the HMA Contractor that must be approved by TxDOT. It serves to establish the job-mix formula for the HMA project in terms of design gradation and asphalt binder content.

2.1.1.2 Quality control and quality assurance (QC/QA)

QC/QA data collected during construction on HMA projects are input into a separate Excel template, previously TX2QCQA04 for 2004 specification items and currently TX2QCQA14 for 2014 specification items. Data in this template are collected by both TxDOT and Contractor on a sub-lot basis. Each template represents data collected on a lot comprising upto four sub-lots. The aggregate-related properties collected during QC/QA are primarily the gradation of the mix and calculated deviations from the design target. Besides master gradation bands, TxDOT specifications also apply tolerance limits that impose restrictions on allowable variations in aggregate gradation from a target during production. Tolerance restrictions between TxDOT and Contractor are also in place. Non-conformance results in suspension of production until corrected.

The QC/QA templates are used to calculate the Contractor's payment for the lot, which for some specification items is multiplied by a payment adjustment factor. This payment adjustment factor is a function of (i) the production quality assessed on the basis of laboratory-molded density of the mix and (ii) the placement quality assessed on the basis of the air void content of the field samples. Contractors can receive a bonus of up to 5% for Item 341, 7.5% for Item 344, and 10% for Item 346 for high quality construction based on the average values of the aforementioned metrics. The question remains, however, whether this payment system translates to consistent production and placement quality and/or whether Contractors with inconsistent quality can still receive bonuses because of the way the metrics are averaged over sub-lots.

2.1.2 Pavement Management Information System (PMIS)

In order to monitor the condition of roads in Texas, TxDOT annually collects distress data on the entire roadway network. The data for HMA or flexible pavements include annual automated and visual assessments of rutting, cracking, and roughness. The PMIS database also includes friction measurements covering the network on a biennial basis. This is a comprehensive database with historical data of roadway performance. In addition to performance measures, PMIS also provides a breakdown of the traffic on the road sections for which performance is reported in terms of average annual daily traffic and 18-kip axles, which is a surrogate measure of all of the different axle loads. PMIS reports performance measures along the roadway in roughly 0.5-mile segments defining data collection sections. The location of these PMIS sections on the TxDOT network is identified by reference marker or specifically Texas reference marker. Every performance measure in PMIS is accompanied with a beginning and ending reference marker and associated displacements from these reference markers. Data are further reported for each roadbed along the network. For undivided roadways this could be a single measure every 0.5 miles; in the case of divided roadways, it could be multiple measures to cover the left and right main-lanes and left and right frontage roads.

2.1.3 Design and Construction Information System (DCIS)

The DCIS database provides a listing of all contracted projects constructed by TxDOT. This is an administration tool used to track the progress and payments made on projects and to track change orders. It also provides location information that can be used to link SiteManager data with performance information in the DCIS. All contracted projects in DCIS are given a unique identifier, called the control-section-job or CSJ number. This CSJ number is also used to identify projects in SiteManager. DCIS furthermore reports the reference marker extents of projects that allows locating the extents of the project within the PMIS database. A major shortcoming of DCIS is that it does not always indicate the roadbed on which the project is located and in the past it did not always provide Texas reference marker limits, which then had to be manually identified based on textual descriptions of the “from” and “to” limits of a project. While the latter issue has been addressed, not knowing the roadbed of a construction project negates the link between SiteManager and PMIS. DCIS only manages contracted projects. In-house projects undertaken by TxDOT crews are managed with the Compass database system (previously the Maintenance Management Information System).

2.2 DATA MINING GOALS

TxDOT regularly collects and stores the above mentioned information regarding the design, construction, materials, and performance of all components of all pavements in these database systems. Querying information for one pavement type, HMA in this case, from databases that are continuously updated for multiple materials and components is inefficient and resource intensive. In order to facilitate efficient management of HMA production and placement related data for analysis, a replica database was created that was used exclusively for the purposes of this project. In addition to efficiency of performing large data intensive queries, this dedicated database was also necessary to prevent accidental loss of data from the original TxDOT systems. This replica database comprises data mainly from the SMGR database pertaining to HMA design, production, and placement. Other information, such as bid price, actual quantity placed, and change orders from the DCIS database system are also included.

2.2.1 SMGR database

The following data pertaining to HMA were imported from the SMGR database and replicated in a separate location for each specification Item (341/344/346) from the 2004 and 2014 specification books to cover every HMA mix:

- constituent properties, e.g., aggregate gradation, binder, and additives;
- volumetric properties, including percent density, air voids, asphalt content (AC), and voids in mineral aggregates (VMA);
- laboratory performance test data in terms of Hamburg Wheel Tracking (HWT) and indirect Tensile Strength (ITS); and
- payment information in terms of payment adjustment factors, quantities, and unit costs.

2.2.2 DCIS databde

Similar to SMGR data, bid quantities, change order quantities, and actual completion dates of the specification Items from a construction project were queried. Results obtained from the above queries were linked to the SMGR data for corresponding projects via the control-section-job identifier. This allowed access to the actual quantity of the mix that was placed and the total cost of the different mix types with different HMA constituent and volumetric properties across the entire state from different specification Items.

2.3 PRELIMINARY EXAMINATION OF DATA

Results obtained from preliminary data analysis are summarized in this section. These results include specification Items 341, 344, and 346. This section contains three subsections based on the mix properties and design parameters controlling quality of the production and placement of HMA.

2.3.1 Mix types and characteristics

After data mining, the first step was to examine the general trends of different mixes used in Items 341, 344, and 346. For example, Items 341, 344, and 346 each cover several different mix types. Figures 2.1 through 2.3 illustrate the different mix design types under each Item along with the number of mixes that were placed for each mix type. For Item 341, fine base (B), coarse surface (C) and fine surface (D) mixes were the most commonly produced mixes. More than 10,000 lots were produced with each of these mixes. Under Item 344 the most commonly used mixes were surface (C) and fine mixes (D) accounting for nearly 8000 of 9686 lots. Similarly for Item 346, about 2,000 lots were produced with medium D mix. whereas coarse (C) and fine (F) accounted for 300 and 200 lots, respectively. Another 200 lots were produced with fine mixes containing crumb rubbers.

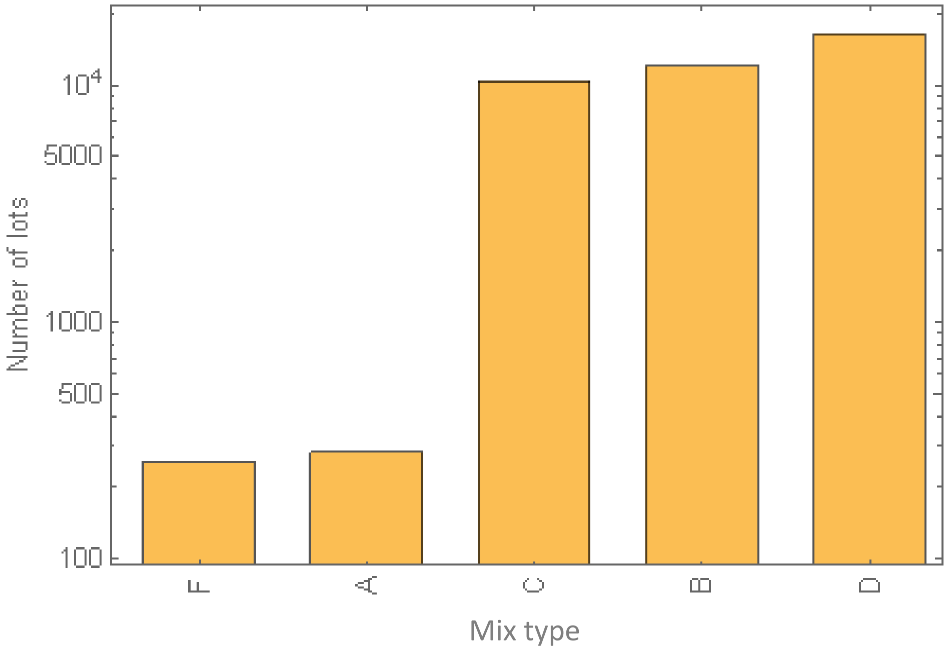


Figure 2.1. Distribution of mix types (A - base, B - fine base, C - coarse surface, D - fine surface, and F - fine mixes) for Item 341.

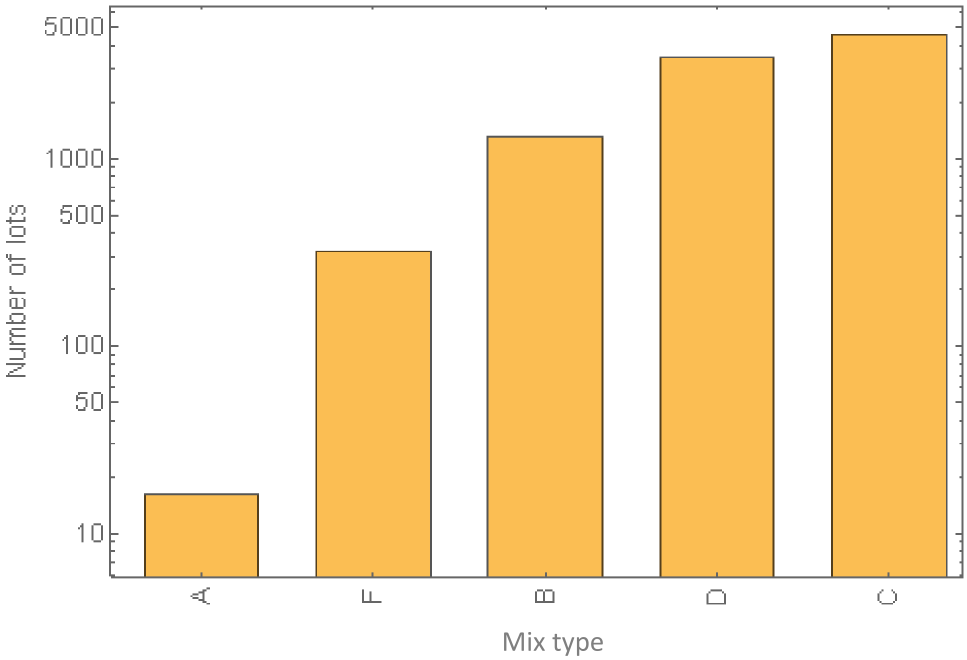


Figure 2.2. Distribution of mix types (A - base, B - intermediate, C - surface, D - fine mixes, and F - fine surface (2004 specification only)) for Item 344.

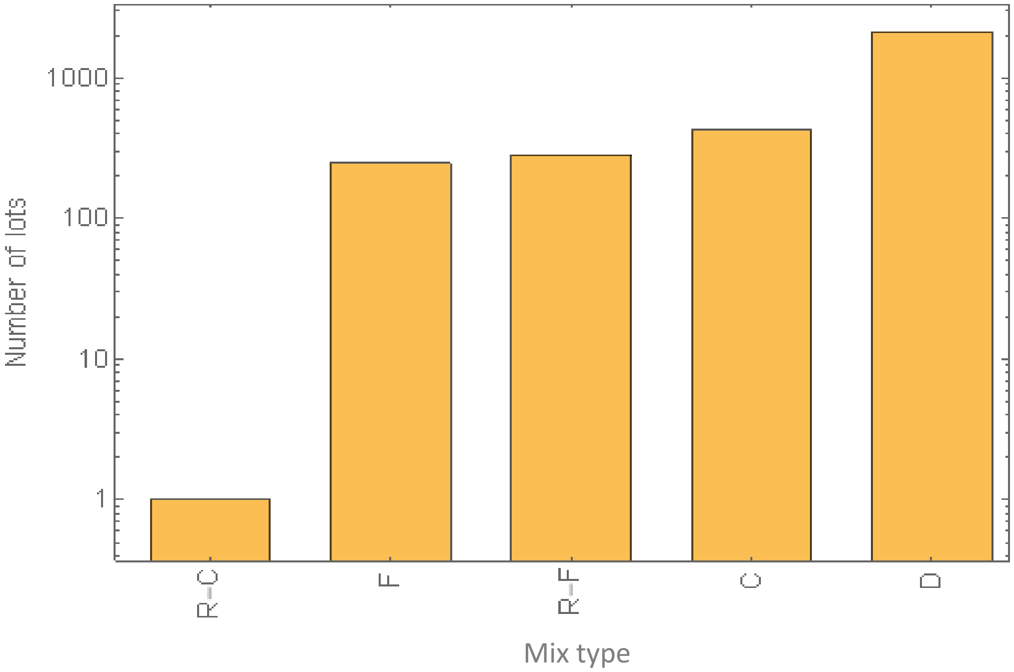


Figure 2.3. Distribution of mix types (C - coarse, D - medium, and F - fine, R-C - coarse with crumb rubber, R-F - fine with crumb rubber mixes) for Item 346.

In addition to the type of mix, it is also of interest to examine the general characteristics of the mix such as the binder content, recycled binder ratio, performance in the Hamburg Wheel tracking device (HWTD), and performance with respect to the indirect tensile strength (IDT) test. These characteristics were examined for all mixes pooled together for each Item. The correlations between these factors will be analyzed in a subsequent chapter.

Figures 2.4 through 2.6 show the typical distributions for the asphalt content of all the mixes for each Item. Based on the aggregated data it appears that for Item 341, about 20,000 lots were produced with a JMF AC between 4 and 6% and about 1,500 had asphalt content between 3 - 4 and 6 - 7%. Around 72% lots were produced with an asphalt content between 5%-6% for Item 344. Among 3,098 lots compacted with stone matrix asphalt (SMA) (Item 346), about 2,500 were produced with JMF AC that falls between 6 and 7%. Rest of the lots contained AC between 7.5 and 8.75%, except a few with AC lower than 6%. Given the emphasis on recycling, it is also important to examine the recycled binder ratio used in asphalt mixes. Figures 2.7 through 2.9 show the typical distributions for the recycled binder ratio. These data show that, around 6500 lots were allowed to produce with 20-30% and around 5500 lots were produced with 30-40% maximum recycled binder ratio. These values were much less for Items 344 and 346, as expected. About three fourth (7,500 out of 9,679) of Superpave lots (item 344) had some recycled binder, most of which was limited to a maximum recycled binder of 20%. For Item 346, approximately one third of the produced lots included a maximum of 15% recycled binder ratio.

In terms of their performance characteristics, the mixes also showed a broad range of rutting values and strength as recorded by the IDT. Figures 2.10 through 2.15 illustrate these distributions. As indicated earlier, the correlations of these factors with other attributes such as air void content in the field have been evaluated in more detail in the following chapter. Per the specification limits, most mixes produced were susceptible to a rut depth that falls within the range of 3 - 10 mm. Likewise, all indirect tensile strength tests appeared to be within the specification limits of 85 - 200 psi.

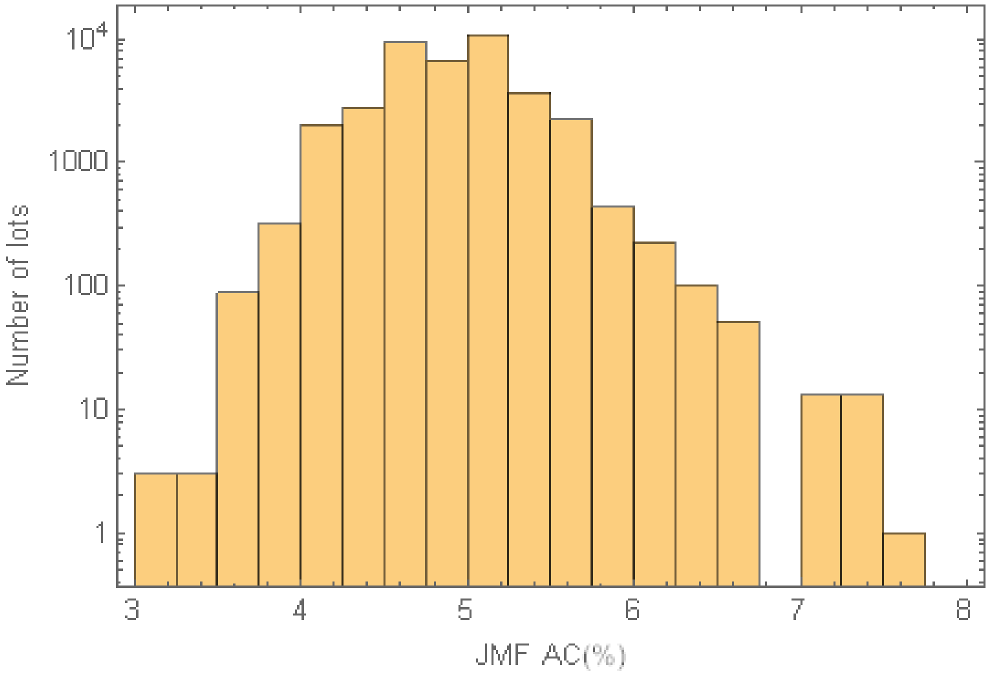


Figure 2.4. Distribution of asphalt content for Item 341.

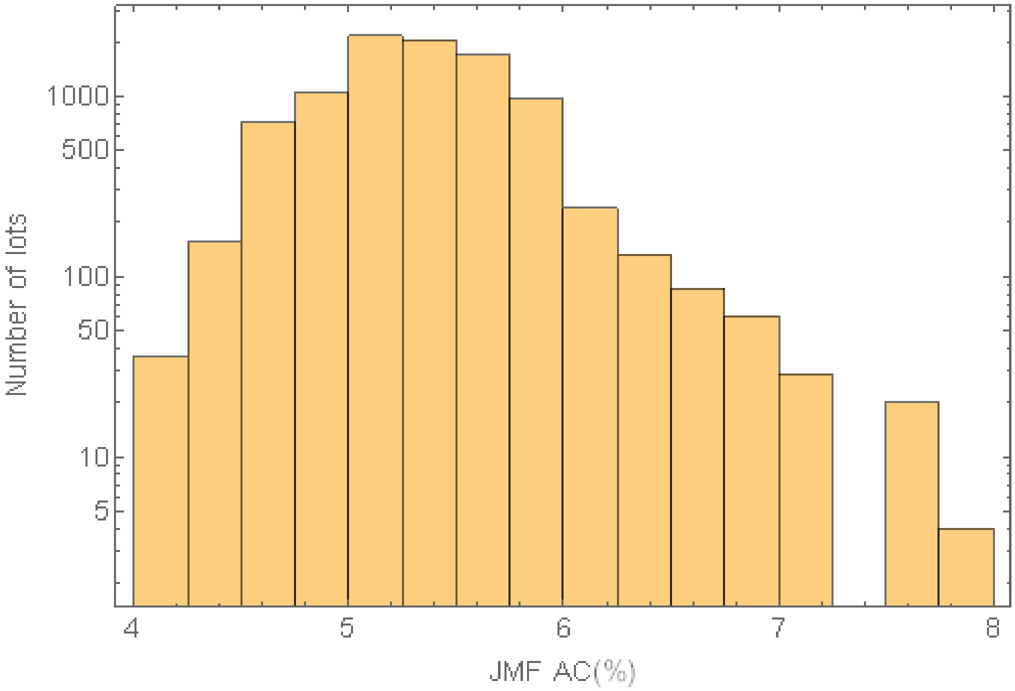


Figure 2.5. Distribution of asphalt content for Item 344.

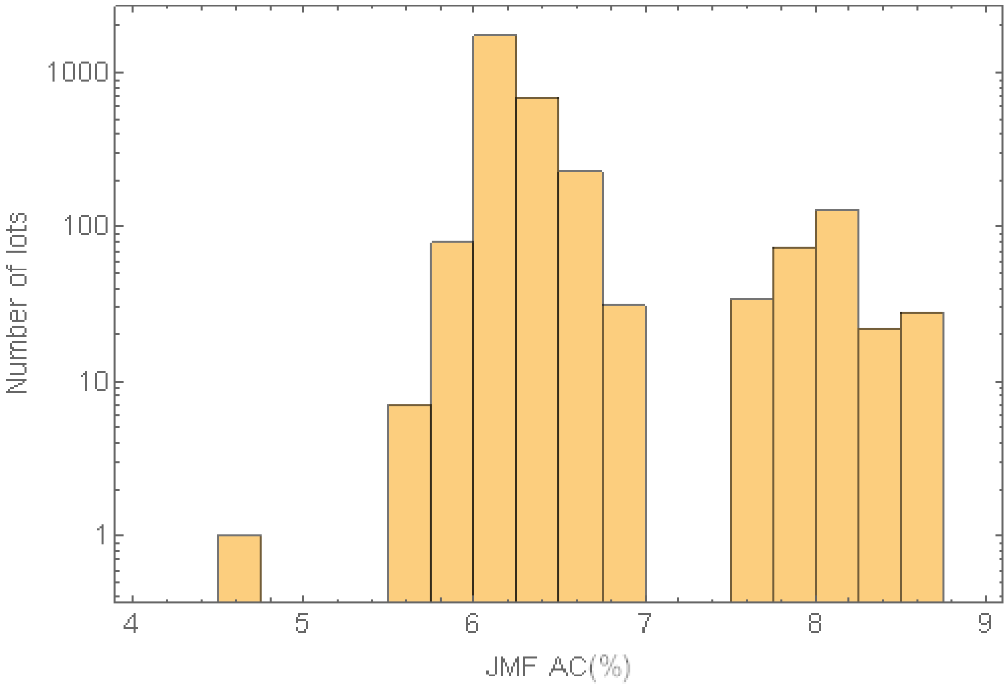


Figure 2.6. Distribution of asphalt content for Item 346.

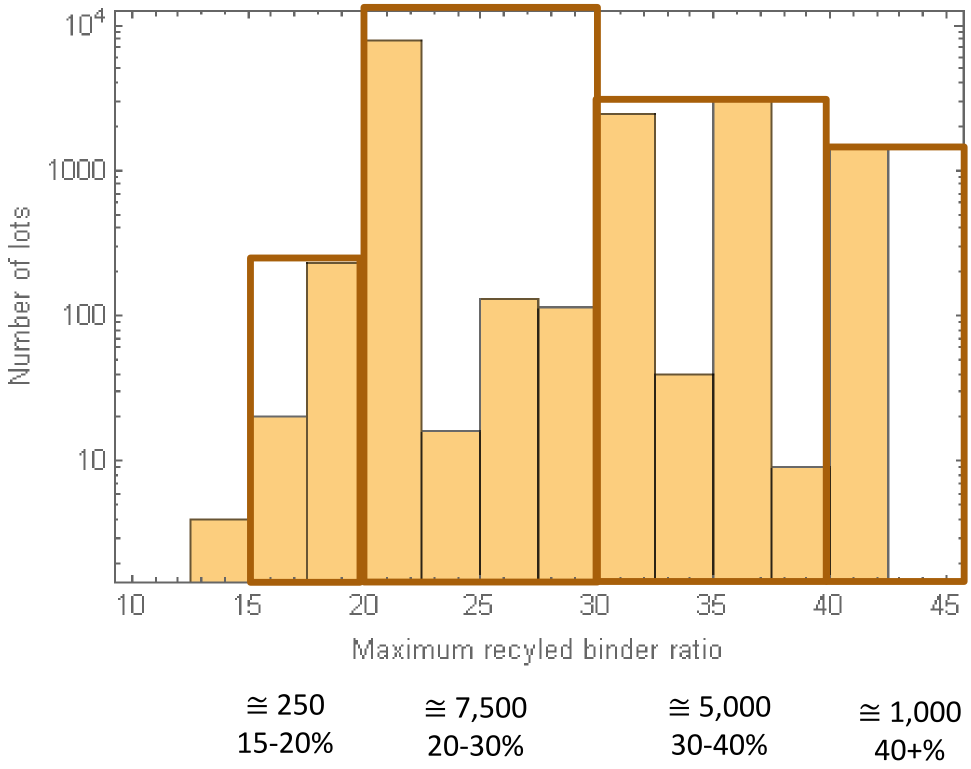


Figure 2.7. Distribution of maximum recycled binder ratio for Item 341.

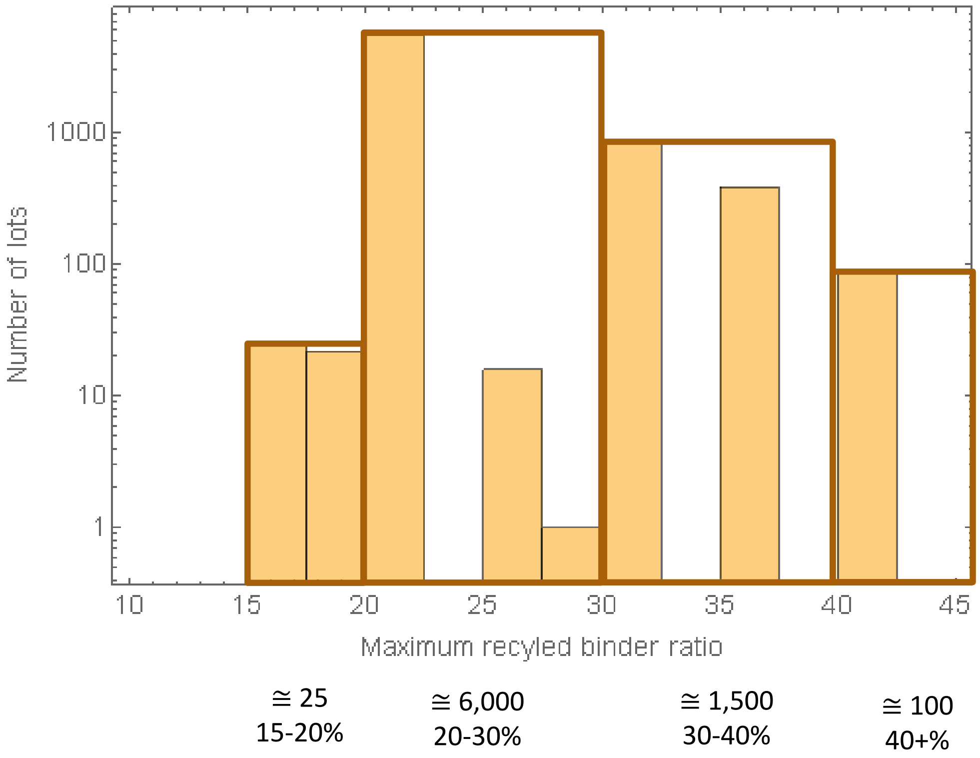
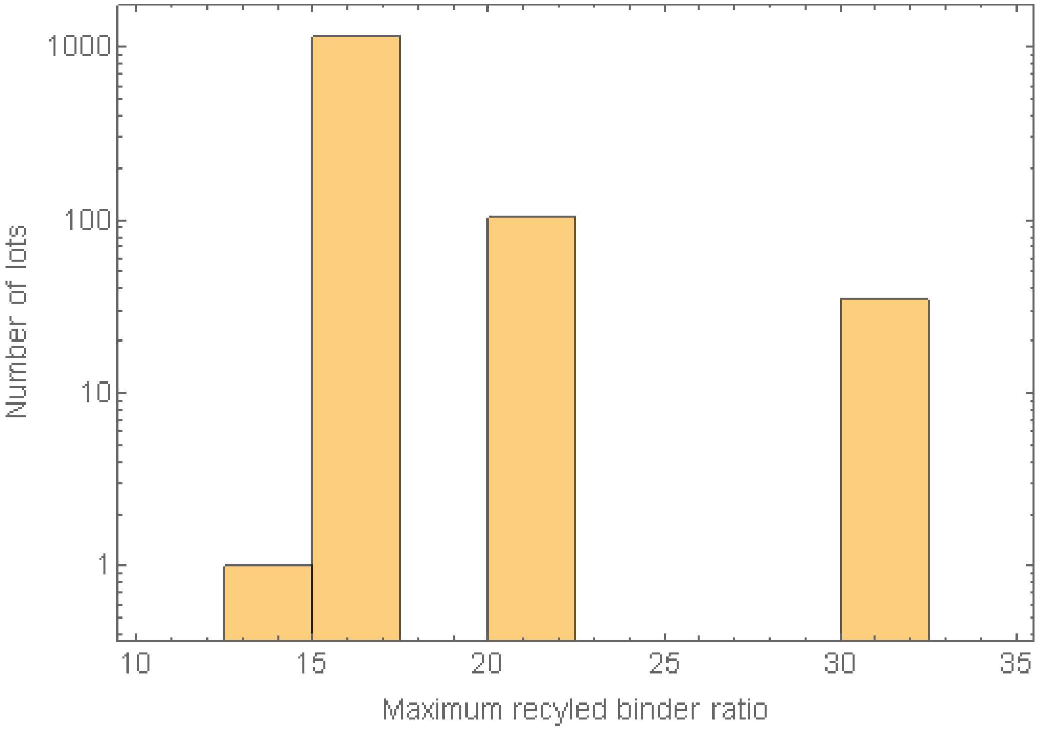


Figure 2.8. Distribution of maximum recycled binder ratio for Item 344.



≅ 1,000
15%

≅ 100
20%

≅ 30
30%

Figure 2.9. Distribution of maximum recycled binder ratio for Item 346.

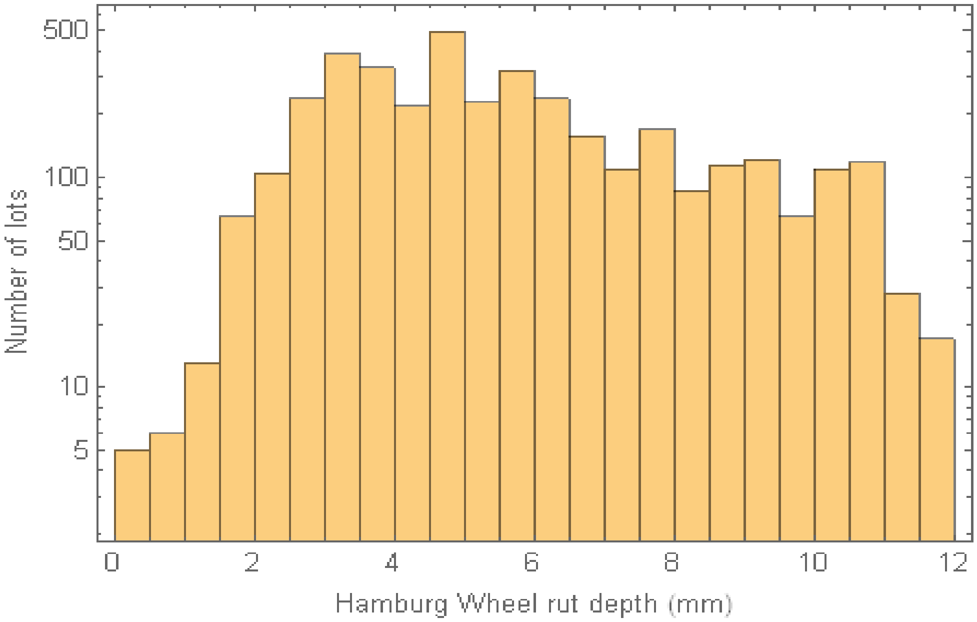


Figure 2.10. Distribution of Hamburg Wheel test results for Item 341 prduction.

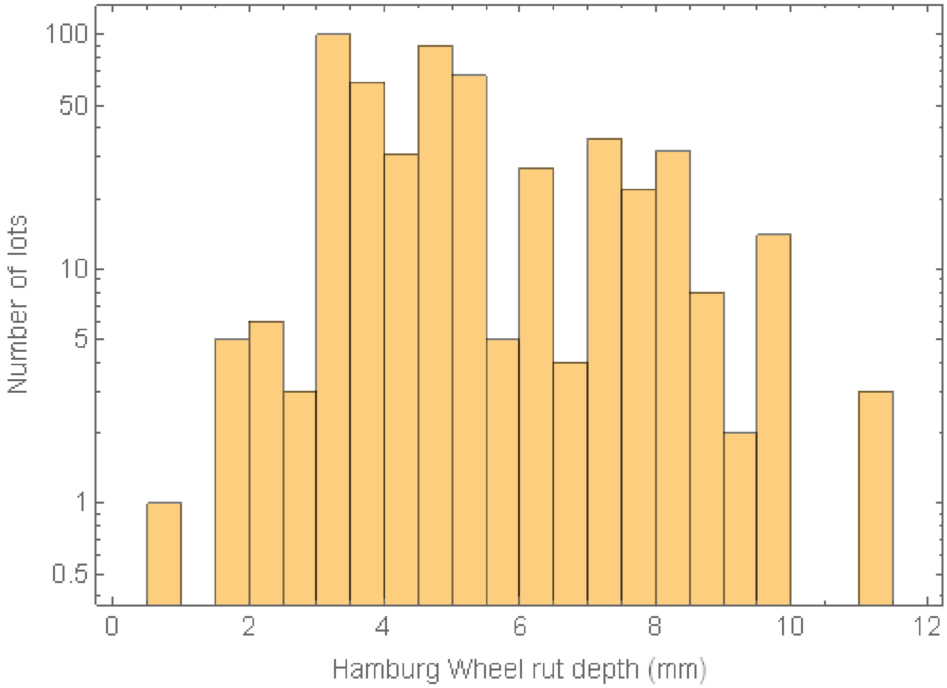


Figure 2.11. Distribution of Hamburg Wheel test results for Item 344 production.

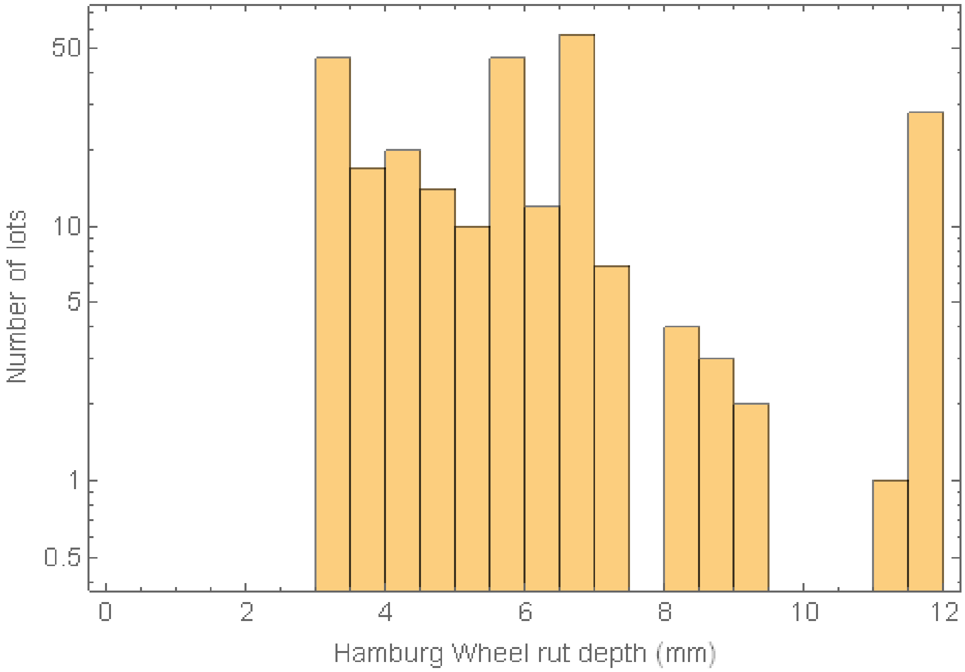


Figure 2.12. Distribution of Hamburg Wheel test results for Item 346 production.

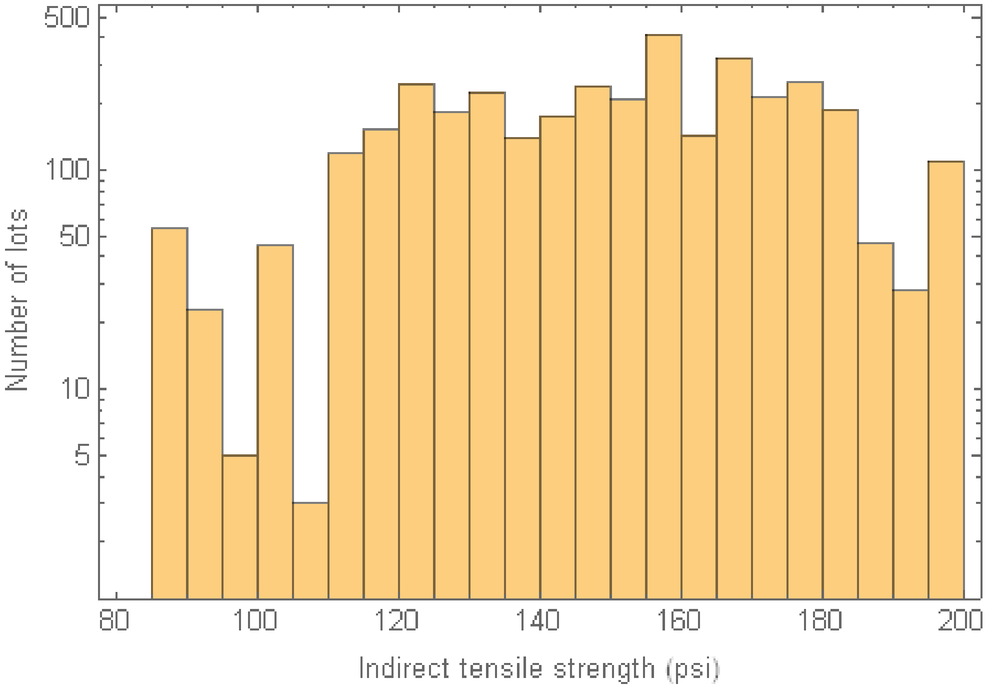


Figure 2.13. Distribution of indirect tensile strength test results for Item 341 production.

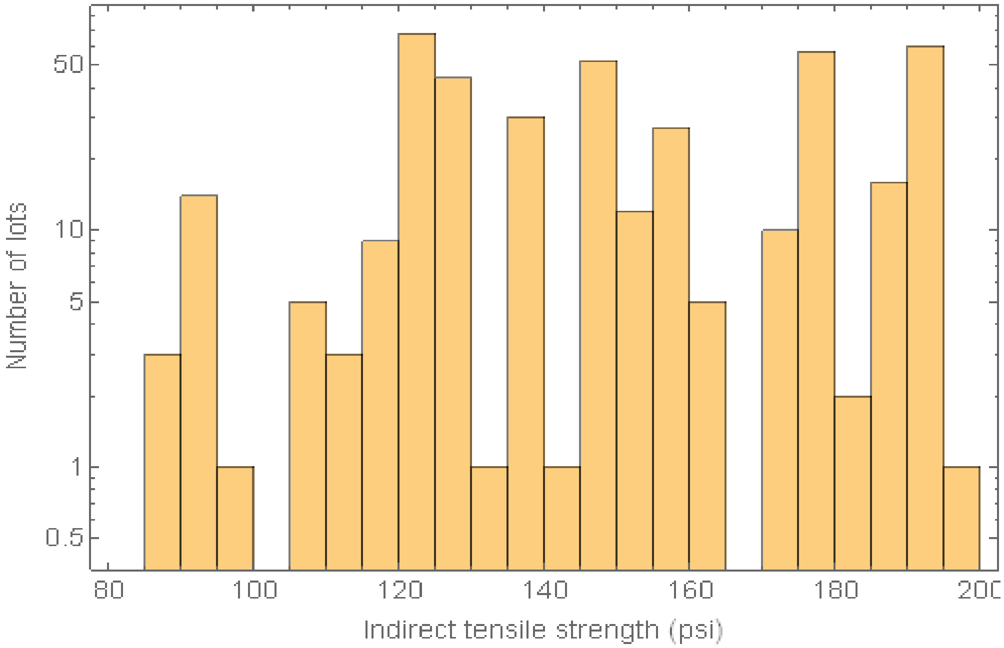


Figure 2.14. Distribution of indirect tensile strength test results for Item 344 production.

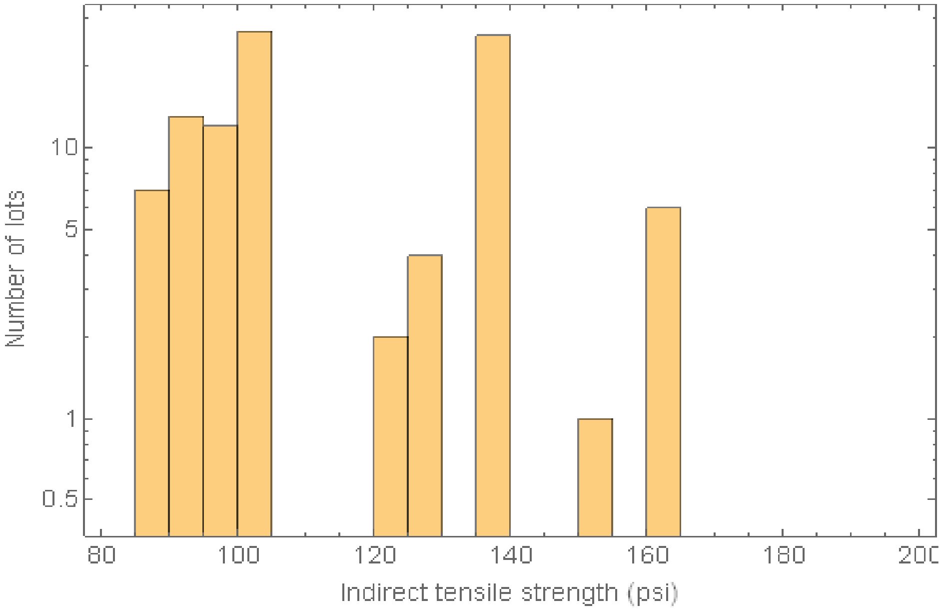


Figure 2.15. Distribution of indirect tensile strength test results for Item 346 production.

2.3.2 Laboratory-molded density and in-place air voids

The main goal of this study was to examine the current method used to calculate payment adjustment factors for production and construction. The first chapter of this report used Item 341 to illustrate two scenarios where both Contractors received a bonus or payment adjustment factor greater than 1.0 for production and placement. However, in one case there was higher variability in production and placement whereas in the other case the variability was much lower. Prior to developing alternative models for the payment adjustment factor (PAF), it is important to first examine the two crucial pieces of data that are used to compute the PAF for production and placement: laboratory-molded density and in-place air voids.

Figure 2.16 shows the distribution for the laboratory-molded density for specification Item 341. These data include the average of the laboratory-molded densities from all available sublots tested by the Engineer. Figures 2.17 shows the same distribution but using a cumulative density plot. This figure also highlights the median or 50th percentile and 90th percentile value for the laboratory-molded density. Similarly, Figures 2.18 through 2.21 show the distribution of laboratory-molded density for Items 344 and 346, and Figures 2.22 through 2.27 show these data for in-place air voids. 50% of the produced lots compacted with Item 341 had density of 96.8%. For Items 344 and 346, this value is about 96.1%.

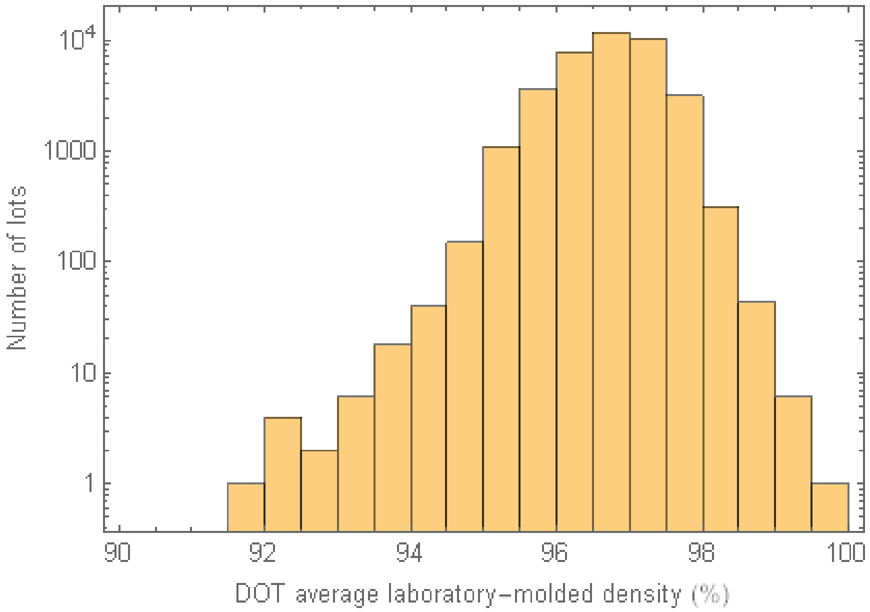


Figure 2.16. Distribution of laboratory-molded density for Item 341. Data include the average of density values from all available sublots and tested by the Engineer.

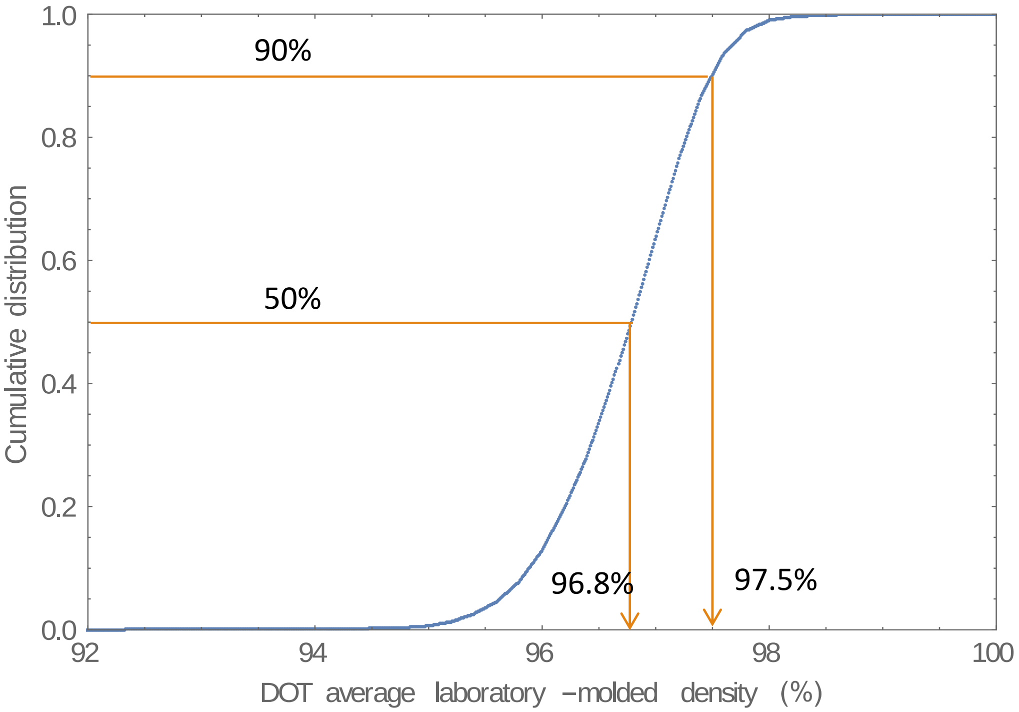


Figure 2.17. Cumulative distribution of DOT tested average laboratory-molded density for Item 341.

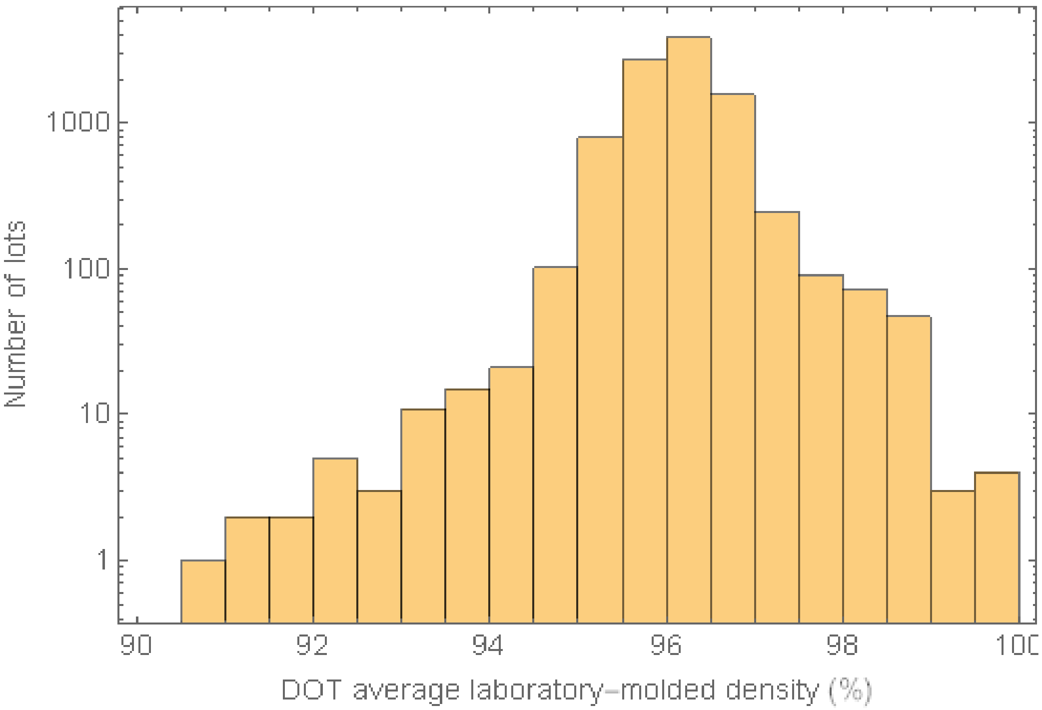


Figure 2.18. Distribution of DOT tested average laboratory-molded density for Item 344. Figure 13 from TM3 for 344.

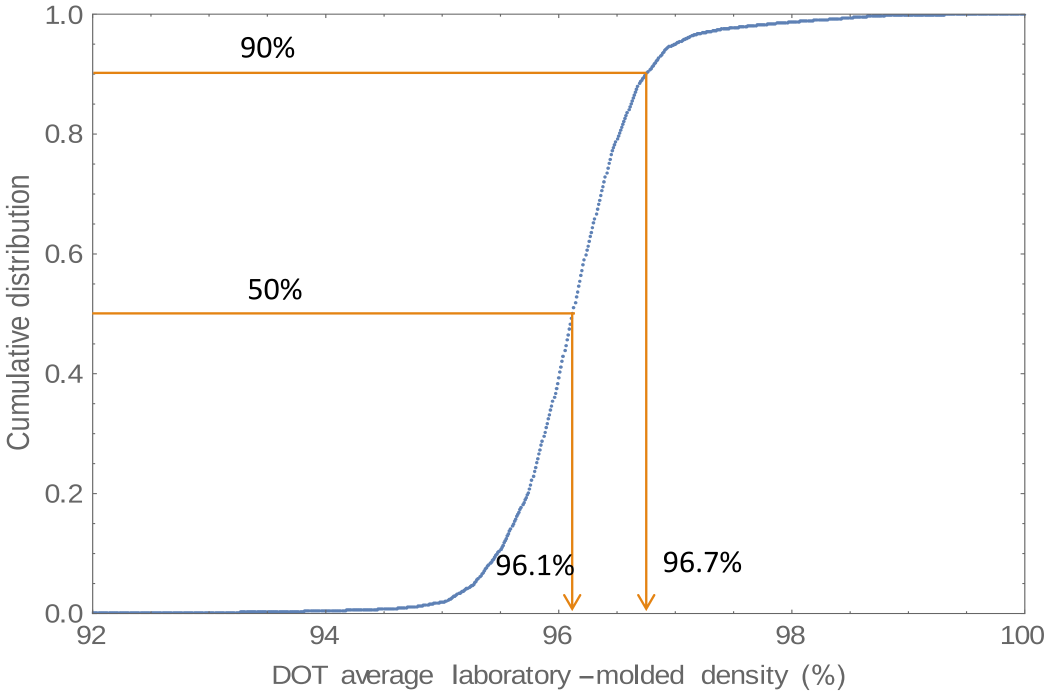


Figure 2.19. Cumulative distribution of DOT tested average laboratory-molded density for Item 344.

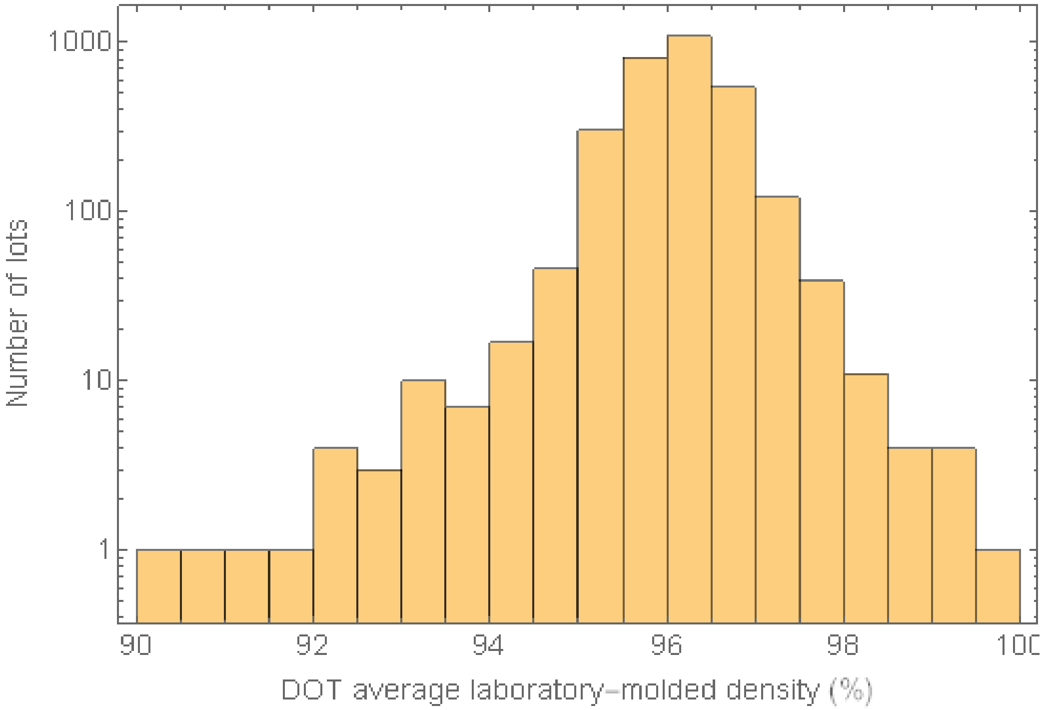


Figure 2.20. Distribution of DOT tested average laboratory-molded density for Item 346.

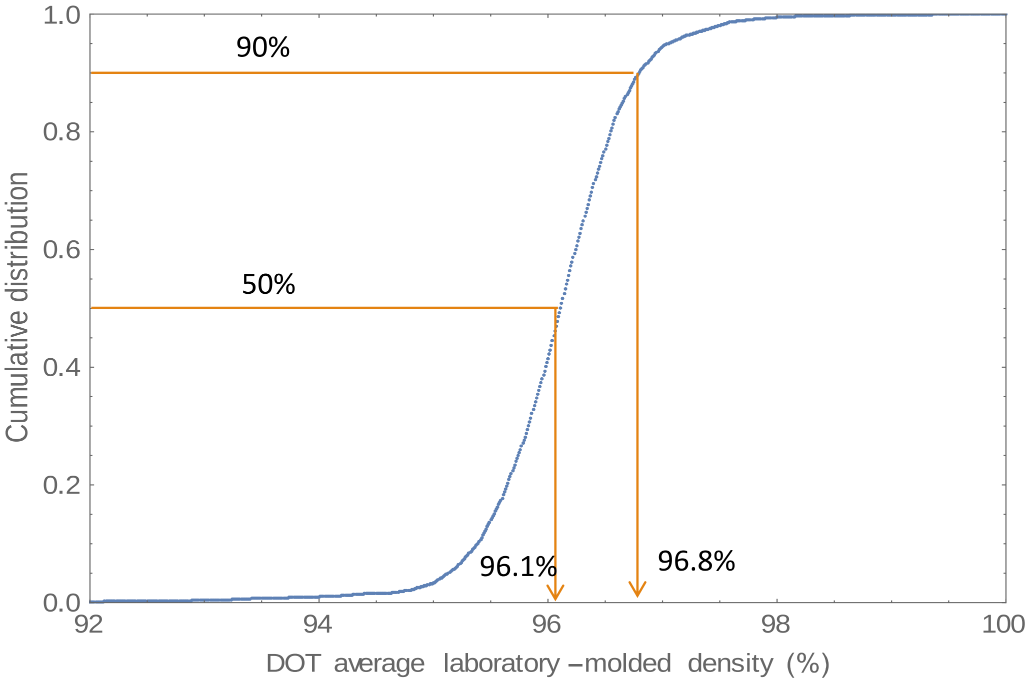


Figure 2.21. Cumulative distribution of DOT tested average laboratory-molded density for Item 346.

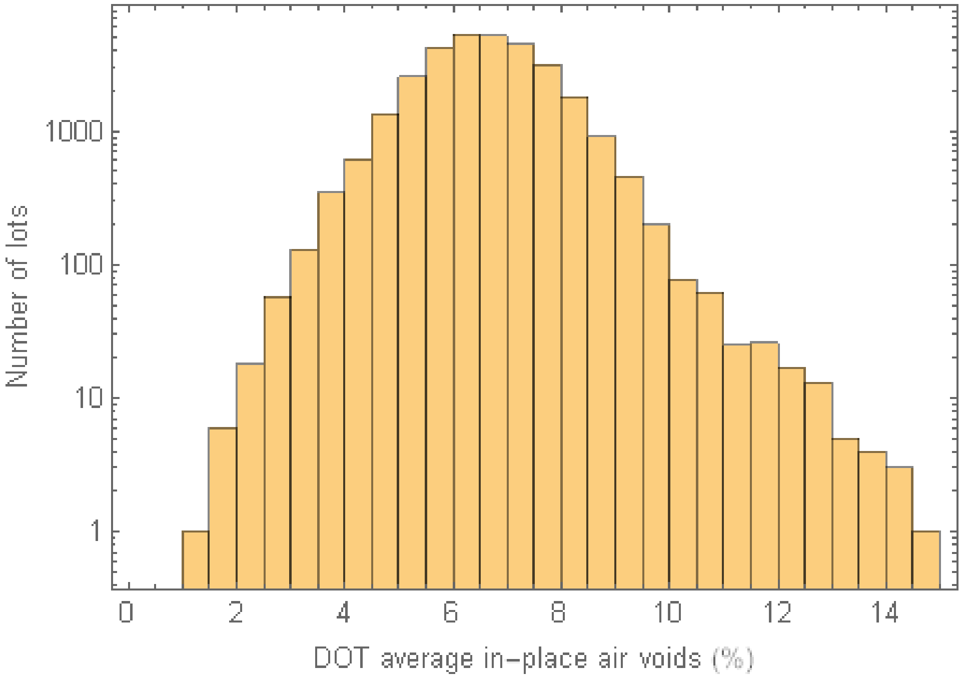


Figure 2.22. Distribution of DOT tested average in-place air voids for Item 341.

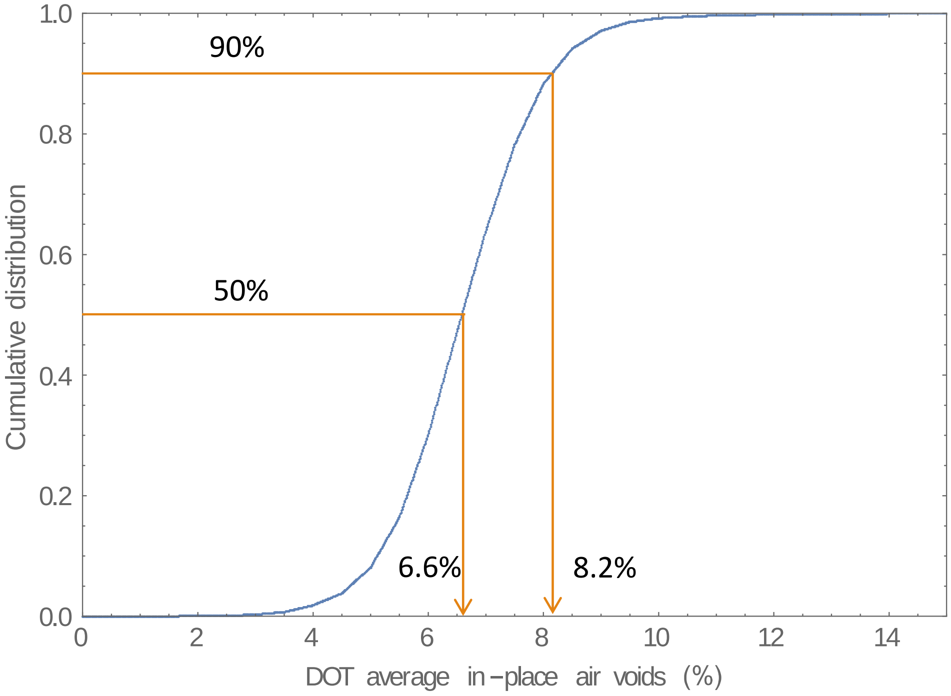


Figure 2.23. Cumulative distribution of DOT tested average in-place air voids for Item 341.

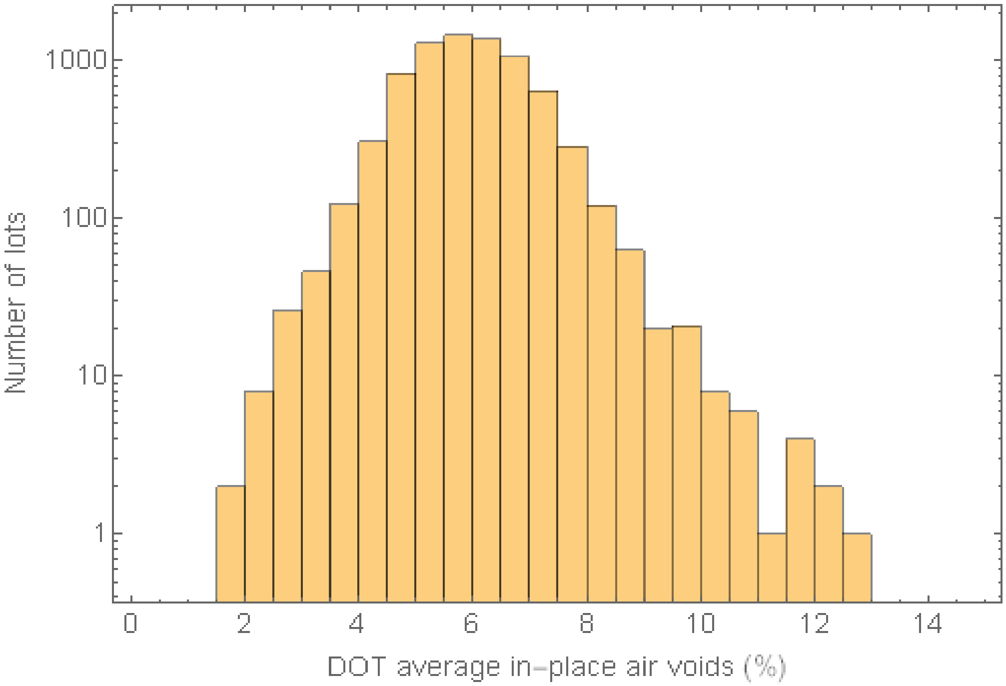


Figure 2.24. Distribution of DOT tested average in-place air voids for Item 344.

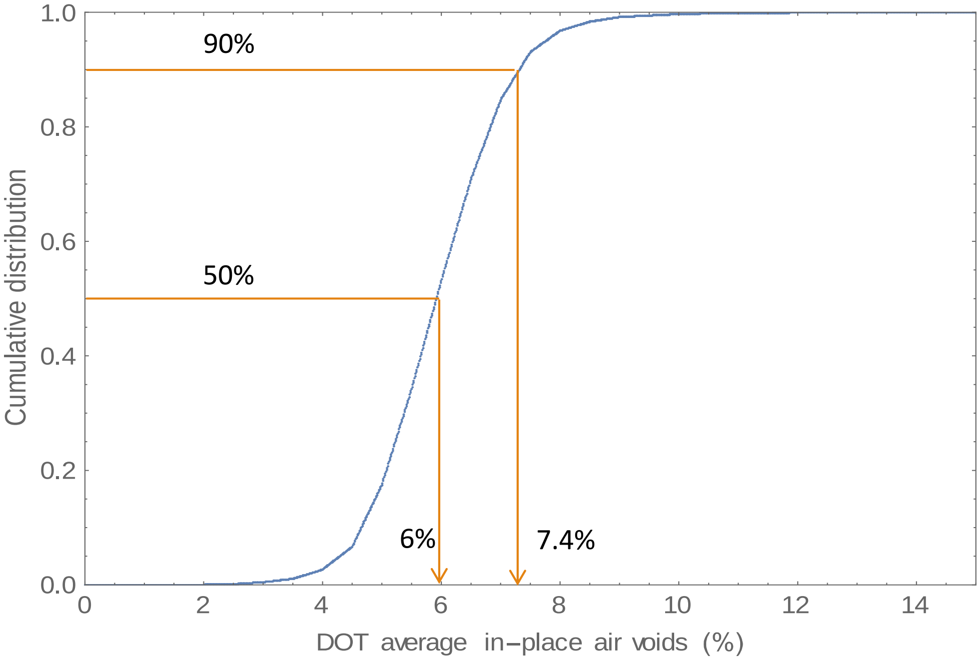


Figure 2.25. Cumulative distribution of DOT tested average in-place air voids for Item 344.

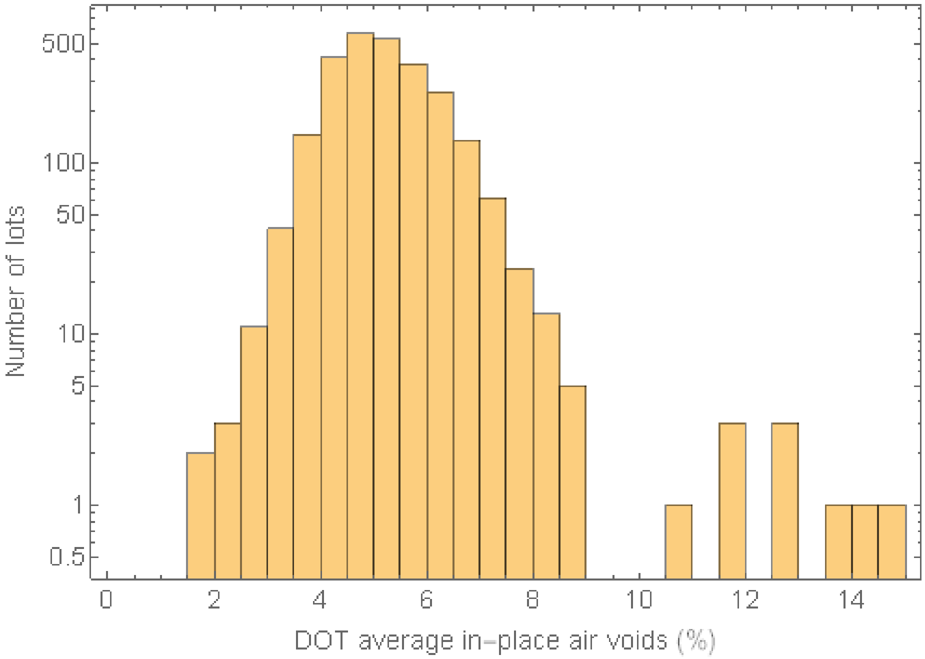


Figure 2.26. Distribution of DOT tested average in-place air voids for Item 346.

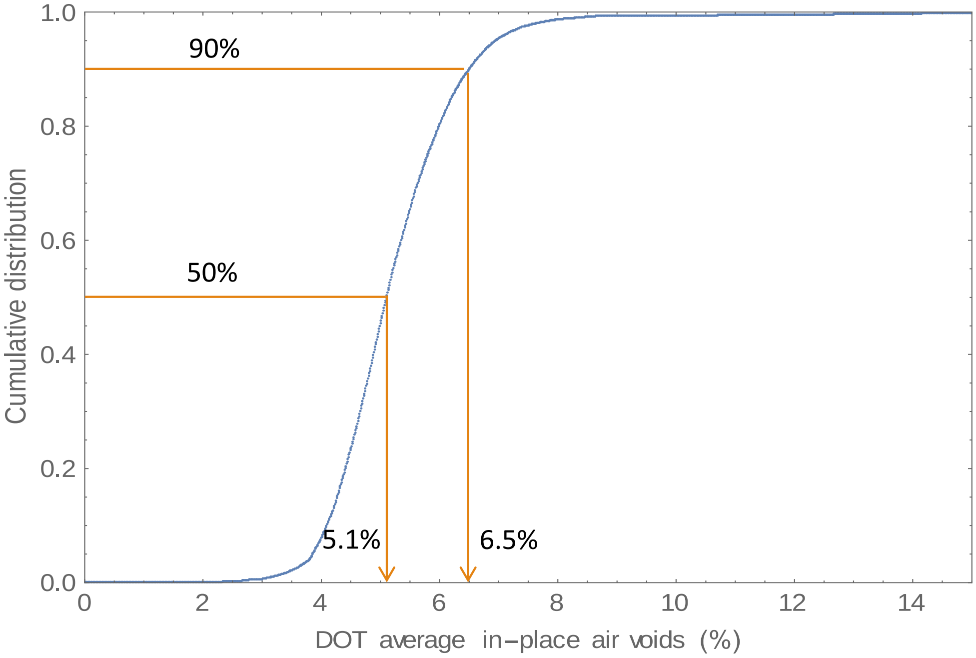


Figure 2.27. Cumulative distribution of DOT tested average in-place air voids for Item 346.

2.3.3 Payment adjustment factors

The laboratory-molded density and in-place air voids are then used to determine the payment adjustment factors (PAF) for production and placement, respectively, as prescribed in Items 341, 344 and 346. Figures 2.28 through 2.30 show the production PAFs for Items 341, 344, and 346. Similarly, Figures 2.31 through 2.33 show the placement PAFs for Items 341, 344, and 346. Some of the key observations from these figures are as follows. It appears that the distribution is skewed to the right showing the high percentage of bonuses that were granted to the Contractors for the production and placement of all three items. This skewed trend is particularly noticeable for Item 341 where more than 90% lots qualified for full pay or more, implying either of the two following hypothesis. Either, the quality of the production and placement of these lots is very consistent, or materials with poor inconsistencies are being produced i.e. the current specifications are not adequate enough to detect inconsistent production and placement of HMA. To verify the later, further analyses are performed, and the findings are presented in the following section.

As discussed earlier, a weakness of the current PAF is that it does not take into account the consistency in production and placement. In simplistic terms, it is possible for a Contractor to receive a PAF that is greater than 1.0 even when the quality of the production and/or placement is not consistent and has a high degree of variability. Figures 2.37 through 2.36 illustrate the standard deviation of the metric used to compute the PAF and the PAF. For example, Figure 2.37 compares the standard deviation, which is a measure of variability, in the laboratory-molded density with the production PAF for Item 341. These data can be divided into four zones. On the bottom left are cases with PAFs that are less than 1.0 indicating failure to meet target values but with low standard deviations. The data in this region correspond to cases that are consistent in production but do not meet the target values. The data on the bottom right correspond to cases that have received a PAF of greater than 1.0 and also have consistent production quality with a low standard deviation. This may be interpreted as the best case or the most desirable scenario. Data on top left region correspond to cases that are inconsistent and also miss the target values. This is perhaps the least desirable scenario because it could potentially translate into overall poor performance as well as high variability in the performance. Finally, the region of interest for this study is on the top right, i.e., data that correspond to cases where a PAF of greater than 1.0 was awarded but the quality of the mix produced was relatively less consistent based on the standard deviation values.

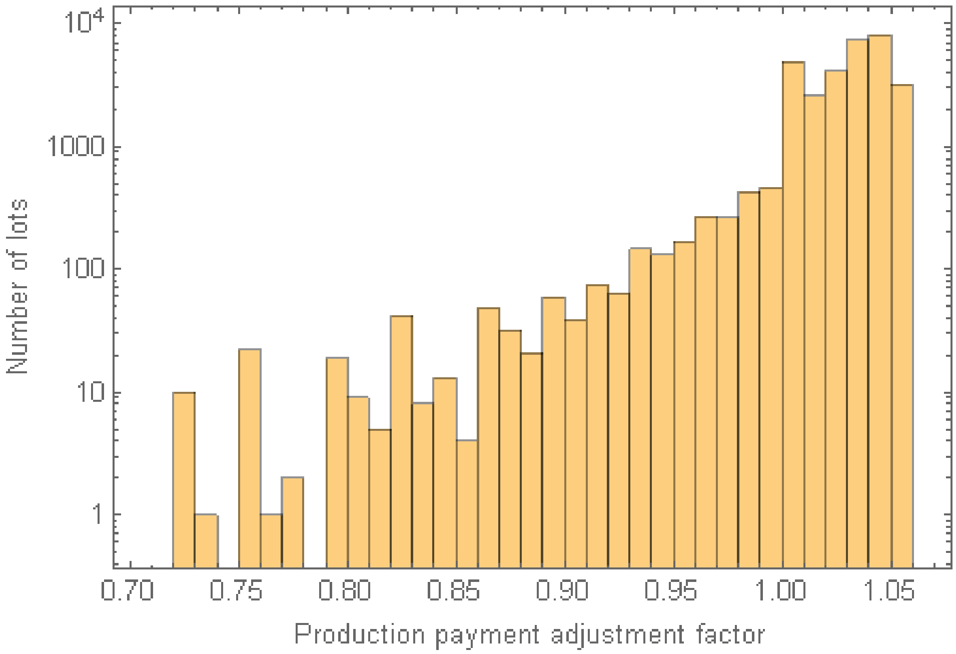


Figure 2.28. Distribution of average production payment adjustment factor for Item 341.

In the case of laboratory-molded density the standard deviation was computed using the density values recorded as 1(1), 1(2), 1(3), 2(1), 2(2), 2(3) ... 4(3). The prefix (1, 2, 3, or 4) corresponds to the sublot and the suffix (1, 2, or 3) corresponds to the three replicate specimens for Texas Gyratory Compactor (TGC) (two replicates for Superpave Gyratory Compactor (SGC)) from the given sublot. Note that the standard deviation in these figures was computed using data from all available sublots which could be as low as one sublot (with 3 replicates for TGC or 2 for SGC) or as high as four sublots (with 12 replicates). Similarly, for in-place air voids, the standard deviation was computed using values recorded as 1(A), 1(B), 2(A), 2(B) .... 4(B). The prefix (1, 2, 3, or 4) corresponds to the sublot and the suffix (A or B) corresponds to the two replicate specimens from the given sublot. As before, the standard deviation included data from all available sublots which could be as low as one sublot (with 2 replicates) or as high as four sublots (with 8 replicates). A more detailed analysis of these data that also accounts for the number of replicates will be discussed in the subsequent chapters.

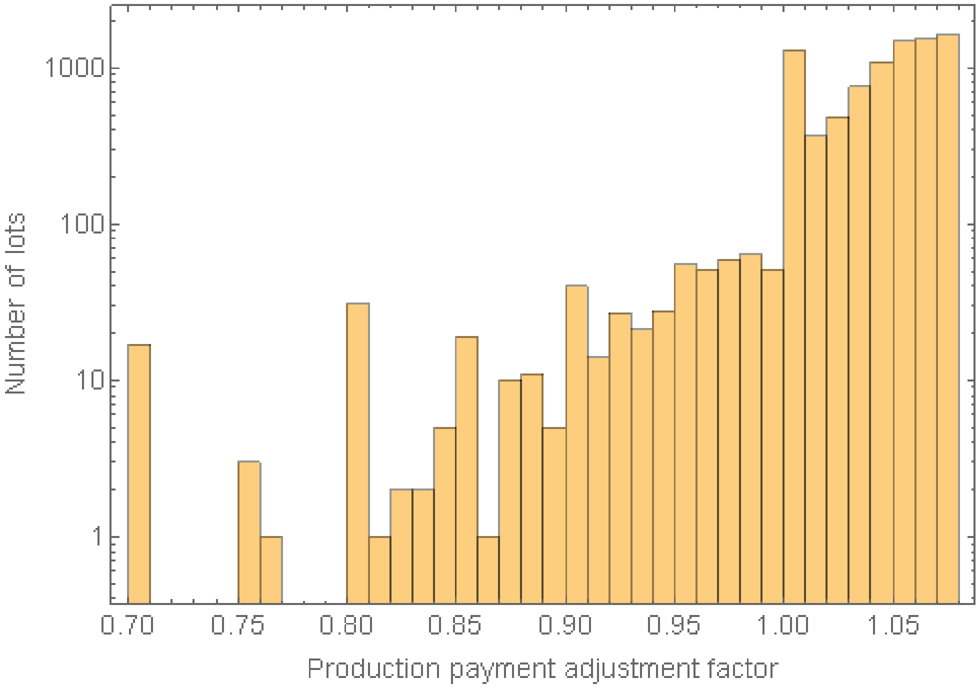


Figure 2.29. Distribution of average production payment adjustment factor for Item 344.

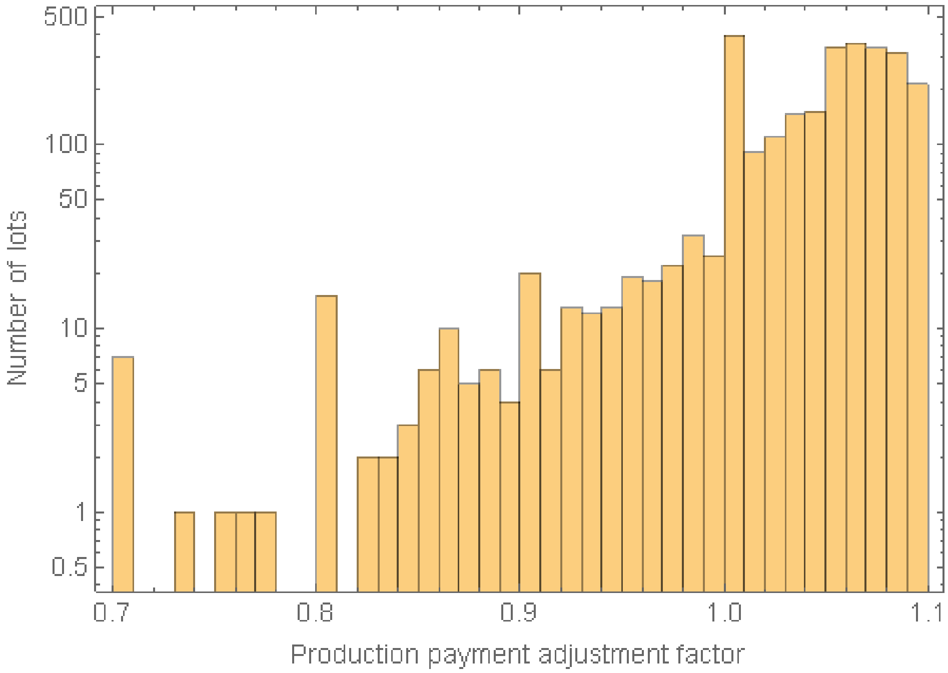


Figure 2.30. Distribution of average production payment adjustment factor for Item 346.

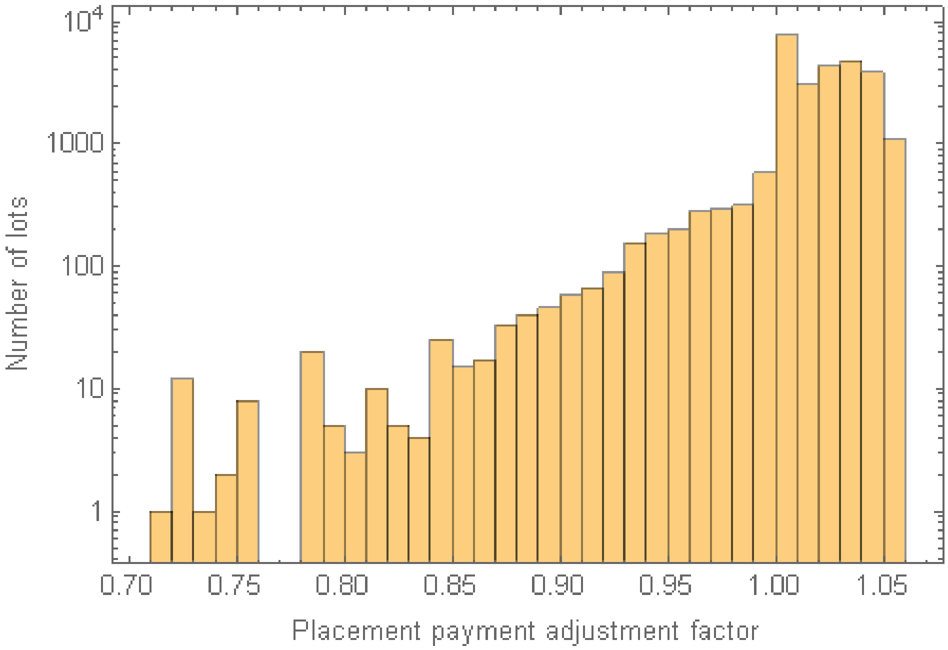


Figure 2.31. Distribution of average placement payment adjustment factor for Item 341.

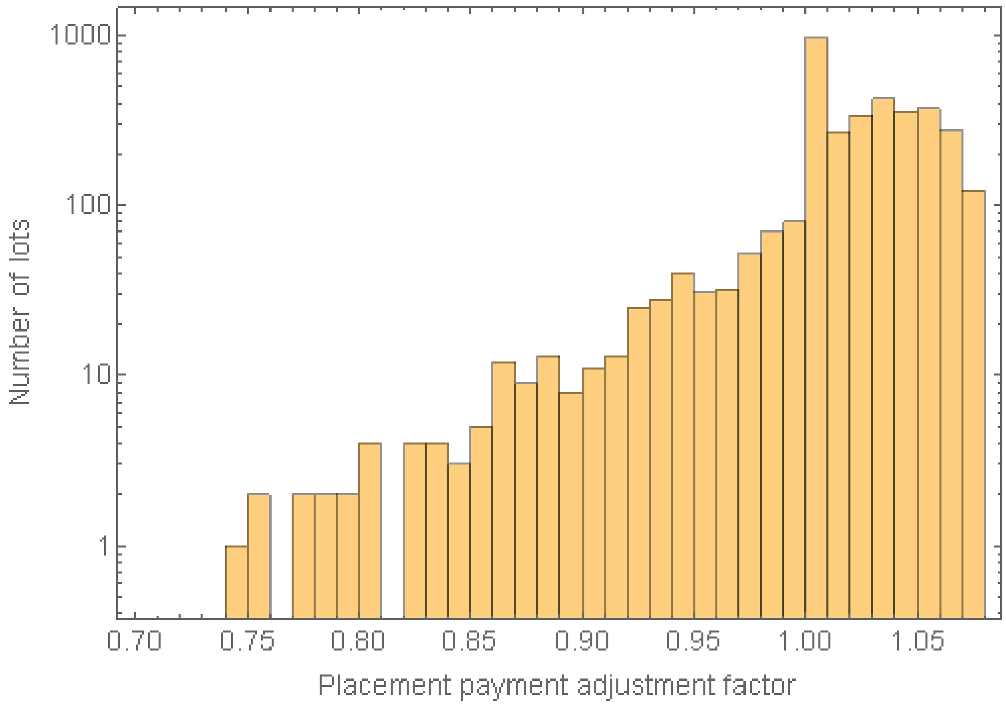


Figure 2.32. Distribution of average placement payment adjustment factor for Item 344.

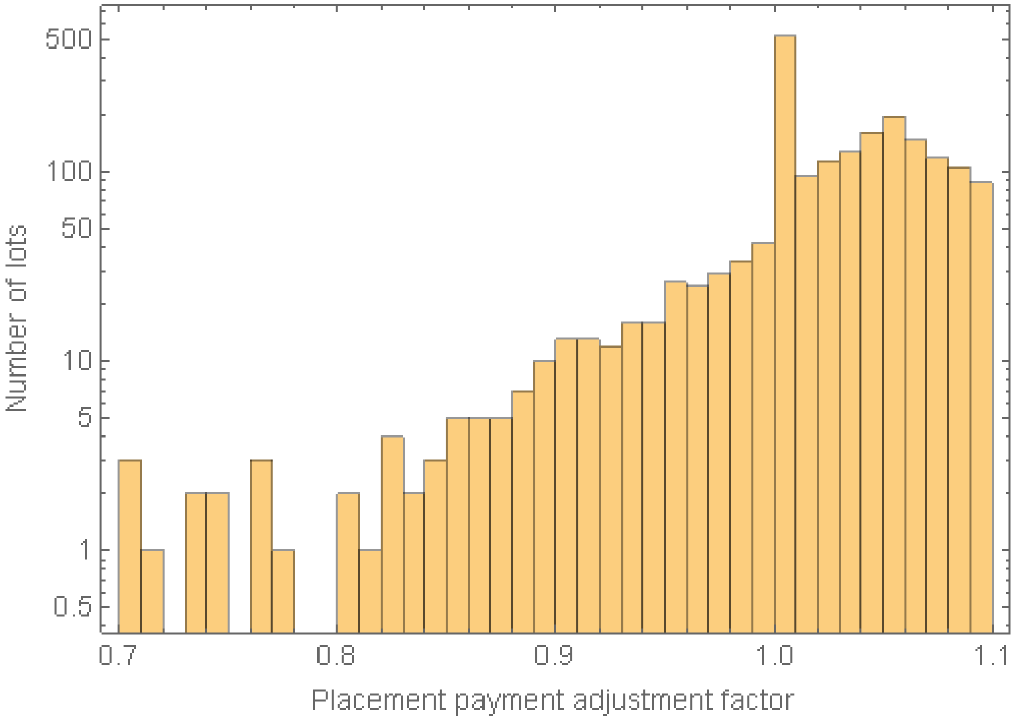


Figure 2.33. Distribution of average placement payment adjustment factor for Item 346.

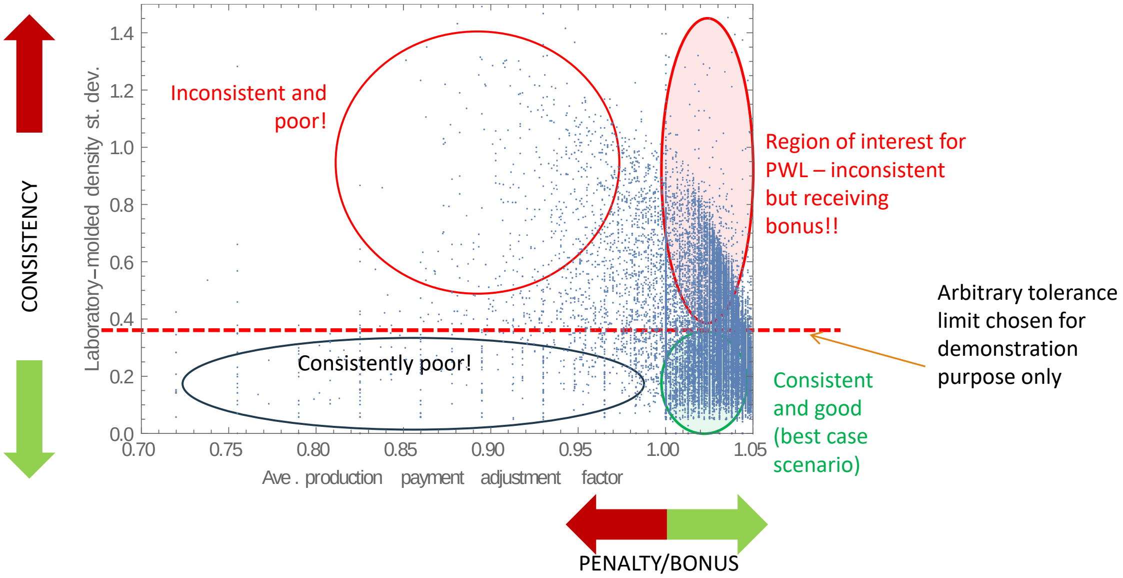


Figure 2.34. Laboratory-molded density standard deviation across the average production payment adjustment factors for Item 341. The red dashed line is an arbitrary line chosen to differentiate between consistent and inconsistent performances.

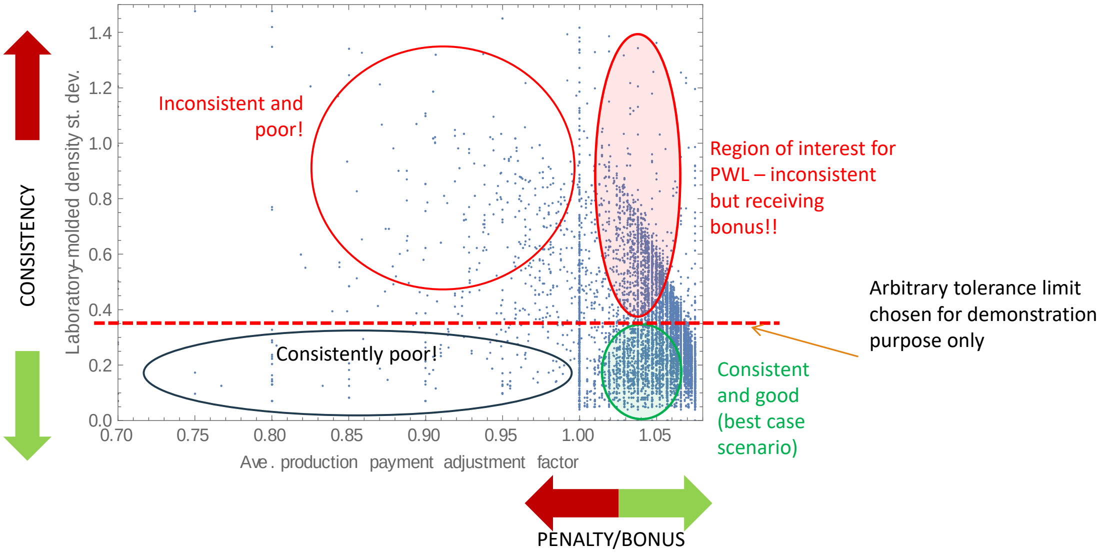


Figure 2.35. Laboratory-molded density standard deviation across the average production payment adjustment factors for Item 344. The red dashed line is an arbitrary line chosen to differentiate between consistent and inconsistent performances.

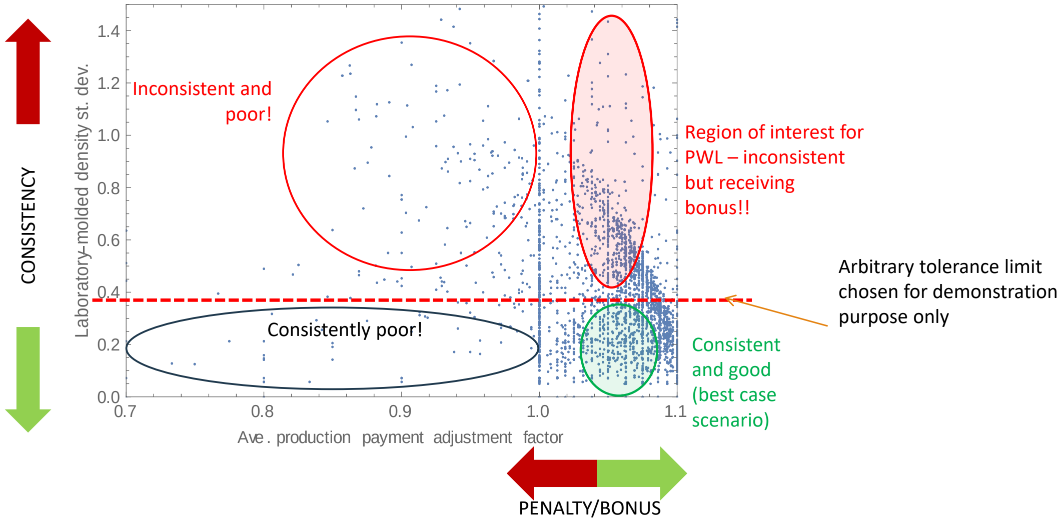


Figure 2.36. Laboratory-molded density standard deviation across the average production payment adjustment factors for Item 346. The red dashed line is an arbitrary line chosen to differentiate between consistent and inconsistent performances.

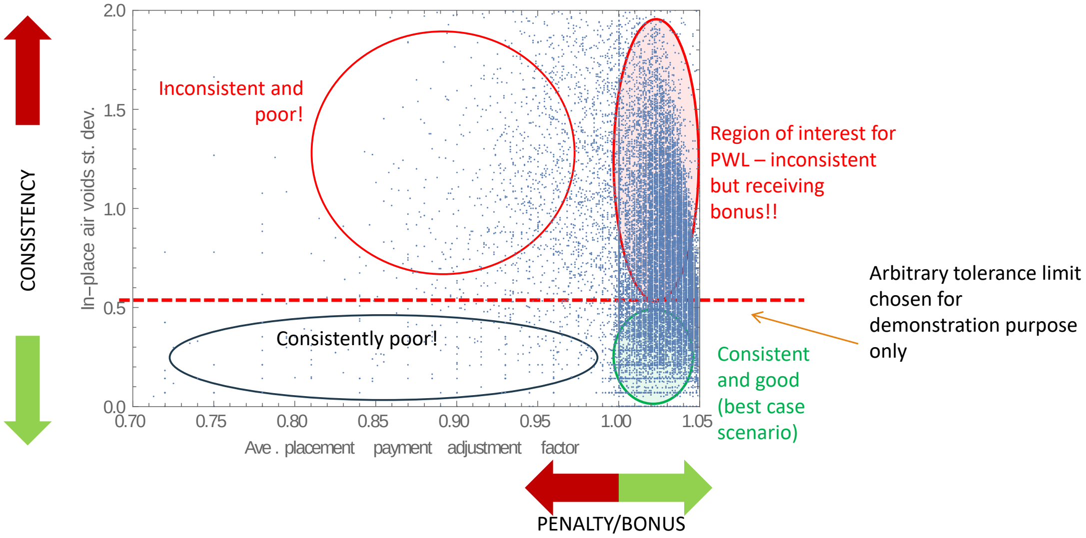


Figure 2.37. In-place air voids standard deviation across the average placement payment adjustment factors for Item 341. The red dashed line is an arbitrary line chosen to differentiate between consistent and inconsistent performances.

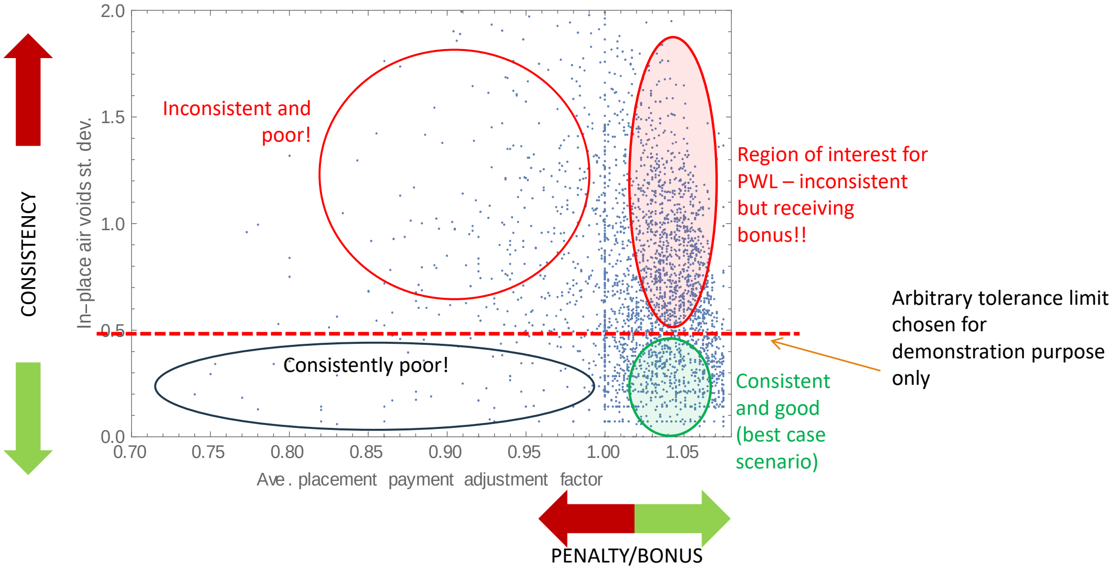


Figure 2.38. In-place air voids standard deviation across the average placement payment adjustment factors for Item 344. The red dashed line is an arbitrary line chosen to differentiate between consistent and inconsistent performances.

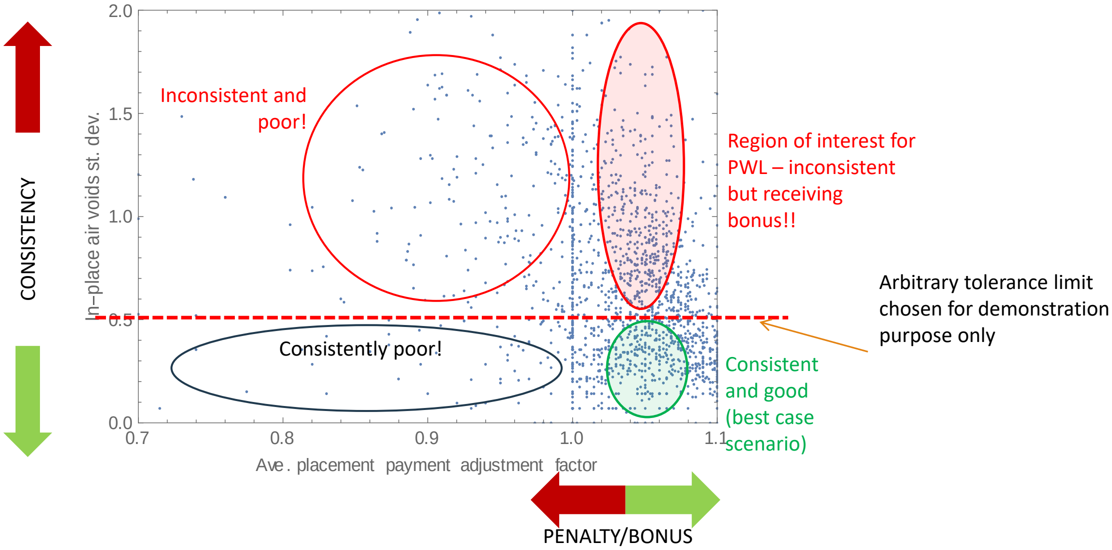


Figure 2.39. In-place air voids standard deviation across the average placement payment adjustment factors for Item 346. The red dashed line is an arbitrary line chosen to differentiate between consistent and inconsistent performances.

2.3.4 Summary

The aim of this chapter was to summarize the work that was done to examine various TxDOT databases and more importantly set up a replicate database that could be used for the purposes of this study. This chapter also summarizes the preliminary evaluation of HMA constituent material properties, design parameters that control the quality of the production and placement of HMA, and the existing payment adjustment factors to lay out the foundation for future analysis in the subsequent chapters. Distributions of material properties including asphalt content, mix type, recycled binder content, performance matrices in terms of Hamburg Wheel rut depth and indirect tensile strength, design parameters such as laboratory-molded density and in-place air voids, and the associated payment adjustment factors were analyzed. Dependency of the design parameters and performance matrices on mix type and properties is further explored in the following chapter. Furthermore, the spread of standard deviation of laboratory-molded density and that of in-place air void across the existing payment adjustment factor range were also explored to demonstrate the variability of performance overlooked by the current practice. Addressing this gap was also the main goal of this study.

CHAPTER 3. PAYMENT ADJUSTMENT FACTOR AND MIX PROPERTIES

3.1 OVERVIEW

The previous chapter demonstrated the typical distributions observed for mix properties (e.g. asphalt content, recycled binder ratio, aggregate gradation), mix performance (e.g. rutting resistance as indicated by the Hamburg Wheel tracking device or HWTD and fracture resistance as indicated by the indirect tensile strength test or IDT), and mix production and placement quality (i.e. laboratory-molded density and in-place air voids). The previous chapter also demonstrated the observed variability in the two key parameters (laboratory-molded density and in-place air voids) that are used to compute the payment adjustment factor or PAF (albeit without accounting for variability in its current format). The focus of this chapter is to examine these data in more detail. Specifically, this chapter examines whether or not any correlations exist between apparently independent variables. The following paragraphs explain why this analysis is important in the context of this study.

One of the goals of this study was to evaluate the possibility of using a metric for consistency, such as standard deviation, in order to compute the PAF. A possible means to achieve this is by setting a maximum allowable standard deviation for the laboratory-molded density and in-place air voids while computing the PAF. Another means to achieve this, that is also being used by several other state agencies, is to compute the percent of tests that are expected to meet the lower and upper specification limits, i.e. percent within limits or PWL. In either case, the standard deviation of the test results from the sample is a metric that will influence the final payment adjustment factor. This would ensure that the contractor is rewarded not only for producing and placing a mix with desired characteristics on an average but also for doing so in a consistent manner. However, the question then arises whether mix related factors, e.g. recycled binder ratio, use of warm mix technology, use of softer binder grades, influence the variability in material production and placement. If so, then it would be necessary to incorporate these factors in establishing the threshold for the standard deviation or establishing different payment adjustment factor schedules based on factors that influence the consistency in production and placement. Specifically, this would make it necessary to have a higher threshold to accommodate mix attributes that tend to increase variability and conversely a lower threshold to accommodate mix attributes that tend to reduce variability in mix production and/or placement. In order to address this question, the influence of the following factors on production, placement, and performance was evaluated:

- optimum asphalt content (to assess influence of the quantity of binder),
- binder grade (to assess the influence of the quality of the binder),
- recycled binder ratio (to assess the influence of RAP and/or RAS content),
- mix type (to assess the influence of aggregate structure, which is reflected in its gradation) and
- warm mix asphalt (to assess the impact of using warm mix asphalt).

The influence of the aforementioned factors was evaluated on the following mix production, placement, and performance characteristics:

- standard deviation for the laboratory-molded density,
- standard deviation for the in-place air voids,
- rutting resistance as reflected in the HWTD results, and
- cracking resistance as reflected in the IDT results.

Note that the first two attributes (standard deviation for the laboratory-molded density and in-place air voids) are directly related to the core of this study, i.e. to examine the influence of mix properties on consistency in mix production. The last two attributes are intended to examine the influence of these mix properties on overall performance. Although not directly related to the main goal of this study, it is expected that this information will provide further insight into the relationship between key mix properties and performance.

3.2 ANALYSIS

3.2.1 Data processing

The HMA mix design, production, placement, and performance parameters are recorded as Field Values, under a single column header named Field Number, in the Result (rslt) table on the SiteManager or SMGR database. In order to examine how the mix performance and quality are affected by the mix design characteristics such as asphalt content, recycled binder content, aggregate gradation, and mix type, the parameters under consideration first need to be queried independently and then linked together with separate column headers for statistical analysis. Therefore, a unique identifier is required to connect the separately queried parameters. Such a unique identifier would be the Sample ID which is generated by the SMGR for each sample. However, the queried results from previous tasks show that multiple samples, even from multiple projects, may have the same Sample ID. This probably happened when multiple samples were uploaded at the same time without generating different Sample IDs for different samples. Ideally, when one sample is uploaded, a unique identifier is generated, and before proceeding to the next sample, another Sample ID needs to be generated which was not done for these particular samples or lots to be more specific. In order to facilitate efficient assembly among different parameters, two additional identifiers are used, Test Method and Sample Test Number. Test Method describes the template used for testing the sample which is either TX2QCQA04 or TX2QCQA14 depending on the specification year that was followed to collect the QC/QA data. Sample Test Number is a user defined value which denotes the lot number for which the test is performed. However, even with three unique identifiers, thousands of duplicates were retrieved. Upon further investigation it was revealed that multiple samples were uploaded to the SMGR with two Test Methods or specification years, i.e. multiple samples contained the same Sample ID, Test Method, and Sample Test Number. As a result, when multiple parameters were joined together, the parameters were permuted several times exponentially spawning thousands of duplicate data points. These duplicates were then deleted by post-processing the data using a Mathematica code, which deletes a data set (data for a specific lot) that matches another data set in its entirity (i.e. all values for all parameters are the same; the likelihood of this happening for two different lots is exceedingly low).

The next step in processing data was to replace or filter out the missing values. For missing recycled binder from mix design fields, zero was assigned as it is very likely that blank cell actually represents the case where no recycled binder is used. Likewise, the 2014 template allows blank cells for mixes that do not use WMA, and it is apparent that blank fields actually refer to such mixes. Data sets containing missing binder grade, mix type, laboratory-molded density, in-place air voids, Hamburg Wheel rut depth, and IDT strength test values were deleted since there is no reasonable rationale to replace these missing data, and replacing the missing cells with any other values (e.g. mean from similar mixes) would create an unwanted bias in the statistical analysis. This exercise was repeated for data obtained for each of the three Items, 341, 344, and 346.

3.2.2 Method of analysis

To analyze the dependence of the mix quality and performance on the mix design characteristics, analysis of variance (ANOVA) was performed. ANOVA is a statistical tool to analyze variances among means. It is particularly useful when the effect of multiple levels of one or more factors and multiple observations at each level are available. Since this study presents itself with multiple levels of multiple factors, two-way crossed ANOVA was performed where interactions among multiple factors were also considered. Two estimates of the population variance were calculated: the mean square error (MSE) based on the differences among observations within the factors, and the mean square between (MSB) factors based on the differences among the factor means. The null hypothesis, i.e. that the population means for all factors are the same, is rejected when the MSB is much larger than the MSE. On the other hand, if the two mean square estimates were about the same, then the data were consistent with the null hypothesis that the population means are equal. For this study, ANOVA was performed using the IBM SPSS software. The Type III Univariate General Linear Model was used to calculate sums of squares. This method calculates the sums of squares of an effect in the design as the sums of squares, adjusted for any other effects that do not contain the effect, and orthogonal to any existing effects. The Type III sums of squares are invariant with respect to the cell frequencies as long as the general form of estimability remains constant and are applicable for an unbalanced model with no missing cells.[1]

The sum of squares were then divided by the degrees of freedom (df) or the number of observations in each factor to determine the mean square. The Mean Square Between (MSB) was then divided by the Mean Square Error (MSE) to calculate the F ratio. After the F ratio was calculated, significance was determined from the the F distribution curve. If F ratio is very large then it is reasonable to conclude that the probability of obtaining identical means for two population is very low, since the probability is defined as the area on the right-hand tail of the F distribution shown in Figure 3.1.

3.3 RESULTS

A typical output of an ANOVA analysis by IBM SPSS is shown in Figures 3.2 and 3.3. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt. For these analyses, JMF asphalt content and recycled binder content were treated as scaled variables and binder grade, mix type, and use of WMA were treated as categorical or nominal variables. Very low significance, in this case defined as less than (0.05) has been

---

[1] Source: IBM SPSS

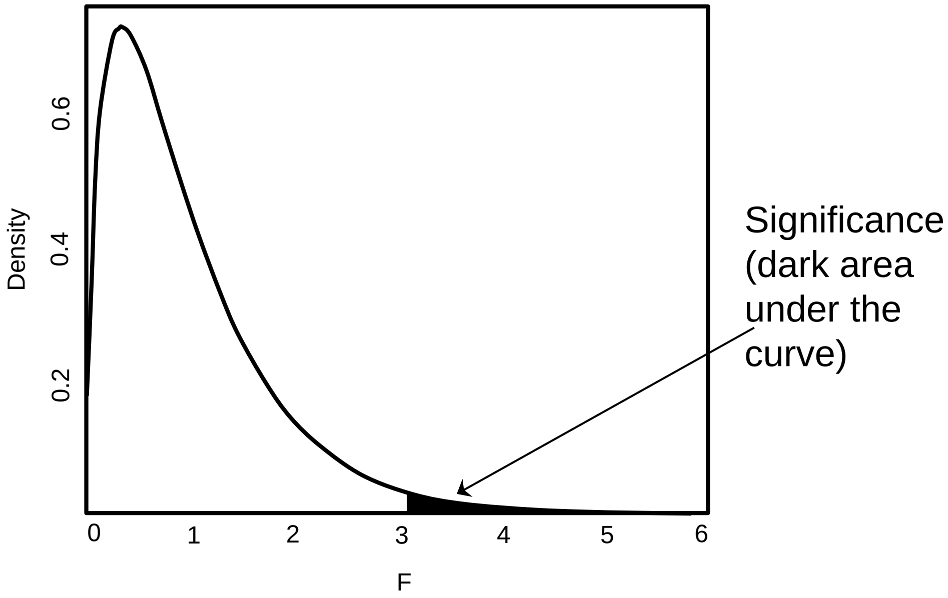


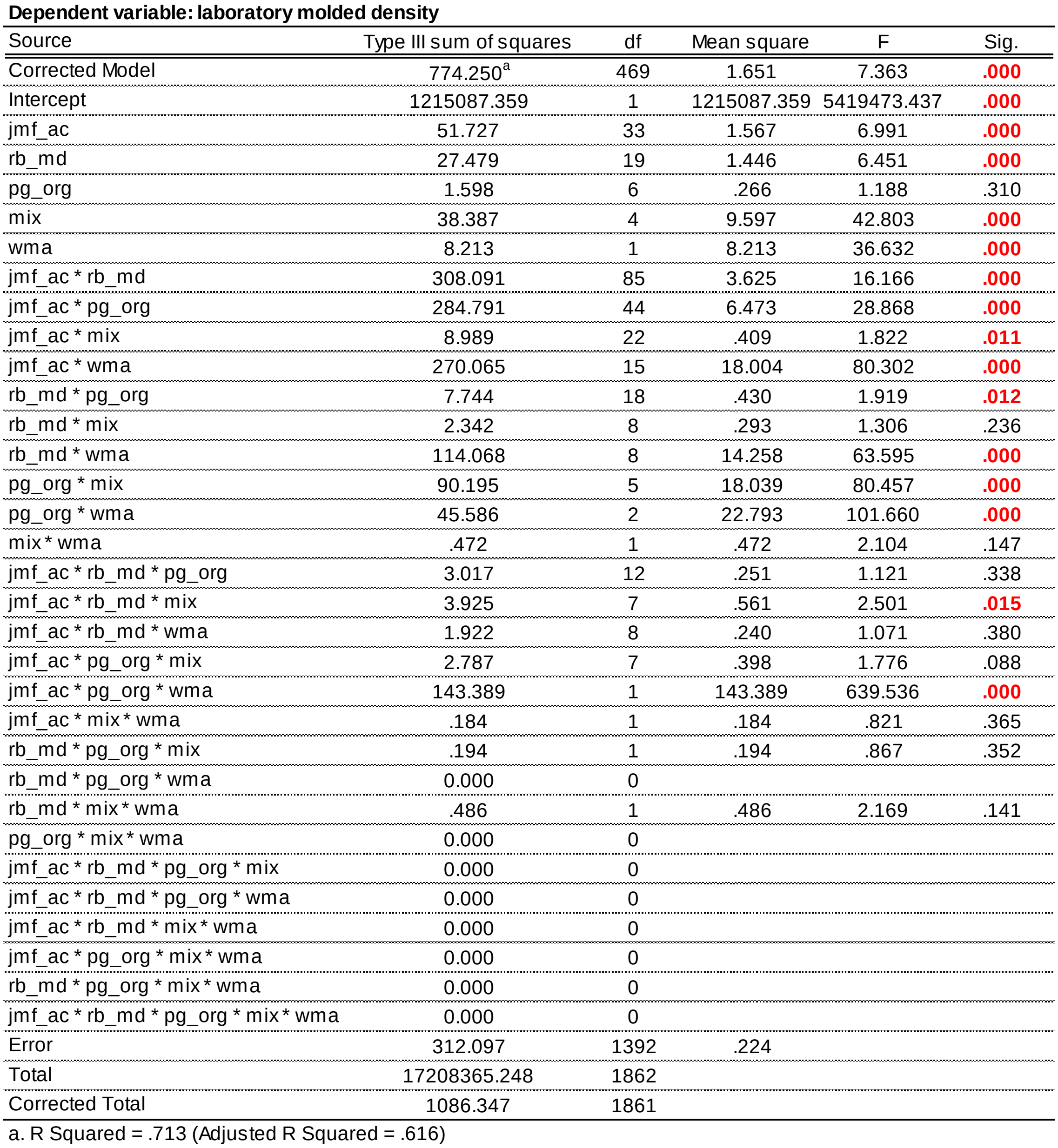
Figure 3.1. Distribution of F and significance.

highlighted in red bold numeric showing that the null hypothesis of equal mean can be rejected for those factors. For example, a change in asphalt content or recycled binder has a significant influence on the average value of the laboratory-molded density. On the other hand, aggregate structure and binder grade do not appear to have a significant influence on the average laboratory-molded density. In other words, it is possible to achieve similar laboratory-molded densities using different mix types and binder grades but not with similar asphalt content and recycled binder ratio. Significance of 0.05 implies that there is a 5% probability that the two different asphalt contents will result in the same density whereas a significance of 0.31 indicates that there is 31% probability that binder grades do not influence the average density achieved in the laboratory. The General Linear Univariate model of SPSS also calculates the R squared value and an adjusted R squared value based on the residual degrees of freedom to show the goodness of fit. The residual degrees of freedom is the number of response values minus the number of fitted coefficients estimated from the response values.

Between-subjects factors

Factors	Value label	N	Factors		Value label	N
jmf_ac	4	5	rb_md	0		604
	4.1	4		0.3		6
	4.2	1		0.4		10
	4.3	12		0.5		87
	4.4	24		0.6		104
	4.5	37		0.7		122
	4.6	1		0.8		222
	4.6	53		0.9		294
	4.7	29		1		186
	4.8	91		1.1		87
	4.9	134		1.2		20
	5	135		1.3		36
	5.1	123		1.4		15
	5.2	158		1.5		40
	5.3	204		1.6		19
	5.4	200		1.7		1
	5.5	107		1.8		2
	5.6	125		2.5		1
	5.7	97		13.2		4
	5.8	130		14		1
	5.9	55		19.2		1
	6	10	pg_org	1	PG 58-28	1
	6.1	25		2	PG 64-22	361
	6.2	14		5	PG 70-22	1004
	6.3	24		6	PG 70-28	142
	6.4	7		7	PG 70-34	4
	6.5	11		8	PG 76-22	347
	6.6	14		9	PG 76-28	3
	6.7	1	mix	1	A	4
	6.8	10		2	B	235
	6.9	3		3	C	879
	7	5		4	D	666
	7.1	6		5	F	78
	7.5	6	wma	1	YES	442
	7.7	1		2	NO	1420

Figure 3.2. Typical ANOVA output: part 1. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.



Dependent variable: laboratory molded density

Source	Type III sum of squares	df	Mean square	F	Sig.
Corrected Model	774.250[a]	469	1.651	7.363	**.000**
Intercept	1215087.359	1	1215087.359	5419473.437	**.000**
jmf_ac	51.727	33	1.567	6.991	**.000**
rb_md	27.479	19	1.446	6.451	**.000**
pg_org	1.598	6	.266	1.188	.310
mix	38.387	4	9.597	42.803	**.000**
wma	8.213	1	8.213	36.632	**.000**
jmf_ac * rb_md	308.091	85	3.625	16.166	**.000**
jmf_ac * pg_org	284.791	44	6.473	28.868	**.000**
jmf_ac * mix	8.989	22	.409	1.822	**.011**
jmf_ac * wma	270.065	15	18.004	80.302	**.000**
rb_md * pg_org	7.744	18	.430	1.919	**.012**
rb_md * mix	2.342	8	.293	1.306	.236
rb_md * wma	114.068	8	14.258	63.595	**.000**
pg_org * mix	90.195	5	18.039	80.457	**.000**
pg_org * wma	45.586	2	22.793	101.660	**.000**
mix * wma	.472	1	.472	2.104	.147
jmf_ac * rb_md * pg_org	3.017	12	.251	1.121	.338
jmf_ac * rb_md * mix	3.925	7	.561	2.501	**.015**
jmf_ac * rb_md * wma	1.922	8	.240	1.071	.380
jmf_ac * pg_org * mix	2.787	7	.398	1.776	.088
jmf_ac * pg_org * wma	143.389	1	143.389	639.536	**.000**
jmf_ac * mix * wma	.184	1	.184	.821	.365
rb_md * pg_org * mix	.194	1	.194	.867	.352
rb_md * pg_org * wma	0.000	0			
rb_md * mix * wma	.486	1	.486	2.169	.141
pg_org * mix * wma	0.000	0			
jmf_ac * rb_md * pg_org * mix	0.000	0			
jmf_ac * rb_md * pg_org * wma	0.000	0			
jmf_ac * rb_md * mix * wma	0.000	0			
jmf_ac * pg_org * mix * wma	0.000	0			
rb_md * pg_org * mix * wma	0.000	0			
jmf_ac * rb_md * pg_org * mix * wma	0.000	0			
Error	312.097	1392	.224		
Total	17208365.248	1862			
Corrected Total	1086.347	1861			

a. R Squared = .713 (Adjusted R Squared = .616)

Figure 3.3. Typical ANOVA output: part 2. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

3.3.1 Influence of mix properties on production characteristics

As of July, 2018, around 39,277 samples were posted to SMGR for specification Item 341, 9685 samples for Item 344, and 3111 samples for Item 346. Full population was used to analyze global correlation using ANOVA for all these Items.

Figures 3.4 through Figure 3.6 present results from the ANOVA for the standard deviation of the laboratory-molded density for Items 341, 344, and 346. Bars in this figures represent different mix properties or the combination of these properties.

ANOVA for items 341 through 346 shows high significance values for all individual factors implying that different population consisting of the different independent factors are unlikely to generate different standard deviations. In other words, it is reasonable to say that standard deviation of laboratory-molded density for different pools of lots produced with different mix types while keeping other variables (asphalt content, binder grade, recycled binder content, and WMA technology) constant may be equal. However, there is one instance where risk of obtaining different standard deviation is indicated i.e. the plotted significance is less than the 5% significance level. For instance, Figure 3.4 for Item 341 shows that recycled binder content, binder grades and mix types, when interact with each other, may influence the standard deviation of the laboratory-molded density. Such correlations observed for only one instance and for a combination of multiple factors may not be adequate to justify the adaptation of different standard-deviation thresholds for different mix types with different binder grades or construction practices.

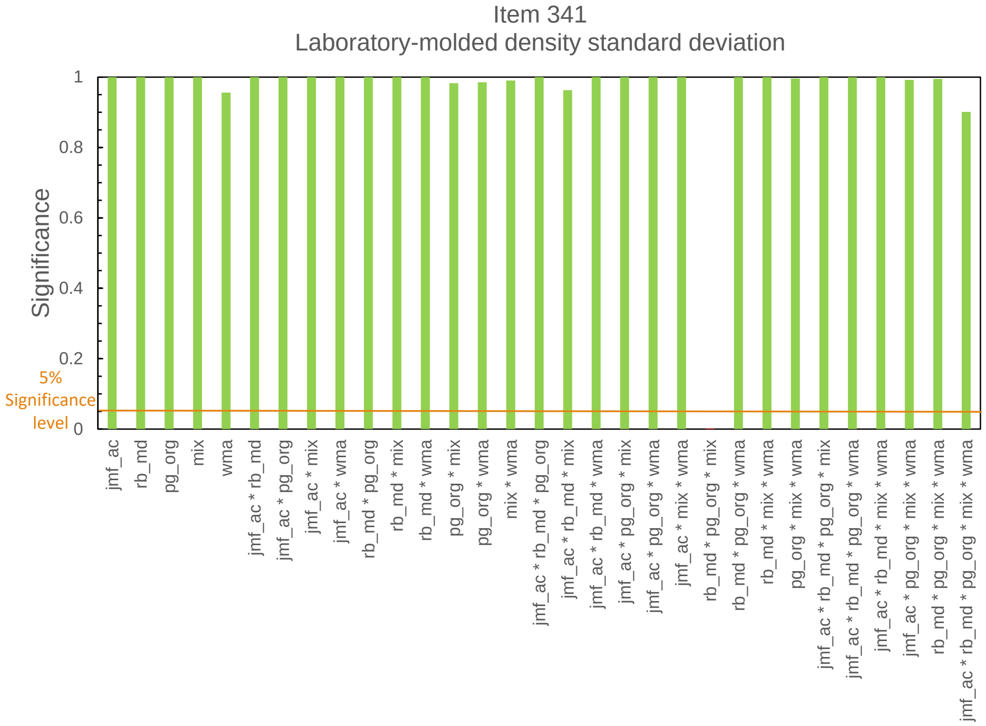


Figure 3.4. ANOVA for standard deviation of Item 341 laboratory-molded density. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

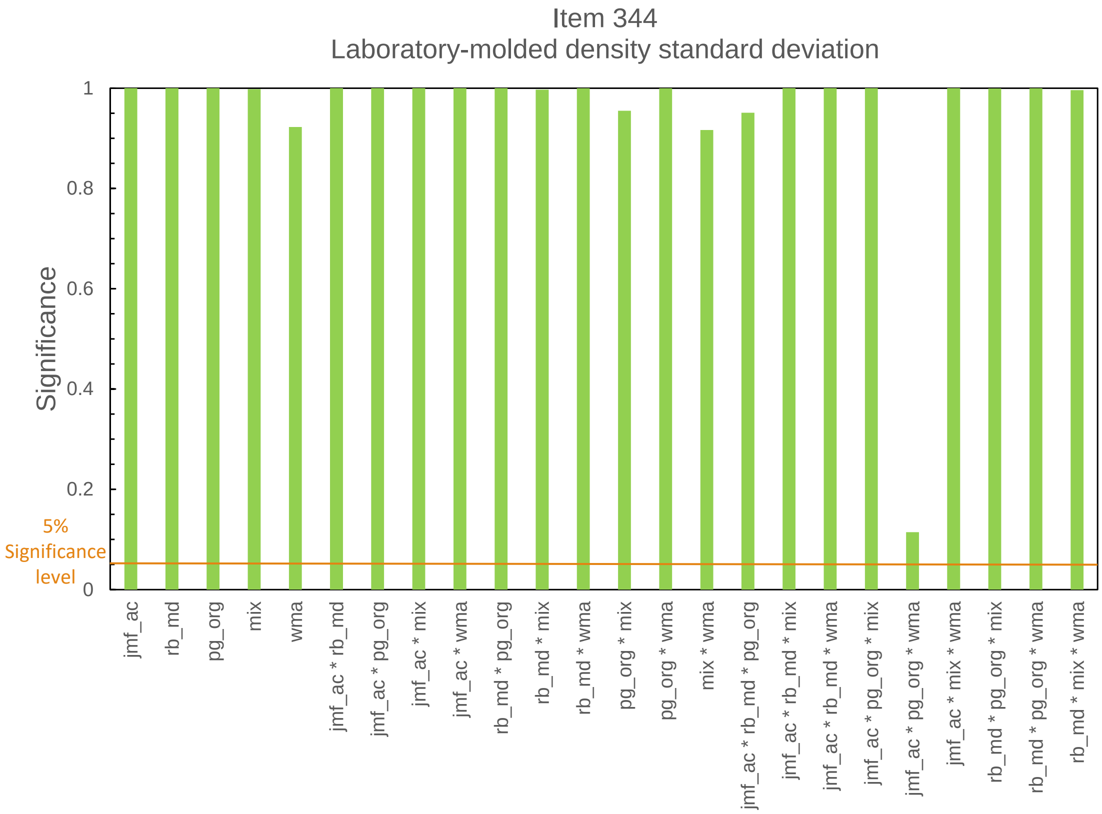


Figure 3.5. ANOVA for standard deviation of Item 344 laboratory-molded density. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

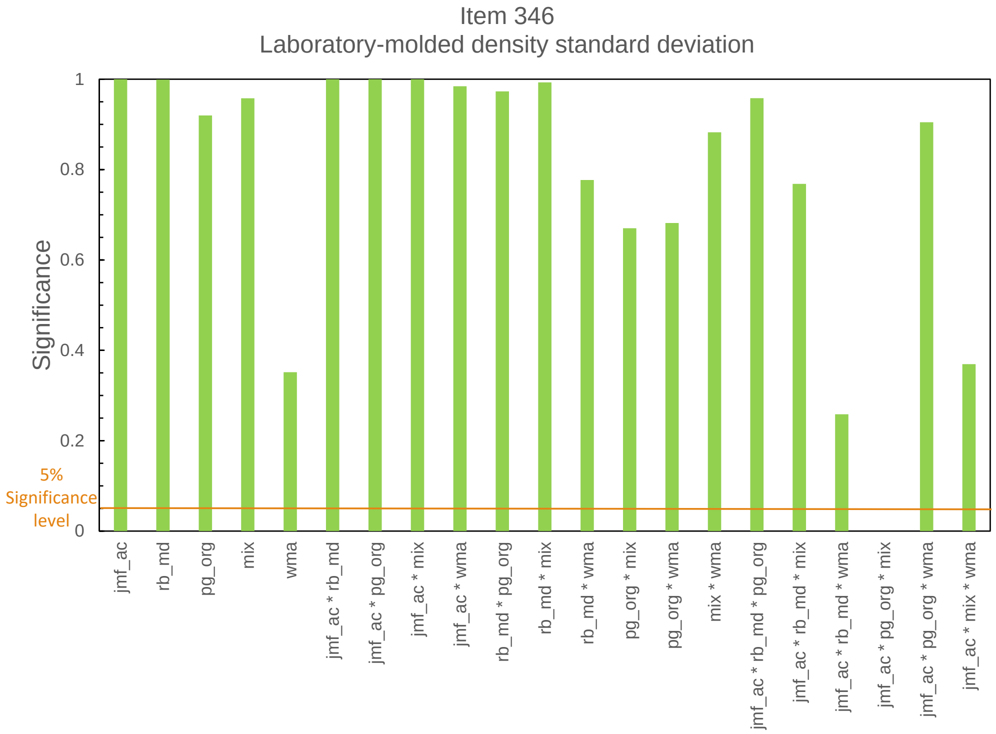


Figure 3.6. ANOVA for standard deviation of Item 346 laboratory-molded density. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

3.3.2 Influence of mix properties on placement characteristics

Figures 3.7 through 3.9 show the ANOVA results for the standard deviation of in-place air voids for Items 341, 344, and 346. Similar to the above analysis, it is unlikely that any single mix design related factor would influence the standard deviation of the in-place air voids of two separate lots produced with different asphalt contents while all other factors (i.e. aggregate gradation, binder grade, recycled binder content, and warm mix technology) are similar. However, when several of these factors are considered together for Item 341, the standard deviations could be changed as highlighted by the lower significance than 5% significance level, shown with red bars in Figure 3.7. Even though this is the case where all these material properties are changed simultaneously in field, further analysis needs to be performed before multiple standard deviation thresholds for different mix types within the same specification Item are considered, which in turn will make the QC/QA overly complicated and unrealistic for practical implementation.

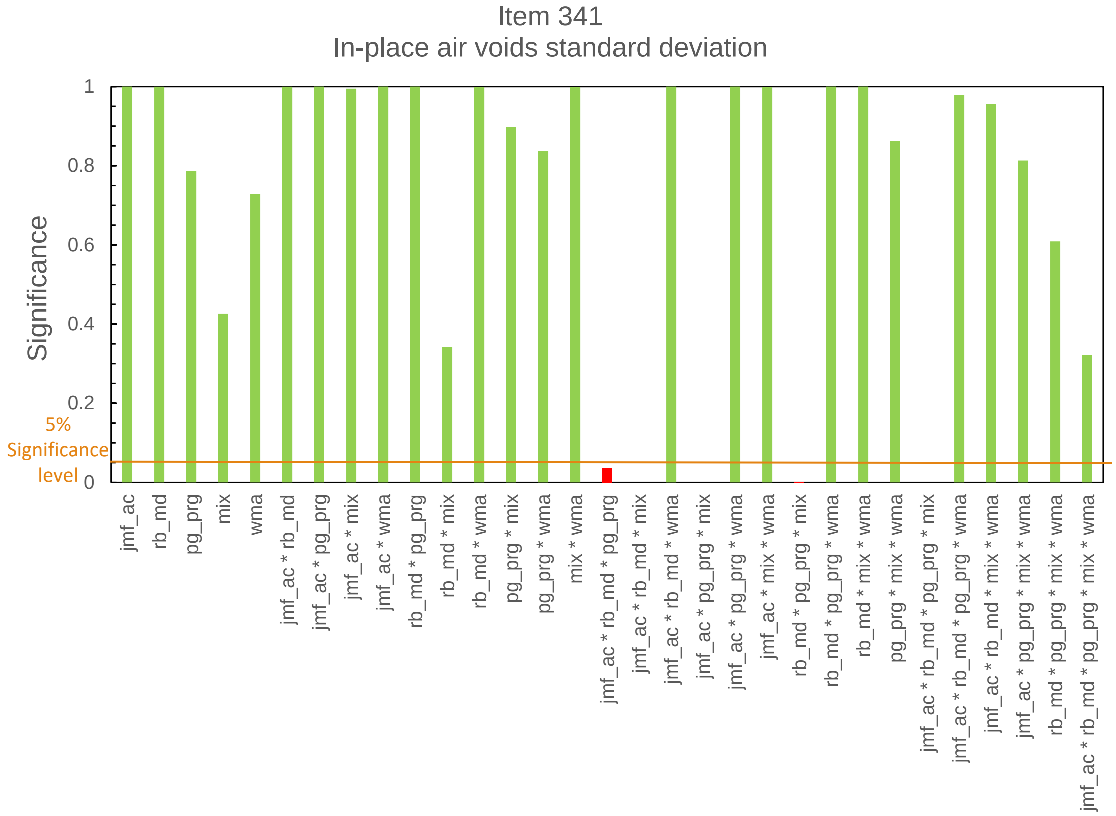


Figure 3.7. ANOVA for standard deviation of Item 341 in-place air voids. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

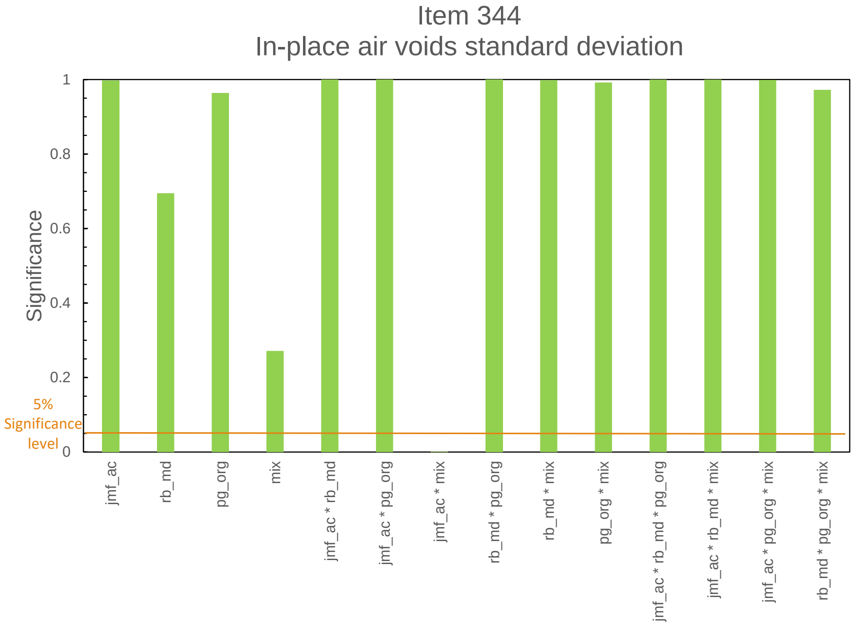


Figure 3.8. ANOVA for standard deviation of Item 344 in-place air voids. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

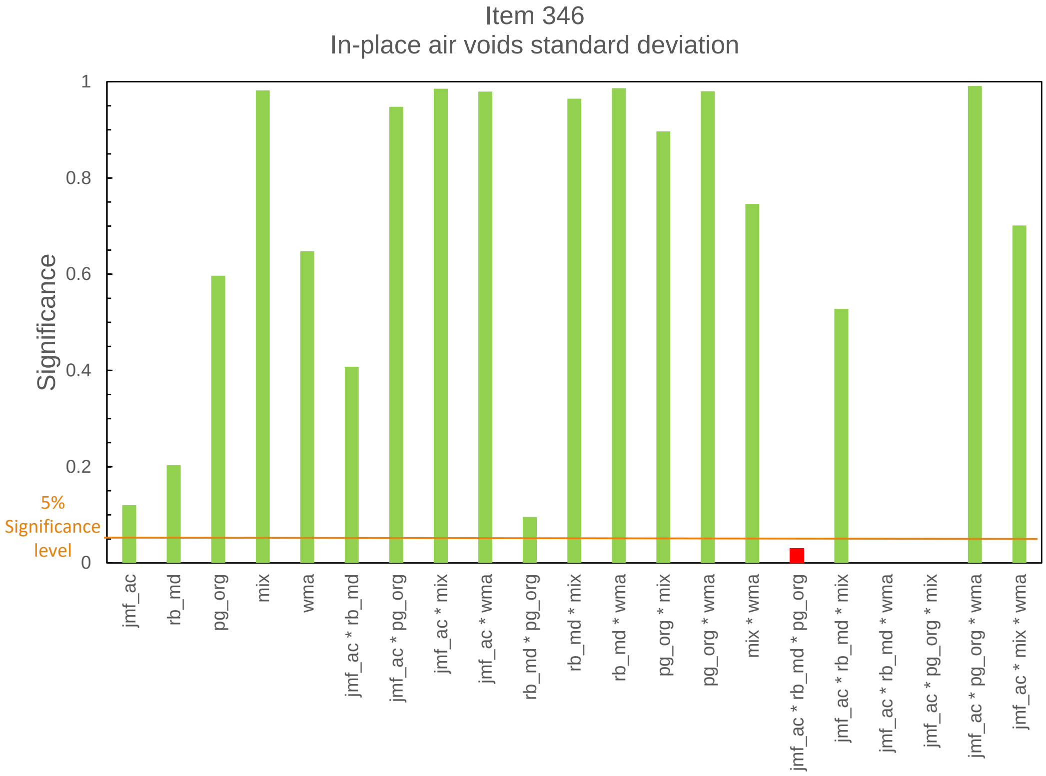


Figure 3.9. ANOVA for standard deviation of Item 346 in-place air voids. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

3.3.3 Influence of mix properties on laboratory mix performance

ANOVA results for the Hamburg Wheel rut depth and IDT tensile strength for Items 341, 344, and 346 are shown in Figure 3.10, 3.11, and 3.12, respectively. The entire population for all Items have been analyzed. Values in red shows significance lower than the 0.05 % significance level suggesting dependence on the mix properties. In contrast to the production and placement characteristics, strong dependence of the performance characteristics on parameters including asphalt content, recycled binder content, original binder grade, mix type, and application of warm mix asphalt was found. This is expected because all of these factors influence the overall mix performance. Interactions among these variables were also observed for specification Item 341. Similar dependence was found for Item 344. On the contrary, no significant dependence of the IDT strength on the mix characteristics other than the asphalt content and recycled binder was observed for Item 346

Dependent variable	Hamburg rut depth	IDT strength
jmf_ac	**.000**	**.000**
rb_md	**.000**	**.000**
pg_org	**.000**	**.000**
mix	**.000**	**.000**
wma	**.000**	.720
jmf_ac * rb_md	**.000**	**.000**
jmf_ac * pg_org	**.000**	**.000**
jmf_ac * mix	**.000**	**.000**
jmf_ac * wma	**.000**	**.000**
rb_md * pg_org	**.000**	**.000**
rb_md * mix	**.000**	**.000**
rb_md * wma	1.000	**.000**
pg_org * mix	1.000	**.041**
pg_org * wma	1.000	**.000**
mix * wma	.871	.955
jmf_ac * rb_md * pg_org	**.000**	**.000**
jmf_ac * rb_md * mix	**.000**	**.000**
jmf_ac * rb_md * wma	1.000	**.000**
jmf_ac * pg_org * mix	.957	**.000**
jmf_ac * pg_org * wma	1.000	**.000**
jmf_ac * mix * wma	.995	**.000**
rb_md * pg_org * mix	.560	**.000**
rb_md * pg_org * wma	1.000	**.007**
rb_md * mix * wma	.274	.278
pg_org * mix * wma	.994	.334
jmf_ac * rb_md * pg_org * mix	1.000	**.000**
jmf_ac * rb_md * pg_org * wma	1.000	**.000**
jmf_ac * rb_md * mix * wma	.987	**.000**
jmf_ac * pg_org * mix * wma	.972	
Adjusted R Squared	.729	.597
Population size	11764	10134

Figure 3.10. ANOVA for item 341 performance characteristics. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

Dependent variable	Hamburg rut depth	IDT strength
jmf_ac	**.000**	**.000**
rb_md	**.000**	**.000**
pg_org	**.000**	**.000**
mix	**.000**	**.000**
wma	**.000**	.610
jmf_ac * rb_md	**.000**	**.000**
jmf_ac * pg_org	**.000**	**.000**
jmf_ac * mix	**.000**	
jmf_ac * wma	**.000**	**.000**
rb_md * pg_org	.411	**.000**
rb_md * mix	**.000**	**.000**
rb_md * wma	**.001**	**.000**
pg_org * wma	.547	
mix * wma	**.000**	.216
jmf_ac * rb_md * pg_org	.713	
Adjusted R squared	0.87	0.952
Population size	1911	1585

Figure 3.11. ANOVA for item 344 performance characteristics. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

Dependent variable	Hamburg rut depth	IDT strength
jmf_ac	**0.000**	**0.020**
rb_md	**0.000**	**0.000**
pg_org	**0.000**	0.802
mix	**0.000**	0.053
wma	0.276	
jmf_ac * rb_md	**0.000**	**0.000**
jmf_ac * pg_org	0.217	0.849
jmf_ac * mix	**0.000**	0.074
rb_md * mix	**0.000**	
jmf_ac * rb_md * mix	**0.036**	
Adjusted R squared	1	0.612
Population size	603	243

Figure 3.12. ANOVA for item 346 performance characteristics. Here, jmf_ac denotes the design asphalt content specified in the job mix formula (JMF), rb_md denotes the recycled binder percentage from the mix design, pg_org is the original binder grade, mix is the mix type or aggregate gradation, and wma is warm mix asphalt.

3.3.4 Summary

As discussed earlier, the primary goal of this study was to evaluate the possibility of using a metric for consistency, such as standard deviation or percent within limits (PWL), computed using both the mean and the standard deviation, in order to compute the payment adjustment factor or PAF. The purpose of this task was to examine whether variability, determined in terms of standard deviation, in mix production and placement is dependent on mix properties such as asphalt content, mix type, binder type, recycled binder ratio, use of warm mix asphalt, or binder grade. The analysis of variance of payment criteria on mix properties, shows that it is unlikely that any single mix characteristic by itself can influence the variability in HMA production and placement. Instances, where a combination of several different mix properties influence the variability, would require different standard deviation thresholds for different mix types for a robust QC/QA scheme. However, such scheme will be extremely complex in nature and impractical to implement.

CHAPTER 4. OVERVIEW OF USING PWL AS A QUALITY MEASURE

4.1 OVERVIEW

The ANOVA analyses from the previous chapter showed that the standard deviations for the laboratory-molded density and in-place air voids were not influenced by any single mix characteristic by itself such as the asphalt content, recycled binder content, binder grade, mix type, and warm mix asphalt. This was true for all cases, i.e. laboratory-molded density and in-place air voids for Items 341, 344, and 346. The variability in some of these cases may be correlated with a combination of several different factors but each factor by itself did not show any significant correlation. Based on these analyses, it is reasonable to conclude that it is possible to achieve consistency in both production and placement regardless of the aggregate gradation, binder content, binder grade or the mix production technology (e.g. WMA) used. In other words, these factors by themselves do not significantly influence the variability in production or placement. Consequently, a new quality measure, such as PWL, can potentially be used in conjunction with an appropriate payment adjustment factor scheme to reward contractors that demonstrate more control and consequently achieve higher levels of consistency in production and placement of asphalt mixes.

As discussed in the previous chapters, PWL is increasingly being used as a quality measure for acceptance and payment to contractors by various states. The overall approach of using PWL is to determine the probability that a test sample will fall within specification limits based on the test results from a field sample. This approach inherently incorporates both the average and variability or standard deviation from the sample test results. The PWL achieved by a Contractor is then used with a concomitant payment adjustment factor scheme for payment to the Contractor.

Chapter 1 presented the common terminology that is used when following the PWL approach. The remainder of this chapter exemplifies the steps involved in computing PWL for production or placement. The following chapter goes into more details on the different payment adjustment factor schemes or schemes that can be used based on the PWL achieved by the Contractor as well as the risks to the agency and Contractor while setting different parameters that drive the payment adjustment factor scheme.

4.2 METHOD TO DETERMINE PWL

PWL can be computed for any quality characteristic that is of interest. For example, the state of Oklahoma computes PWL for four parameters: asphalt content, roadway density, laboratory air voids, and voids in mineral aggregates or VMA. A payment adjustment factor is then calculated for the corresponding PWL for each parameter using Equation 1.1. A combined payment adjustment factor (Equation 1.2) is finally determined by taking the weighted average of the four characteristics, with a weight factor of 4, 3, 2, and 1 for roadway density, laboratory air voids, asphalt content, and VMA, respectively. Since, the existing specifications for TxDOT use a production payment adjustment factor that is based on the laboratory-molded density and a placement payment adjustment factor that is based on the in-place air voids, these two parameters are considered here to demonstrate the steps to compute PWL. The step by step process for a preliminary PWL model, is explained below and summarized in Figure 4.1.

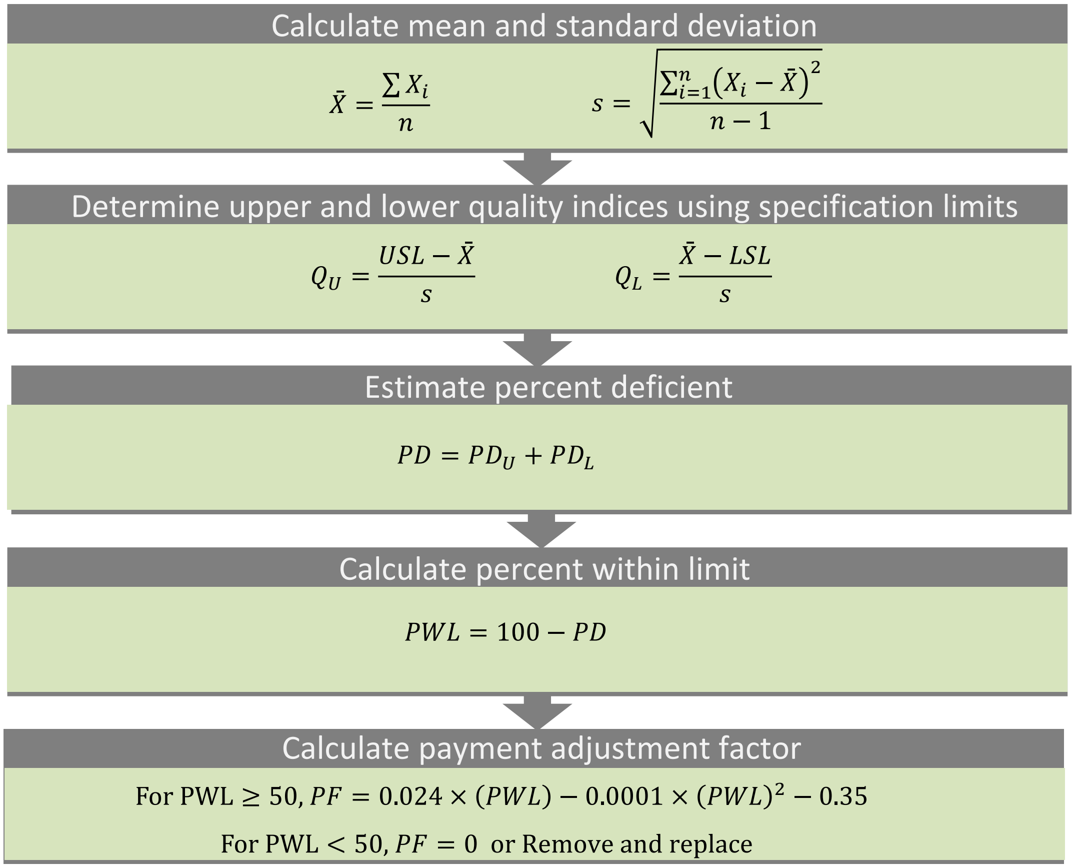


Figure 4.1. Preliminary PWL model.

Step 1. Calculate mean and standard deviation

The sample mean, $\overline{X}$, and the sample standard deviation, s, of the test results are calculated as

$$\overline{X} = \frac{\Sigma(X_i)}{n} \tag{4.1}$$

and

$$s = \sqrt{\frac{\Sigma(X_i - \overline{X})^2}{n-1}}, \tag{4.2}$$

where $\overline{X}$ = mean, X_i = individual test result , n = number of test specimens or results, and s = sample standard deviation.

Step 2. Determine quality indices using specification limits

Since PWL with respect to a particular quality characteristic is the amount of materials and construction which falls within the specified limits, the upper specification limit (USL) and lower specified limit (LSL) need to be established for that particular characteristic a priorí. For the purposes of this example and in the case of the laboratory-molded density the specification limits are chosen from the existing operational tolerance specified by TxDOT specification. Recall that these tolerances were developed based on the sensitivity of the parameter and potential impact to the quality and performance of the product. Therefore, it is rational to continue and preserve these tolerance values from the existing specifications while computing PWL in this demonstration. Similarly, for the in-place air void specification limits, the existing cut-off values for bonuses are chosen from the same specification. These limits are determined in the same way for all items, 341, 344, and 346. Table 4.1 presents a summary of these limits. The upper and lower quality indices, (Q_U) and (Q_L), respectively, are then calculated using these limits as given by Equations 4.3 and 4.4,

$$Q_U = \frac{USL - \overline{X}}{s} \tag{4.3}$$

and

$$Q_L = \frac{\overline{X} - LSL}{s}, \tag{4.4}$$

where Q_U = upper quality index, Q_L = lower quality index, USL = upper specification limit, and LSL = lower specification limit.

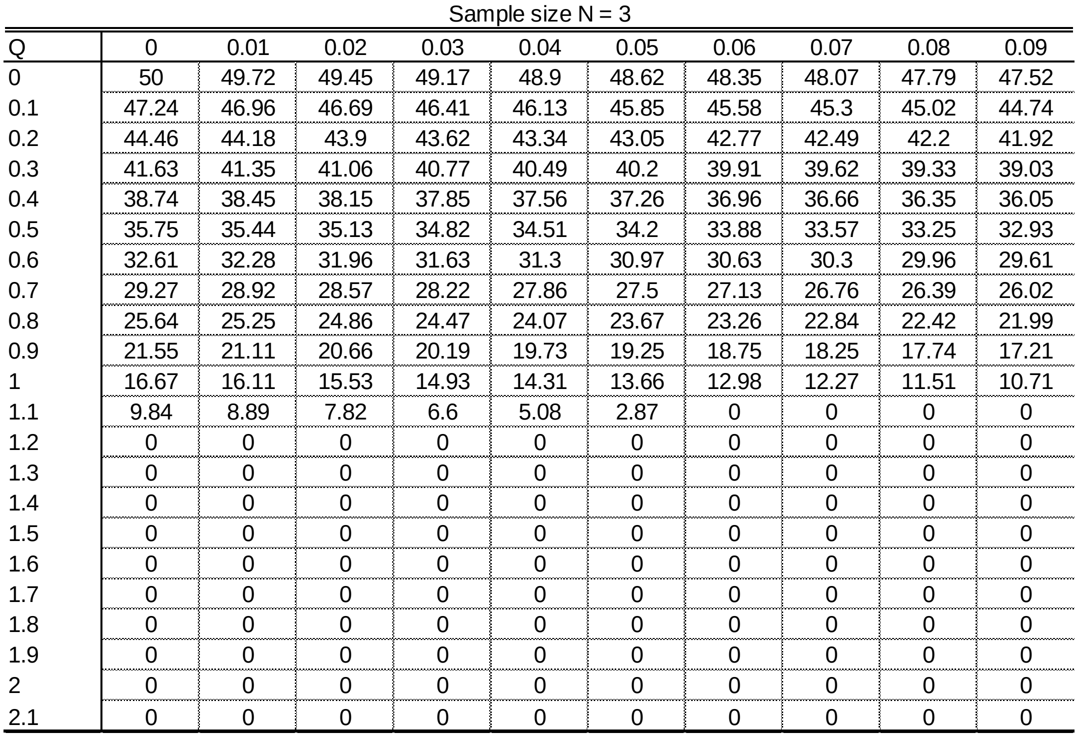
Table 4.1. Specification limits used for laboratory- molded density and in-place air voids for Items 341, 344 and 346.

	Laboratory-molded density		In-place air voids	
Spec Item	USL	LSL	USL	LSL
341	JMF target + 1	JMF target - 1	8.6	3.7
344	JMF target + 1	JMF target - 1	7.6	3.6
346	JMF target + 1	JMF target - 1	7.1	3.6

Step 3. Estimate percent deficient

For the calculated (Q_U) and (Q_L), upper percent deficient PD_U and lower percent deficient PD_L are estimated from the appropriate percent deficient table for the corresponding sample size. These values can be directly estimated from interpolating the tabulated values or using beta distribution. Note that these calculations are made assuming a symmetric distribution of the test results. The lower percent deficient PD_L is the probability of finding a test result that is less than the LSL. The upper percent deficient PD_U is the probability of finding a test result that is higher than the USL. Some qualitative observations can be made here. For example, if the sample mean is exactly between the upper and lower specification limits then the distribution is symmetric and the PD_U and PD_L values are the same. For a symmetric distribution, if the sample mean is closer to the USL compared to the LSL, then the probability of a test result exceeding the USL is higher than the probability of a test result being lower than the LSL and vice-versa. Finally, for a given location of the sample mean, the PD_L and PD_U values increase as the sample standard deviation increases. An example PD table for a sample size of 3 (3 test results) is shown in Figure 4.2. The total percent deficient PD is then calculated by summing PD_U and PD_L as shown in Equation 4.5.

$$PD = PD_U + PD_L. \tag{4.5}$$



Sample size N = 3

Q	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	50	49.72	49.45	49.17	48.9	48.62	48.35	48.07	47.79	47.52
0.1	47.24	46.96	46.69	46.41	46.13	45.85	45.58	45.3	45.02	44.74
0.2	44.46	44.18	43.9	43.62	43.34	43.05	42.77	42.49	42.2	41.92
0.3	41.63	41.35	41.06	40.77	40.49	40.2	39.91	39.62	39.33	39.03
0.4	38.74	38.45	38.15	37.85	37.56	37.26	36.96	36.66	36.35	36.05
0.5	35.75	35.44	35.13	34.82	34.51	34.2	33.88	33.57	33.25	32.93
0.6	32.61	32.28	31.96	31.63	31.3	30.97	30.63	30.3	29.96	29.61
0.7	29.27	28.92	28.57	28.22	27.86	27.5	27.13	26.76	26.39	26.02
0.8	25.64	25.25	24.86	24.47	24.07	23.67	23.26	22.84	22.42	21.99
0.9	21.55	21.11	20.66	20.19	19.73	19.25	18.75	18.25	17.74	17.21
1	16.67	16.11	15.53	14.93	14.31	13.66	12.98	12.27	11.51	10.71
1.1	9.84	8.89	7.82	6.6	5.08	2.87	0	0	0	0
1.2	0	0	0	0	0	0	0	0	0	0
1.3	0	0	0	0	0	0	0	0	0	0
1.4	0	0	0	0	0	0	0	0	0	0
1.5	0	0	0	0	0	0	0	0	0	0
1.6	0	0	0	0	0	0	0	0	0	0
1.7	0	0	0	0	0	0	0	0	0	0
1.8	0	0	0	0	0	0	0	0	0	0
1.9	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
2.1	0	0	0	0	0	0	0	0	0	0

Figure 4.2. Percent deficient table for sample size, $n = 3$.

Step 4. Calculate percent within limit

PWL is then calculated from the total percent deficient by subtracting it from 100 as shown in Equation 4.6. For the same value of the sample mean, as the sample standard deviation increases, the *PD* increases and the PWL values decreases thus reflecting not just the sample average but also the sample variability.

$$PWL = 100 - PD. \tag{4.6}$$

Step 5. Calculate payment adjustment factor

The payment adjustment factor is then finally calculated according to a payment adjustment factor scheme that is based on the estimated PWL and not on the sample average. Several different models can be used to establish the payment adjustment factor scheme. The payment adjustment factor scheme can also have a rejectable and acceptable quality levels; the production and placement will be rejected below the rejectable quality level or RQL and will receive full pay or more above the acceptable quality level or AQL. Setting up these levels and establishing the payment adjustment factor schemes are a crucial part of achieving PWL based QA approach and are discussed in more detail in the subsequent chapter. For demonstration purposes, Equation 4.7 is used to compute the payment adjustment factor using the PWL determined from the previous step. This scheme is adopted from the OKDOT specification (Equation 1.1) and is only used here as an example payment adjustment factor scheme. If the PWL is less than 50 percent, zero is assigned assuming that the Engineer may not require removal and replacement of the lot.

$$PF = \begin{cases} 0.024 * (PWL) - 0.0001 * (PWL)^2 - 0.35 & PWL \geq 50 \\ 0 & PWL < 50 \end{cases}$$

(4.7)

4.3 SUMMARY

The model shown here is just a preliminary model and various aspects of PWL need to be considered while adopting a robust payment adjustment factor scheme. These aspects, including establishing the sample size, AQL, and RQL are discussed in more detail in the following chapter. In addition, the following items must also be considered during the final adoption of this method.

- One of the main factors that affect the results from PWL analysis is the calculation and rounding procedure used. When the payment adjustment factor is calculated based on PWL, this issue can be a point of conflict. This is because the values from the tables can be rounded up, rounded down, or interpolated. A Contractor usually prefers the rounding up because it increases the PWL achieved. The method of rounding must be specified prior to implementation to ensure that there are no conflicts regarding this issue. As a result, it is important to stipulate the calculation process, including number of decimal places to be carried in the calculations, as well as the exact manner in which the PWL is determined (interpolation from table or computation from beta distribution formula).

- It is also important to identify outliers that may penalize the Contractor unexpectedly. Clear methods to identify outliers need to be established.
- Transitioning the use of two existing metrics or quality characteristics, i.e. laboratory-molded density and in-place air voids from a PAF scheme that is based on average test results to a PAF scheme that is based on PWL is a significant shift in contracting practice and will require substantial effort for implementation. However, the agency may also consider the possibility of incorporating other characteristics such as asphalt content, aggregate gradation, and VMA in calculating payment adjustment factors using PWL in the future.
- It is crucial to establish clear methods of calculating PAF from PWL, (e.g. the form of equation or interpolation table) and averaging technique to obtain a combined payment adjustment factor, that complement the existing specification requirement of material production and placement.

CHAPTER 5. PAYMENT ADJUSTMENT FACTOR SCHEMES, RISK ANALYSIS, AND COMPARISON WITH CURRENT PRACTICE

5.1 OVERVIEW

The previous chapter reviewed PWL as a quality measure for two existing payment criteria, laboratory-molded density and the in-place air voids. As summarized in Figure 4.1, PWL accounts for both mean ($\bar{X}$) and standard deviation (s) of a normally distributed sample to measure upper and lower quality indices, Q_U and Q_L. These quality indices essentially measure the deviation of the mean from the upper specification limit, USL, and the lower specification limit, LSL, normalized by the standard deviation. These indices, along with percent deficient or PD tables for a specific sample size, are then used to determine upper and lower percent deficient, PD_U and PD_L values. These values are combined to calculate PWL, an estimate of the percentage of material or construction that is within the specification limits. This PWL can then be utilized to determine payment adjustment factor or PAF using an appropriate PAF scheme specified by the agency.

In this chapter, several PAF schemes have been proposed that the agency can adopt depending on their requirements and goals. Risks associated with different quality levels have also been analyzed. Finally, several examples with different specification limits have been presented. In these examples, a PWL based PAF scheme has been applied to past TxDOT projects to demonstrate the effect of such PAF scheme on bonuses and penalties associated with HMA production and placement.

5.2 PAF SCHEME

The AASHTO R-9, "Standard Recommended Practice for Acceptance Sampling Plans for Highway Construction" [1], suggests two provisions for an effective payment adjustment factor scheme when using PWL as a quality measure:

(1) The PAF should be 1.00 (100 percent) when the calculated value of PWL is exactly at the AQL.

(2) For the average pay to be 1.00 at the AQL there must be an incentive that allows pay above 1.00 to offset for lower PAFs from estimated quality levels below the AQL.

In the context of laboratory-molded density and in-place air voids, the PWL may be computed on a per lot basis by pooling data from sublots (lot based PWL) or PWL may be computed for a project by pooling data from all lots and sublots (project based PWL). The examples in this chapter are based on a lot based PWL. However, it must be recognized that when PWL is computed on a project basis, some lots may exceed AQL while others may not. In a scenario where there is no provision for an incentive for PWL values that exceed AQL, even a single lot with a PWL below the AQL will result in an average PAF that is less than 1.00.

A survey, conducted as a part of NCHRP 10-79 study, among 37 state highway agencies showed that, 31 states use incentives ranging from 1% to 15%, 18 of which use a maximum incentive of 5%, i.e. a maximum PAF of 1.05 [5]. Typically, the 15% incentives are restricted to ride quality. The AASHTO R-9 advocates a continuous pay scheme as shown in equation 5.1 with a straight-line relationship between the PAF and PWL. This equation, as plotted in Figure 5.1, yields 105% pay at 100 PWL, 100% pay at 90 PWL (AQL), and 80% pay at 50 PWL (RQL).

$$PAF = 0.55 + 0.005PWL \tag{5.1}$$

where,
PAF = payment adjustment factor
PWL = estimated percent within limits.

A continuous PAF scheme, similar to Equation 5.1 is more popular than a stepped or tabular PAF schemes because of its straightforward and continuous nature. It also helps avoid disputes that may arise when the estimated quality level falls short on one side of a large step in a stepped PAF scheme. Similar other linear continuous payment schemes are provided in Equations 5.2 - 5.3, and the estimated PAF plots for these PAF schemes are presented in Figure 5.2. Equation 5.2 provides 50% pay for an RQL of 50 and yields a lower PAF than 1.00 for an AQL of 90, which does not satisfy the second AASHTO R-9 pay provision. If a 100% pay, conforming to the AASHTO R-9 pay provision, is adopted for AQL = 90, then a single linear PAF scheme, Equation 5.3 with a PAF of 0.6 for RQL = 50, yields a bonus of 10% for 100 PWL.

$$PAF = 0.011PWL - 0.05 \tag{5.2}$$

$$PAF = 0.01PWL + 0.1 \tag{5.3}$$

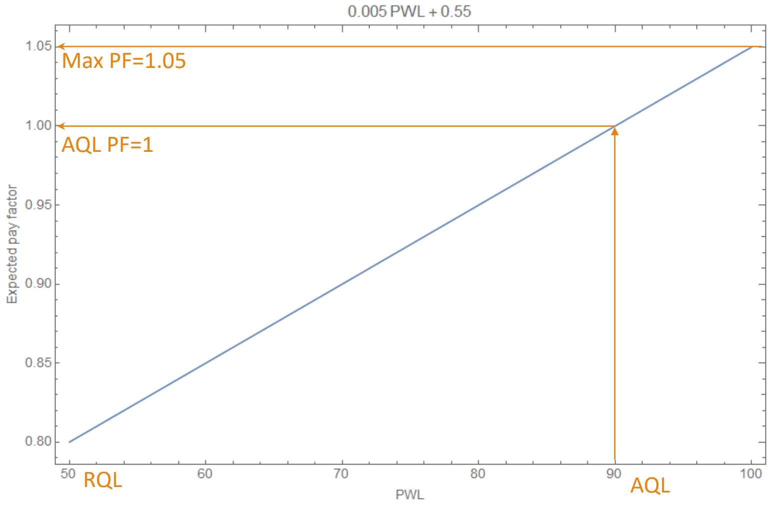


Figure 5.1. AASHTO R-9 recommended PAF scheme.

However, if the agency decides to limit the maximum pay at a certain amount, for instance, the existing 105% pay for Item 341, a quadratic equation presenting a curvilinear relationship between the PAF and PWL can be utilized. An example of such a relationship would be the Oklahoma DOT (ODOT) PAF scheme, Equation 5.4, as presented in Figure 4.1. The corresponding expected PAF plot is provided in Figure 5.3. A quadratic equation, similar to Equation 5.4, penalizes the Contractor less when the estimated PWL is close to AQL than that close to RQL.

An alternative to a quadratic relationship is a series of straight-line PAF equations with different slopes that define different disincentives as indicated in Equations 5.5 and 5.6. The expected PAF plot for such a pay system is shown in Figure 5.4. This system essentially replaces the quadratic PAF scheme with a couple of straight lines with two different slopes and yields the same PAFs for RQL, AQL and 100 PWL materials as Equation 5.4.

$$PAF = -0.0001PWL^2 + 0.024PWL - 0.35 \tag{5.4}$$

Both the quadratic PAF equation and series of linear PAF equations make the disincentive harsh as the quality level decreases with high deviation from the target specified limit. Should a linear, quadratic, or series of linear PAF schemes be adopted, risk analysis to various sample size and quality levels need to be performed to help make informed decisions.

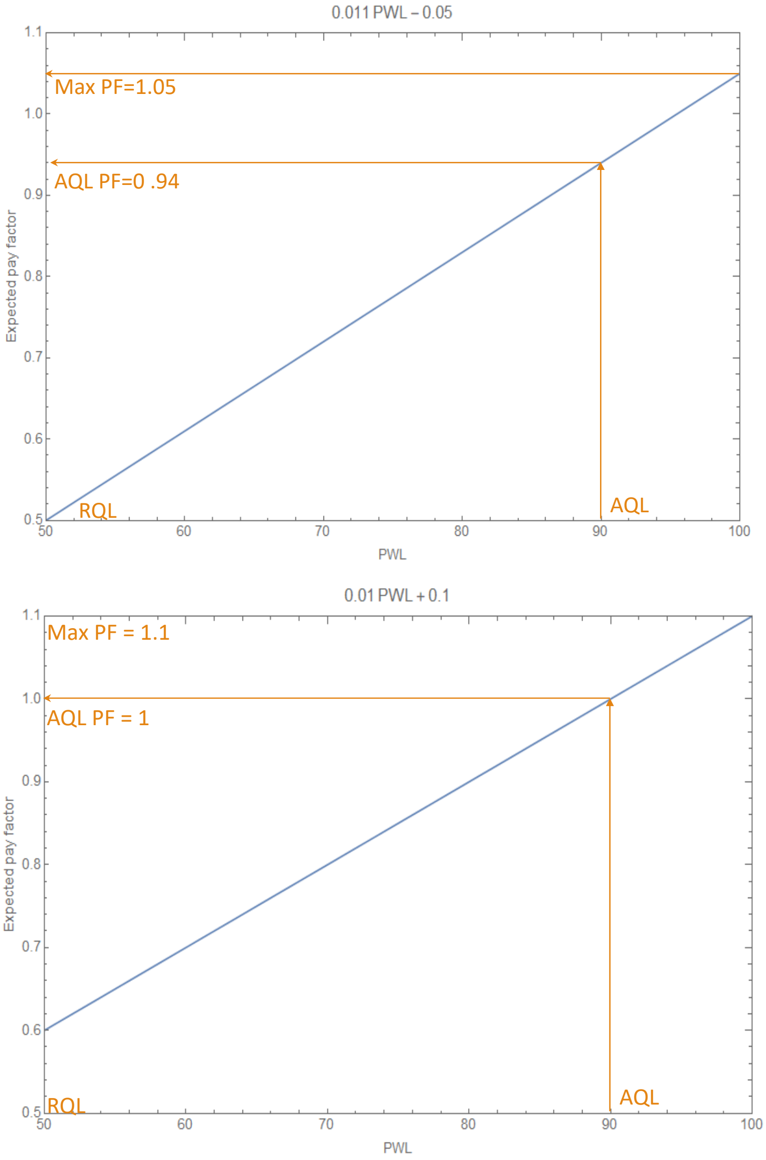


Figure 5.2. Linear PAF scheme- the rate of change in disincentives with the change in PWL is constant.

$$PWL \geq (AQL = 90) : PAF = 0.005PWL + 0.55 \tag{5.5}$$

$$(RQL = 50) \geq (PWL < AQL) = 90 : PAF = 0.05PWL + 0.1 \tag{5.6}$$

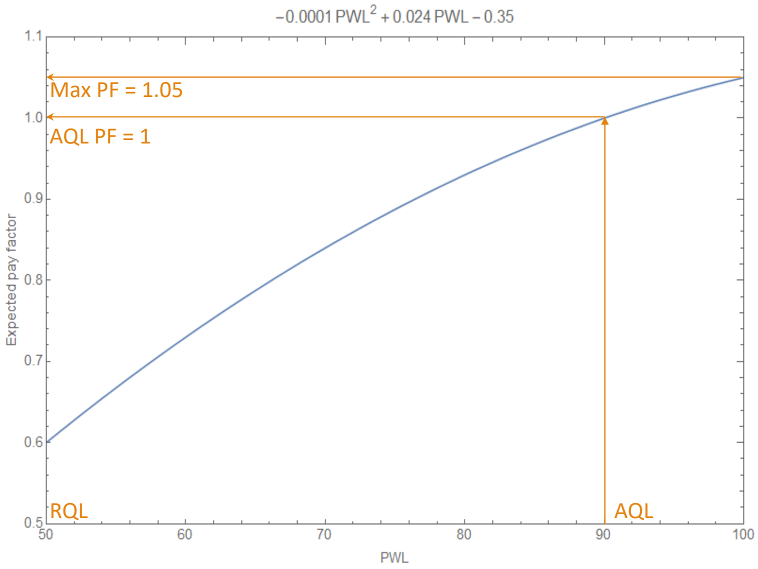


Figure 5.3. Quadratic PAF scheme adopted by Oklahoma DOT with low disincentives for quality levels near AQL that increase by higher rates as the quality approaches AQL.

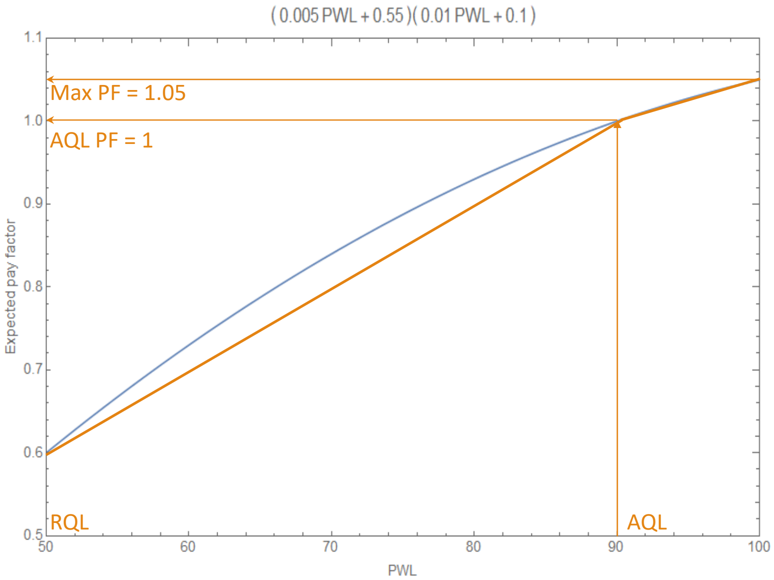


Figure 5.4. A series of linear PAF schemes to facilitate different disincentive rates for different levels of PWL.

5.3 RISK EVALUATION

Queried data from the SiteManager database reveal that sample size for the laboratory-molded density tests can be anywhere from 3 to 12 and between 2 to 8 for in-place air void tests for the production and placement PAF calculations, respectively. It is important to note (and perhaps address in near future) that samples are randomly collected from each sublot, and sometimes the random sample location may fall on one of the areas that may not be subject to testing. Cores, for instance, are not taken from "Miscellaneous Areas" or locations that are designated on plans as areas not subject to in-place air void determination, and a PAF of 1.000 is assigned to that sublot. Examples of "Miscellaneous Areas" are areas that typically involve significant handwork, such as temporary detours, driveways, crossovers, spot level-up areas, and similar areas. When implementing PWL as a measure of quality, it is important to analyze the effect of sample size on the PAF calculation. One way to determine this effect would be analyzing risk associated with the acceptable and rejectable quality levels specified for an acceptance plan. Furthermore, determination of the limits for the acceptable and rejectable materials is crucial to ensure construction quality and to provide sufficient opportunity for contractors to achieve this quality. Too restrictive limits may deprive the contractor of a reasonable scope of meeting the specification, whereas too lenient limits may not be effective in achieving the desired levels of quality and consistency. To ensure the acceptable and rejectable quality levels serve both these purposes, risks associated with the proposed schemes to the agency and contractors must be evaluated and weighed against each other.

The two types of risks associated with an acceptance plan are the seller's risk and the buyer's risk. Seller's risk (α) or Type I risk, according to TRB glossary, is the probability that an acceptance plan will erroneously reject acceptable quality level (AQL) material or construction with respect to a single acceptance quality characteristic. It is the risk the contractor takes in getting rejected for producing a material or placing it on site, which should otherwise be acceptable.

The TRB glossary defines the buyer's risk (β) or the Type II risk as the probability that an acceptance plan will erroneously fully accept (100 percent or greater) rejectable quality level (RQL) material or construction with respect to a single acceptance quality characteristic. It is the risk the agency takes in accepting a material or construction which should have otherwise been rejected.

The α and β risk levels depend on the material and construction process that are involved. AASHTO R-9 suggests risk levels of $\alpha = 0.05$ and $\beta = 0.005$ for critical construction practices where superior quality material and construction are required for the long term performance of the structure. While α and β risks define the risks of rejecting an AQL material or accepting an RQL material, operating characteristic (OC) curves provide indications of the risk over a wide range of quality levels. According to the TRB glossary, OC curve is a graphical representation of an acceptance plan that shows the relationship between the actual quality of a lot and either (1) the probability of its acceptance (for accept/reject acceptance plan) or (2) the probability of its acceptance at various payment levels (for acceptance plans that include pay adjustment provisions). OC curves for various sample sizes (n) are plotted in Figure 5.5, where, out of the plotted sample size, one test result is allowed to fall outside the specification limits. According to this figure, one can clearly see that for the same RQL of 50%, as the sample size decreases from 12 to 3, the risk of fully accepting a rejectable material, as depicted by β, increases significantly from 0.3% to 50%. It is also worth noting that to attain the AASHTO R-9 recommended risk level of 0.005 for critical construction at a RQL of 50%, at least 12 samples are needed for testing. The contractor's risk, α, on the other hand increases as sample size increases. Hence, it is recommended to determine an optimum sample size that will serve both parties with the lowest risk possible. Both of these risks need to be analyzed carefully before the agency decides on the minimum sample size required to achieve target acceptable and rejectable quality levels.

Figure 5.6 shows influence of accepting a maximum of 0, 1, 2, or 3 defective samples on the risk that the agency takes in accepting an RQL material. As more test results fall outside the specification limits, the higher risk the agency takes in erroneously accepting rejectable materials. To keep risk level within the AASHTO R-9 recommended value, only one out of twelve samples can be allowed to fall outside the operational tolerances.

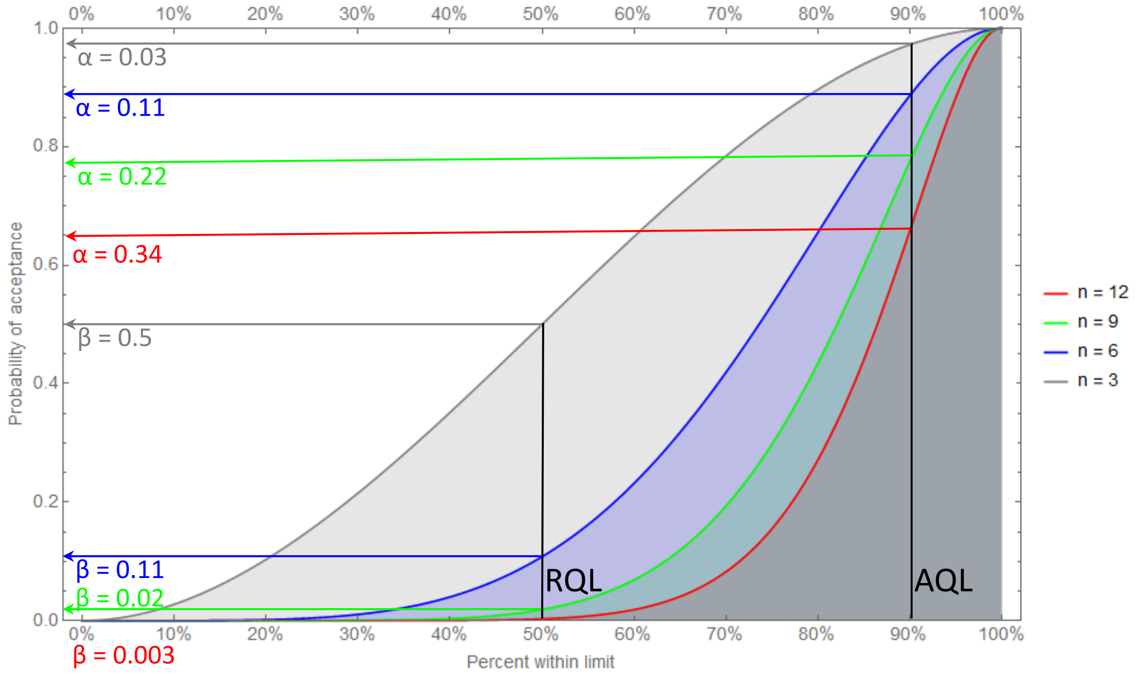


Figure 5.5. Agency's risk (β) increases with decreasing sample size.

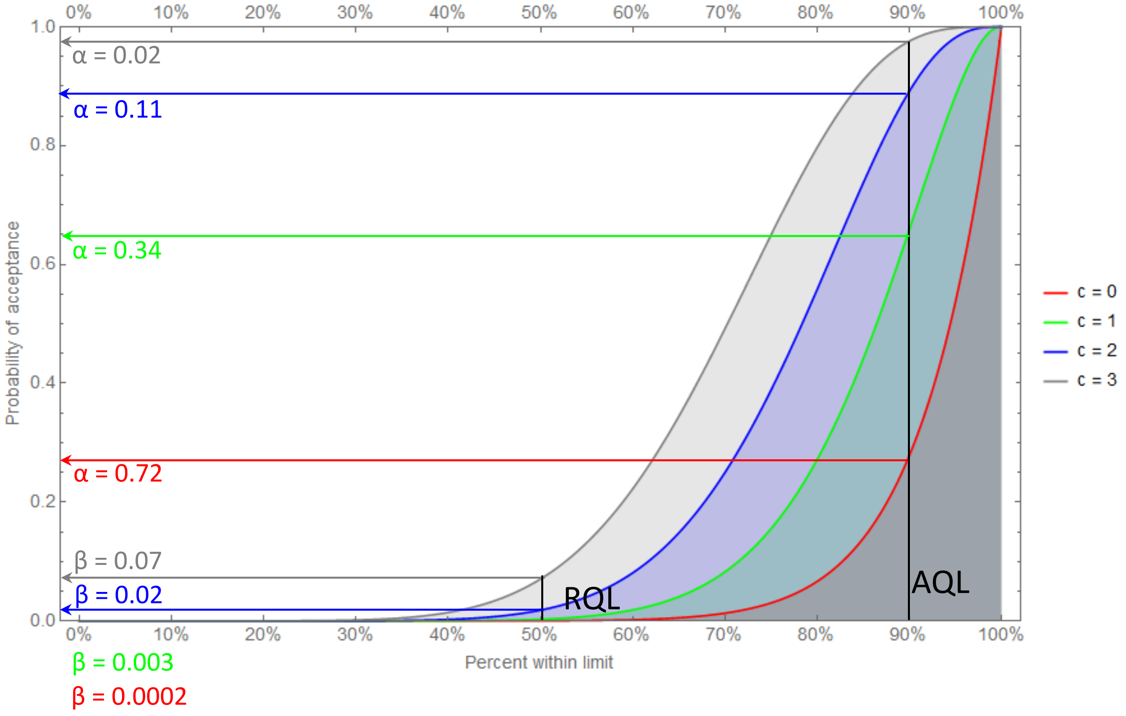


Figure 5.6. Agency's risk (β) increases with an increasing number of defective samples allowed.

5.4 PRELIMINARY INVESTIGATION OF EXISTING DATA USING PWL BASED PAF SCHEMES

To apply PWL, laboratory-molded density and in-place air void test results for past projects were queried. As the first example, a lot with twelve laboratory-molded density test results, as presented in Figure 5.7, was selected to calculate the production PAF. One of the reasons for selecting this specific lot was that the production was been approved for a bonus pay of 2.4% based on existing specifications while data showed a high standard deviation of 1.254 for the laboratory-molded density. In particular, the density test results for sublot 1 varied substantially from 95.7% to an unrealistic value of 100.6%. The current practice determines the average laboratory-molded density values for each sublot, calculates absolute deviation from the target, assigns a PAF to each sublot from a tabulated PAF scheme based on the absolute deviation, and finally takes the average over all sublots to calculate the overall production PAF for that specific lot. As a consequence, despite the high variability in the measured values, the calculations resulted in a favorable average density estimate of 97.8%, placing it slightly above the queried target density of 97%. The TxDOT 2014 specification for Item 341 also prohibits production and requires immediate corrective action when the Engineer's laboratory-molded density on any sublot falls outside ±1 of the target value. Even with two test results falling outside the existing operational tolerances, the average for sublot 1, in this example, resulted in an absolute deviation of only 0.8, and earned a bonus pay of 1.3% (i.e. PAF of 1.013).

In order to demonstrate the ability of PWL to overcome the aforementioned limitation, the same data set was analyzed using a PWL based PAF scheme. The existing specification tolerances of JMF target ±1 were used as the upper and lower specification limitsn (USL and LSL), as shown in Figure 5.8. The upper and lower quality indices Q_U and Q_L were calculated to be 0.55 and 1.04, respectively. In other words, these values imply that the mean for this particular lot departs from the upper specification limit (JMF target + 1) by about 0.55 standard deviation and lower specific limit (JMF target - 1) by about 1.04 standard deviation. Using these quality indices and a percent deficient or PD table for a sample size of 12, a PD_U of 29 and PD_L of 14, were obtained, i.e. the probability of finding a test result that was greater than the USL is 29% and the probability of finding a test result that was less than the LSL is 14%. It should be noted here that the estimated value of PD was rounded down to the nearest integer in Contractor's favor. These two PD values were then combined to yield an overall PWL of 57 for the entire lot, i.e. the probability of finding a test result that was within the LSL and USL bounds was 57%. Finally, this PWL value was inserted in a PAF scheme (equation 5.4), and a PAF of 0.69 is obtained. Note that the aforementioned PAF was computed using just one of many different possible PAF schemes. Thus, the PWL approach accounts for both the average and standard deviation to calculate the probability that test results corresponding to the lot are within the specification limits and penalizes construction practices that deviate away from these limits. This high penalty of 31% compared to the bonus of 2.4% based on the current practice exemplifies the need to reevaluate the current practice that is based on the average values without incorporating the influence of variability. In other words, production or placement with very high variability can still result in average values that are amenable to receiving a bonus. Moreover, as discussed in the previous section (Figure 5.6), with three test results being outside the specification limits, the agency takes a high risk of accepting this lot. The risk of erroneously accepting such a lot with three defective samples is about 7%, far exceeding the AASHTO R-9 recommended risk level of 0.5% for critical materials. This risk was calculated based on all twelve samples including the test results that fell outside the specification limits.

Another example is provided in Figure 5.9, where the specification limits were made more lenient by lowering the lower specification limit (LSL) to 95% from 96%, according to the operational tolerances of 2014 specification 6.1 for Item 341, production payment adjustment factors. When this specification limit was applied, the resulting PWL increased to 69 resulting in a higher PAF of 0.83 (i.e. a penalty of 17%) than that obtained previously and shown in Figure 5.8, but still avoided rewarding a bonus of 2.4%.

A close examination of the data set also revealed that almost all values were clustered between 95.7 and 98.1 except 100.6, which might be an outlier. A single outlier can be identified using the test criterion T_c according to ASTM E178-16a [2]. Per this test, T_n is calculated as

$$T_n = \frac{|X - \bar{X}|}{s}, \tag{5.7}$$

since the value could lie on either side of the mean. T_n is then compared to the critical value T_c, for a specific sample size and significance. If T_n is greater than T_c with a specific significance, then that value can be deemed as an outlier by that significant amount. Here, X is the test result under consideration, $\bar{X}$ is the mean, and s is the standard deviation.

For this specific case, T_n was calculated to be

$$T_n = \frac{|100.6 - 97.3|}{1.254} = 2.63. \tag{5.8}$$

From ASTM E178, T_c, for a sample size of 12, was found to be 2.55 with 1% significance, which was smaller than T_n. That is, there was a 99% chance that 100.6 may be an outlier and reasonably eliminated from the data set. In this case, the sample size was reduced to 11, and the new PWL calculation for this sample size and dataset was shown in Figure 5.10. According to this analysis, the new PWL excluding the outlier was found to be 84, resulting in a PAF of 0.96 (i.e. penalty of 4%), in contrast to the bonus of 2.4% awarded to this lot.

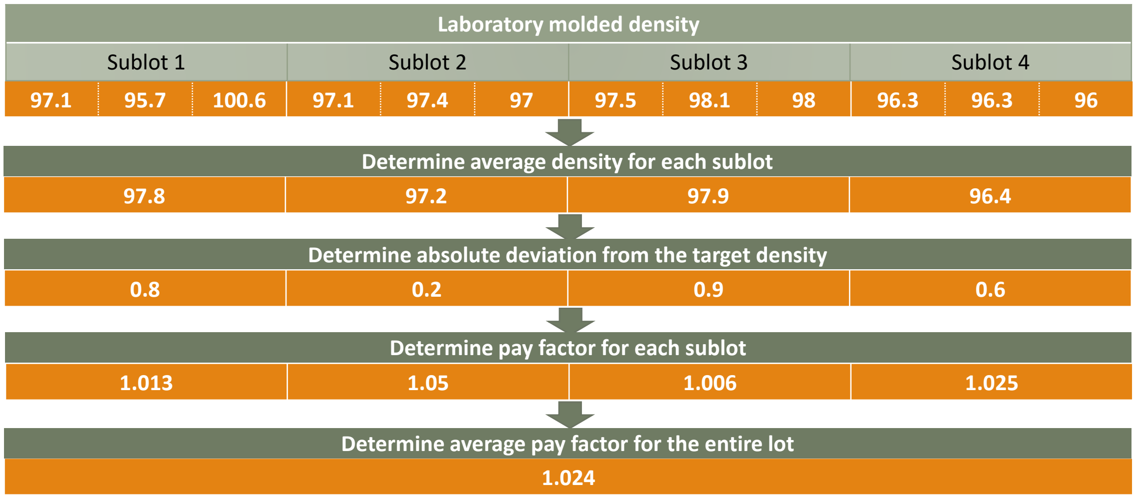


Figure 5.7. Example showing the current specification for the determination of production payment adjustment factor (PAF) for Item 341. The production PAF is determined for each sublot based on the average absolute deviation from the target laboratory-density. These PAFs for all the sublots are then averaged to obtain a production PAF of 1.024 for this lot, with a bonus of 2.4%.

Another example for production PAF is shown in Figures 5.11 and 5.12. According to this example, PWL yielded a penalty of 21% to a lot that was approved for a full pay. According to the risk analysis of Figure 5.5, the agency took 30% risk of accepting this lot with a full pay (sample size 6). This risk would be much lower if additional samples were tested. Both of these examples show that the use of either mean (e.g. in-place air void) or the average absolute deviation from the target quality characteristic (e.g. laboratory-molded density) allows test results outside the operational tolerances, and possibly rewards materials and construction practices with full and/or bonus pay based on the average test results even when the variability in the quality of the material is very high.

Figures 5.13 and 5.14 show the application of the PWL approach to placement PAF. Current specification requires that a maximum of eight test samples, taken from four sublots (two from each sublot), are tested. Individual PAF for each sublot is then determined from Table 17 of TxDOT 2014 specification based on the average in-place air void test results. HMA placement with air voids greater than 9.9 and less that 2.7 are entitled to removal and replacement or be left in place without payment. Similar to previous examples, this lot also had two results that are outside the operational tolerances, but was eligible for a bonus pay of 1.2%. In order to compute the PWL, upper and lower specification limits were taken to be 3.8 and 8.5 for Item 341, since these were the two limiting values that qualified for full pay per the current specification. The PWL calculation for this lot is shown in Figure 5.14. For a sample size of 8, PD_U and PD_L were found to be 43 and 15, respectively. These high PD values, in turn, resulted in a PWL of 42. It is customary to reject a lot that has more than 50% material failing the specification limits. If a PWL value of 50 was specified to be the rejectable quality level (RQL), then this lot would be subject to removal and replacement or be left with zero pay, and not a bonus of 1.2% (i.e. PAF of 1.012).

Figures 5.15 and 5.16 compare the existing quality measure and PWL for HMA placement. In contrast to the previous examples, this example suggests that PWL can also reward materials or construction practices that are consistent in performance. If the PWL and hypothetical PAF scheme described above were implemented, this lot would be awarded a bonus of 5% for very consistent quality characteristics, which was higher than the bonus of 3.9% (i.e. PAF of 1.039) that was assigned to this lot.

It must be emphasized that the examples provided above are intended to demonstrate the potential gap when only average and not the variability of the material produced or placed is considered in the calculation of the PAF. This existing approach can sometimes result in payment bonuses to the contractors even when the variability is high. Also, in some cases, the existing approach may not provide a contractor with adequate bonus for producing and placing a material with a very high degree of consistency.

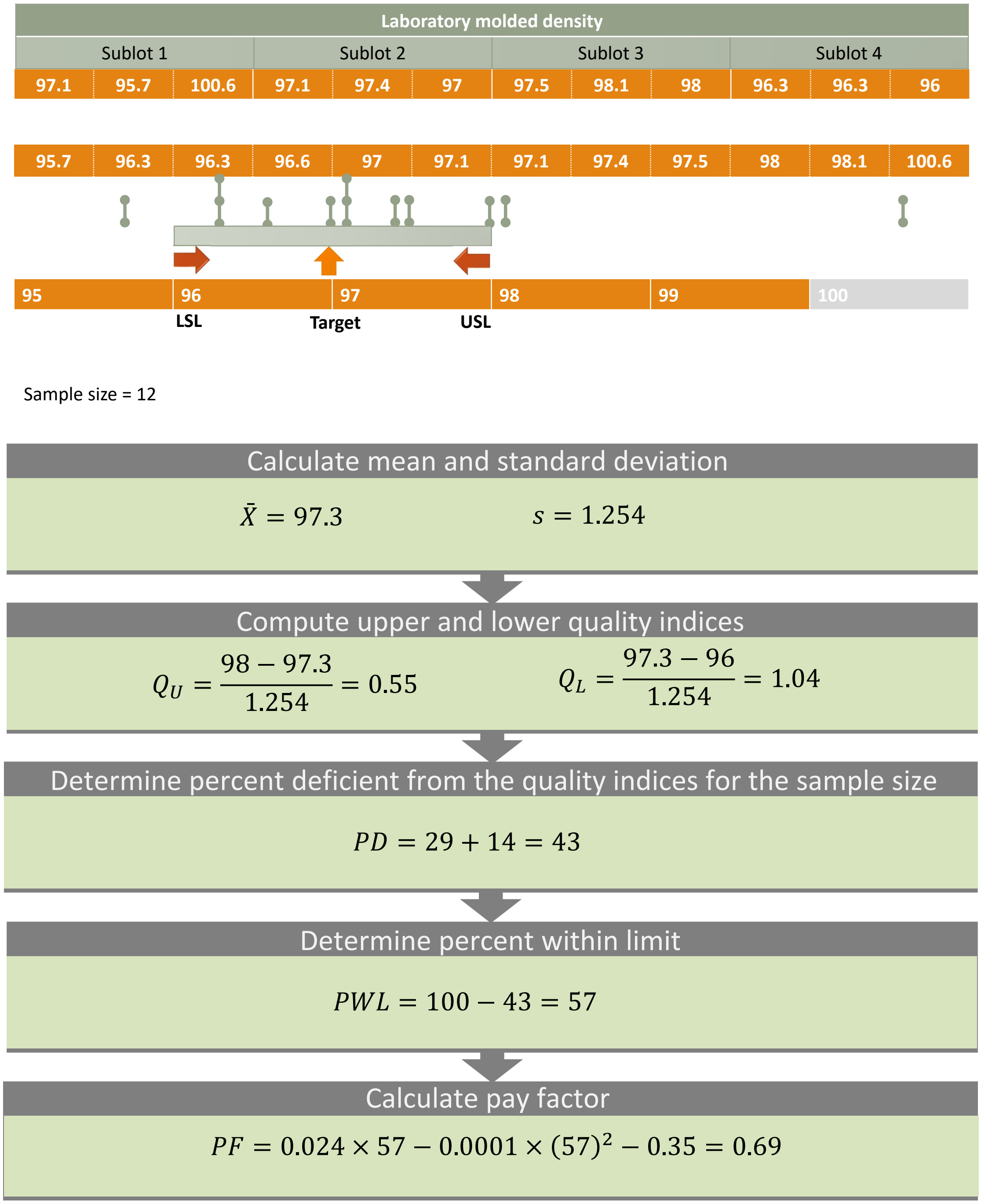
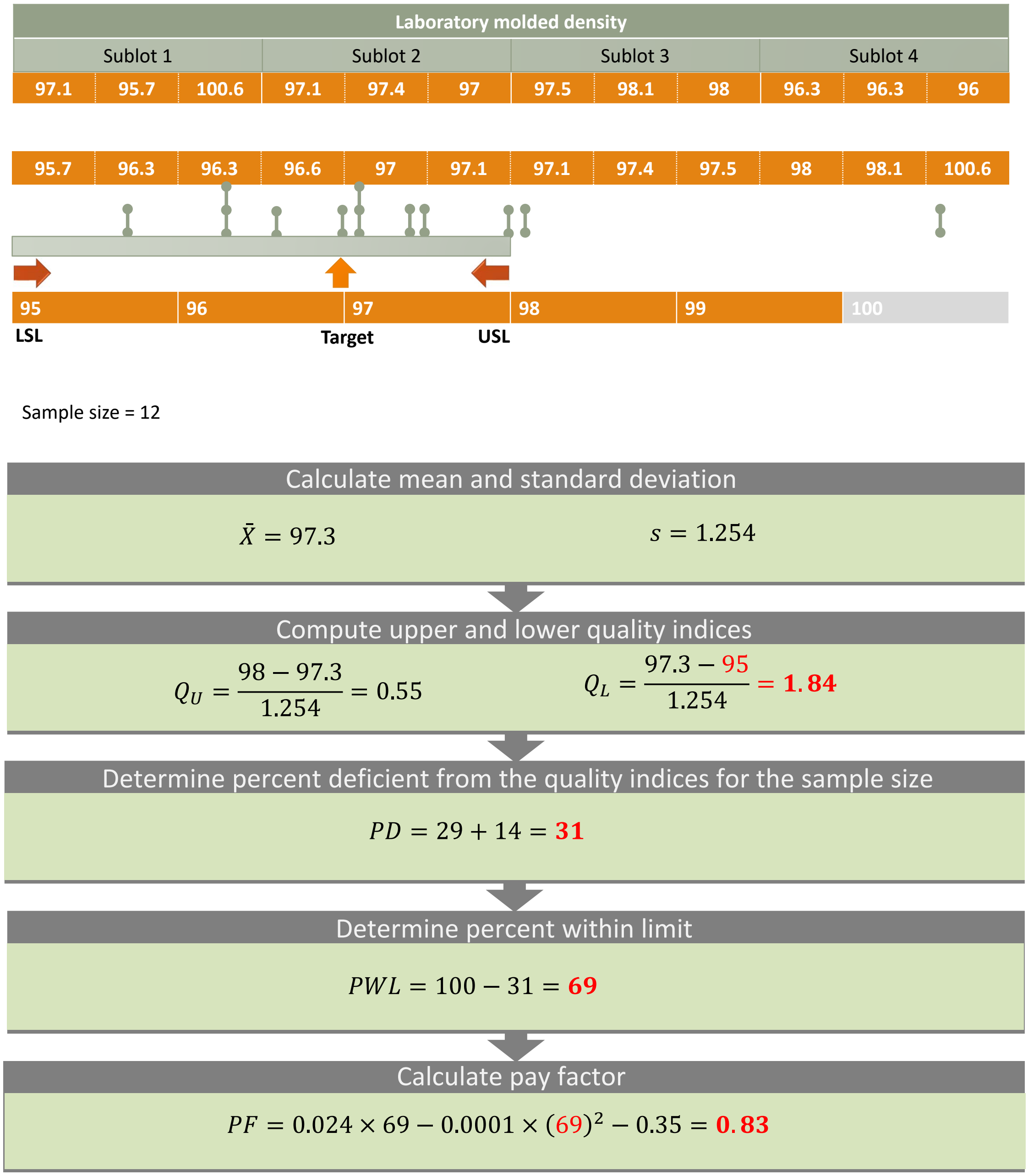


Figure 5.8. PWL applied to a lot's production quality characteristic presented in Figure 5.7. The PWL approach takes the average value and standard deviation into account to statistically determine the variability of the material produced. The PWL for this particular lot was estimated to be 69 which was below the AQL threshold of 90 and qualified for penalty rather a bonus of 2.4%.



Calculate mean and standard deviation

$\bar{X} = 97.3$ $s = 1.254$

Compute upper and lower quality indices

$Q_U = \frac{98 - 97.3}{1.254} = 0.55$ $Q_L = \frac{97.3 - 95}{1.254} = \mathbf{1.84}$

Determine percent deficient from the quality indices for the sample size

$PD = 29 + 14 = \mathbf{31}$

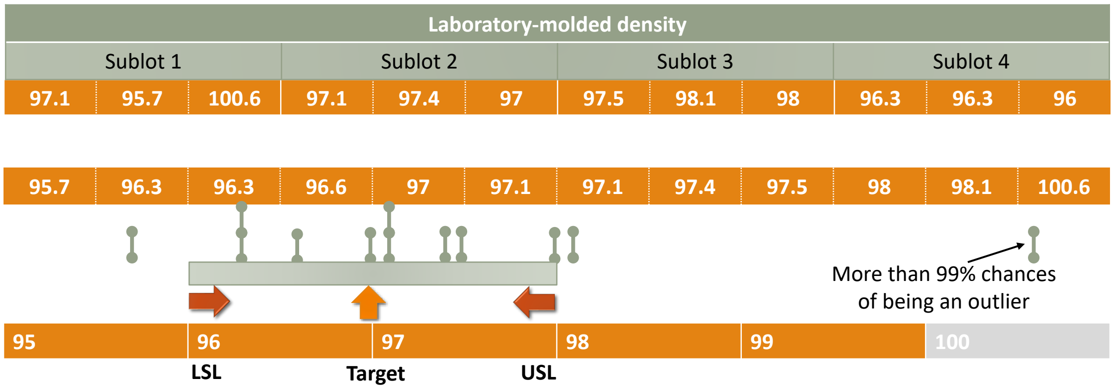
Determine percent within limit

$PWL = 100 - 31 = \mathbf{69}$

Calculate pay factor

$PF = 0.024 \times 69 - 0.0001 \times (69)^2 - 0.35 = \mathbf{0.83}$

Figure 5.9. PWL applied to a lot production quality characteristic presented in Figure 5.7 with wider specification limits than Figure 5.8. Even with wider specification limits the the existing practice, the PWL was calculated to be 69, which was way below the AASHTO recommended threshold for acceptable quality level.



Sample size = 11

Calculate mean and standard deviation

$\bar{X} = \mathbf{97.0}$ $\quad s = \mathbf{0.740}$

Compute upper and lower quality indices

$Q_U = \frac{98 - 97}{0.74} = \mathbf{1.34}$ $\quad Q_L = \frac{97 - 96}{0.74} = \mathbf{1.36}$

Determine percent deficient from the quality indices for the sample size

$PD = 8 + 8 = \mathbf{16}$

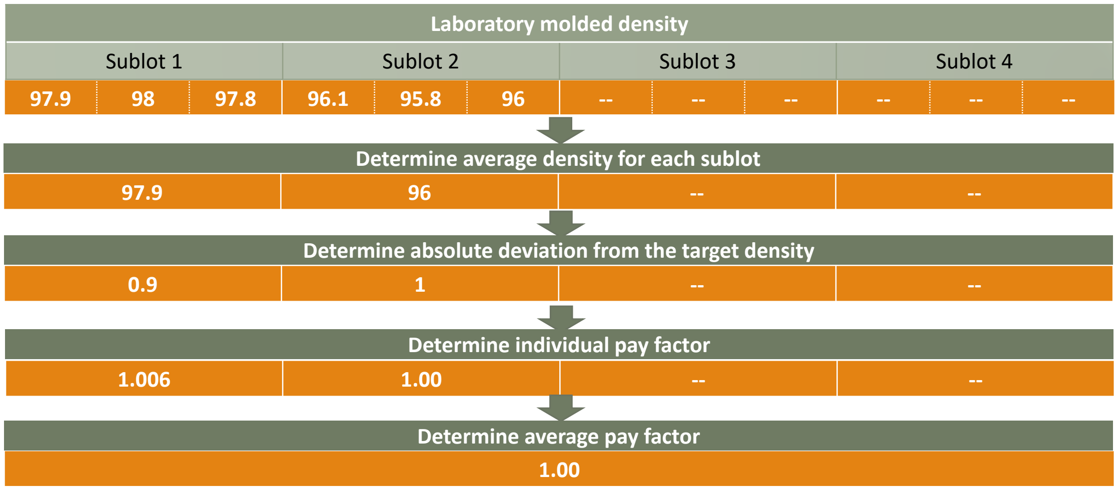
Determine the percent within limit

$PWL = 100 - 16 = \mathbf{84}$

Calculate pay factor

$PF = 0.024 \times 84 - 0.0001 \times (84)^2 - 0.35 = \mathbf{0.96}$

Figure 5.10. PWL calculation for the same example shown in Figure 5.8 but excluding the outlier. Since the third test value (100.6) from sublot 1 was determined to be outlier, it was excluded from the PWL estimation.



Laboratory molded density

Sublot 1			Sublot 2			Sublot 3			Sublot 4		
97.9	98	97.8	96.1	95.8	96	--	--	--	--	--	--

Determine average density for each sublot

97.9	96	--	--

Determine absolute deviation from the target density

0.9	1	--	--

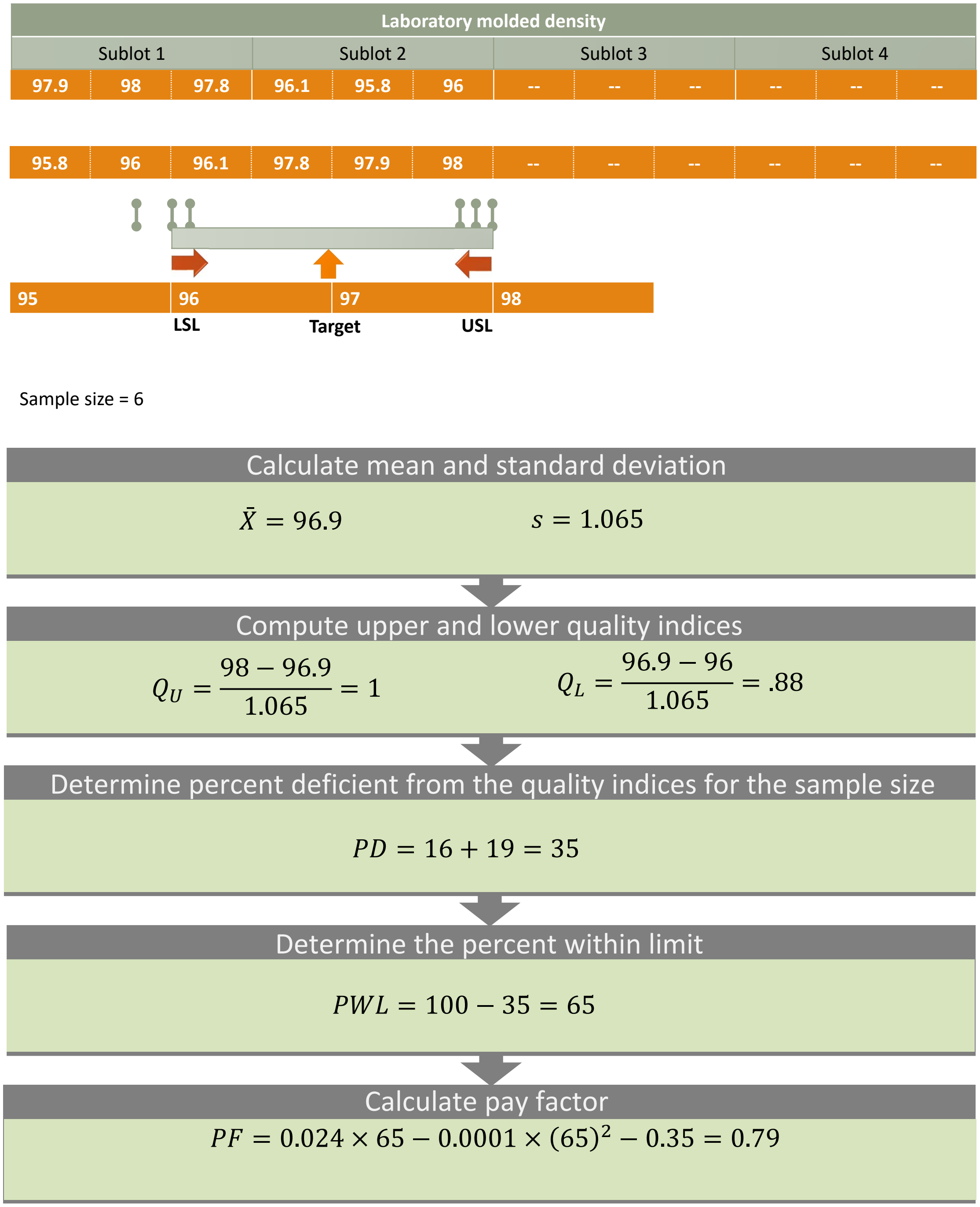
Determine individual pay factor

1.006	1.00	--	--

Determine average pay factor

1.00

Figure 5.11. Example showing the calculation for the existing production payment adjustment factor (PAF) for test results obtained from two sublots (total six samples). According to the queried data and calculation shown here, this lot received a full pay.



Calculate mean and standard deviation

$\bar{X} = 96.9$ $s = 1.065$

Compute upper and lower quality indices

$Q_U = \frac{98 - 96.9}{1.065} = 1$ $Q_L = \frac{96.9 - 96}{1.065} = .88$

Determine percent deficient from the quality indices for the sample size

$PD = 16 + 19 = 35$

Determine the percent within limit

$PWL = 100 - 35 = 65$

Calculate pay factor

$PF = 0.024 \times 65 - 0.0001 \times (65)^2 - 0.35 = 0.79$

Figure 5.12. PWL applied to the lot presented in Figure 5.11. A standard deviation (s) of 1.065 resulted in a low PWL of 65 which falls way below the acceptable quality level (AQL) of 90, recommended by AASHTO.

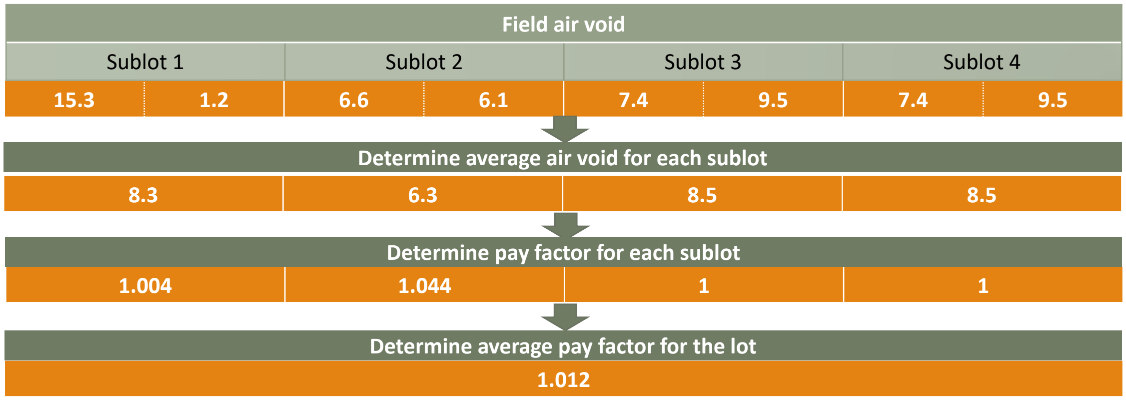
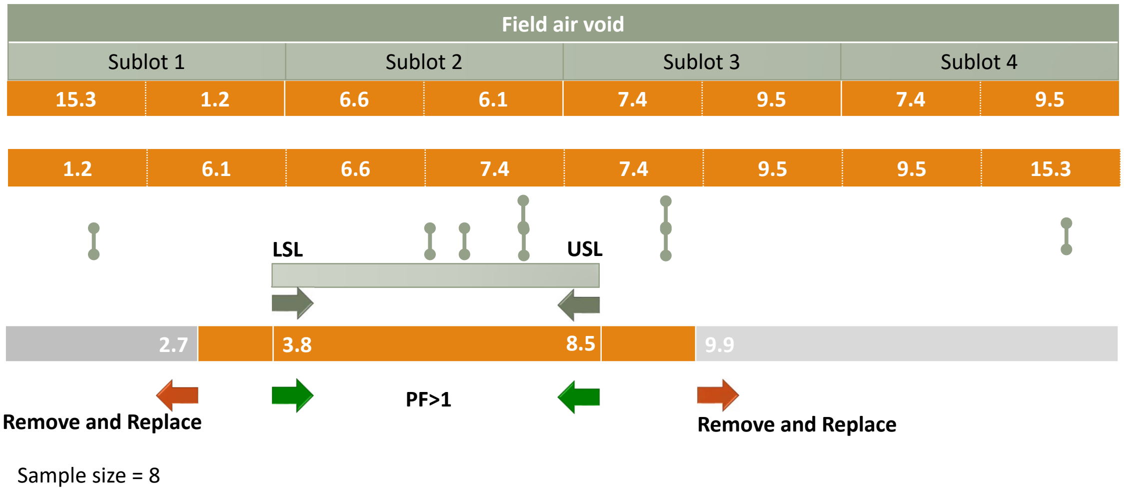


Figure 5.13. Current practice for the determination of placement payment adjustment factor (PAF) based on the average in-place air voids. The placement PAF iss determined for the average air void value for each sublot. The lot placement PAF is then calculated by taking the average of the PAFs for all sublots.



Calculate mean and standard deviation

$\bar{X} = 7.875$ $\quad s = 3.967$

Compute upper and lower quality indices

$$Q_U = \frac{8.5 - 7.875}{3.967} = 0.16 \qquad Q_L = \frac{7.875 - 3.8}{3.967} = 1.03$$

Determine percent deficient from the quality indices for the sample size

$$PD = 43 + 15 = 58$$

Determine the percent within limit

$$PWL = 100 - 58 = 42$$

Calculate pay factor

For PWL $\geq$ 50, $PF = 0.024 \times (PWL) - 0.0001 \times (PWL)^2 - 0.35$

For PWL $<$ 50, $PF = 0$ or Remove and replace

Figure 5.14. PWL accounts for variability and penalizes inconsistent test results of Figure 5.13. The current PAF system awarded a bonus of 1.2% with three test results failing the specification limits. Whereas, the PWL approach identified this lot to be rejectable as the estimated PWL (42) fell below the common threshold of rejectable quality level of 50.

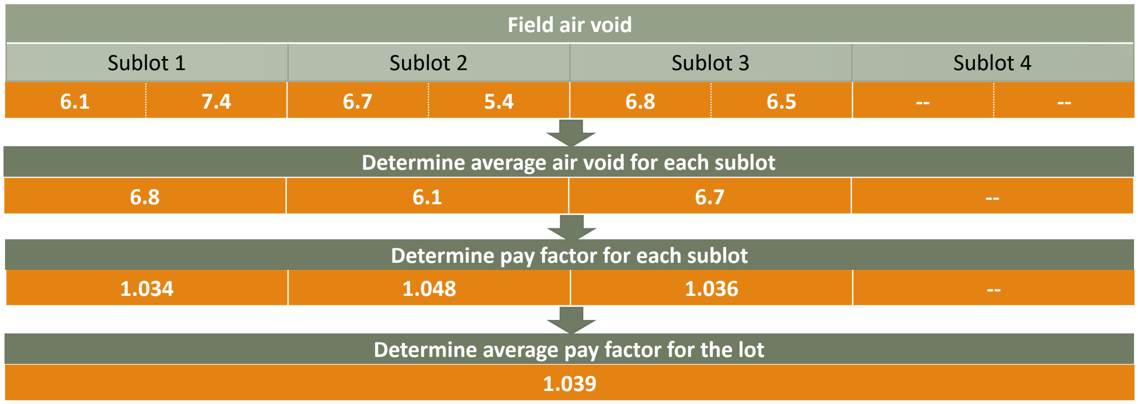


Figure 5.15. Example showing the placement PAF calculation for a lot that received a bonus of 3.9% according to the current practice.

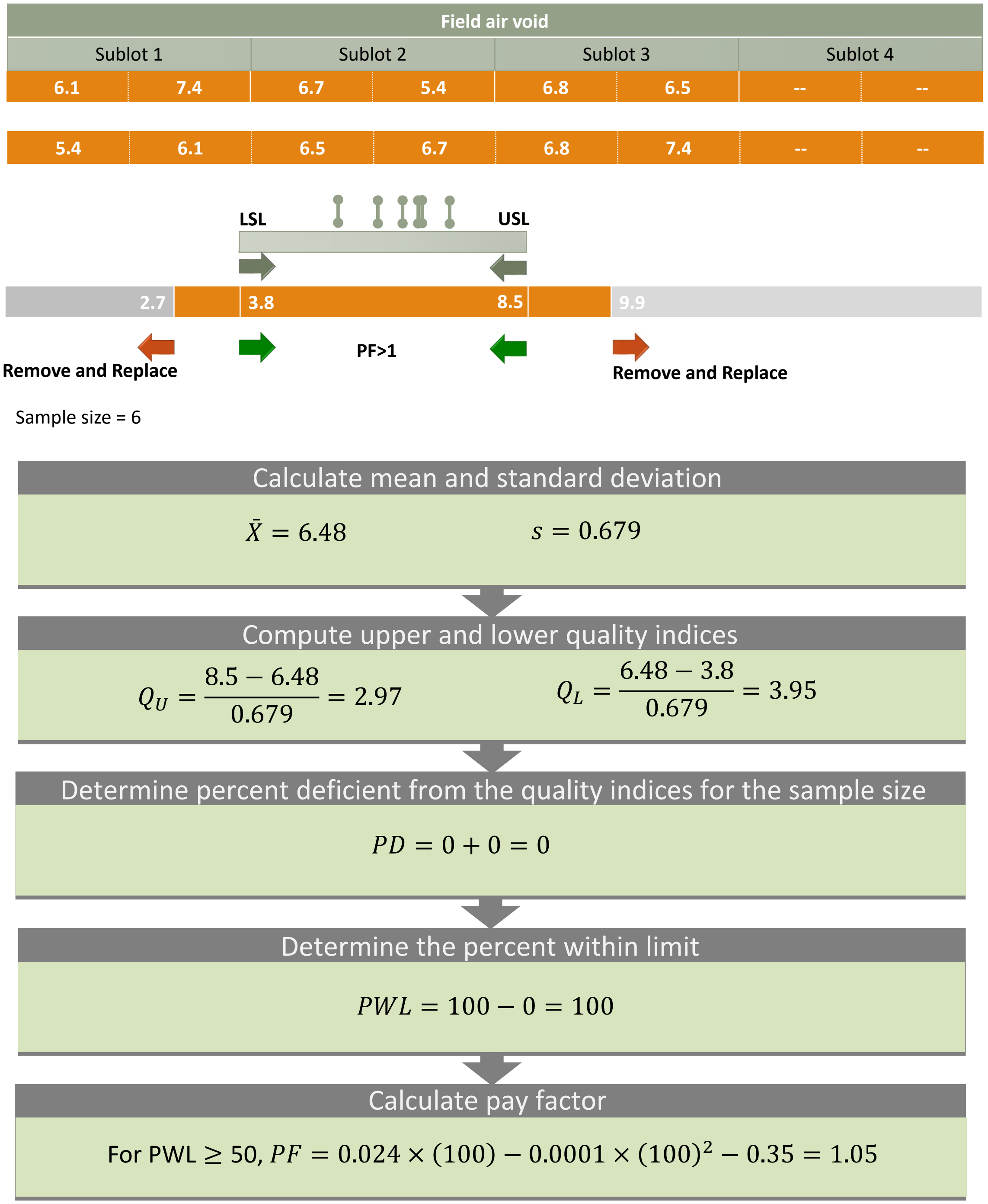


Figure 5.16. Example of PWL scheme that rewarded consistent test results of Figure 5.15 with a pay incentive of 5%.

5.5 SUMMARY

In summary, as evidenced by the illustrations presented here as well as its use by several state highway agencies, PWL based PAF scheme is an effective quality measure that accounts for both central tendency and variability. Several variables must be considered in establishing a PAF scheme that is based on the PWL approach. These variables depend on specification limits that satisfy the agency's requirements, objectives, and goals. There are several aspects of PWL that an agency can control to make the system suitable to its needs. Figure 5.17 summarizes different characteristics of PWL that the agency can control in order to establish an efficient quality measure for the production and placement of HMA.

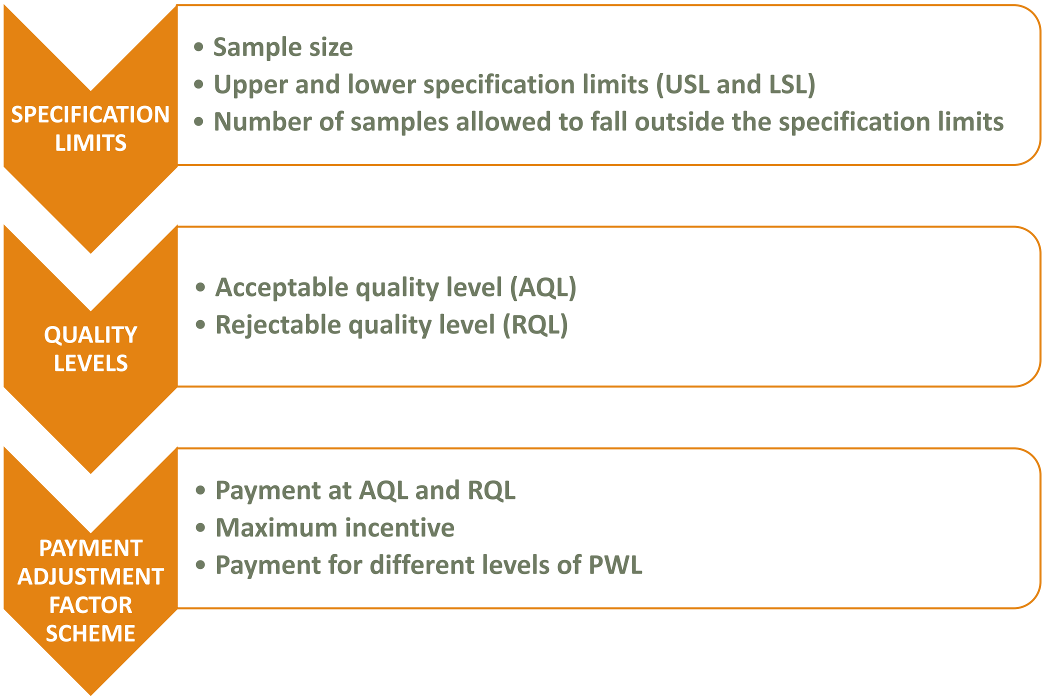


Figure 5.17. Characteristics of the PWL quality measure controlled by the agency.

CHAPTER 6. HYPOTHETICAL COST CALCULATION

6.1 OVERVIEW

The previous chapter demonstrated hypothetical PWL calculations using payment adjustment factor schemes that incorporate the current TxDOT specification limits for production and placement of asphalt mixes. This chapter extends these schemes as example to compute the difference and overall hypothetical costs to the agency if these schemes were being enforced in lieu of the current payment adjustment factors. *It must be emphasized here that this calculation is based on a hypothetical PWL based PAF scheme and utilizes typical values used in other states for factors such as acceptable quality level (AQL), rejectable quality level (RQL), and shape of the payment adjustment factor scheme. Therefore, any changes to these variables will influence the outcome of this calculation. It is also important to emphasize and recognize that such calculations do not take into account any expected changes to the cost of the material being procured.* For example, in a scenario where the Contractor bears more risk, the cost of the product will likely increase and potential cost savings realized from reduced bonus payments may be offset by such cost increases.

In order to demonstrate the impact of adopting the PWL scheme on the bonus / penalties paid to the Contractors, the hypothetical costs for TxDOT Specification Items 341, 344, and 346 were computed and compared to the actual payments that were made based on the 2014 specifications. The following section provides a brief description of the approach used for this analysis. As mentioned above, there are several variables that impact the payment adjustment factors computed using the PWL based PAF scheme. This chapter only demonstrates such a calculation using one set of variables. However, in order to facilitate implementation, a graphical user interface tool was developed that accounted for all variables that influenced the payment adjustment factor calculation using the PWL approach.

6.2 ANALYSIS

To calculate and compare the cost to the agency, the bid price, actual quantity placed, and the current production and placement payment adjustment factors were queried from the SiteManager database. The bid quantity and the net change order quantity for a given project were queried from the Sample (smpl) table and summed to obtain the total quantity placed, while the bid price and current production payment adjustment factor for an individual lot were queried from the Result (rslt) table. Although quantities are stored in the Result (rslt) table through the QC/QA template, different field numbers capture quantities of different specification Items for different productions. Consequently, queries from the Result table may fail to provide the true quantity if they do not include the relevant field numbers. To avoid this, quantities placed under a given project were queried from the Sample (smpl) table instead of the Result (rslt) table. Since bid price and payment adjustment factor are logged in via the QC/QA template for individual lots (and not projects), quantities queried from the Sample table needed to be sorted by lots. In order to facilitate this connection, projects from the Sample table were linked to corresponding lots from the Item (itm) table through common sample identification numbers. This query was then ultimately used to generate a table containing the bid price, quantity placed, and current payment adjustment factor for a specific lot of a project.

Production or placement pay was then calculated by multiplying the bid price with the actual quantity placed (including all change orders) and the corresponding payment adjustment factor. The current payment adjustment factors were queried from the Result table whereas, the proposed payment adjustment factors were calculated from the PWL based PAF schemes. Production pay and placement pay were computed separately and combined to provide the total adjusted pay (TAP). The production pay does not include quantity left in place without pay. Likewise, the placement pay excludes quantity placed in miscellaneous areas or quantity that does not require air void testing.

Since the purpose of this task was to compare the hypothetical payment for a sample PWL based PAF scheme with the payment made with current specifications, it was necessary to take the referee data into account as referee data ultimately replaced DOT data in calculating the total adjusted pay. Furthermore, as the referee results were averaged over a sublot and a maximum of four sublot values were reported, the queried lab density and air void values were also averaged for a sublot furnishing a maximum of four samples per lot. Finally, since a minimum of three samples were required to calculate standard deviation, the minimum number of sublots was restricted to three for analysis. Data pertaining to scenarios with just one or two sublots were not included in these calculations.

6.2.1 Hypothetical analysis for specification Item 341

For specification Item 341, a total of 2799 data sets were analyzed for construction projects over a period of approximately five years. Results from a subset of 100 lots from these

2799 data sets showing the relative impact of applying PWL based payment adjustment factors for production and placement are presented in Figures 6.1 and 6.2, respectively for Item 341. When this analysis was carried out for all the 2799 data sets using the PWL based payment adjustment factor scheme, a total reduction of 3.56% in production pay was estimated for the 2014 specification lots, and a 14.63% reduction was estimated in the placement pay. This percentage reduction was calculated by subtracting the total proposed pay from the total current pay and dividing it by the total current pay for all 2014 specification lots. Much of this reduction in placement pay was due to inconsistent performances that were buffered by the averaging technique and received full pay, even bonuses in some cases, based on the current specification, which would otherwise qualify for no pay if the PWL based PAF schemes were adopted. When these pays were combined to calculate the total adjusted pay, as shown in 6.3, the total adjusted pay was further reduced to 15.5%, since the current specification assigned zero pay to lots that qualified for no pay for either production or placement. An illustration of five random samples showing the effect on production, placement, and total adjusted pay is provided in Figure 6.4. It shows that the third lot production that qualifies for more pay than the current pay may be eligible for zero total adjusted pay as the placement pay satisfies the removal and replace criterion per the example PWL based payment adjustment factor scheme. Such reduction in the payment adjustment factor is consistent with the field observations in Wisconsin, where the PWL approach resulted in more "remove and replace" conditions than the previous method. This high reduction in the total adjusted pay emphasizes the need for implementing PWL and ensuring consistent performance by Contractors. As mentioned before, *it is emphasized that this outcome is based on typical values assigned to the different variables while using the PWL based payment adjustment factor scheme (e.g. AQL and RQL limits).* Any changes to the values of these variables will result in different outcomes of this analysis.

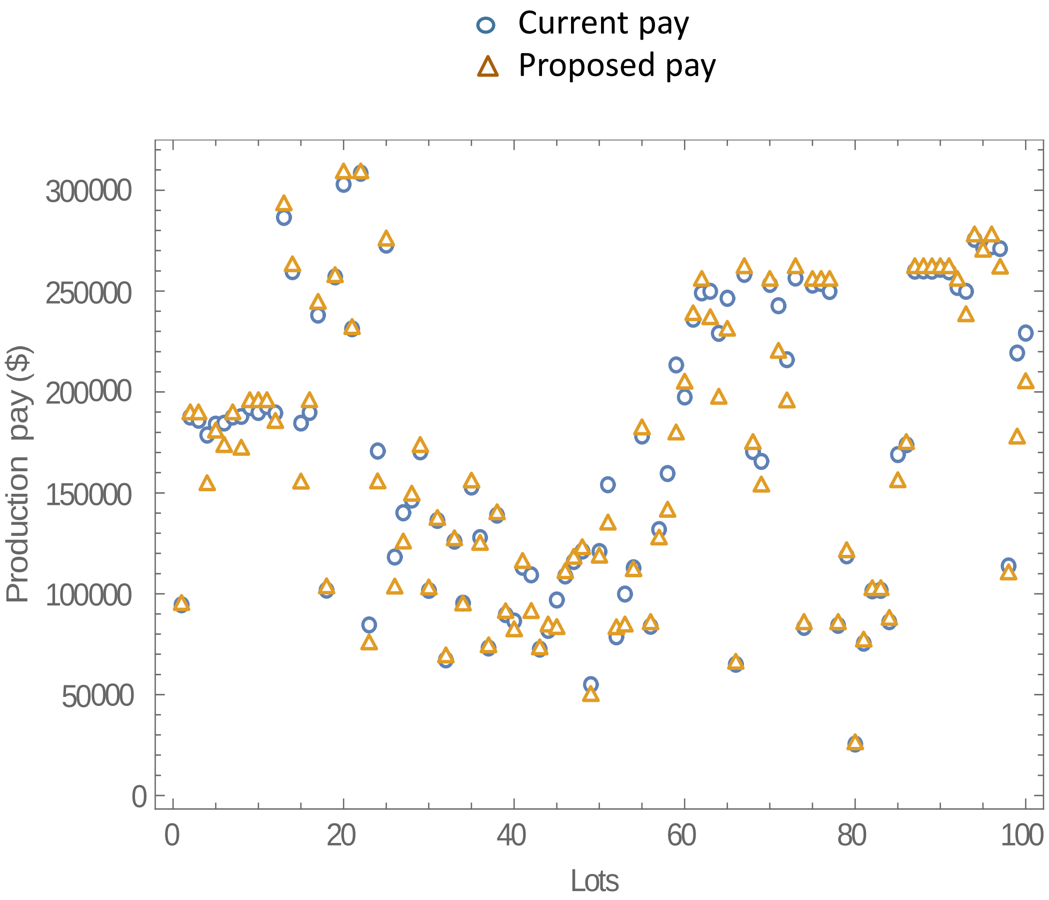


Figure 6.1. Sample production data comparing the hypothetical cost to the agency with the existing cost for item 341.

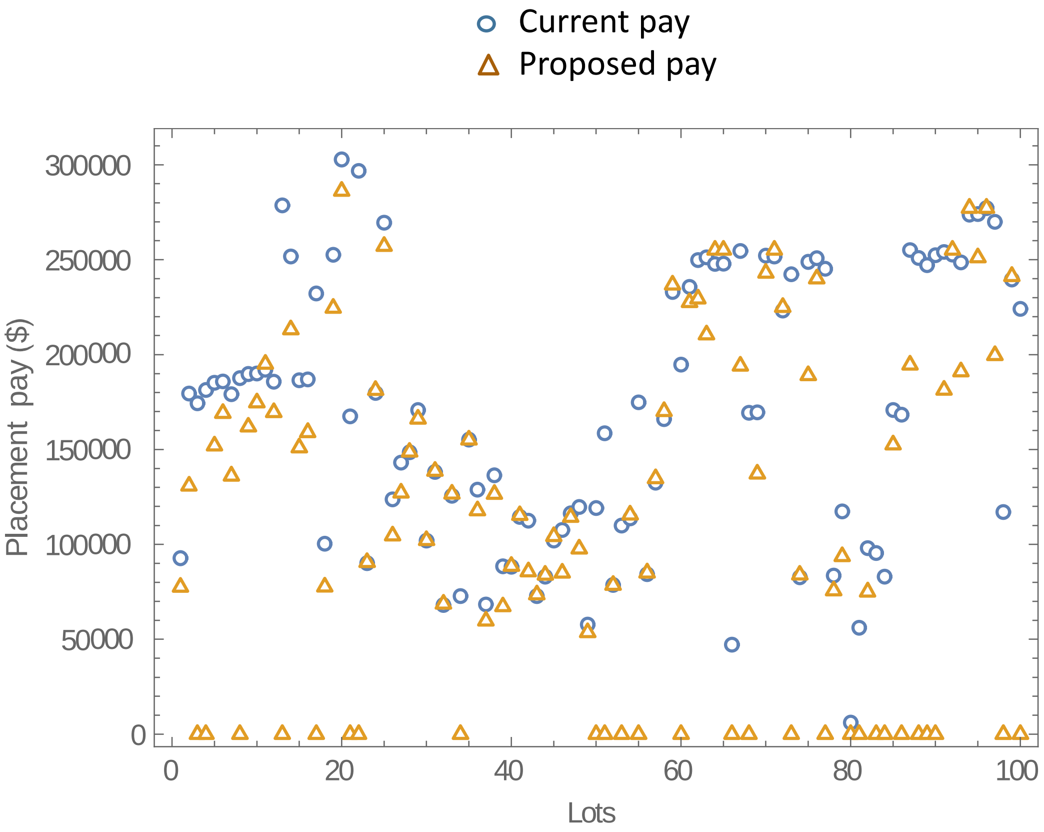


Figure 6.2. Sample placement data comparing the hypothetical cost to the agency with the existing cost for Item 341.

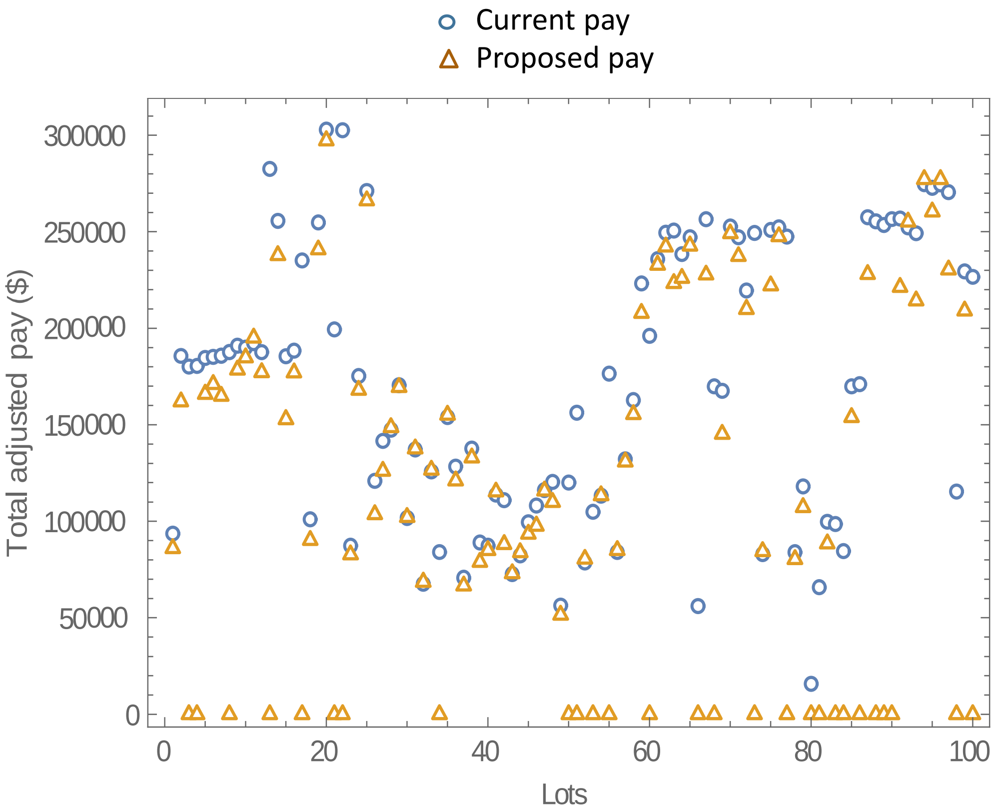


Figure 6.3. Sample data comparing the hypothetical cost to the agency with the existing cost for Item 341.

6.2.2 Hypothetical analysis for specification Item 344

Similar to Item 341, Item 344 data included referee data with values available for at least three sublots for the laboratory molded densities and field air voids. A total of 2926 data sets qualified based on the aforementioned criterion. Results for a subset of 100 data sets are provided in Figures 6.5, 6.6, and 6.7. For all the analyzed 2926 lots from Item 344, a total reduction of 1.27% was achieved for the total adjusted pay with 1.02% and 1.54% reduction in the production and placement pay, respectively.

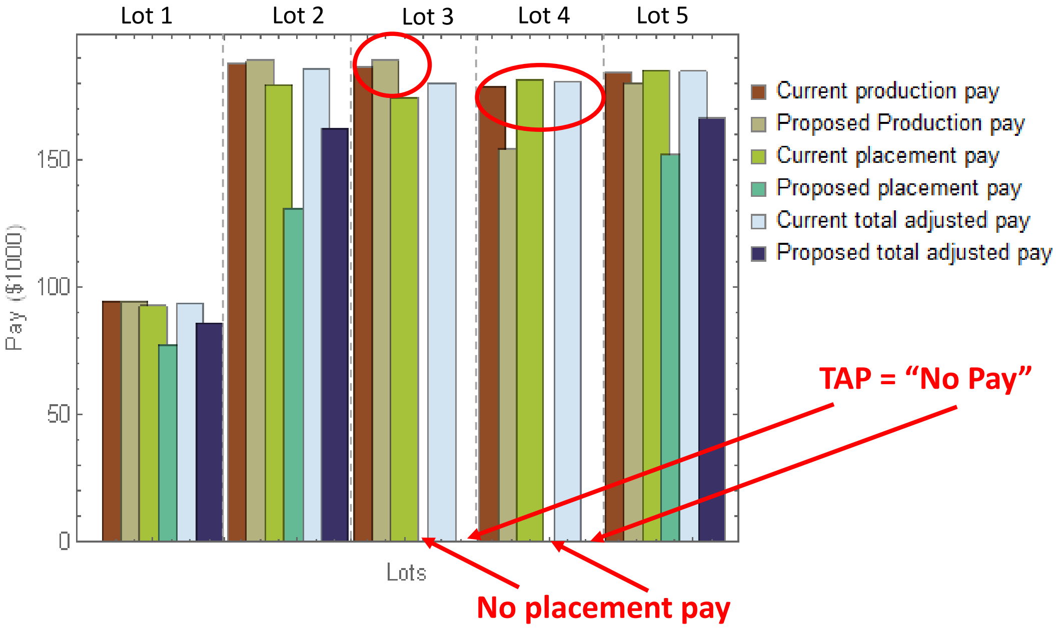


Figure 6.4. Sample data showing total adjusted pay (TAP) for Item 341.

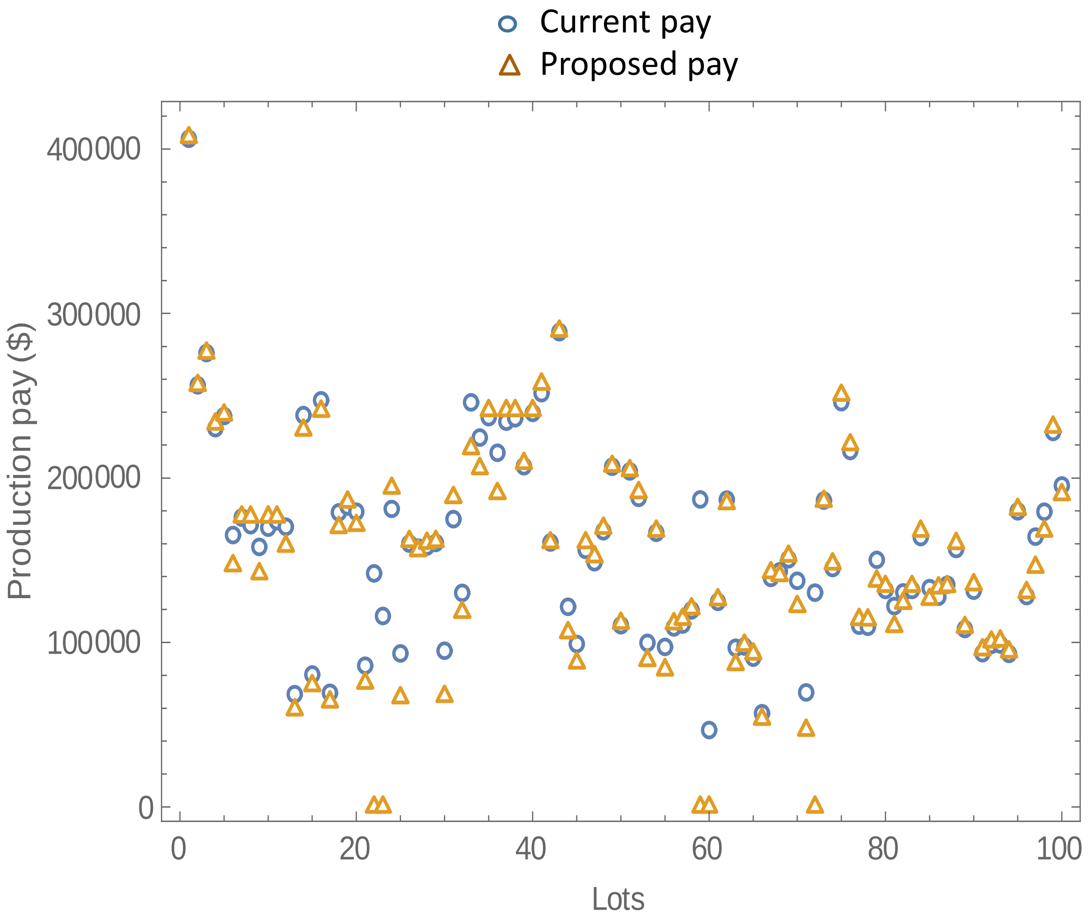


Figure 6.5. Sample production data comparing the hypothetical cost to the agency with the existing cost for Item 344.

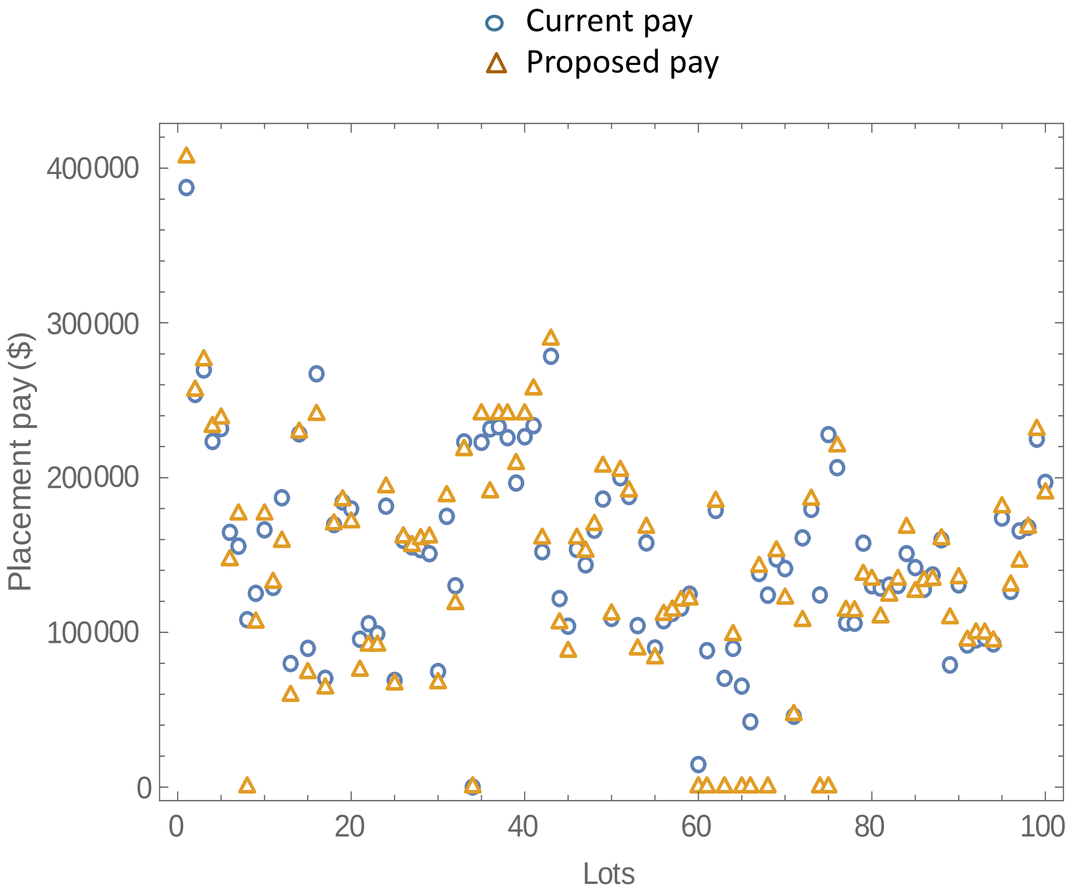


Figure 6.6. Sample placement data comparing the hypothetical cost to the agency with the existing cost for Item 344.

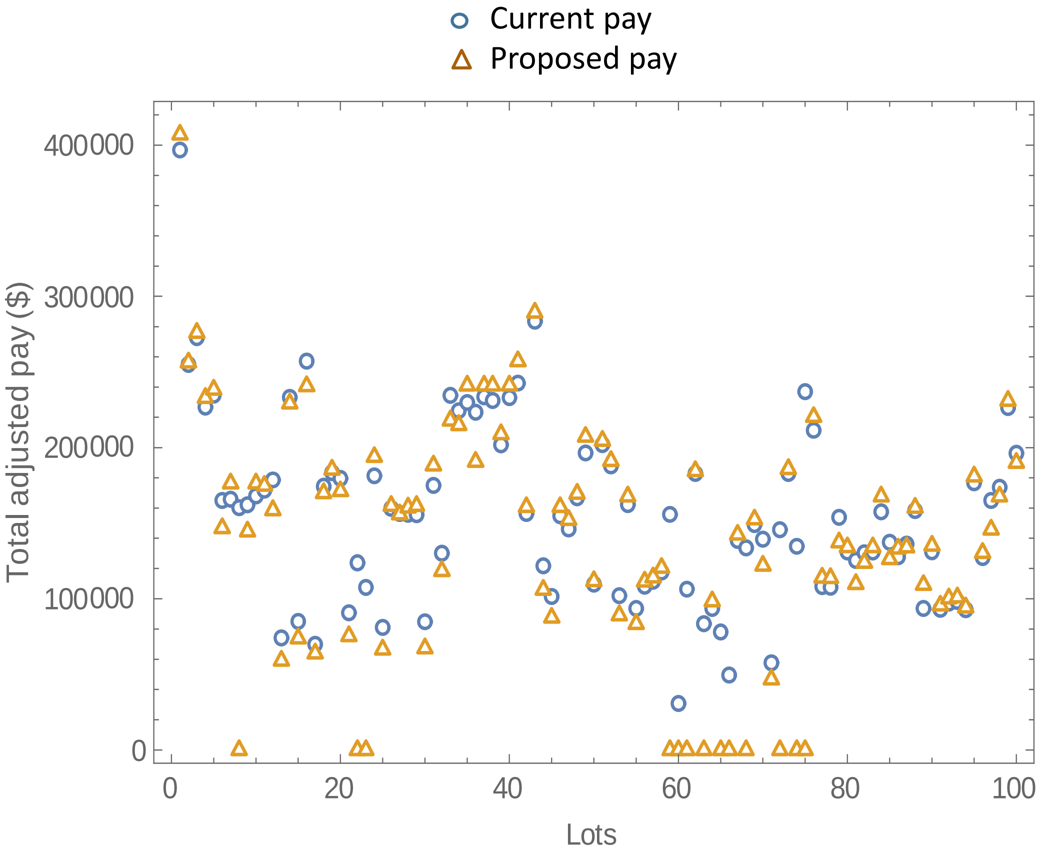


Figure 6.7. Sample data comparing the hypothetical total adjusted pay with the existing total adjusted pay for Item 344.

6.2.3 Hypothetical analysis for specification Item 346

Contrary to Items 341 and 344, Item 346 showed less reduction in placement pay than the production pay. Among the 495 analyzed data sets, a reduction of 0.31% in placement and 4.14% in production were obtained for this Item with a total adjusted pay reduction of 4.06%. This result implies that for Item 346, more consistent performance is expected in placement when compared to production. In other words, Contractors are able to achieve consistent densities while placing and compacting the mix. A subset of 100 data representing Item 346 was used to demonstrate the effect of PWL on the production, placement, and total adjusted pay in Figures 6.8 to 6.10, respectively.

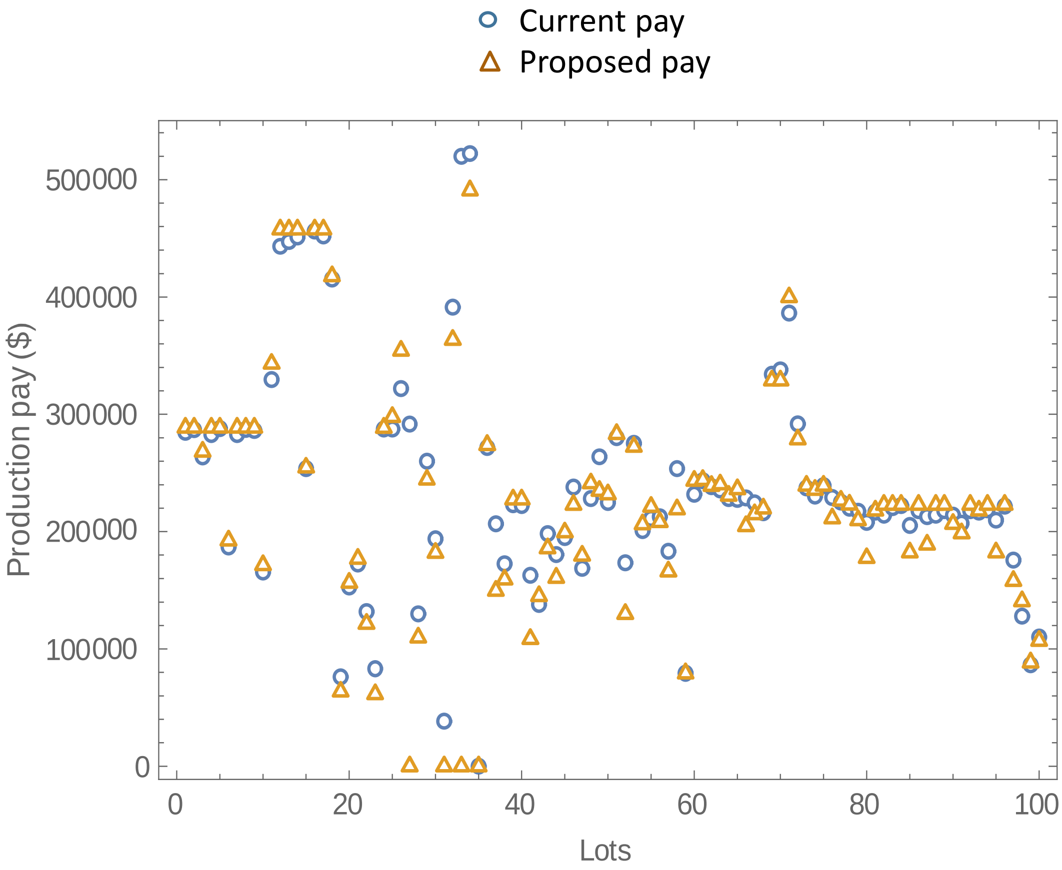


Figure 6.8. Sample production data comparing the hypothetical cost to the agency with the existing cost for Item 346.

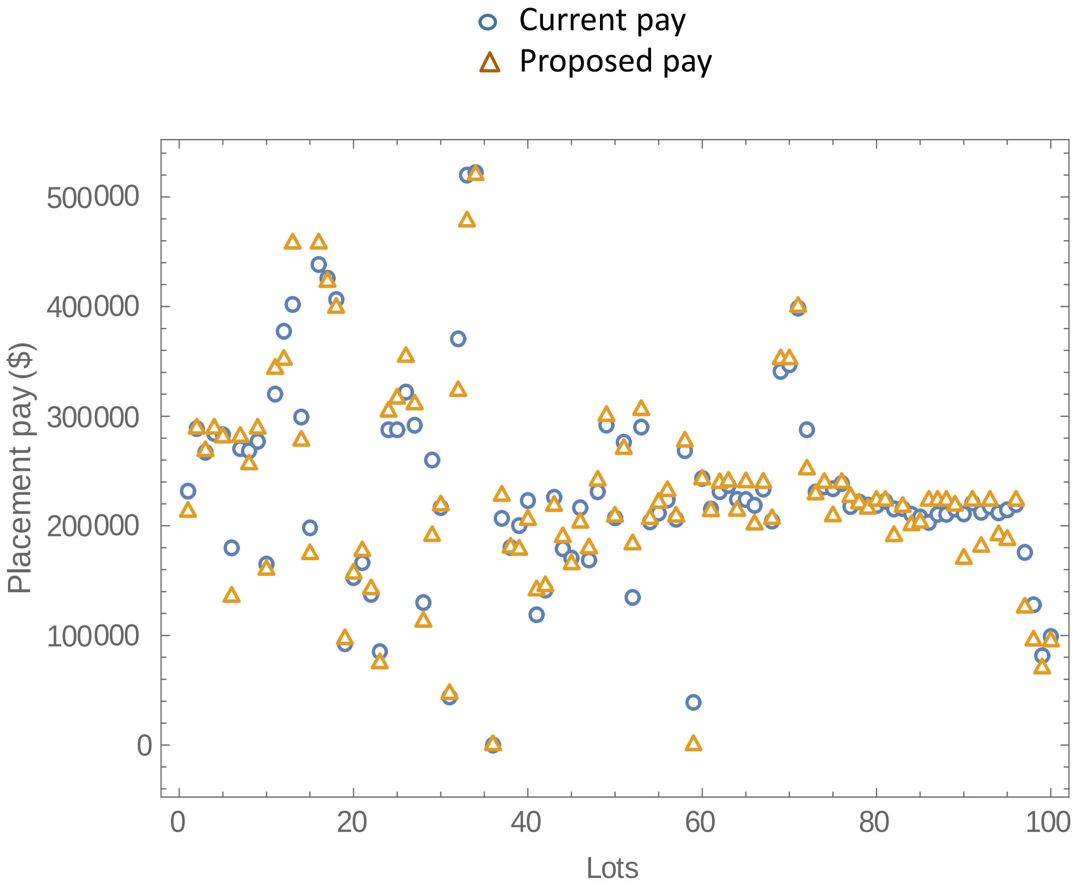


Figure 6.9. Sample placement data comparing the hypothetical cost to the agency with the existing cost for Item 346.

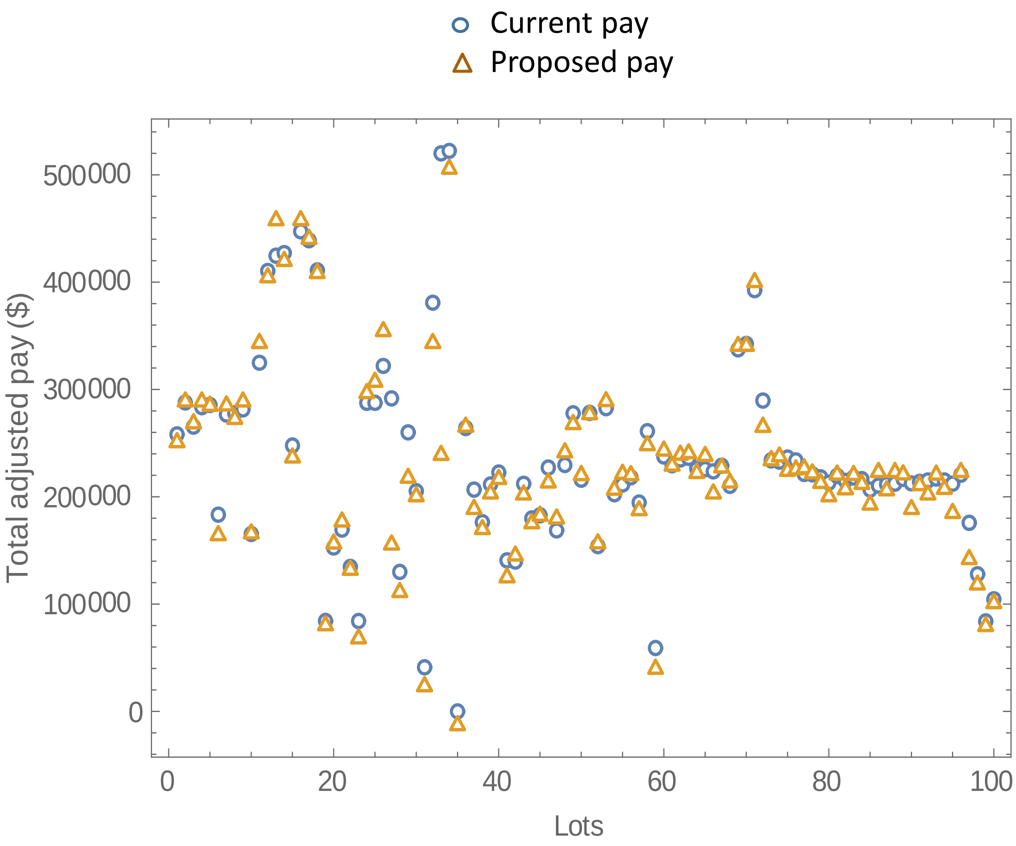
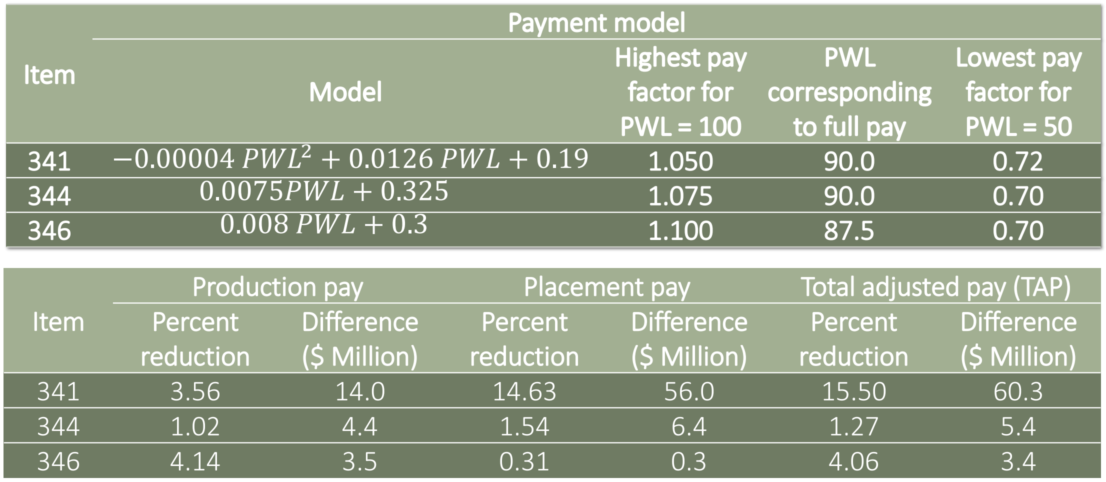


Figure 6.10. Sample data comparing the hypothetical total adjusted cost to the agency with the existing cost for Item 346.

A summary of the PWL based PAF schemes along with the resulting estimated cost impact is provided in Figure 6.11. A quadratic PWL based PAF scheme, similar to ODOT's PWL based PAF scheme, was used for the specification Item 341, whereas linear PAF schemes were used for Items 344 and 346. The AQL (acceptable quality level) values used for this analysis were 90 for Items 341 and 344, and 87.5 for Item 346. The rejectable quality level of 50 was used for all specification Items, and the corresponding lowest pay factors are presented in Figure 6.11. The maximum and minimum PAFs were identical to the maximum and minimum PAFs specified for the current specification Items. The upper and lower specification limits (USL and LSL) for Items 341, 344, and 346 were provided in Table 4.1.



Item	Payment model			
	Model	Highest pay factor for PWL = 100	PWL corresponding to full pay	Lowest pay factor for PWL = 50
341	$-0.00004\,PWL^2 + 0.0126\,PWL + 0.19$	1.050	90.0	0.72
344	$0.0075PWL + 0.325$	1.075	90.0	0.70
346	$0.008\,PWL + 0.3$	1.100	87.5	0.70

Item	Production pay		Placement pay		Total adjusted pay (TAP)	
	Percent reduction	Difference ($ Million)	Percent reduction	Difference ($ Million)	Percent reduction	Difference ($ Million)
341	3.56	14.0	14.63	56.0	15.50	60.3
344	1.02	4.4	1.54	6.4	1.27	5.4
346	4.14	3.5	0.31	0.3	4.06	3.4

Figure 6.11. Summary of the example PWL based PAF scheme and the corresponding hypothetical cost reductions for Item 341, 344, and 346.

6.3 HYPOTHETICAL COST CALCULATION TOOL

As previously indicated, the aforementioned analysis demonstrated the potential impact for just one set of variables in the PWL based payment adjustment factor scheme (as shown in Figure 6.11 for each specification Item. In order to facilitate the implementation of this scheme it would be necessary to analyze the results from several different scenarios. To achieve this, a graphical user interface capable of calculating the hypothetical cost to the agency was developed. A screen-shot of this interface is presented in Figure 6.12. It must be noted that the tool used a static data set and was not intended for routine use. Rather, the tool was only intended for use by TxDOT to evaluate different scenarios in developing a specification.

This tool queries laboratory molded density data for production and field air void data for placement from a static image of the database stored in the same directory for a sample size specified by the user. The user also provides the acceptable and rejection quality level (AQL and RQL) as well as the upper and lower specification limits (USL and LSL). Finally, the user has the ability to design a payment adjustment factor scheme that defines the payment adjustment factor as a function of the PWL when the PWL is above the RQL. In other words, the payment adjustment factor scheme defines the multiplier based on PWL (a measure of the consistency and being within specification limits) for both production and placement as long as the material is above the rejection threshold. Below the rejection threshold, the tool is programmed to assign a zero pay. The tool then calculates risks associated with the acceptable and rejectable quality levels for that sample size, the PWL for each data set based on the specified limits, assigns pay using the defined pay-factor scheme, and computes the net change in pay with respect to the current pay queried from the database. The tool presents three graphics displaying the PWL based PAF scheme defined by the user, a distribution of affected lots, and the probability of being affected by the proposed PWL based PAF scheme (in combination with the AQL, RQL, LSL, and USL).

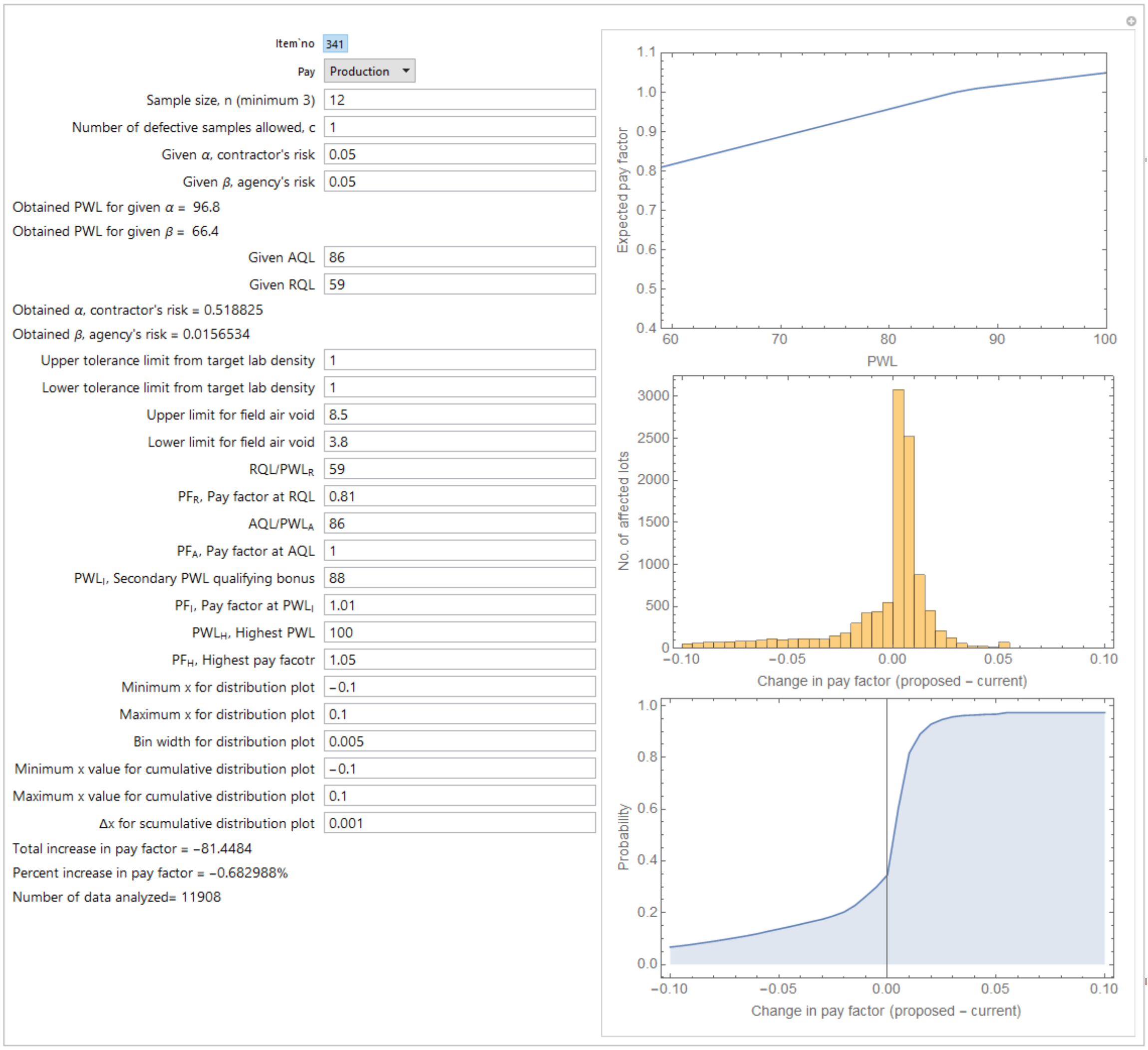
The user can select either production or placement pay from a dropdown list and specify a sample size, n, to be analyzed. For instance, for Item 341, the maximum n for the current specification for production is 12 for laboratory molded density when determined by Texas Gyratory Compactor (TGC) and 8 for Superpave Gyratory Compactor (SGC). For the current placement specification, the sample size can be 8 at best, as two cores are collected from each sublot with the total maximum of 8 cores per lot. This tool also determines the risks to the agency (β) and Contractor(α) for a given set of acceptable and rejectable quality levels, (AQL) and (RQL), and sample size n. It is also possible for the user to define these risks (α and β) and query the tool to compute the AQL and RQL that are required to achieve these risk goals.

In compliance with the current specification, the user can specify operational tolerances with respect to the target laboratory-molded density for production pay. Similarly, limits for the field air voids can also be controlled by the user for placement pay. The payment adjustment factor scheme can be manipulated by controlling four PWL input variables, AQL, RQL, secondary PWL qualifying bonus, and the highest possible PWL. The desired payment adjustment factors corresponding to these limiting PWL values need to be specified as well. The program uses these limiting payment adjustment factors and boundary PWL values to generate a mathematical PWL based payment adjustment factor scheme. Once it determines PWL for each data set for the upper and lower specification limits for production / placement pay, this code utilizes the generated PWL based payment adjustment factor scheme to assign pay to the corresponding PWL for each data set. Finally, it computes the change in pay with respect to the current payment queried from the database and reports it as the increase in payment adjustment factor and the percent increase in payment adjustment factor . A negative value in these fields imply that the agency would save money or pay less to the Contractors if the user defined PWL based payment adjustment factor scheme was used. It also reports the total number of data analyzed by the program for a given sample size.

For instance, according to Figure 6.12, for item 341 production pay with sample size of 12, Contractor's risk for the given AQL of 86 was 51.9% and agency's risk associated with RQL of 59 was 1.6%. That is, if a PWL value of 86 was set as the minimum for full pay, there was a 51.9% chance that the agency would reject the material with no pay. This risk could be reduced to 5% if the AQL was set at 96.8 PWL or higher. On the contrary, if a PWL value of 59 was set for the lowest allowable quality level below which the construction qualified for zero pay, the agency had 1.6% chance of erroneously accepting a rejectable quality material with full pay.

If the upper and lower tolerance limits were specified as target density±1, for an AQL of 86 and RQL of 59, respectively, and a maximum bonus possibility of 1.05 was set, the reduction in cost to the agency would be 0.68% for 2014 specification lots. For this analysis the program queried 11908 datasets that have all twelve laboratory molded densities reported on SiteManager.

Depending on the defined boundary PWL values, the code generated a tri-linear payment adjustment factor scheme as shown in Figure 6.12. This scheme essentially satisfied the conditions whereby full pay was assigned to 86 PWL (AQL), and material was rejected with no pay for PWL below the RQL of 59. At 59 PWL, the Contractor would be subject to 81% of the full pay. In addition, a bonus of 1% could be attainable with a PWL of 88 with a maximum bonus of 5% for 100 PWL. The results provided by the tool also included a distribution of lots that were negatively or positively affected by the sample scheme. Any lots to the right of the zero line along the horizontal axis were positively affected by the scheme, or qualified for higher pay than the current system, whereas lots to the left of the zero line received lower pay based on this scheme compared to the existing specification. According to the simulated distribution, for a sample size of 12, about 3000 lots out of 11908 may not be affected at all by the specified PWL based PAF scheme. The probability distribution curve implied that there was less than 30% chance that a lot would be penalized, i.e. would receive less pay compared to the existing specification, if this scheme was implemented and all twelve samples were tested.



Item`no	341
Pay	Production
Sample size, n (minimum 3)	12
Number of defective samples allowed, c	1
Given α, contractor's risk	0.05
Given β, agency's risk	0.05

Obtained PWL for given α = 96.8
Obtained PWL for given β = 66.4

Given AQL	86
Given RQL	59

Obtained α, contractor's risk = 0.518825
Obtained β, agency's risk = 0.0156534

Upper tolerance limit from target lab density	1
Lower tolerance limit from target lab density	1
Upper limit for field air void	8.5
Lower limit for field air void	3.8
RQL/PWL_R	59
PF_R, Pay factor at RQL	0.81
AQL/PWL_A	86
PF_A, Pay factor at AQL	1
PWL_I, Secondary PWL qualifying bonus	88
PF_I, Pay factor at PWL_I	1.01
PWL_H, Highest PWL	100
PF_H, Highest pay facotr	1.05
Minimum x for distribution plot	−0.1
Maximum x for distribution plot	0.1
Bin width for distribution plot	0.005
Minimum x value for cumulative distribution plot	−0.1
Maximum x value for cumulative distribution plot	0.1
Δx for scumulative distribution plot	0.001

Total increase in pay factor = −81.4484
Percent increase in pay factor = −0.682988%
Number of data analyzed= 11908

Figure 6.12. A hypothetical cost computing tool with graphical user interface showing controlling variables (inputs) and typical outputs.

CHAPTER 7. SUMMARY AND CONCLUSIONS

7.1 OVERVIEW

The payment for production and placement of asphalt mixes typically requires that the Contractor meet requirements for several metrics including gradation, binder content, laboratory-molded density, segregation, longitudinal joint density, thermal profile and in-place air voids. Of these, payment adjustment factors are computed to adjust the payment awarded to the Contractor based on test results for laboratory-molded density (production) and in-place air voids (placement or construction). The payment adjustment factors are computed separately for production and placement and averaged for the final payment. The payment adjustment factors can increase the Contractor payment by as much as 5%, 7.5%, and 10% as a bonus for the specification Items 341, 344, and 346, respectively or decrease by 28% for Item 341 and 30% for Items 344 and 346 as a penalty. As an extreme case, the Contractor may be required to remove or replace or forfeit the payment if production and placement specifications are not met.

A characteristic feature of the current payment adjustment factors for production and placement is that these factors are based on average values of the in-place air voids or the absolute deviations from the target laboratory-molded density. These averages are obtained using measurements made over typically four sublots for each lot of production. In some cases, this practice results in a scenario where the Contractor delivers inconsistent quality but still receives a bonus because of the average performance. The goal of this project was to evaluate the use of a payment adjustment factor that is not based solely on the average test results for production and placement but is based on a measure of quality that reflects both the average and the variability in production and placement.

A review of the existing literature and practice in other states and recommended practice by the Federal Highway Administration (FHWA) revealed that this can be achieved using a metric that is referred to as the Percent Within Limits (PWL). In simple terms, PWL provides an estimate for the probability that a test result from a lot falls within the upper and lower specification limits. The PWL is calculated using the upper specification limit, the lower specification limit, the sample mean, the sample standard deviation, and an appropriate probability distribution table. A minimum value of this probability or PWL, referred to as the acceptable quality level or AQL, is typically prescribed as a part of the specification for the Contractor to receive full payment. A payment adjustment factor can be developed to award a bonus for exceeding this AQL and a penalty for falling below this AQL value. A rejectable quality level or RQL can also be prescribed as the lower threshold below which the Contractor will not receive any payment or may be required to remove or replace the work. The main difference between the current and the PWL based payment adjustment factor scheme is that the latter accounts for variability and consistency in quality, whereas the former does not.

As a part of this study, several different quality measures including the PWL were reviewed. Data from TxDOT's SiteManager database were collected for both 2004 and 2014 specifications for Items 341, 344, and 346. Nearly 39,545, 9,685, and 3,109 data sets, covering all projects from all districts, were queried for Items 341, 344, and 346, respectively. Each data set represented one lot of production and placement. These data were analyzed to determine the variability in production and placement metrics. Results showed a diversity in the variability for both of these factors when compared with the payment adjustment factors that were awarded to these projects. Specifically there were four different scenarios (example Figure 2.34):

- lots that received a bonus and had very low variability in production and/or placement (this is the desirable outcome),
- lots that received a penalty and had very high variability in production and/or placement (this is acceptable although not desirable from a quality point of view),
- lots that received a penalty but had very low variability in production and/or placement (this is acceptable although not desirable from a quality point of view; this also suggests that the Contractors are consistent in their production and/or placement and can make some amendments to adjust the average target values), and
- lots that received a bonus but had very high variability in production and/or placement (this is not acceptable because it ignores inconsistency in quality; fixing this gap is also the main goal of this study).

The use of PWL as an approach to rectify the aforementioned gap was investigated. Practices from several states and recommendations from FHWA were reviewed. Typical schemes for the payment adjustment factor were selected for further examination. These schemes were then used with existing data from the SiteManager database to determine the new pay factors for the 2014 specification Items 341, 344, and 346, assuming such a hypothetical scheme were in force for a period of approximately five years. These analyses were used to demonstrate the efficacy of using a payment adjustment factor based on PWL instead of sample averages. Finally, a calculation tool that is based on the different variables that go into the PWL and payment adjustment factor scheme was developed. This tool can be used to explore the implications of setting different values for the variables for implementation.

7.2 SUMMARY OF BROADER RECOMMENDATIONS

The following are some of the key recommendations and findings from this study:

- A review of the literature and findings from this study clearly demonstrate that PWL is much superior measure of quality compared to sample mean that is currently used and must be implemented in the future or through provisional specifications. This approach overcomes the gap wherein a lot can receive a bonus payment based on the average quality without consideration of the consistency in quality of production or placement.
- The PWL approach can be more readily adopted for production and placement based on test results for laboratory-molded density and in-place air voids, respectively. This is because these two metrics are already being used with payment adjustment factors in the current specification. However, other metrics such as binder content and VMA (voids in mineral aggregate) must also be considered for future implementation using the PWL approach.
- The PWL approach is better suited when the sample size used for testing is large. The current specifications call for a sample size of eight (four pairs) for in-place air voids and twelve for laboratory-molded density for TGC mixes or eight for laboratory-molded density for SGC mixes for a lot. The PWL approach can be implemented for all lots that call for these sample sizes and a flat payment adjustment factor may be prescribed for lots that result in fewer sublots and smaller sample sizes.
- In making measurements for in-place air voids, the sample locations are generated randomly. In certain situations, these locations may fall in an area that does not qualify for payment adjustment factor calculation (e.g. miscellaneous areas including temporary detour, driveways, spot level-up areas). For such cases, a payment adjustment factor of 1.000 is assigned to those sublots. It is preferable that this practice be modified to identify a different randomly selected location for testing and incorporation in payment adjustment factor calculation.
- The implementation of PWL requires establishing several different variables. A list of these variables along with recommended values is presented in the following section.

Implementation of PWL is certainly an improvement over the existing specification that is based only on average values. This approach also ensures that the quality of material produced or placed is consistent and rewards Contractors for meeting the specification limits consistently. However, it must also be pointed out that PWL is only a mechanism that ensures that the material produced and placed matches the specified mix design. The use of PWL therefore ensures *consistent* performance but by itself this is not synonymous with long-term durability or "improved" performance. The latter can only be achieved by ensuring that material specifications and mix design are based on performance metrics (e.g. balanced mix design).

7.3 SUMMARY OF PWL PARAMETERS FOR IMPLEMENTATION APPROACH AND SUGGESTED VALUES

The following is a summary of the key factors and variables that must be considered in the implementation of the PWL approach along with proposed values that can be considered for adoption.

7.3.1 Specification limits

The use of PWL requires that the upper and lower specification limits be established (USL and LSL). These values can be adopted from the current specifications. For example, in the current specifications, laboratory-molded density is allowed to vary by ±1 from the target value in the JMF. In this case the USL can be set to target laboratory-molded density + 1 and the LSL can be set to target laboratory-molded density - 1. Similarly for in-place air voids, the current payment adjustment factors allow full payment when the air voids are between 3.8 and 8.5, 3.7 and 7.5 and 3.7 and 7.0 for Items 341, 344, and 346, respectively. These values can be used as the LSL and USL for the respective items.

7.3.2 Sample size

It is possible to extend the sampling requirements and lot definitions under the current specifications to the proposed method. The current specifications require eight samples for in-place air voids over four sublots and twelve samples for TGC mixes (or eight samples for SGC mixes) for laboratory-molded density over four sublots. These requirements may be continued. However, the lot size is allowed to vary from 1000 to 4000 tons. In most cases, this varies from 1000 to 3000 tons. One possibility for implementation is to fix a sublot size and sampling requirements per sublot and invoke PWL and payment adjustment factors only for a minimum number of sublots. This also ensures that a minimum number of samples being used with the PWL method.

7.3.3 Acceptable quality level

This is the minimum value for the PWL that must be achieved in order to be eligible for full payment. The Agency (TxDOT) may choose to set this value at any level, however, a commonly used value for AQL is 90, i.e. based on the analysis a PWL of 90 must be achieved to qualify for full payment. This is also a recommendation in the AASHTO Quality Assurance Guide Specification (1995).

7.3.4 Rejectable quality level

This is the minimum value for the PWL that must be achieved in order to be eligible for any payment at all. In other words, a PWL below this value will trigger remove and replace or no payment. As with AQL, the Agency (TxDOT) may choose to set this value at any level, however, typically most states set the RQL between 30 and 70 PWL based on sample sizes. An RQL of 50 is also commonly used by many state agencies. It is, however, recommended to establish this value after fully evaluating the associated risk to the agency.

7.3.5 Payment adjustment factor scheme

The AQL and RQL define specific boundaries for the payment adjustment factor (PAF) table. At or above the AQL, the PAF is 1.000 or higher. Based on the existing specifications and PAF tables, the maximum PAF is 1.050 for Item 341, 1.075 for Item 344, and 1.100 for Item 346. These values can also be used to define the proposed PAF adjustment table based on PWL. One option would be for the PAF to increase linearly from 1.0 to 1.05 for Item 341 as the PWL increases from 90 (which is the AQL) to 100. Existing specifications and PAF tables have a “plateau” region for the highest PAF of 1.05. A second option would be to accommodate something similar to existing specifications. For example, the PAF can increase linearly from 1.0 to 1.05 as the PWL increases from 90 (which is the AQL) to 95 for Item 341. A PWL at or above 95 would qualify for a PAF of 1.05.

On the penalty side, the current PAF schedules have pay factors that go as low as around 0.720 for Item 341 and 0.700 for Items 344 and 346 . This lower boundary can be continued with the PWL scheme by setting the PAF to linearly decrease from 1.000 to 0.720 as the PWL values decreases from 90 (which is the AQL) to 50 (which is the RQL) for the specification Item 341, for instance.

7.3.6 Number of allowable failed samples

As the number of failed samples increases for a given sample size, the risk to the agency of erroneously accepting rejectable materials also increases. It is recommended to specify the maximum number of test results that can fall outside the specification limits to keep agency's risk within the recommended risk level. For example, if twelve samples were collected for determining laboratory-molded density compacted with TGC, one sample could be allowed to fall outside the specification limits that would keep the agency's risk within the AASHTO recommended risk level of 0.005% for critical construction. The test results that fail the specification limits should be included in the estimation of PWL to reflect the variability of the material and construction unless these values are deemed to be outliers.

7.3.7 Rounding

One of the main factors that affect the results from PWL analysis is the calculation and rounding procedure used. When the payment adjustment factor is calculated based on PWL, this issue can be a point of conflict. This is because the values from the tables can be rounded up, rounded down, or interpolated. A Contractor usually prefers the rounding up approach because it increases the PWL achieved. The method of rounding must be specified prior to implementation to ensure that there are no conflicts regarding this issue. As a result, it is important to specify the calculation process, including number of decimal places to be carried in the calculations, as well as the exact manner in which the PWL is determined (interpolation from table or computation from beta distribution formula).

In conclusion, implementation of PWL requires establishing the aforementioned parameters as a part of the specification. Values for several of these parameters can be adopted from existing specifications or practices recommended by AASHTO or used by other states. Implementation of these specifications may require input from stake holders. An analysis tool has been provided that will allow the Agency (TxDOT) to create any hypothetical scheme and evaluate its implications using past data.

REFERENCES

[1] AASHTO. *AASHTO R-9, Standard Practice for Acceptance Sampling Plans for Highway Construction*, January 2005.

[2] ASTM International. *ASTM E178-16a, Standard Practice for Dealing With Outlying Observations*, January 2016.

[3] J. L. Burati, R. M. Weed, C. S. Hughes, and H. S. Hill. Optimal Procedures for Quality Assurance Specifications, Federal Highway Administration. Technical report, 2003.

[4] S. A. Cross and J. A. Frantzen. Evaluating Percent within Limits (PWL) Specifications, Volume 1: HMA Specifications. Technical report, 2003.

[5] C. S. Hughes, J. S. Moulthrop, S. Tayabji, R. M. Weed, and J. L. Burati. Guidelines for Quality-Related Pay Adjustment Factors for Pavements. Technical report, 2011.

[6] Indiana Department of Transportation. *Standard Specifications*, 2018.

[7] Kansas Department of Transportation. *Standard Specifications for State Road and Bridge Construction*, 2015.

[8] Oklahoma Department of Transportation. *Standard Specifications*, 2009.

[9] State of Louisiana Department of Transportation and Development. *Louisiana Standard Specifications for Roads and Bridges*, 2016.

[10] Texas Department of Transportation. *Standard Specifications for Construction and Maintenance of Highways, Streets, and Bridges*, 2004.

[11] Texas Department of Transportation. *Standard Specifications for Construction and Maintenance of Highways, Streets, and Bridges*, 2014.

[12] Wisconsin Department of Transportation. *Construction and Materials Manual*, 2017.

[13] Wisconsin Department of Transportation. *Standard Specifications for Highway and Structure Construction*, 2019.