EVALUATION OF PERCENT-WITHIN-LIMITS SPECIFICATIONS

Final Report

Steven M. Trost, Ph.D., P.E.
STRATEGIC SOLUTIONS

Prepared for the

OKLAHOMA DEPARTMENT OF TRANSPORTATION
in cooperation with the
FEDERAL HIGHWAY ADMINISTRATION

June 2003

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.

4. TITLE AND SUBTITLE EVALUATION OF PERCENT-WITHIN-LIMITS SPECIFICATIONS: FINAL REPORT	5. REPORT DATE June 2003 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Steven M. Trost, Ph.D., P.E.	8. PERFORMING ORGANIZATION REPORT NO. SS-0210202
9. PERFORMING ORGANIZATION NAME AND ADDRESS Strategic Solutions 2823 West 28th Avenue Stillwater, OK 74074	10. WORK UNIT NO. 11. CONTRACT OR GRANT NO. EC #725
12. SPONSORING AGENCY NAME AND ADDRESS Oklahoma Department of Transportation 200 N. E. 21st Street Oklahoma City, OK 73105	13. TYPE OF REPORT AND PERIOD COVERED Final Report 10/02 - 6/03 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

In 1996, the Oklahoma Department of Transportation (ODOT) drafted a set of special provisions to address the quality control and quality assurance (QC/QA) of new asphalt concrete (AC) pavements, portland cement concrete (PCC) pavements, and concrete bridge decks using "Percent-within-Limits" (PWL) specifications. Those QC/QA special provisions have never been implemented by ODOT. However, in 1998, the Oklahoma Transportation Authority (OTA) adapted the AC and PCC versions of those special provisions for use on over $150 million worth of turnpike paving projects. In 2002, as the final OTA projects were nearing completion, ODOT commissioned this study to evaluate the implementation of OTA's QC/QA PWL specifications. The aim of the study was to address the following broad objectives:

- Provide a means for objective assessment by ODOT of the implementation of percent-within-limits specifications.
- Investigate alternatives for dealing with quality characteristics having non-normal distributions.
- Investigate alternatives for limiting contractor exposure under percent-within-limits specifications due to variability in test methods and procedures.
- Identify suitable adjustments to percent-within-limits specification limits for various quality characteristics.
- Provide guidelines and recommendations concerning statistical methods for quality assurance testing.

In addition to the above objectives, the following key conclusions are discussed:

- Overall, the contractors' processes demonstrated high capability with respect to the specified quality levels.
- Greater emphasis is needed on statistical process control.
- Non-normality was present, but not a significant problem.
- Testing precision throughout the industry appears marginal at best—across-the-board improvements are needed.
- Minor changes to the specification limits are recommended.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Quality Control; Quality Assurance; Percent-within-Limits; PWL; Lot-Percent-Defective; Percent-Defective; LPD; Statistical Process Control;	No restrictions.

19. SECURITY CLASSIF. (OF THIS REPORT)	20. SECURITY CLASSIF. (OF THIS PAGE)	21. NO. OF PAGES	22. PRICE
Unclassified	Unclassified	121	

Form DOT F 1700.7 (8-72)

[This page intentionally left blank]

Table of Contents

List of Figures

List of Tables

Executive Summary

The use of *Percent-Within-Limits* (PWL) specifications by state highway agencies throughout the United States continues to increase. The implementation of PWL specifications by the Oklahoma Transportation Authority (OTA) successfully increased the level of quality achieved on over $150 million worth of paving projects. This study, commissioned by the Oklahoma Department of Transportation, examined the OTA's implementation of PWL specifications. Five key conclusions can be drawn from this study as follows:

- Overall, the Contractors' Processes Demonstrated High Capability with Respect to the Specified Quality Levels
- Greater Emphasis is Needed on Statistical Process Control
- Non-Normality was Present, but Not a Significant Problem
- Testing Precision throughout the Industry appears Marginal at Best—Across-the-Board Improvements are Needed
- Minor Changes to the Specification Limits are Recommended

Each of these topics is covered in-depth in the pertinent sections of this report. Summary discussions of these key conclusions are provided below.

Overall, the Contractors' Processes Demonstrated High Capability with Respect to the Specified Quality Levels

The "Acceptable Quality Level" (or AQL) included in the OTA PWL specifications was ninety percent-within-limits, meaning that contractors producing at 90 PWL should receive full payment for their work. Almost without exception, the contractors for the projects included in this study demonstrated that they are highly capable of achieving this level of quality (with 127 lots out of 161 having composite pay factors greater than 100%, representing exceptional quality). Even though overall process capability was quite high, many of the processes lingered near the outer limits of the specification. To increase the emphasis on consistent production of *on-target* materials, the use of target-adjusted standard deviations is recommended. The procedures and sample specifications presented in this report incorporate the use of target-adjusted standard deviations for calculating PWL.

Greater Emphasis is Needed on Statistical Process Control

Statistical process control (SPC) charting represents an extremely powerful tool for quickly identifying and correcting external sources of variation in a process. Strictly speaking, if a process is "not in statistical control," PWL and process capability calculations lose their significance. If high levels of quality are to be achieved, the following three conditions must be consistently satisfied:

1. The process is in statistical control (meaning that only common-cause variation is present).
2. The process is capable (meaning that the common-cause variation for the process is small enough to allow the process to consistently remain within the specified limits).
3. The process is on target (meaning that the process is performing at or near the specified target).

The recommended procedures and sample specifications set forth in this report are meant to assist the Department in ensuring that these three criteria are consistently met.

Non-Normality was Present, but Not a Significant Problem

The calculations that form the basis for determining PWL and the resulting pay factors assume that the population from which the data are sampled follow a normal distribution. Although evidences of non-normality were observed with some of the data, the levels and types of non-normality do not suggest that the resulting estimates of PWL were negatively influenced. Two recommendations are given for dealing with the issue of non-normality in the context of PWL specifications:

- Increase the resolution of the reported measurements for certain quality characteristics (such as gradation percent-passing).
- Use statistical process control (SPC) charting to identify whenever a significant change occurs that effectively results in a "new" process; treat the new process as a separate population.

Testing Precision throughout the Industry appears Marginal at Best—Across-the-Board Improvements are Needed

A general lack of precision currently exists with respect to the available test methods for measuring pavement quality. This lack of precision stretches industry-wide. Although this lack of precision is alarming, it does not negate the usefulness of PWL specifications. Indeed, the problems associated with using marginal measurement systems to distinguish "good" products from "bad" products remains, irrespective of the acceptance-procedure employed. A number of procedures are set forth in this report to help alleviate the difficulties associated with testing variability. The key recommendations are as follows:

- Establish and implement routine procedures for identifying and correcting biases between the Department's and Contractor's test methods.
- Rely upon the Contractor's test methods whenever they are demonstrated to be unbiased.
- Establish research funding opportunities that encourage innovation in the development of improved measurement methods and equipment.

Minor Changes to the Specification Limits are Recommended

As mentioned previously, the process capability demonstrated by the contractors was high. The following changes from the OTA PWL specification limits are recommended for implementation on the upcoming ODOT PWL pilot projects:

- Establish target limits for all PWL quality characteristics.
- Decrease the allowable deviation (from Job Mix Formula) for AC Air Voids from +/- 2.5 to +/- 2.0 (%).
- Decrease the allowable deviation (from Job Mix Formula) for AC Asphalt Cement Content from +/- 0.6 to +/- 0.4 (%).
- Revise the Composite Pay Factor equation for AC from 30/30/30/10 to 40/30/20/10 (for Roadway Density / Air Voids / Asphalt Cement Content / Gradation).
- Establish a lower specification limit for PCC Class AP Concrete Strength (e.g. 3,000 psi).
- For PCC gradation, use the percent-passing the No. 200 sieve (applied separately to both fine and coarse aggregates).
- Revise the Pay Factor equation to provide for an increase in the maximum potential incentive (going from 102 to 104% pay at 100 PWL) and a reduction in pay at the rejectable quality level (going from 60 to 50% pay at 50 PWL).

Introduction

Statistical quality control / quality assurance (QC/QA) specifications are intended to place responsibility on the Contractor for *controlling* the quality of the final product, while the highway agency retains responsibility for *acceptance*, *rejection* and *price adjustment* based on the level of quality actually delivered. These specifications, also commonly referred to as "Lot-Percent-Defective" (LPD) or "Percent-Within-Limits" (PWL) specifications, relate payment to actual performance based on the percentage of materials and construction that fall within the specified limits (PWL) or the percentage falling outside those limits (LPD).

In 1996, the American Association of State Highway and Transportation Officials (AASHTO) published two documents to aid state highway agencies (SHAs) in the development of QC/QA specifications (AASHTO 1996a,b). Currently, at least 40 of the nation's SHAs are implementing statistical QC/QA specification programs in one form or another. Of those 40+ SHAs, 90% provide incentive payments to the contractors when higher-than-specified levels of quality are achieved (Butts and Ksaibati 2002).

In 1996, utilizing the pre-published draft versions of the aforementioned AASHTO documents, Ahmed developed a set of draft QC/QA special provisions for the Oklahoma Department of Transportation (ODOT) to address the quality control and quality assurance of new asphalt concrete (AC) pavements, portland cement concrete (PCC) pavements, and concrete bridge decks (Ahmed 1996a,b,c). Those QC/QA special provisions, which utilized Lot-Percent-Defective terminology, have never been implemented by ODOT. However, in 1998, the Oklahoma Transportation Authority (OTA) adapted the AC and PCC versions of the ODOT draft LPD special provisions for use on over $150 million worth of turnpike paving projects. At the conclusion of those projects, OTA's on-site representatives reported that the contractors provided greater attention to and focus on quality as a result of the QC/QA specifications, yielding a higher level of overall quality. Table 1 lists the OTA projects included in this study.

Table 1 – PCC and AC Paving Projects included in this Study

Contract	Contractor	AC/PCC	Total Contract Amount ($ millions)
CKT-1151	Koss	PCC	11.1
CKT-1354	Duit	PCC	19.6
CKT-1355	Keck	PCC	11.0
CKT-1356	Western Plains	PCC	11.9
CKT-1451	Duit	PCC	2.7
CKT-1452	Duit	PCC	12.6
CKT-1453	APAC	AC	6.9
CKT-1453A	TTK	PCC	5.4
HEB-1551	Broce	AC	7.4
HEB-1553	Haskell-Lemon	AC	8.3
JKT-1251	Wittwer Paving	PCC	3.0
JKT-1253	Duit	PCC	8.7
JKT-1256	Duit	PCC	11.2
JKT-1258	Duit	PCC	9.3
WR-MC-70	Glover	AC	10.6
WR-MC-71	APAC	AC	3.5
WR-MC-72	Cummins	AC	8.8
WR-MC-76C	Cummins	AC	4.9

In 2002, as the final OTA projects were nearing completion, ODOT commissioned this study to evaluate the implementation of OTA's QC/QA specifications. The aim of the study was to address the following broad objectives:

- Provide a means for objective assessment by ODOT of the implementation of percent-within-limits specifications.
- Investigate alternatives for dealing with quality characteristics having non-normal distributions.
- Investigate alternatives for limiting contractor exposure under percent-within-limits specifications due to variability in test methods and procedures.
- Identify suitable adjustments to percent-within-limits specification limits for various quality characteristics.
- Provide guidelines and recommendations concerning statistical methods for quality assurance testing.

This report presents recommendations and data-analysis results in conjunction with these objectives. Throughout this report, the original ODOT draft special provisions and the subsequent OTA special provisions will both be referred to as "PWL specifications" even though they were written using LPD terminology. This change in terminology is intended to shift the focus toward the positive aspects of achieving specified quality levels rather than on the negative aspects of failure to do so.

Identifying and Dealing with Non-Normality

Whereas *Percent-Within-Limits* (PWL) calculations assume *normally*-distributed data, estimating PWL for quality characteristics exhibiting *non-normal* distributions can either overstate or understate the true PWL. Improper estimation of the true PWL is most pronounced when the deviation from normality is due to multimodality or excessive skewness or kurtosis. Multimodality represents the condition where the distribution has more than one mode, or peak. Typically, multimodality is an indication that two or more distinct populations are present, each with its own mode. Skewness is the measure of asymmetry of the distribution. Kurtosis represents the "flatness" or "peakedness" of the distribution, with relatively flat distributions known as "platykurtic" and highly-peaked distributions referred to as "leptokurtic." In the early 1900s, Pearson demonstrated that the shape of most unimodal continuous distributions can be adequately described by four parameters—the first four "moments about the mean"—which are the mean, standard deviation, skewness and kurtosis. The normal distribution is a special-case unimodal distribution where the skewness and kurtosis are both equal to zero.

However, a distribution can have zero skewness and kurtosis, yet still be non-normal. This can be the case when the data are discrete rather than continuous. Discrete distributions, such as binomial or poisson distributions, are typically the result of processes that involve "count" data, (e.g. the number of people attending a meeting, number of defective parts, etc.). However, discrete data can also be observed as a result of measurement rounding, even if the underlying distribution is continuous. This occurs when the resolution of the measurement system is inadequate, thus forcing the would-be continuous values into discrete bins. Figure 1 shows the histogram and descriptive statistics for a sample of N = 100 observations taken randomly from a population having a normal distribution with mean = 10 and standard deviation = 1.

The Anderson-Darling (A-D) normality test for this sample (a common test used to evaluate the normality of a distribution) has a p-value = 0.840 (meaning that there is an 84% likelihood that the data came from a normal population). Figure 2 depicts the *exact same* data, except that the observations have been rounded to the nearest whole number (thus simulating a low-resolution measurement system). The corresponding A-D normality test for the rounded data has a p-value < 0.001, meaning there is less than one chance in 1,000 that the observations came from a normal population (even though they actually did—they were simply rounded to the nearest whole number).

For these two hypothetical data sets, Figures 3 and 4 provide graphical representations of their calculated process capability (assuming lower and upper specification limits of 8 and 12). As can be seen, the capability indices (C_p, P_p, etc.) and calculated "expected performance" values (which are actually *Lot-Percent-Defective* values) are relatively unchanged, despite the irrefutable non-normality of the distribution based on rounded measurements. The true overall *Lot-Percent-Defective* (LPD) for these hypothetical examples is 4.6%. The sample based on high-resolution measurements (shown in Figures 1 and 3) estimated this value at 3.2%. The estimated LPD from the discrete (rounded) sample (Figures 2 and 4) was 4.3%.

The hypothetical scenarios described above show how non-normal data can still provide a decent estimate of *Lot-Percent-Defective* (or *Percent-Within-Limits*) when the non-normality is due to

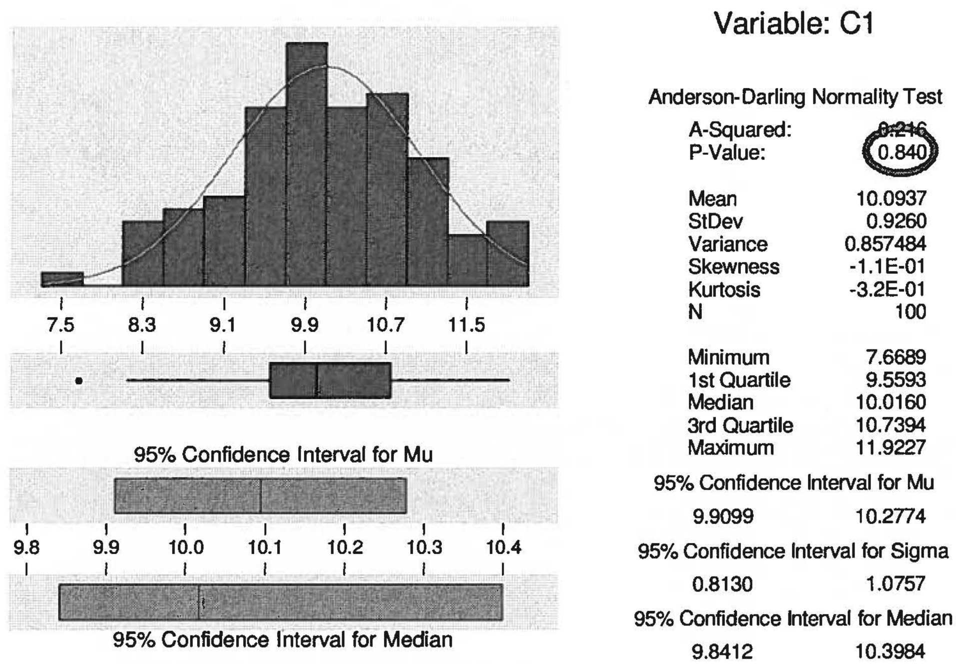


Figure 1 – Descriptive Statistics for Random Data
(N = 100; Mean = 10; Standard Deviation = 1)

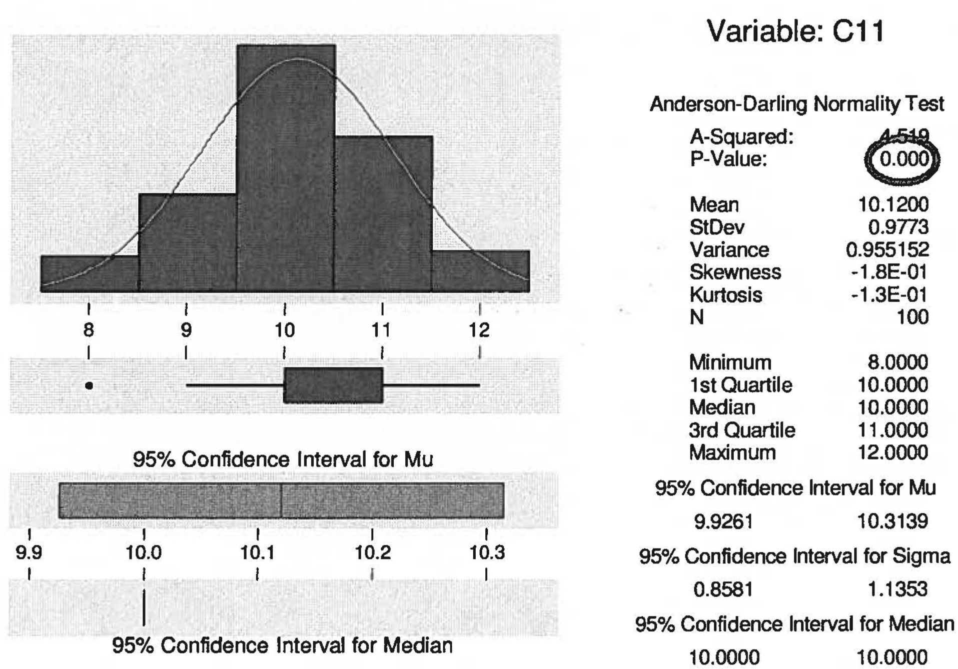


Figure 2 – Descriptive Statistics for Rounded Random Data
(N = 100; Mean = 10; Standard Deviation = 1)

rounded measurements rather than deviations from normality due to the shape of the distribution. Nonetheless, for acceptance testing performed under *Percent-Within-Limits* specifications, the measured quality characteristics should be reported with higher resolutions than are currently used within the industry. For example, ASTM C136-01 (*Standard Test Method for Sieve Analysis of Fine and Coarse Aggregates*) recommends reporting gradation percent-passing values to the nearest whole percentage (ASTM 2002). ASTM's standard test methods commonly recommend gross rounding in an effort to limit false perceptions about the precision of the reported measurements. However, when the process operates over a relatively narrow range (e.g. 95 – 100 for percent-passing the 1-inch sieve for PCC coarse aggregate), the resulting standard deviation estimates (using rounded measurements) can be somewhat erratic. The extreme case occurs when all the measured values within a lot end up rounded to the same whole number, causing the estimated standard deviation to be zero. As a rule of thumb, the individual test measurements should be reported to a resolution such that at least twenty (and preferably fifty or more) unique values are possible within the expected operating range for the process. As such, percent-passing the No. 200 sieve for PCC coarse aggregate (having a 0 – 2 specification range) should be reported to the nearest 0.1 or 0.01. Day (1999) aptly addresses the issue of rounding of test results when he states:

> It is bad practice to round calculations before the very last step. The strength of the individual specimen **used** to the be last step, but now we have hopefully realized that this should no longer be the case. Action on compressive strength results should always be based on the analysis of groups of test results, effectively ignoring individual test results. So it is the mean and standard deviation of a number of results which has significance.

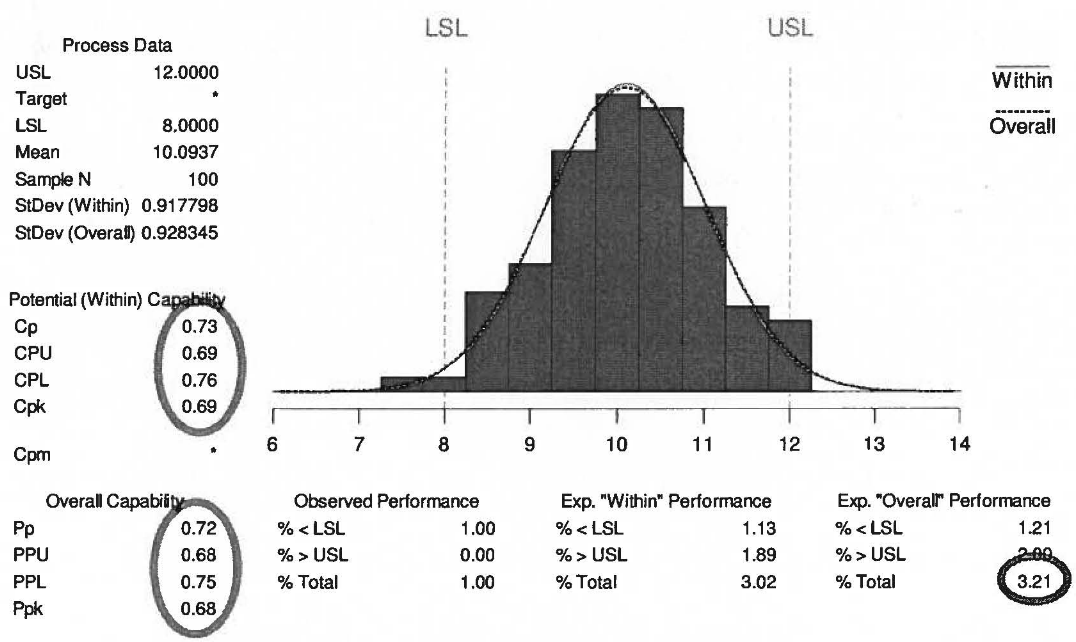


Figure 3 – Capability Analysis for Random Data
(N = 100; Mean = 10; Standard Deviation = 1; LSL = 8; USL = 12)

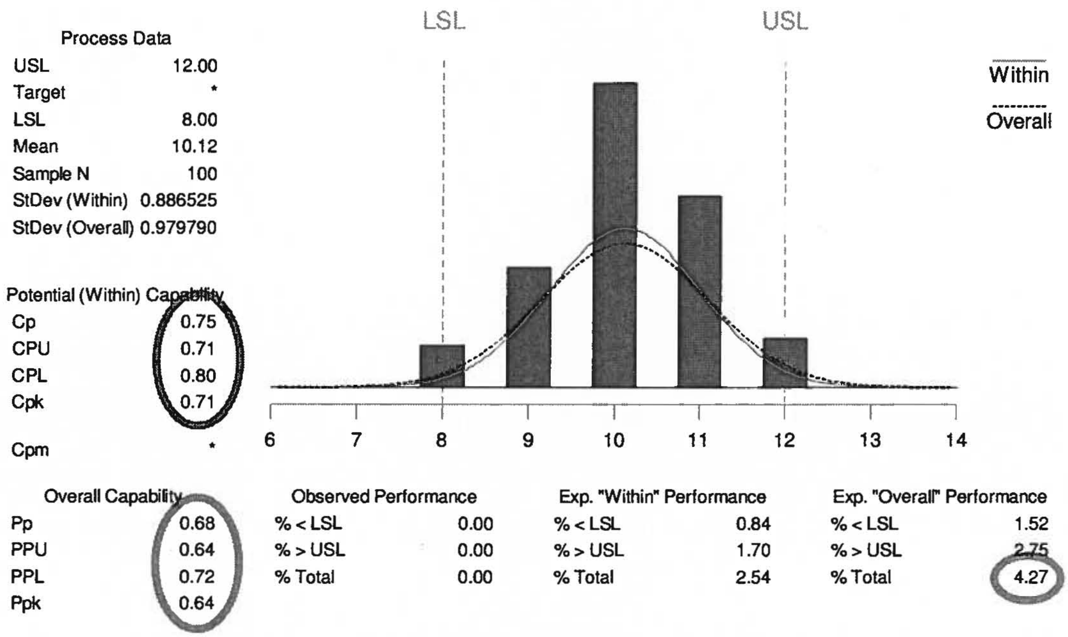


Figure 4 – Capability Analysis for Rounded Random Data
(N = 100; Mean = 10; Standard Deviation = 1; LSL = 8; USL = 12)

Figure 6 provides the results of the non-normality analysis of the compressive-strength data from the PCC projects included in this study. The results are presented as "p-values" based on the Anderson-Darling normality test. Similarly, Figure 7 shows the non-normality test results for percent-passing the 1-inch sieve for PCC coarse aggregates. Appendix C provides charts (similar to Figures 6 and 7) identifying the non-normality test results for all the PCC and AC quality characteristics included in this study.

For the PCC strength data (Figure 6), using $\alpha = 0.05$ leads us to conclude that in five instances (out of 22), the data are likely to have come from non-normal populations. [NOTE: Alpha (α) is the probability, or risk, of wrongly concluding that the data are not normal even if they actually are normal.] However, in all five instances, analysis of run charts of the data suggest that the non-normality was the result of special-cause variations occurring within each respective paving process. Figure 5 provides an example run chart (for Class A concrete on CKT-1452), which shows significant instances where distinct changes occurred with the concrete production process, thus explaining the likely cause of the observed non-normality.

The normality tests on PCC coarse-aggregate gradation data for percent-passing the 1-inch sieve (Figure 7) suggest that 19 of the 22 data sets are from populations having non-normal distributions. However, each of these 19 data sets exhibited the effects of measurement rounding, which has a profound effect on the results of the Anderson-Darling normality test (causing otherwise "normal" data to appear grossly non-normal).

In an attempt to show the instances where the observed non-normality was likely due to measurement-system resolution, the non-normality charts (shown in Figures 6 and 7 and in

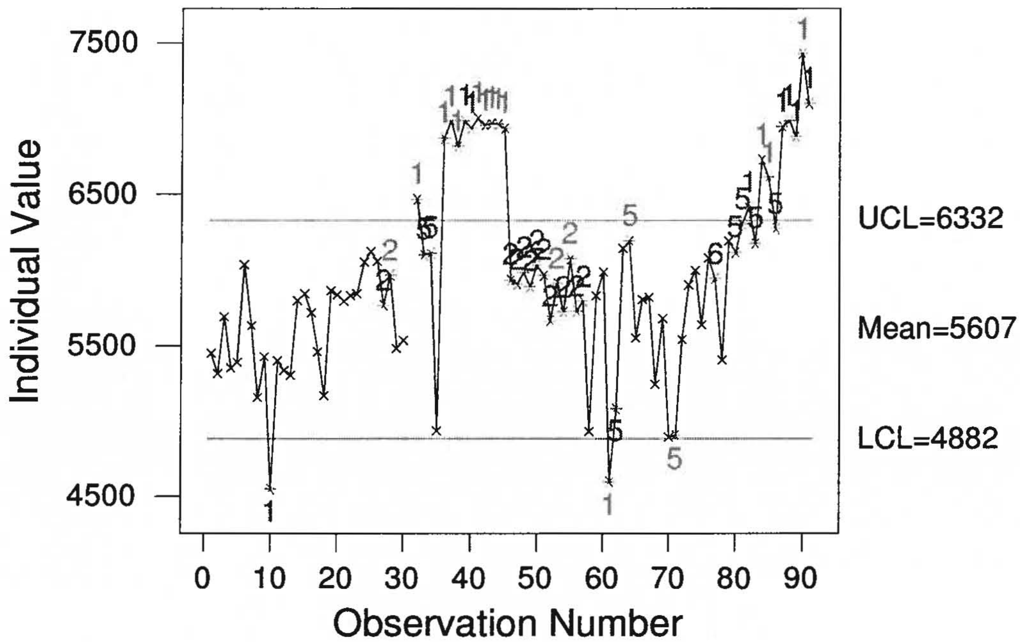


Figure 5 – Individuals Control Chart for PCC Strength (Class A) for Contract CKT-1452

Appendix C) also provide the "measurement resolution" for each sample, where the measurement resolution is reported as the percentage of observed values that are truly unique. As an example, if the sample includes the following four observations: (3, 3, 4 and 4), only 50% of the observed values are unique (3 and 4 – two out of the four values). Similarly, a sample including the following ten observations: (2, 2, 2, 3, 3, 3, 3, 4, 5, and 6) would also have 50% unique values (2, 3, 4, 5 and 6 – five out of the ten values). Whenever the measurement resolution is below about 30%, the data are *very* likely to fail the normality test even if the true shape of the distribution is essentially normal.

Instances of non-normality resulting from special-cause variations was observed with much of the quality-control data collected. This suggests that statistical process control charts were used sparingly, incorrectly, or not at all during the execution of the projects. Of those data sets that "failed" the normality test, the majority exhibited very low measurement resolution and/or significant process changes observable as special-cause variations. As such, the data do not suggest that the PWL or the pay-factor calculations for the projects were poorly estimated due to violations of the normality assumptions.

As can be seen by the charts in Appendix C, the aggregate gradation data for the projects exhibited considerable non-normality (168 of 256 instances for PCC and 79 of 204 instances for AC, based on $\alpha = 0.05$). However, much of the aggregate gradation data exhibited the effects of rounding described above. After removing data sets that exhibited considerable non-normality due to rounding, those numbers drop to 38 of 112 (for PCC) and 65 of 188 (for AC). For the non-gradation quality characteristics included in the study, 81 of 242 data sets exhibited non-normality (using $\alpha = 0.05$, ignoring data sets with measurement resolution less than 30%).

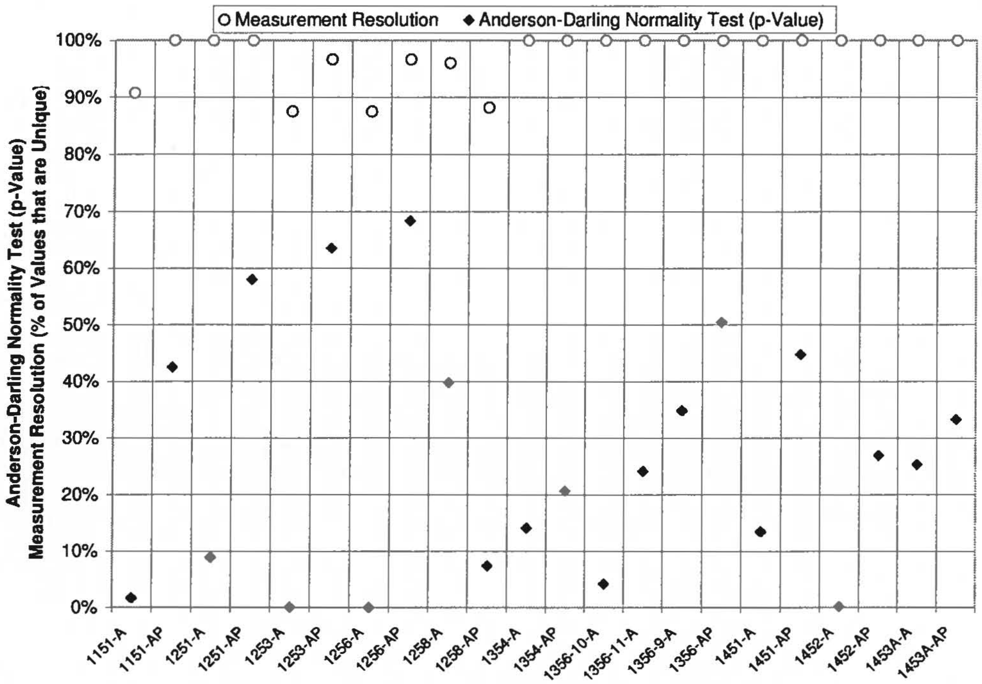


Figure 6 – Normality Test on PCC Strength Data

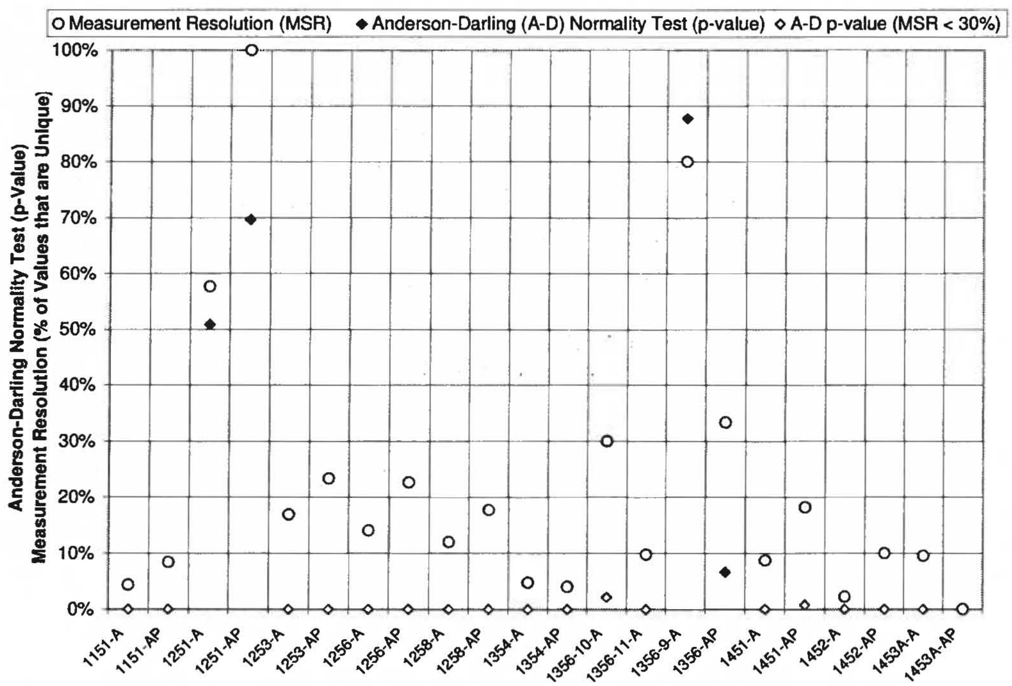


Figure 7 – Normality Test on PCC 1-inch Coarse Aggregate Data

Even though the estimates of PWL did not appear to be adversely affected by non-normality issues, various means are available for dealing with non-normal data. Table 2 presents five methods for transforming or otherwise negating the impacts of the various causes of non-normality. The researcher recommends implementing Methods 1 and 2 for future PWL projects.

Table 2 – Methods for Transforming Non-Normal Data

<table>
<tr><th>Type of Non-Normality</th><th>Methods for Dealing with the Non-Normality</th></tr>
<tr><td>Discrete observations due to rounding</td><td>1. Report the measurements using additional significant digits (e.g. to the nearest tenth or hundredths place for gradation percent-passing)</td></tr>
<tr><td>Bimodal distributions (having more than one peak)</td><td>2. Use statistical process control (SPC) charting to identify each time the process changes to a "new" process; when the process changes, treat the new process as a separate process and, therefore, as a separate population.</td></tr>
<tr><td>Distributions having a physical limit</td><td rowspan="4">3. Increase the frequency of the measurements, then use the averages from multiple measurements, rather than the individual measurements themselves. As the number of measurements averaged and taken as a single measurement increases, the population of averages will, in accordance with the Central Limit Theorem, approach a normal distribution irrespective of the shape of the original distribution.

4. Utilize Pearson curve-fitting by moments to define the shape of the distribution, then perform capability analysis calculations based on the curve-fit distribution.

5. Transform the data using a power- or logarithmic-transformation (e.g. Box-Cox).</td></tr>
<tr><td>Skewed distributions (asymmetric)</td></tr>
<tr><td>Platykurtic distributions (flat)</td></tr>
<tr><td>Leptokurtic distributions (peaked)</td></tr>
</table>

Testing Variability

Testing variability is a significant issue within the construction industry, particularly as it relates to *Percent-Within-Limits* (PWL) specifications. PWL specifications seek to reward consistency (i.e. low variability), yet high testing variability increases the *overall* observed variability. Whereas the PWL specifications are unable to distinguish between testing variability and materials variability, *all* the observed variability ends up being attributed to the *materials* themselves. This occurs even though a considerable portion of the observed variability may be the result of repeatability and reproducibility errors associated with the measurement systems rather than variability with the materials. A common method for reducing testing variability involves the use of average values of repeat tests (which reduces the testing variability by a factor of one over the square root of n, where n equals the number of individual tests that are averaged and considered as a single test). This is the reason compressive-strength testing typically involves the average of two or three cylinder specimens considered as a single test. This approach is valid as long as the specimens are sampled from the exact same materials (e.g. from the same batch and the same proximity within the batch). Figure 8 shows the reduction in testing error resulting from sampling based on the averages of repeat tests.

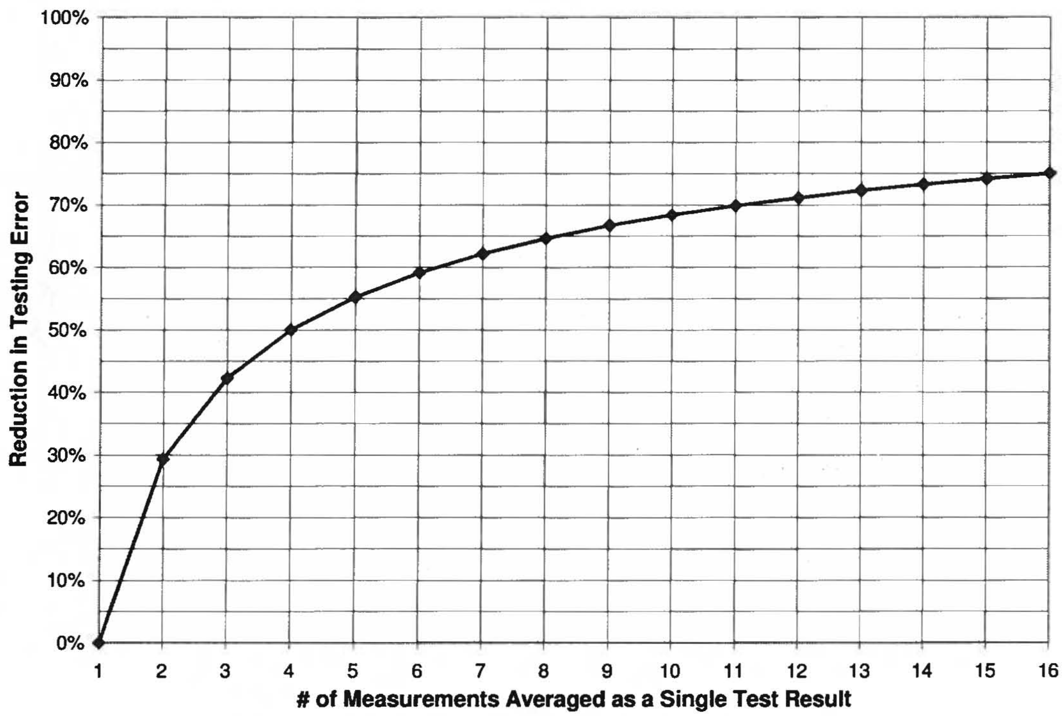


Figure 8 – Potential Reduction in Testing Error by using Average Results from Repeat Tests

With respect to PWL specifications, averaging test results should not be performed if the test specimens are from different batches of materials or from different locations within a lot or sublot. For example, the Oklahoma Department of Transportation's original draft PWL special provision for PCC states the following:

Thickness determination: three cores per sublot with additional cores taken as necessary in accordance with Subsection 414.04(t)-2, of the Standard Specifications. The average of the three thickness measurements is considered as one individual test.

The wording of this draft special provision suggests that the thickness measurements from three different cores be averaged and considered as a single test. As stated above, averaging across different batches or different locations is not recommended in the context of PWL specifications. This approach in effect reduces the measured variability due to the *materials* as well as the variability due to testing error.

Tables 3 – 5 provide materials and testing standard deviations and variances for several AC quality characteristics (including gradation). Table 6 shows similar data for three PCC quality characteristics (unit weight, air content, and compressive strength). The AC data are from a research study conducted for the Oklahoma Department of Transportation aimed at quantifying the materials and testing variability associated with AC pavement construction (Ahmed 1997). The PCC data are from a "Gauge R&R" study conducted as part of the current research study. These tables (Tables 3 – 6) show the standard deviation of the testing error in red whenever the value is 50% higher than the related ASTM multilaboratory precision statement. Nine of the 28 test procedures presented in Tables 3 – 6 showed standard deviations thusly higher than the corresponding ASTM values. These higher standard deviations suggest that precision-improvements should be attainable for those particular test procedures and, as such, increased scrutiny over those procedures is warranted. Whenever the observed testing variability was less than the ASTM values, the corresponding cells in the tables have been highlighted in light green (5 instances) and whenever the observed values were more than 50% below the ASTM values, the cells have been shown bright green (5 instances), thus indicating precision levels better than the industry norms.

Table 3 – Testing Variability Comparisons for AC Gradation Testing (Mixture Analysis)

	Gradation – Mixture Analysis (Devation from JMF)								
	1"	3/4"	1/2"	3/8"	No. 4	No. 10	No. 40	No. 80	No. 200
Standard Deviation	(%-passing)								
Between Projects	0.31	0.11	1.78	3.56	2.15	1.64	2.27	1.40	1.32
Within Project (Between Lots)	0.28	0.75	1.58	1.77	1.67	1.69	1.03	0.47	0.38
Within Lot (Between Sublots)	0.16	0.33	0.82	1.65	2.18	1.97	0.96	0.69	0.57
Within Sublot (Between Tests)	0.59	1.52	2.77	3.49	3.42	3.01	3.00	0.82	0.49
ASTM Single-Operator Precision (D 5444)	0.50	1.00	1.00	1.00	1.00	0.70	0.70	0.30	0.20
ASTM Multilaboratory Precision (D 5444)	0.50	1.20	1.20	1.20	1.20	0.90	0.90	0.60	0.40
TOTAL	**0.75**	**1.73**	**3.74**	**5.54**	**4.89**	**4.30**	**4.02**	**1.83**	**1.56**

Variance	$(\%\text{-passing})^2$								
Between Projects	0.10	0.01	3.16	12.65	4.64	2.70	5.15	1.97	1.74
Within Project (Between Lots)	0.08	0.57	2.49	3.15	2.78	2.86	1.06	0.22	0.15
Within Lot (Between Sublots)	0.03	0.11	0.67	2.73	4.77	3.88	0.92	0.47	0.32
Within Sublot (Between Tests)	0.35	2.31	7.67	12.16	11.69	9.06	9.03	0.67	0.24
TOTAL	**0.56**	**3.00**	**13.99**	**30.70**	**23.89**	**18.50**	**16.17**	**3.34**	**2.45**

Table 4 – Testing Variability Comparisons for AC Gradation Testing (Cold Feed Analysis)

	Gradation -- Cold Feed Analysis (Devation from JMF)								
	1"	3/4"	1/2"	3/8"	No. 4	No. 10	No. 40	No. 80	No. 200
Standard Deviation	(%-passing)								
Between Projects	0.32	1.48	2.77	2.18	2.01	2.40	2.57	0.89	0.66
Within Project (Between Lots)	0.21	1.27	0.88	1.04	1.38	1.49	0.88	0.59	0.39
Within Lot (Between Sublots)	0.39	0.88	1.39	1.88	1.60	0.99	0.64	0.36	0.23
Within Sublot (Between Tests)	**1.03**	**2.22**	**2.73**	**2.75**	2.24	1.87	**0.79**	**0.62**	**0.38**
ASTM Single-Operator Precision (C 136)	0.32	0.81	2.25	2.25	1.32	0.83	0.54	0.36	0.37
ASTM Multilaboratory Precision (C 136)	0.35	1.37	2.82	2.82	1.97	1.41	1.10	0.73	0.65
TOTAL	**1.17**	**3.08**	**4.22**	**4.12**	**3.68**	**3.53**	**2.90**	**1.28**	**0.89**

Variance	$(\%\text{-passing})^2$								
Between Projects	0.10	2.18	7.67	4.74	4.06	5.78	6.60	0.79	0.44
Within Project (Between Lots)	0.04	1.61	0.78	1.09	1.89	2.21	0.77	0.35	0.15
Within Lot (Between Sublots)	0.15	0.77	1.93	3.54	2.56	0.99	0.41	0.13	0.05
Within Sublot (Between Tests)	1.06	4.93	7.46	7.58	5.02	3.51	0.63	0.39	0.14
TOTAL	**1.36**	**9.49**	**17.83**	**16.96**	**13.53**	**12.49**	**8.41**	**1.65**	**0.78**

Table 5 – Testing Variability Analysis for Various AC Quality Characteristics

	AC Content (Extraction)	AC Content (Nuclear Gauge)	Air Voids	Hveem	Roadway Density (Core)	Roadway Density (Nuclear Gauge)
Standard Deviation	(% by weight)	(% by weight)	(% by volume)	(%)	(% max theor.)	(% max theor.)
Between Projects	0.33	0.26	1.44	3.40	3.24	1.76
Within Project (Between Lots)	0.14	0.17	0.85	10.49	0.72	1.06
Within Lot (Between Sublots)	0.15	0.14	0.65	3.75	1.38	1.28
Within Sublot (Between Tests)	**0.16**	**0.11**	**0.42**	**4.24**	**0.77**	**0.67**
ASTM Single-Operator Precision	0.21	0.16	0.58	9.00	0.55	0.44
ASTM Multilaboratory Precision	0.23	0.23	1.11	21.00	1.20	0.44
TOTAL	**0.42**	**0.36**	**1.84**	**12.40**	**3.68**	**2.51**

Variance	$(\%\text{ by weight})^2$	$(\%\text{ by weight})^2$	$(\%\text{ by volume})^2$	$(\%)^2$	$(\%\text{ max theor.})^2$	$(\%\text{ max theor.})^2$
Between Projects	0.11	0.07	2.08	11.53	10.51	3.11
Within Project (Between Lots)	0.02	0.03	0.72	110.05	0.52	1.13
Within Lot (Between Sublots)	0.02	0.02	0.42	14.05	1.91	1.64
Within Sublot (Between Tests)	0.02	0.01	0.18	18.01	0.59	0.45
TOTAL	**0.17**	**0.13**	**3.39**	**153.65**	**13.53**	**6.32**

Table 6 – Testing Variability Analysis for Various PCC Quality Characteristics

	Unit Weight	Air Content	Compressive Strength
Standard Deviation	(pcf)	(% by volume)	(psi)
Between Batch	0.92	0.26	220
Between Test	1.08	0.38	**343**
ASTM Single-Operator Precision	0.65	N/A	N/A
ASTM Multilaboratory Precision	0.82	0.28	158
TOTAL	**1.42**	**0.46**	**408**

Variance	$(\text{pcf})^2$	$(\%\text{ by volume})^2$	$(\text{psi})^2$
Between Batch	0.85	0.07	48,340
Between Test	1.17	0.14	117,808
TOTAL	**2.02**	**0.21**	**166,149**

When performing measurement-systems analyses (MSA), four common metrics for evaluating the suitability of the measurement systems are as follows:

- Precision-Over-Tolerance Ratio (P/T)
- Precision-to-Total-Variance Ratio (P/TV)
- Measurement-System Variance as a Percentage of Total Variance ($\%TV_{MS}$)
- Signal-to-Noise Ratio (SNR)

These measures are calculated according to Equations 1 – 4.

$$P/T = \frac{6 \cdot \sigma_{MS}}{USL - LSL} \qquad \text{(Equation 1)}$$

$$P/TV = \frac{\sigma_{MS}}{\sigma_{Total}} \qquad \text{(Equation 2)}$$

$$\%TV_{MS} = (P/TV)^2 = \frac{\sigma^2_{MS}}{\sigma^2_{Total}} = \frac{\sigma^2_{MS}}{\sigma^2_{Product} + \sigma^2_{MS}} \qquad \text{(Equation 3)}$$

$$SNR = \frac{\sigma_{Product}}{\sigma_{MS}} \approx \frac{\sqrt{\sigma^2{}_{Product} - \sigma^2{}_{MS}}}{\sigma_{MS}} \qquad \text{(Equation 4)}$$

where

P/T	=	Precision-over-tolerance ratio
T	=	Tolerance, or specification range (= $USL - LSL$)
σ_{MS}	=	Measurement-system standard deviation
USL	=	Upper specification limit
LSL	=	Lower specification limit
P/TV	=	Precision-to-total-variance ratio (measurement-system standard deviation as a percentage of total standard deviation)
σ_{Total}	=	Total standard deviation for the process (includes product variation plus measurement-system error)
$\%TV_{MS}$	=	Measurement-system variance as a percentage of total variance
σ^2_{MS}	=	Measurement-system variance
σ^2_{Total}	=	Total observed variance
$\sigma^2_{Product}$	=	Product variance
SNR	=	Signal-to-noise ratio
$\sigma_{Product}$	=	Product standard deviation

In their standard *Practice for Determining the Precision Over Tolerance (P/T) Ratio of Test Equipment*, Semiconductor Equipment and Materials International (SEMI) (1996) states the following:

> In general, one prefers that the value of the measurement precision *P*, be much smaller than the specification range, *T*. A test instrument is usually deemed to be suitable for the purpose if *P/T* lies below 10%. If *P/T* is greater than 30%, the test instrument is not likely to be suitable for the purpose. Cases for which *P/T* lies between 10% and 30% must be judged on an individual basis, depending on the requirements being placed on the measurement system.

Concerning Signal-to-Noise Ratio (SNR), Eastman (1995) states:

> It is desirable to have SNR as large as possible: greater than 10 implies that the instrument can be used to distinguish levels of quality of a product. An instrument with SNR less than 3 or 4 would generally be unsuitable for use.

Concerning the issue of when to use SNR versus when to use P/T, Eastman (1995) states:

> If the tolerance has been arbitrarily defined, P/T may be meaningless. The signal-to-noise ratio (SNR) may be more helpful in assessing the suitability of the measurement tool.

Table 7 summarizes the levels of acceptability suggested by SEMI (1996) and Eastman (1995) for P/T and SNR. In addition, Table 7 provides levels of acceptability for P/TV and $\%TV_{MS}$. These P/TV and $\%TV_{MS}$ acceptability levels are based on direct conversions from the acceptability levels for SNR suggested by Eastman (1995).

Table 7 – Acceptability Levels using P/T, P/TV, $\%TV_{MS}$, and SNR Metrics

	Unacceptable	**Marginal**	**Acceptable**
Precision-Over-Tolerance (P/T)	> 30%	10% to 30%	< 10%
Precision-to-Total Variance (P/TV)	> 32%	10% to 32%	< 10%
Measurement-System Variance as % of Total Variance ($\%TV_{MS}$)	> 10%	1% to 10%	< 1%
Signal-to-Noise Ratio (SNR)	< 3	3 to 10	> 10

Tables 8 – 11 present calculations for P/T, P/TV, $\%TV_{MS}$, and SNR for various quality characteristics. For each quality characteristic, all four MSA metrics are presented for each of the following three determinations of measurement-system variability—the "between test" values reported in Tables 3 – 6, as well as the applicable ASTM single-operator precision statements, and the ASTM multilaboratory precision statements. The P/T, P/TV, $\%TV_{MS}$, and SNR values in Tables 8 – 11 are color-coded based on the acceptability levels presented in Table 7.

As can be seen from Tables 8 – 11, the measurement systems for all the quality characteristics tested are marginal at best, with all the test methods falling into the "unacceptable" category for at least one MSA metric. For $\%TV_{MS}$, using the 1997 and 2002 study data, 21 of the 27 test methods were "unacceptable" and the other 6 were "marginal." For $\%TV_{MS}$, using the ASTM multilaboratory precision statements, 19 were unacceptable and 8 were marginal. This suggests that even the current ASTM standard practices warrant investigation into ways to further reduce the testing error associated with those procedures. It bears noting that the observed testing errors for sieve analyses as presented by Ahmed (1997) and reported herein also inherently include sampling error. As such, the observed errors could have been significantly influenced by the procedures employed for splitting the samples in addition to the actual testing errors.

Two of the objectives outlined for this study specifically focused on the issues of sampling and testing:

- Investigate alternatives for limiting contractor exposure under percent-within-limits specifications due to variability in test methods and procedures.
- Provide guidelines and recommendations concerning statistical methods for quality assurance testing.

Procedures have been developed as a part of this study to help meet these objectives. The procedures are presented in detail in Appendices H and I, with further discussion included in the section entitled *Statistical Quality Assurance*. The procedures provide a framework wherein considerable up-front side-by-side testing would be performed by the Department and the Contractor to establish whether or not any substantial biases exist between the Contractor's and the Department's test equipment or routine test procedures. If no discernable bias exists for a given quality characteristic, the frequency of side-by-side testing can be reduced. After the initial determination of a lack of significant bias, statistical process control (SPC) charting would be performed on the calculated differences between the Contractor's and the Department's subsequent side-by-side test results. As long as no "out of control" conditions are reported by the SPC procedures, the Department can rest assured that no significant change to the Contractor's testing procedures has occurred, thus further assuring that the initial no-bias conditions remain valid.

Although the procedures detailed in Appendices H and I do not in and of themselves reduce the testing error, they allow the Contractors to retain a measure of control over the equipment and technicians to be used for acceptance testing. And, whereas the Contractors have a financial incentive to reduce variability within the context of *Percent-Within-Limits* (PWL) specifications, they will also have an incentive to take measures to reduce testing variability as well (through the use of better equipment, better maintenance on existing equipment, increased focus on technician training, etc.). The obvious objection to this scenario relates back to the classic "fox guarding the hen house" analogy. Obviously, the Contractor will have a significant financial stake in the test results and, therefore, may try to exert pressure on the testing technicians to behave unethically. However, with PWL specifications, the standard deviation of the test results within a given lot can have a significant impact on the final pay adjustment. And, if at least one sublot per lot is performed under supervised side-by-side assurance testing, the potential payoff from unethical behavior will be substantially diminished.

Other means for dealing with testing variability include the use of average rather than individual test values (as discussed previously) or the development of altogether new testing equipment or procedures. The cost of taking and averaging multiple measurements can be considerable. In addition, this procedure follows the law of diminishing returns. Figure 8 shows the reduction in testing error associated with taking the average of multiple tests. As can be seen, four measurements can reduce testing error by 50% (perhaps a good return on the extra effort). However, it takes another twelve (for a total of sixteen) to reduce that error by yet another 50%

(resulting in a final testing error equal to 25% of the original—not likely to be worth the sixteen-fold effort!).

In recognition of the added costs and diminishing returns associated with reliance upon averages from multiple repeat measurements, true improvements to testing error will require the development of more robust procedures and equipment. The gravity of this need is reinforced by the fact that, in many instances, the precision levels published by ASTM are marginal at best. And, for that matter, the reference-testing that forms the basis for the published ASTM precision values is usually performed by laboratories and technicians who *know* that their test results will be directly compared to other laboratories and technicians, thus potentially leading to overstatements of the true precision of the test methods.

Table 8 – Measurement-System Analyses using P/T, P/TV, $\%TV_{MS}$, and SNR for AC Gradation Testing (Mixture Analysis)

	Gradation -- Mixture Analysis (Devation from JMF)								
	1"	3/4"	1/2"	3/8"	No. 4	No. 10	No. 40	No. 80	No. 200
Standard Deviation	**(%-passing)**								
Within Sublot (Between Tests)	0.59	1.52	2.77	3.49	3.42	3.01	3.00	0.82	0.49
ASTM Single-Operator Precision (D 5444)	0.50	1.00	1.00	1.00	1.00	0.70	0.70	0.30	0.20
ASTM Multilaboratory Precision (D 5444)	0.50	1.20	1.20	1.20	1.20	0.90	0.90	0.60	0.40
TOTAL	**0.75**	**1.73**	**3.74**	**5.54**	**4.89**	**4.30**	**4.02**	**1.83**	**1.56**
Precision Over Tolerance (P/T)	**(% of Specification Tolerance)**								
(from 1997 AC Variability Study)	30%	76%	139%	174%	171%	201%	200%	55%	73%
(using ASTM Single-Operator Precision)	25%	50%	50%	50%	50%	47%	47%	20%	30%
(using ASTM Multilaboratory Precision)	25%	60%	60%	60%	60%	60%	60%	40%	60%
Signal-to-Noise Ratio (SNR)	**Ratio**								
(from 1997 AC Variability Study)	0.8	0.5	0.9	1.2	1.0	1.0	0.9	2.0	3.1
(using ASTM Single-Operator Precision)	0.9	0.8	2.5	4.3	3.5	4.4	3.8	5.4	7.4
(using ASTM Multilaboratory Precision)	0.9	0.7	2.1	3.6	2.9	3.4	3.0	2.7	3.7
Precision-to-Total-Variance (P/TV)	**(% of TOTAL Standard Deviation)**								
(from 1997 AC Variability Study)	79%	88%	74%	63%	70%	70%	75%	45%	31%
(using ASTM Single-Operator Precision)	67%	58%	27%	18%	20%	16%	17%	16%	13%
(using ASTM Multilaboratory Precision)	[illegible]	69%	32%	22%	25%	21%	22%	33%	26%
Variance	**(% of TOTAL Variance)**								
Between Projects	18%	0%	23%	41%	19%	15%	32%	59%	71%
Within Project (Between Lots)	14%	19%	18%	10%	12%	15%	7%	7%	6%
Within Lot (Between Sublots)	5%	4%	5%	9%	20%	21%	6%	14%	13%
Within Sublot (Between Tests)	63%	77%	55%	40%	49%	49%	56%	20%	10%
Measurment-System Variance ($\%TV_{MS}$)	**(% of TOTAL Variance)**								
(from 1997 AC Variability Study)	[illegible]	77%	55%	40%	49%	49%	56%	20%	10%
(using ASTM Single-Operator Precision)	45%	33%	7%	3%	4%	3%	3%	3%	2%
(using ASTM Multilaboratory Precision)	[illegible]	48%	10%	5%	6%	4%	5%	11%	7%

Table 9 – Measurement-System Analyses using P/T, P/TV, $\%TV_{MS}$, and SNR for AC Gradation Testing (Cold Feed Analysis)

	Gradation – Cold Feed Analysis (Devation from JMF)								
	1"	**3/4"**	**1/2"**	**3/8"**	**No. 4**	**No. 10**	**No. 40**	**No. 80**	**No. 200**
Standard Deviation	**(%-passing)**								
Within Sublot (Between Tests)	1.03	2.22	2.73	2.75	2.24	1.87	0.79	0.62	0.38
ASTM Single-Operator Precision (C 136)	0.32	0.81	2.25	2.25	1.32	0.83	0.54	0.36	0.37
ASTM Multilaboratory Precision (C 136)	0.35	1.37	2.82	2.82	1.97	1.41	1.10	0.73	0.65
TOTAL	**1.17**	**3.08**	**4.22**	**4.12**	**3.68**	**3.53**	**2.90**	**1.28**	**0.89**
Precision Over Tolerance (P/T)	**(% of Specification Tolerance)**								
(from 1997 AC Variability Study)	52%	111%	137%	138%	112%	125%	53%	41%	57%
(using ASTM Single-Operator Precision)	16%	41%	113%	113%	66%	55%	36%	24%	56%
(using ASTM Multilaboratory Precision)	18%	69%	141%	141%	99%	94%	73%	49%	98%
Signal-to-Noise Ratio (SNR)	**Ratio**								
(from 1997 AC Variability Study)	0.5	1.0	1.2	1.1	1.3	1.6	3.5	1.8	2.1
(using ASTM Single-Operator Precision)	1.7	2.6	1.4	1.4	2.2	3.6	5.2	3.1	2.2
(using ASTM Multilaboratory Precision)	1.6	1.6	1.1	1.1	1.5	2.1	2.5	1.5	1.2
Precision-to-Total-Variance (P/TV)	**(% of TOTAL Standard Deviation)**								
(from 1997 AC Variability Study)	88%	72%	65%	67%	61%	53%	27%	48%	43%
(using ASTM Single-Operator Precision)	27%	26%	53%	55%	36%	23%	19%	28%	42%
(using ASTM Multilaboratory Precision)	30%	44%	67%	68%	54%	40%	38%	57%	73%
Variance	**(% of TOTAL Variance)**								
Between Projects	7%	23%	43%	28%	30%	46%	78%	48%	56%
Within Project (Between Lots)	3%	17%	4%	6%	14%	18%	9%	21%	19%
Within Lot (Between Sublots)	11%	8%	11%	21%	19%	8%	5%	8%	7%
Within Sublot (Between Tests)	78%	52%	42%	45%	37%	28%	8%	23%	18%
Measurment-System Variance ($\%TV_{MS}$)	**(% of TOTAL Variance)**								
(from 1997 AC Variability Study)	78%	52%	42%	45%	37%	28%	8%	23%	18%
(using ASTM Single-Operator Precision)	8%	7%	28%	30%	13%	6%	3%	8%	17%
(using ASTM Multilaboratory Precision)	9%	20%	45%	47%	29%	16%	14%	32%	54%

Table 10 – Measurement-System Analyses using P/T, P/TV, $\%TV_{MS}$, and SNR for Various AC Quality Characteristics

	AC Content (Extraction)	AC Content (Nuclear Gauge)	Air Voids	Hveem	Roadway Density (Core)	Roadway Density (Nuclear Gauge)
Standard Deviation	(% by weight)	(% by weight)	(% by volume)	(%)	(% max theor.)	(% max theor.)
Within Sublot (Between Tests)	0.16	0.11	0.42	4.24	0.77	0.67
ASTM Single-Operator Precision	0.21	0.16	0.58	9.00	0.55	0.44
ASTM Multilaboratory Precision	0.23	0.23	1.11	21.00	1.20	0.44
TOTAL	**0.42**	**0.36**	**1.84**	**12.40**	**3.68**	**2.51**
Precision Over Tolerance (P/T)	**(% of Specification Tolerance)**					
(from 1997 AC Variability Study)	94%	64%	51%	N/A	92%	80%
(using ASTM Single-Operator Precision)	126%	96%	69%	N/A	66%	53%
(using ASTM Multilaboratory Precision)	138%	138%	133%	N/A	144%	53%
Signal-to-Noise Ratio (SNR)	**Ratio**					
(from 1997 AC Variability Study)	2.5	3.2	4.3	2.7	4.7	3.6
(using ASTM Single-Operator Precision)	1.8	2.1	3.1	1.3	6.5	5.5
(using ASTM Multilaboratory Precision)	1.7	1.5	1.6	0.6	3.0	5.5
Precision-to-Total-Variance (P/TV)	**(% of TOTAL Standard Deviation)**					
(from 1997 AC Variability Study)	38%	30%	23%	34%	21%	27%
(using ASTM Single-Operator Precision)	51%	45%	31%	73%	15%	18%
(using ASTM Multilaboratory Precision)	55%	65%	60%	169%	33%	18%
Variance	**(% of TOTAL Variance)**					
Between Projects	61%	53%	61%	8%	78%	49%
Within Project (Between Lots)	12%	23%	21%	72%	4%	18%
Within Lot (Between Sublots)	13%	15%	12%	9%	14%	26%
Within Sublot (Between Tests)	14%	9%	5%	12%	4%	7%
Measurment-System Variance ($\%TV_{MS}$)	**(% of TOTAL Variance)**					
(from 1997 AC Variability Study)	14%	9%	5%	12%	4%	7%
(using ASTM Single-Operator Precision)	26%	20%	10%	53%	2%	3%
(using ASTM Multilaboratory Precision)	31%	42%	36%	287%	11%	3%

Table 11 – Measurement-System Analyses using P/T, P/TV, %TV_{MS}, and SNR for Various PCC Quality Characteristics

	Unit Weight	Air Content	Compressive Strength
Standard Deviation	**(pcf)**	**(% by volume)**	**(psi)**
Between Test	1.08	0.38	343
ASTM Single-Operator Precision	0.65	N/A	N/A
ASTM Multilaboratory Precision	0.82	0.28	158
TOTAL	**1.42**	**0.46**	**408**
Precision Over Tolerance (P/T)	**(% of Specification Tolerance)**		
(from 2002 Gauge R&R Study)	65%	75%	N/A
(using ASTM Single-Operator Precision)	39%	N/A	N/A
(using ASTM Multilaboratory Precision)	49%	56%	N/A
Signal-to-Noise Ratio (SNR)	**Ratio**		
(from 2002 Gauge R&R Study)	0.9	0.7	0.6
(using ASTM Single-Operator Precision)	1.4	N/A	N/A
(using ASTM Multilaboratory Precision)	1.1	0.9	1.4
Precision-to-Total-Variance (P/TV)	**(% of TOTAL Standard Deviation)**		
(from 2002 Gauge R&R Study)	76%	83%	84%
(using ASTM Single-Operator Precision)	46%	N/A	N/A
(using ASTM Multilaboratory Precision)	58%	61%	39%
Variance	**(% of TOTAL Variance)**		
Between Batch	42%	32%	29%
Between Test	58%	68%	71%
Measurment-System Variance (%TV_{MS})	**(% of TOTAL Variance)**		
(from 2002 Gauge R&R Study)	58%	68%	71%
(using ASTM Single-Operator Precision)	21%	N/A	N/A
(using ASTM Multilaboratory Precision)	33%	38%	15%

Process Capability and Contractor Performance

Process capability charts for each of the quality characteristics included in the study are presented in Appendix D. Figure 9 provides a sample process capability chart. Three process capability indices (C_p, P_p, and C_{pm}) are reported for each contract item for each paving contract included in the study. C_p represents short-term capability; P_p relates to long-term capability; and C_{pm} is the target-sensitive short-term capability (sometimes referred to as the "Taguchi Index"). Both C_p and P_p are measures of the "best-case" capability, in that they present the capability of the process assuming it is perfectly centered within the specification limits, with the difference between the two being that C_p is calculated based on the pooled, *short-term* standard deviation of the process and P_p is calculated from the overall, or *long-term*, standard deviation. C_{pm}, on the other hand, takes into consideration the deviation of the process mean from the process target (where the target need not be centered within the specification limits). To that end, C_{pm} provides an adjustment to the short-term standard deviation based on the distance of the process mean from the process target. If the process mean is at the process target and the process target is in the middle of the specification limits, C_p and C_{pm} will be identical. Large differences between C_p and C_{pm} indicate a process that may have good potential (a relatively small short-term standard deviation) but is not currently operating near the process target. This was definitely the case with AC Roadway Density, as can be seen in Figure 9. All the C_p and P_p capability indices are above 1.0 (and thus safely above the "bonus/penalty" threshold of 0.55, which equates to 90 percent-within-limits). However, over half the C_{pm} values are below 1.0 and eight (of the 33) are even below 0.55.

The standard calculations for C_p, P_p and C_{pm} are given in Equations 5 – 7. Equation 8 shows the calculation for the target-adjusted standard deviation used by C_{pm}.

$$C_p = \frac{USL - LSL}{6 \cdot \sigma_{ST}} \qquad \text{(Equation 5)}$$

$$P_p = \frac{USL - LSL}{6 \cdot \sigma_{LT}} \qquad \text{(Equation 6)}$$

$$C_{pm} = \frac{USL - LSL}{6 \cdot \sigma''_{ST}} \qquad \text{(Equation 7)}$$

$$\sigma''_{ST} = \sqrt{\sigma^2_{ST} + (\mu - T)^2} \qquad \text{(Equation 8)}$$

where

USL	=	Upper specification limit
LSL	=	Lower specification limit
σ_{ST}	=	Short-term process standard deviation (or short-term *sample* standard deviation if the true short-term process standard deviation is not known)

σ_{LT} = Long-term process standard deviation (or long-term *sample* standard deviation if the true process standard deviation is not known)

σ''_{ST} = Target-adjusted short-term process standard deviation

μ = Process mean (or *sample* mean if the true process mean is not known)

T = Process target (ideal process mean)

For one-sided specifications, such as PCC strength, Equations 5 and 6 are typically modified by changing the "6" in the divisor to a "3" and replacing "$USL - LSL$" in the numerator with "$USL - \mu$" or "$\mu - LSL$" depending on whether the specified limit is an upper limit or a lower limit. C_{pm} is not normally calculated for one-sided specifications.

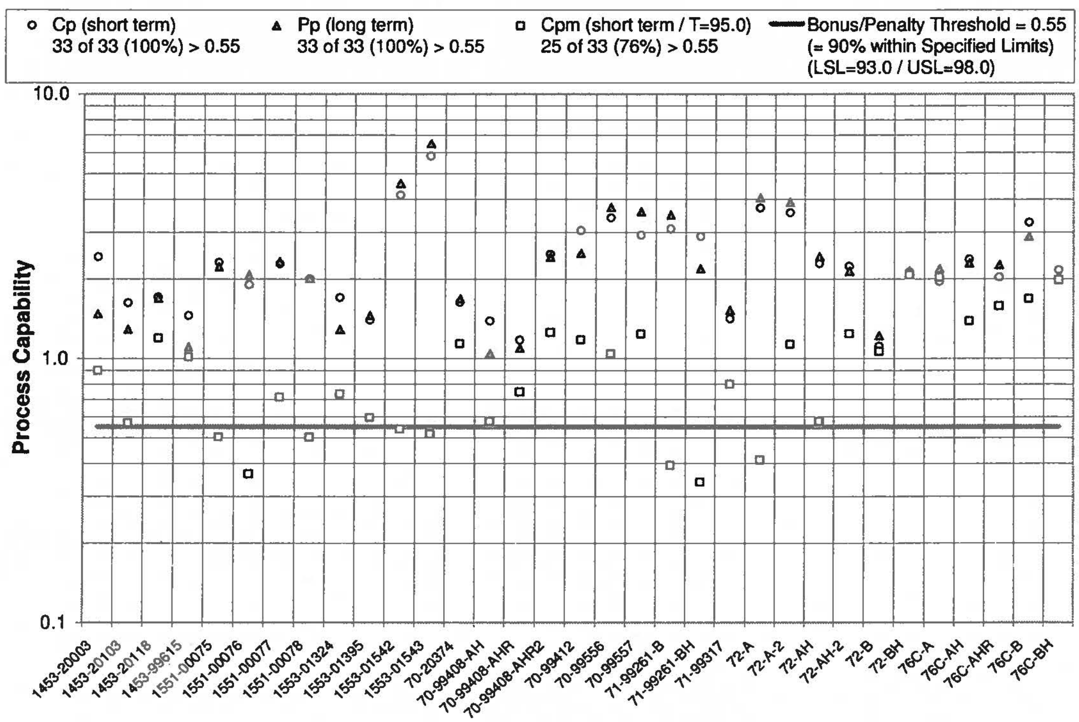


Figure 9 – Sample Process-Capability Chart (for AC Roadway Density)

Statistical Process Control (SPC)

Although the issue of process control was not directly included in the stated objectives for this study, its importance to the successful implementation of *Percent-Within-Limits* specifications must not be overlooked. In fact, the American Association of State Highway and Transportation Officials' *Implementation Manual for Quality Assurance* (AASHTO 1996a) states:

> The need for Contractors to use statistical control charts cannot be overemphasized. A control chart provides a visual indication of whether a process is in control. Timely reaction to the QC chart can prevent the production of nonconforming material.

The advent of statistical control charts began when Dr. Walter Shewhart, with Bell Laboratories, developed the procedure that is now commonly referred to as "Statistical Process Control" (SPC) charting (Shewhart 1931). SPC charting involves the sequential plotting of test results for a given characteristic of a given process. Dr. Shewhart recognized that the overall variation in a process can be broken down into two distinct categories, or components—common-cause and special-cause. Common-cause variation represents the variation that is inherent to the process. In other words, common-cause variation is variation that cannot be eliminated from the process unless something within the process itself is changed. Special-cause variation, on the other hand, is variation caused by external, or non-random, forces acting upon the process. Dr. Shewhart recognized that special-cause variation could be readily identified by applying a series of rules every time a new measurement or set of measurements is taken. The rules essentially represent a set of ongoing hypothesis tests, where the null and alternate hypotheses are as follows (where "in statistical control" is defined as the existence of common-cause variation only):

H_o: the process is in statistical control

H_a: the process is not in statistical control

Whenever special-cause variation enters the process, the process is said to be "out of statistical control" or "not in statistical control." It bears noting that whether or not a process is "in statistical control" is not an indication of whether or not the process is meeting the specified tolerances. "Process Capability" (as discussed in the previous section) deals with the issue of specification limits and whether or not the process is meeting or can meet the specified tolerances. SPC is a tool for quickly identifying and correcting special-cause variation, whereas capability analysis provides the means for assessing how well the process can or does meet the specification limits. Both are critical to achieving sustained levels of quality. In fact, the highest levels of quality can be achieved only when each of the following three conditions are being consistently satisfied:

1. The process is in statistical control (meaning that only common-cause variation is present).
2. The process is capable (meaning that the common-cause variation for the process is small enough to allow the process to consistently remain within the specified limits).
3. The process is on target (meaning that the process is performing at or near the specified target).

This report has been written so as to encourage and facilitate the implementation of this three-fold approach to quality within the construction industry.

Figures 10 – 12 demonstrate the power associated with the use of SPC charting to identify special-cause variation. The figures also show the inter-relatedness of many of the measured quality characteristics on an AC paving project. As can be seen from the figures, maintaining control charts for multiple quality characteristics at the same time can also help identify cause-and-effect whenever special-cause variations are actually observed.

Figure 10 shows a distinct change in the achieved roadway density on project WR-MC-70 beginning somewhere around observation #30 (circled in red). Concurrently, the control chart for air voids shows a significant "out of control" reduction in the percent-air-voids (also circled in red). This suggests that the increase in roadway density resulted from a change in the mixture characteristics rather than a change in the compaction efforts or a change in weather conditions. This deduction stems from the fact that the air voids measurement is based on applying a standard compaction effort to a sample of asphalt in a laboratory setting. If the change in roadway densities was merely the result of, say warmer weather or an increased number of roller passes, the air voids characteristic would not have changed.

In addition to the nearly-simultaneous change in air voids, we see a change in the percent-passing the No. 80 sieve (as shown in Figure 12). This suggests that the changes to the percent-passing the No. 80 sieve may have led to a denser and/or more-easily-compacted mix (thus causing the changes observed in the air voids and roadway density measurements).

It bears noting with this example that *none* of the changes in roadway density, air voids, or percent-passing the No. 80 sieve exceeded the *specification* limits for those quality characteristics. Yet, the overall quality of the product was certainly affected by these changes (presumably for the better, in this instance). This further highlights the importance of monitoring the common-cause versus special-cause variation in addition to measuring the process capability. Also, this validates the reason why the control limits on control charts should be based on the common-cause variation inherent in the process and SHOULD NOT BE BASED ON THE SPECIFICATION LIMITS (which is a common, but misguided and incorrect procedure). Concerning control limits versus specification limits, Day (1999) aptly states:

> It is important to understand that these [control] limits were not specification limits. Their function was not to indicate whether the result plotted was **acceptable**, but to indicate whether it was **unusual**. The intention was not to decide whether to accept or reject the product represented by the result, but to detect whether there has been any **change in the process** producing the product. This concept has proved very difficult to promote but is still the basis needed to achieve good quality control.

The guidelines for quality control testing for PWL in Appendix F include procedures for implementing process control using SPC charting. The procedures can be readily incorporated into *Percent-Within-Limits* specifications to help the Contractor quickly identify and correct special-cause variations, thus helping to ensure the paving processes remain "in statistical control" throughout the project.

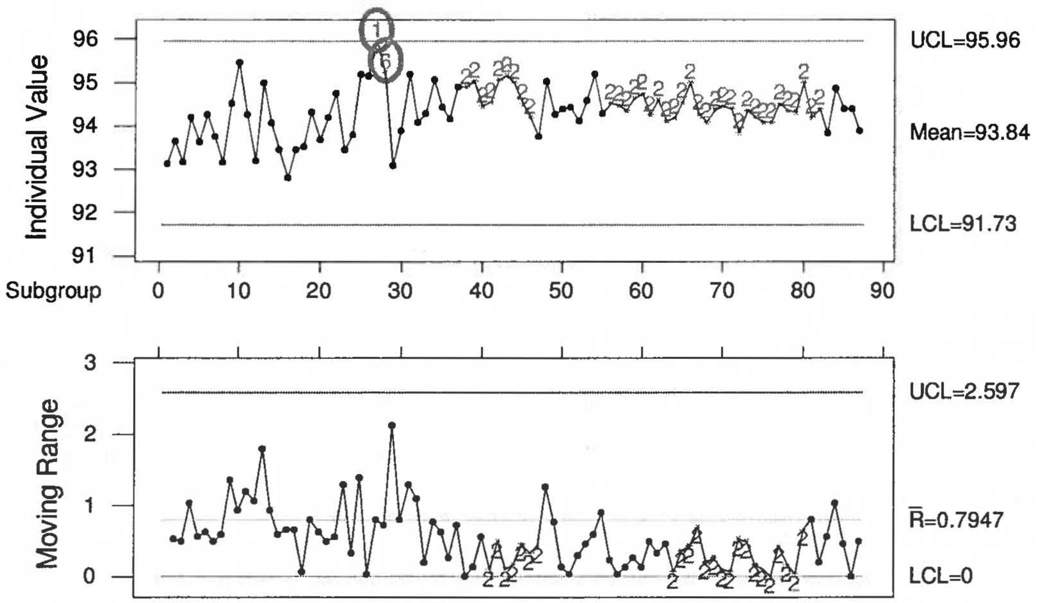


Figure 10 – Control Chart for Roadway Density on WR-MC-70

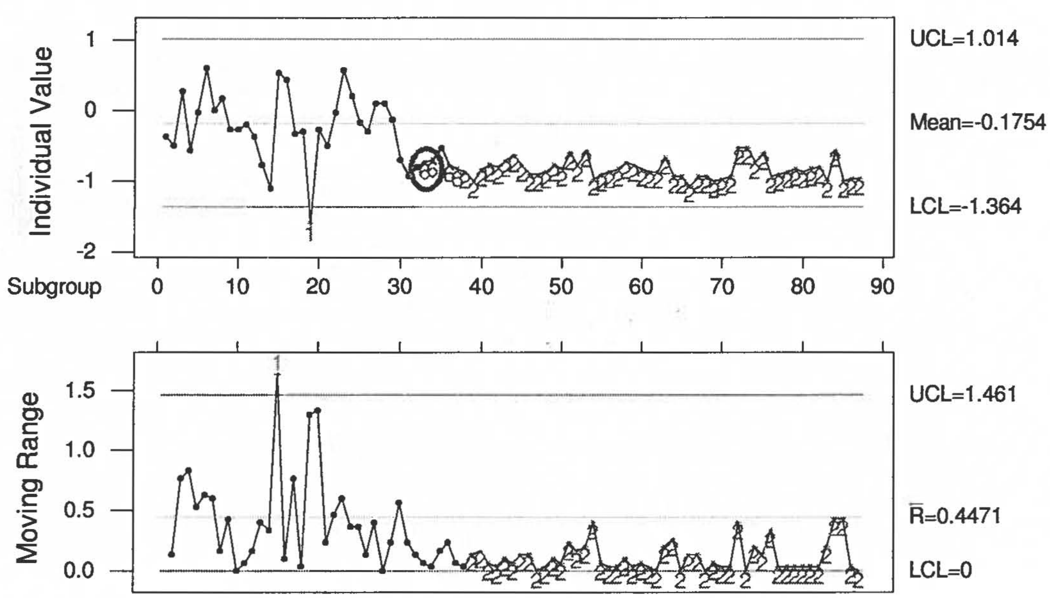


Figure 11 – Control Chart for Air Voids on WR-MC-70

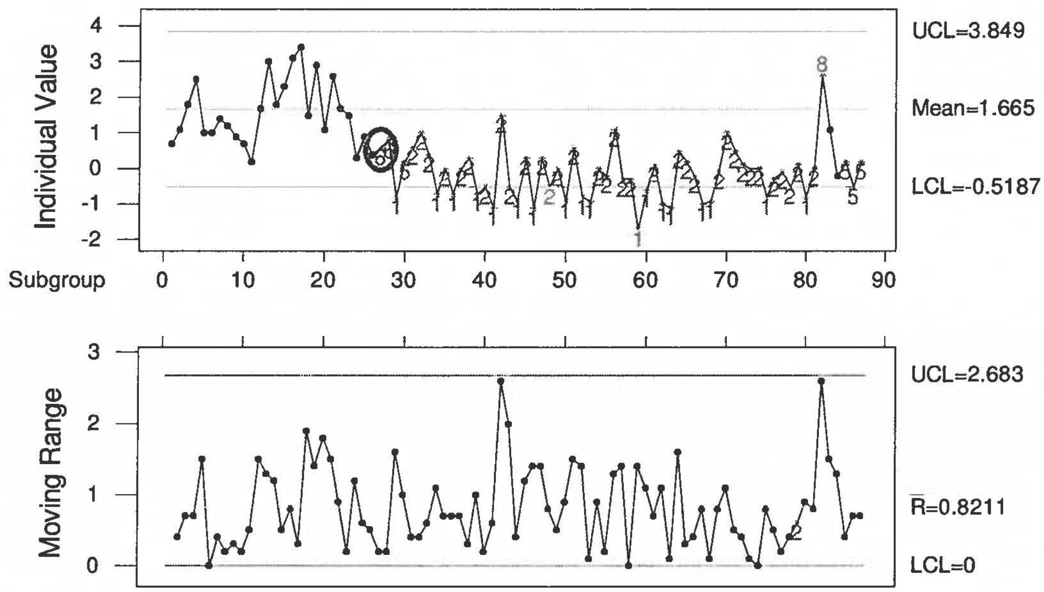


Figure 12 – Control Chart for Percent-Passing the No. 80 Sieve on WR-MC-70

Summary of Pay Adjustments based on the Oklahoma Transportation Authority's PWL Specifications

Table 12 provides a summary of the pay adjustments for the PCC and AC paving projects based on the *Percent-Within-Limits* specifications implemented by the Oklahoma Transportation Authority (OTA). Tables 13 and 14 summarize the overall pay adjustments by contractor for the projects included in the study. As can be seen, the vast majority of the lots for the projects exceeded the minimum "ninety-percent-within-limits" requirements, thus earning bonus payments for the contractors involved. As attested to by various on-site representatives for the projects, the contractors, in most instances, took a more proactive approach to quality and quality-control as a result of the *Percent-Within-Limits* specification requirements. The net result, according to the on-site representatives, was a significant increase in the overall quality of the projects.

Table 12 – Pay Adjustment Summary for Ten PCC and Seven AC Paving Projects

PCC Quality Characteristic	Specified Lower Limit	Specified Upper Limit	# of Lots Receiving Bonus	% of Lots Receiving Bonus	Average Weighted Pay Factor (All Lots)	# of Lots Receiving PF=0	% of Lots Receiving PF=0
Combined Pay Factor (All Characteristics)	Varies	Varies	117 of 142	82%	96.8%	N/A	N/A
Strength (ALL)	Varies	No Limit	123 of 142	87%	96.6%	9 of 142	6%
Strength (Class A)	Varies	No Limit	86 of 103	83%	94.9%	9 of 103	9%
Strength (Class AP)	Varies	No Limit	37 of 39	95%	100.9%	0 of 39	0%
Air Content	4.5	7.5	150 of 154	97%	100.5%	0 of 154	0%
Coarse 1-1/2-inch	100	100	N/A	N/A	N/A	N/A	N/A
Coarse 1-inch	95	No Limit	132 of 136	97%	100.1%	1 of 136	1%
Coarse 1/2-inch	25	60	142 of 152	93%	100.1%	0 of 152	0%
Coarse No. 4	No Limit	10	146 of 147	99%	100.2%	0 of 147	0%
Coarse No. 8	No Limit	5	142 of 145	98%	100.2%	0 of 145	0%
Coarse No. 200	No Limit	2	142 of 153	93%	100.1%	0 of 153	0%
Fine 3/8-inch	100	100	N/A	N/A	N/A	N/A	N/A
Fine No. 4	95	No Limit	118 of 119	99%	100.2%	0 of 119	0%
Fine No. 8	80	No Limit	149 of 150	99%	100.2%	0 of 150	0%
Fine No. 16	50	85	151 of 151	100%	100.2%	0 of 151	0%
Fine No. 30	25	60	147 of 151	97%	100.2%	0 of 151	0%
Fine No. 50	5	30	144 of 150	96%	100.2%	0 of 150	0%
Fine No. 100	No Limit	10	143 of 143	100%	100.2%	0 of 143	0%
Fine No. 200	No Limit	3	136 of 141	96%	99.9%	4 of 141	3%
AC Quality Characteristic	**Specified Lower Limit**	**Specified Upper Limit**	**# of Lots Receiving Bonus**	**% of Lots Receiving Bonus**	**Average Weighted Pay Factor (All Lots)**	**# of Lots Receiving PF=0**	**% of Lots Receiving PF=0**
Combined Pay Factor (All Characteristics)	Varies	Varies	150 of 193	78%	99.5%	N/A	N/A
Density	93	98	163 of 194	84%	99.0%	7 of 194	4%
Air Voids	-2.5	2.5	189 of 193	98%	100.5%	0 of 193	0%
AC Content	-0.6	0.6	193 of 194	99%	100.6%	0 of 194	0%
1-inch	-6	6	99 of 99	100%	100.2%	0 of 99	0%
3/4-inch	-6	6	102 of 117	87%	100.1%	0 of 117	0%
1/2-inch	-6	6	157 of 193	81%	100.0%	0 of 193	0%
3/8-inch	-6	6	38 of 43	88%	100.1%	0 of 43	0%
No. 4	-6	6	167 of 194	86%	100.1%	1 of 194	1%
No. 10	-4.5	4.5	168 of 194	87%	99.9%	3 of 194	2%
No. 40	-4.5	4.5	180 of 193	93%	100.0%	2 of 193	1%
No. 80	-4.5	4.5	192 of 192	100%	100.2%	0 of 192	0%
No. 200	-2	2	172 of 193	89%	100.0%	1 of 193	1%

Table 13 – Pay Adjustment Summary by Contractor (PCC Projects)

Contractor	Sum of Weighted Composite Pay Factors	Total # of Lots	Average Overall Pay Factor
Duit	97.49	96	101.5%
Koss	14.96	21	71.2%
TTK	7.13	7	101.9%
Western Plains	12.17	12	101.5%
Wittwer Paving	5.71	6	95.1%
Grand Total	137.46	142	96.8%

Table 14 – Pay Adjustment Summary by Contractor (AC Projects)

Contractor	Sum of Weighted Composite Pay Factors	Total # of Lots	Average Overall Pay Factor
APAC	25.82	26	99.3%
APAC (Non-PWL)	13.02	15	86.8%
Broce	34.84	35	99.6%
Cummins	55.66	55	101.2%
Glover	29.33	29	101.1%
Haskell-Lemon	33.43	33	101.3%
Grand Total	192.11	193	99.5%

One of the AC projects included in the study (WR-MC-71, by APAC) was constructed under the OTA's Standard Specifications with no *Percent-Within-Limits* requirements. However, the data for WR-MC-71 were collected using the same sampling procedures as required by the OTA's *Percent-Within-Limits* specifications and were included in the study for comparative purposes. In addition, the strength test data for one of the PCC projects (CKT-1151, by Koss) were extremely low, which resulted in a lengthy dispute between the OTA and the contractor. The dispute was eventually resolved through extensive third-party laboratory testing of the in-place concrete and subsequent dispute-resolution proceedings. The third-party analysis ultimately indicated that the in-place strength of the concrete could be considered "adequate" and thusly recommended 100% pay for the work (but zero pay for quality control). Elimination of the disputed CKT-1151 (Koss) data from Tables 12 and 13 results in an increase in the overall pay factor (for all PCC projects) from 96.8 to 101.2%.. Similarly, elimination of the WR-MC-71 (APAC) non-PWL data from Tables 12 and 14 results in an increase in the overall pay factor (for all AC projects) from 99.5 to 100.6%

Target-Adjusted Standard Deviation Calculations

As presented in the section entitled *Process Capability and Contractor Performance*, the Taguchi Index (C_{pm}) uses an adjustment to the short-term standard deviation of the process based on the deviation of the process mean from the process target. As such, C_{pm} is a better measure of the true process performance because it takes into consideration the deviation from target. One of the objectives of this study was to provide recommendations concerning methods to minimize the occurrence of full-incentive payments when the average values deviate from design (or target) levels.

A simple way to achieve this objective involves an adjustment to the within-lot standard deviation (prior to the PWL calculations) which is analogous to the C_{pm} standard-deviation adjustment performed for process capability calculations. A major difference between the C_{pm} target-adjusted standard deviation calculations and the procedure set forth herein involves the use herein of *upper and lower* target limits, rather than reliance upon a single target value. Whereas in manufacturing, a single target value is commonplace (e.g. a single dimension on a part that must fit snugly with another part), the construction industry in general, and the paving industry in particular, does not always lend itself to such clearly-defined target values. To accommodate this need for flexibility with target values, the procedures presented herein allow the specifier to provide a target *range* rather than a single target value. Of course, if a single target value is truly desired for a given quality characteristic, the upper and lower target limits can both be set equal to that singular target value.

Equation 9 represents the current standard deviation calculation as outlined in the original draft *Percent-Within-Limits* special provisions:

$$S' = \sqrt{\frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2}{N-1}} \qquad \text{(Equation 9)}$$

where

- S' = sample standard deviation (within the lot) for the quality characteristic of interest
- X_i = observed (measured) value of the quality characteristic of interest for sublot i
- $\overline{X}$ = sample mean (within the lot)
- N = number of sublots within the lot (i.e. the number of X_i's).

Equation 10 below gives the recommended method for calculating a target-adjusted standard deviation:

$$S'' = \sqrt{S'^2 + \left(X_{target} - \overline{X}\right)^2} \qquad \text{(Equation 10)}$$

where

S'' = target-adjusted standard deviation (within the lot) for the quality characteristic of interest

S' = sample standard deviation (within the lot) as calculated by Equation 9

X_{target} = nearest target limit

$\bar{X}$ = sample mean (within the lot)

It should be pointed out that the target-adjusted standard deviation calculation (Equation 10) should only be applied when the average value of the quality characteristic of interest (for a given lot) falls *within* the specification limits but *outside* the target limits. As such, the adjustment should NOT be performed whenever any of the following three conditions occur:

1. the average value is *outside* the specification limits (i.e. the *Percent-Within-Limits* is less than 50%),

2. the average value is *within* the target limits, or

3. if a one-sided specification is being used and the specified target lies between the average value and the specification limit (i.e. the average value is *better* than the target value).

Sample specifications are provided in Appendix G to assist in the implementation of this target-adjusted standard-deviation procedure within the context of PWL specifications.

Recommended Specification Limits for Oklahoma DOT's Pilot Projects

Despite the overall success of the Oklahoma Transportation Authority's (OTA's) *Percent-Within-Limits* (PWL) experience, the issue of deviations from target values was not fully addressed by the OTA PWL specifications. In fact, this perceived limitation was the impetus for one of the subtasks outlined for this study, which was to:

- Quantify the occurrences of incentive payments that occurred even though average strength values were below design levels.

Concerning this subtask, the project data show that relatively few instances of this actually occurred. Of the 142 PCC paving lots included in the study, 35 experienced average strengths below the target (or design) strength. Of those 35 below-target lots, 17 still received incentive payments related to strength. However, the average deviation from target for those 17 lots was less than 200 psi. By contrast, the remaining 18 below-target lots (those receiving a penalty for strength) averaged over 1,000 psi below target. Even so, the use of target-adjusted standard deviations would have further improved upon this situation, wherein only 15 of the below-target lots would still have received a bonus.

As suggested, this potential limitation with the OTA *Percent-Within-Limits* specifications can be reduced by the use of the target-adjusted standard deviation method discussed in the previous section. The recommend specification limits presented herein along with the specification guidelines presented in Appendix G incorporate the use of the aforementioned target-adjusted method for calculating within-lot standard deviations.

Table 15 provides recommended specification and target limits along with a summary of the impacts that would have been felt on the OTA projects had those specification limits been in force. The pay-factor calculations summarized in Table 15 are based on the revised Pay Factor equation presented in Appendix G in lieu of the OTA Pay Factor equation shown in Appendices A and B. Table 15 can be directly compared to Table 12 (which presented a summary of pay adjustments based on the actual OTA-specified upper and lower limits with no target-based adjustments to standard deviations). Tables 16 and 17 (analogous to Tables 13 and 14) provide summaries by contractor of the data used to generate Table 15.

Removal of the contractor-disputed CKT-1151 (Koss) data from the PCC analysis increases the overall would-be weighted-average pay factor for PCC projects (as shown in Tables 15 and 16) from 100.2 to 102.8%. Similarly, removal of the non-PWL WR-MC-71 (APAC) data from the analyses presented in Tables 15 and 17 results in an increase in the overall would-be weighted-average pay factor for AC projects from 97.8 to 99.2%.

Concerning PCC strength specifications, the OTA PCC projects were constructed based on a target flexural strength of 750 psi with a lower limit of 675 psi (for Class A paving concrete). However, due to the difficulty associated with testing flexural strengths in the field, the OTA required the determination of flexural-to-compressive-strength correlations for each mix design. The target compressive strengths for Class A paving concrete ranged from 4,150 psi to 5,350 psi, with an average target of 4,950 psi. Similarly, the minimum, maximum and average lower-limit compressive strengths for Class A paving concrete were 2,600 psi, 4,350 psi and 3,950 psi respectively. (The minimum, maximum and average target and lower-limit compressive strengths for Class AP concrete were 2,600 psi, 4,075 psi and 3,350 psi respectively and 1,950 psi, 3,450 psi and 2,750 psi respectively). Whereas all the OTA PCC Class A paving mixes were designed for 750 psi flexural strength or better, utilization of the aforementioned target and lower-limits for compressive strength may not be warranted for ODOT projects, where pavement designs are typically based on 650 psi flexural strength. As such, the recommendations presented in Table 15 keep the lower-limit specification of 3,800 psi (from the original ODOT draft PWL special provisions) and add a target compressive strength of 4,500 psi.

Table 15 –Recommended Specification Limits for ODOT Pilot Projects (using Target-Adjusted Standard Deviations) and Summary of Would-be Impacts on the As-Constructed OTA Projects

PCC Quality Characteristic	Specified Lower Limit	Lower Target Limit	Upper Target Limit	Specified Upper Limit	# of Lots Receiving Bonus	% of Lots Receiving Bonus	Average Weighted Pay Factor (All Lots)	# of Lots Receiving PF=0	% of Lots Receiving PF=0
Combined Pay Factor (60% Strength, 30% Air Content, 10% Gradation)	Varies	Varies	Varies	Varies	125 of 141	89%	100.2%	N/A	N/A
Strength (Class A)	3800	4500	No Limit	No Limit	93 of 103	90%	98.7%	5 of 103	5%
Strength (Class AP)	3000	3750	No Limit	No Limit	38 of 39	97%	102.3%	0 of 39	0%
Air Content	4.5	5.5	6.5	7.5	142 of 154	92%	101.0%	0 of 154	0%
Coarse No. 200	No Limit	No Limit	1	2	112 of 153	73%	99.9%	0 of 153	0%
Fine No. 200	No Limit	No Limit	1	3	118 of 141	84%	99.9%	4 of 141	3%

AC Quality Characteristic	Specified Lower Limit	Lower Target Limit	Upper Target Limit	Specified Upper Limit	# of Lots Receiving Bonus	% of Lots Receiving Bonus	Average Weighted Pay Factor (All Lots)	# of Lots Receiving PF=0	% of Lots Receiving PF=0
Combined Pay Factor (40% Density, 30% Air Voids, 20% AC Content, 10% Gradation)	Varies	Varies	Varies	Varies	121 of 187	65%	97.8%	N/A	N/A
Density	93	94	96	97	143 of 193	74%	97.8%	8 of 193	4%
Air Voids	-2	-0.8	0.8	2	181 of 193	94%	100.5%	1 of 193	1%
AC Content	-0.4	-0.16	0.16	0.4	174 of 193	90%	100.5%	0 of 193	0%
1-inch	-6	-2.4	2.4	6	100 of 101	99%	100.4%	0 of 101	0%
3/4-inch	-6	-2.4	2.4	6	97 of 116	84%	100.0%	0 of 116	0%
1/2-inch	-6	-2.4	2.4	6	159 of 192	83%	100.1%	0 of 192	0%
3/8-inch	-6	-2.4	2.4	6	39 of 43	91%	100.3%	0 of 43	0%
No. 4	-6	-2.4	2.4	6	168 of 193	87%	100.2%	1 of 193	1%
No. 10	-4.5	-1.8	1.8	4.5	163 of 193	84%	100.0%	3 of 193	2%
No. 40	-4.5	-1.8	1.8	4.5	178 of 193	92%	100.2%	2 of 193	1%
No. 80	-4.5	-1.8	1.8	4.5	190 of 191	99%	100.4%	0 of 191	0%
No. 200	-2	-0.8	0.8	2	164 of 192	85%	100.1%	1 of 192	1%

Table 16 – Would-be Pay-Adjustment Summary by Contractor (PCC Projects)

Contractor	Sum of Weighted Composite Pay Factors	Total # of Lots	Average Overall Pay Factor
Duit	97.49	96	103.1%
Koss	14.96	21	85.3%
TTK	7.13	7	103.9%
Western Plains	12.17	12	102.8%
Wittwer Paving	5.71	6	96.0%
Grand Total	137.46	142	100.2%

Table 17 – Would-be Pay-Adjustment Summary by Contractor (AC Projects)

Contractor	Sum of Weighted Composite Pay Factors	Total # of Lots	Average Overall Pay Factor
APAC	25.82	26	94.0%
APAC (Non-PWL)	13.02	15	70.0%
Broce	34.84	35	95.3%
Cummins	55.66	55	102.2%
Glover	29.33	29	101.4%
Haskell-Lemon	33.43	33	100.8%
Grand Total	192.11	193	97.8%

Statistical Quality Assurance

As stated in the section entitled *Testing Variability*, testing error can have a profound impact on the pay-adjustment calculations associated with *Percent-Within-Limits* (PWL) specifications. In recognition of this need for valid test results and the assurance thereof, the following objective was included for this study:

- Provide guidelines and recommendations concerning statistical methods for quality assurance testing.

AASHTO's *Implementation Manual for Quality Assurance* (AASHTO 1996a) states:

> If an agency decides to use the Contractor's QC results as a part of the Acceptance Program, there must be a system of checks and balances to ensure the reliability and accuracy of the Contractor's test results. Agency testing is required to form a basis for validating Contractor results. ... [Since] Contractor's test results are a legitimate source of data that can be used in the Acceptance Program, the validation methodology becomes an important decision for the agency.

Detailed procedures for implementing statistical quality assurance in the context of PWL specifications are provided in Appendices H and I. The foundational steps for those statistical quality assurance procedures are as follows:

1. Determine whether or not a statistically-significant bias exists between the Contractor's and the Department's testing equipment, technicians, procedures, etc.
2. If a *statistically*-significant bias exists, determine whether or not the magnitude of the bias is substantial enough to warrant investigation. In other words, is the bias *practically* significant as well?
3. If the testing bias is *statistically*- and *practically*-significant, determine and correct the source of the bias. Default to using the Department's test results for acceptance and pay-factor calculations until the bias is corrected and verified as such.
4. Once the Contractor's test methods have been verified as bias-free, allow the use of the Contractor's test results for acceptance and pay-factor calculations.
5. Continue monitoring the Contractor's test methods via statistical process control (SPC) charting to enable quick identification of any special-cause variations with the Contractor's testing program.
6. Whenever special-cause variation is observed, go back to Step 1.

Operating Characteristic Curves and Pay Factor Equations

Although an analysis of Operating Characteristic (OC) curves was not included in the scope of this study, a non-exhaustive set of simulations performed by the researcher revealed ranges of producer's (i.e. contractor's) risk from 1.0 to 2.5% at the Acceptable Quality Level of 90 PWL. In light of this observation and in recognition of the increased stringency associated with the use of target-adjusted standard-deviation calculations, the researcher recommends consideration of a revised Pay Factor equation such as the one depicted in Figure 13 and further presented in Appendix G. This revised Pay Factor equation would increase the maximum possible incentive payment from 102 to 104% (at 100 PWL) while reducing pay from 60 to 50% at the Rejectable Quality Level (50 PWL). At the Acceptable Quality Level (90 PWL), the revised Pay Factor equation represents an increase from 100 to 101.85%. For comparison purposes, Figure 13 shows the original OTA Pay Factor equation juxtaposed with the revised equation.

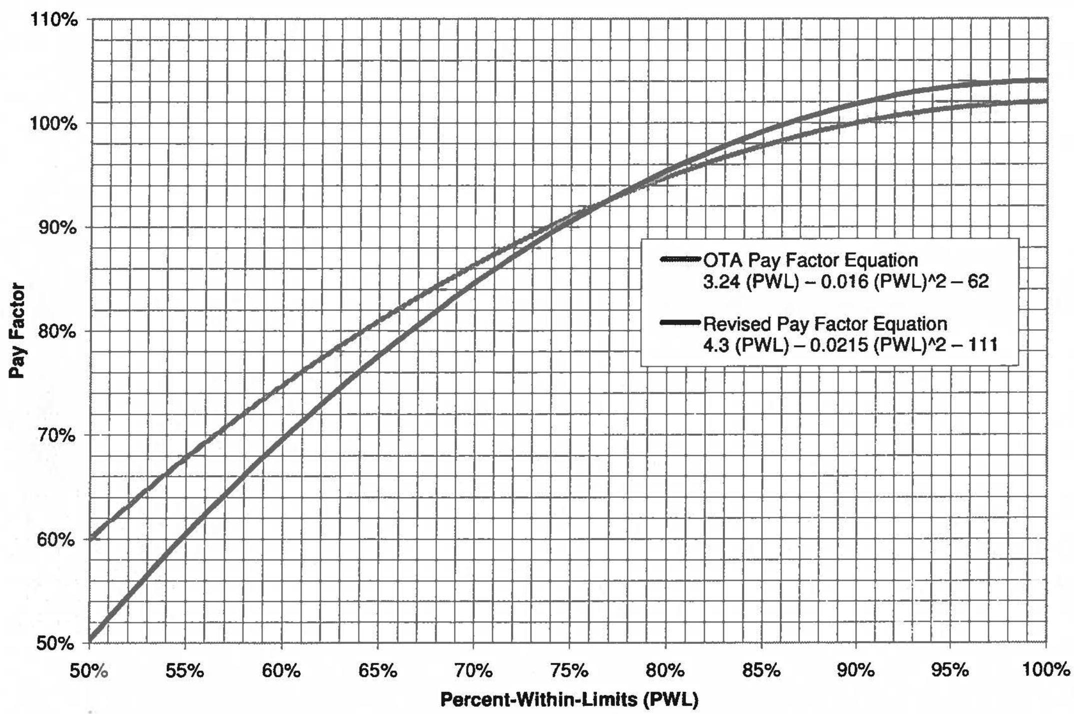


Figure 13 – Comparison of Original OTA versus Revised Pay Factor Equations

Using *Percent-Within-Limits* (PWL) Terminology in Lieu of *Lot-Percent-Defective* (LPD)

The original draft versions of the Oklahoma DOT's *Percent-Within-Limits* special provisions are actually written using *Lot-Percent-Defective* language. The researcher prefers to use *Percent-Within-Limits* terminology, thus focusing on the positive rather than the negative. To switch from *Lot-Percent-Defective* to *Percent-Within-Limits* terminology, the calculations for percent-defective should be carried out the same way, using the same tables as detailed in the original ODOT draft special provisions with the total LPD converted to PWL just prior to the pay factor calculations. The procedures presented in Appendix G demonstrate the implementation of *Percent-Within-Limits* specifications using *Percent-Within-Limits* (in lieu of *Lot-Percent-Defective*) terminology.

REFERENCES

Ahmed, S. A. (1996a). *Oklahoma Department of Transportation Special Provisions for Quality Control and Acceptance Procedures Section 411 Plant Mix Asphalt Concrete.* ODOT Special Provision 411-1QA (10/07/96). Stillwater, OK: Oklahoma State University.

Ahmed, S. A. (1996a). *Oklahoma Department of Transportation Special Provisions for Quality Control and Acceptance Procedures Section 414 Portland Cement Concrete.* ODOT Special Provision 414-1QA (10/07/96). Stillwater, OK: Oklahoma State University.

Ahmed, S. A. (1996a). *Oklahoma Department of Transportation Special Provisions for Quality Control and Acceptance Procedures Section 504 Concrete Bridge Floors.* ODOT Special Provision 504-4QA (10/07/96). Stillwater, OK: Oklahoma State University.

Ahmed, S. A. (1997). *Variability in Bituminous Concrete Pavement Construction: Final Report.* Stillwater, OK: Oklahoma State University.

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

American Association of State Highway and Transportation Officials (AASHTO) (1996b). *Quality Assurance Guide Specification.* Washington, DC: AASHTO.

ASTM C 136 - 01. (2002). "Standard Test Method for Sieve Analysis of Fine and Coarse Aggregates." *2002 Annual Book of ASTM Standards Vol. 04.02.* West Conshohocken, PA: ASTM International.

Butts, N. E. and Ksaibati, K. (2002). "Asphalt Pavement Quality Control/Quality Assurance Program in the United States." *2003 Annual Meeting of the Transportation Research Board.* Washington, DC: Transportation Research Board.

Day, K. W. (1999). *Concrete Mix Design, Quality Control and Specification, Second Edition.* New York, NY: Routledge.

Eastman, S. A. (1995). *Evaluating Automated Wafer Measurement Instruments.* 94112638A-XFR Austin, TX: SEMATECH Technology Transfer

SEMI M27-96 (1996). *Practice for Determining the Precision Over Tolerance (P/T) Ratio of Test Equipment.* San Jose, CA: Semiconductor Equipment and Materials International (SEMI).

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product.* New York, NY: D. Van Nostrand Company, Inc.

Appendix A: OTA's Percent-Within-Limits (PWL) Special Provisions for PCC Pavements

OKLAHOMA TRANSPORTATION AUTHORITY
SPECIAL PROVISIONS
FOR
QUALITY CONTROL AND ACCEPTANCE PROCEDURES
FOR
PORTLAND CEMENT CONCRETE PAVEMENT

These special provisions revise, amend and where in conflict, supersede applicable sections of the *Standard Specifications for Turnpike Construction*, Edition of 1999. These Special Provisions apply to Portland Cement Concrete Pavement.

The Contractor is responsible for the quality of the materials and construction, whereas the Engineer will be responsible for determining the acceptability of such materials and construction.

It is the intent of these Special Provisions that materials and construction of acceptable quality shall receive an average pay factor of one hundred (100) percent; that materials and construction of truly superior quality shall be awarded a bonus payment; and that materials and construction of deficient quality will receive a reduced payment or be removed and replaced.

(Add the following to **Subsection 414.01. "DESCRIPTION."***):*
Contractor's Quality Control and Acceptance Procedures will apply to this work in accordance with the applicable requirements of Section 106 and as herein specified.

(Add the following to **Subsection 414.04.(u) "CONSTRUCTION METHODS. Contractor's Quality Control Testing and Inspection."***):*

3. *Contractor's Testing.* As a minimum, the Contractor's sampling and testing shall comply with the following schedule:

Property	Sampling and Testing Frequency
Gradation (Fine & Course)	1 per 2500 SQ YDS
Air Content & Slump	First load per placement, then 1 per 500 SQ YDS
Strength	2 cylinders per 2500 SQ YDS
Thickness	As needed to control operations
Smoothness	As needed to control operations

Additional sampling and testing to ensure compliance with the Standard Specifications and other Special Provision requirements shall be in accordance with the Contractor's Quality Control Plan.

4. *Control Charts.* Control charts covering as a minimum the characteristics of gradation, strength, air content, slump and thickness shall be maintained by the Contractor and displayed at the plant or job site. The charts shall identify the project number, the contract item number, the characteristics, the date, the lot and sublot numbers, the applicable upper and/or lower specification limits, the Contractor's test results and any other data needed to facilitate control of the process and identify problems before they become serious. Copies of the Contractor's quality control tests shall be provided to the Engineer at time intervals acceptable to the Engineer.

(Add the following to **Subsection 414.04.(v) "CONSTRUCTION METHODS. Contractor's Quality Control Testing and Inspection.** ***Acceptance of Pavement."****):*

While the Contractor shall be fully and exclusively responsible for producing an acceptable product, acceptance responsibility rests with the Engineer. The entire lot of concrete pavement, as defined in Subsection (v)-1.1, will be accepted or rejected and paid for on the basis of acceptance test results. The Engineer may choose to use the Contractor's tests for acceptance after the Contractor's test results have been demonstrated to be consistent with tests taken by the Authority and that they adequately represent the material being evaluated.

1. *Basis for Acceptance and Payment.* The following characteristics will be considered in evaluating materials and construction for acceptance and payment:

   - Gradation
   - Air Content
   - Strength
   - Pavement Structure Thickness (as provided in Section 431)
   - Pavement Surface Smoothness (as provided in Section 430)

1.1 Lot and Sublot Definition

Except for surface smoothness and thickness, acceptance and pay adjustments will be made on a lot-by-lot basis. Each lot of Portland Cement Concrete will be sampled at random and tested for gradation, air content and strength in accordance with the following requirements.

The standard lot size shall consist of six equal sublots of 2500 square yards each. Each sublot will be sampled *at random* to obtain one or more test specimens as follows:

a) Gradation and air content determination: one specimen and one test for each characteristic per sublot.
b) Strength determination: two cylinders per sublot averaged and considered as one individual test.

In the event that operational conditions cause work to be interrupted before the standard lost size has been achieved, the Engineer may redefine the lot size. However, the number of test determinations required to evaluate each lot will be at least four. Each partial lot will be divided into at least four equal sublots, and each sublot will be sampled at random to obtain the required number of test specimens.

Quantities of concrete less than 625 square yards may be accepted by the Engineer upon visual inspection by which the Engineer has reason to believe that the materials and construction are of acceptable quality. At the Engineer's option, this quantity may be treated as a separate lot, combined with the previous lot or combined with the following lot, as the case may warrant. On a multiple project contract, the lots of concrete will carry over from project to project.

1.2 Smoothness and Thickness Acceptance and Pay Adjustments

For smoothness and thickness determination and pay adjustment purposes, the pavement will be tested on an extent-to-extent basis in accordance with Sections 430 and 431. Acceptance and pay adjustment determinations made under Sections 430 and 431 will be completely independent of those made under this Special Provision.

1.3 Criteria for Lot Acceptance and Payment

Except for surface smoothness and thickness, conformance with the specifications will be judged on the basis of the following three criteria:

a) The estimated lot percent defective with respect to gradation, air content and strength. The lot percent defective with respect to a particular quality characteristic is the amount of materials and construction which falls outside the specified limit(s) listed in the following table:

Quality Characteristic	Lower Limit (L)	Upper Limit (U)
Gradation (Fine):		
3/8 inch	100.0%	100.0%
Sieve #4	95.0%	100.0%
Sieve #8	80.0%	100.0%
Sieve #16	50.0%	85.0%
Sieve #30	25.0%	60.0%
Sieve #50	5.0%	30.0%
Sieve #100	0.0%	10.0%
Sieve #200	0.0%	3.0%
Gradation (Coarse):		

1-1/2 inch	100.0%	100.0%
1 inch	95.0%	100.0%
1/2 inch	25.0%	60.0%
Sieve #4	0.0%	10.0%
Sieve #8	0.0%	5.0%
Sieve #200	0.0%	2.0%
Air Content	4.5%	7.5%
Strength (Class A)	Compressive strength equivalent of 575 psi flexural strength unless otherwise specified [see Subsection 414.02(a)]	------
Strength (Class AP)	Compressive strength equivalent of 500 psi flexural strength unless otherwise specified [see Subsection 414.02(a)]	------

b) For the strength characteristic, whenever any individual test result as defined in Subsection (v)-1.1 falls below the corresponding lower-critical-limit listed in the following table, acceptance of the lot will be based on the amount of materials and construction which falls below that critical limit.

Characteristic	Lower-Critical-Limit
Strength (Class A)	Compressive strength equivalent of 550 psi flexural strength unless otherwise specified [see Subsection 414.02(a)]
Strength (Class AP)	Compressive strength equivalent of 450 psi flexural strength unless otherwise specified [see Subsection 414.02(a)]

c) Any load of concrete that is visually unacceptable for reasons of being too wet, excessively segregated, or otherwise obviously deficient will be rejected for use in the work. Furthermore, sections of completed pavement which from visual observation or known deficiencies appear to be seriously inadequate will be extensively tested. The results of such tests will not be used for pay adjustment purposes, but will be used to determine whether the section is totally unacceptable and must be removed. In the event that a section is determined to be unacceptable, its removal and placement shall be at no additional cost to the Authority.

1.4 *Acceptable Quality Level*

A lot shall be considered of acceptable quality with respect to a particular characteristic if the percent defective, as defined in Subsection (v)-1.3(a), is no more than ten (10) percent. In addition, no individual test result on the strength or thickness of an acceptable quality level lot shall fall below the critical limits defined in Subsection (v)-1.3(b).

The Contractor shall perform the necessary quality control sampling and testing to ensure that acceptable quality level requirements are met.

1.5 *Determination of Lot Percent Defective*

The lot percent defective with respect to each of the characteristics of gradation, air content and strength will be determined as follows:

1. Compute the sample mean ($\overline{X}$) and the standard deviation (S) of the N=6 (or N=5 or N=4 if fewer than 6 sublots) test results (X_i):

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

2. Compute the upper quality index (Q_U) and/or the lower quality index (Q_L) corresponding to the upper and/or lower limits listed in Subsection (v)-1.3(a):

$$Q_U = \frac{U - \overline{X}}{S} \qquad Q_L = \frac{\overline{X} - L}{S}$$

3. Using the appropriate portion of Table 1 (for N = actual sample size), determine the percentage of materials and construction

falling outside the specification limits PD_U and/or PD_L associated with Q_U and/or Q_L, respectively. Add these two values to obtain the lot percent defective:

$$PD = PD_U + PD_L$$

4. In the event that an individual test result on strength falls below the corresponding lower critical limit listed in Subsection (v)-1.3(b), compute the lower critical quality index (Q_{LC}):

$$Q_{LC} = \frac{\overline{X} - Lower\_Critical\_Limit}{S}$$

Where all terms are as previously defined. The value of Q_{LC} is then used in Table 1 to determine the percentage of materials and construction falling below the lower critical limit (PD_{LC}).

1.6 *Pay Factors for Lot Quality Characteristics*

Case I: Provided that no individual test result on strength falls below the corresponding lower-critical-limit listed in Subsection (v)-1.3(b), the pay factor for each of the characteristics of strength, air content and for each of the sieves listed in Subsection (v)-1.3 will be determined as follows:

1. Compute the lot percent defective (PD) as described in Subsection (v)-1.5(3).

2. If PD is less than or equal to fifty (50) percent, compute the pay factor percentage for the lot using the following equation:

$$PF = 102 - 0.04\ (PD) - 0.016\ (PD)^2$$

3. If PD is greater than fifty (50) percent, the Engineer may require removal and replacement of the defective lot at the Contractor's expense. If this option is not exercised, the Contractor may elect to replace the lot or leave it in place subject to a pay factor of PF=0%.

Case II: In the event that an individual test result on strength falls below the corresponding lower-critical-limit listed in Subsection (v)-1.3(b), the lot will be re-evaluated for the deficient characteristic by taking cores. Coring shall be performed as directed by the Engineer, and the final disposition of the lot will be based on the core results.

The sampling rate shall be at least two cores per sublot, averaged and considered as one individual test result.

The pay factor for strength will be determined based on the core test results as follows:

1. Compute the percentage of materials and construction falling below the critical limit (PD_{LC}) as described in Subsection (v)-1.5(4).

2. If PD_{LC} is less than or equal to five (5) percent, proceed to Step 4.

3. If PD_{LC} is greater than five (5) percent, the defective lot shall be removed and replaced at no cost to the Authority.

4. Using the core test results, compute the lot percent defective (PD) as described in Subsection (v)-1.5(3).

5. If PD is less than or equal to fifty (50) percent, compute the pay factor percentage for the lot using the following equation:

$$PF = 102 - 0.04\,(PD) - 0.016\,(PD)^2$$

6. If PD is greater than fifty (50) percent, the Engineer may require removal and replacement of the defective lot at the Contractor's expense. If this option is not exercised, the Contractor may elect to replace the lot or leave it in place subject to a pay factor of PF=0%.

1.7 *Pay Adjustment For Lots*

Once a lot has been defined, its identity will be maintained throughout the mixing and placement process. When the lot is completed, the individual pay factors determined in Subsection (v)-1.6 for gradation, air content and strength will be used to calculate a composite pay factor percentage (CPF) and a pay adjustment (PA) for the subject lot as follows:

$$CPF = \frac{PF_G + 3PF_A + 6PF_S}{10}$$

Where:

PF_G = Pay factor percentage for gradation -- the smallest of the individual pay factors for the sieves listed in Subsection (v)-1.3(a),
PF_A = Pay factor percentage for air content,
PF_S = Pay factor percentage for strength, and

The pay adjustment for the completed lot will be determined in accordance with the following formula:

$$PA_{Lot} = (CPF - 1)\,(CUP)\,(Q_{Lot})$$

Where:

PA_{Lot} = Pay adjustment for the lot
CPF = Composite pay factor percentage
CUP = Contract unit price ($/SQ YD), and

Q_{Lot} = Quantity of concrete in the lot (SQ YDS)

1.8 *Pay Adjustments Not Covered in Special Provisions.*
Adjustments in pay, for deviates from specified standards of characteristics other than those described in these Special Provisions (if any) will be made in accordance with Subsection 105.03.

1.9 *Total Pay Adjustment for Entire Project*
The total adjustment in pay for the entire project is the sum of (1) the pay adjustments for individual lots per Subsection (v)-1.7; plus (2) the pay adjustments for smoothness per Section 430; plus (3) the pay adjustments for thickness per Section 431; plus (4) other pay adjustments, if appropriate, per Subsection (v)-1.8.

2. *Conflicts Between Engineer's and Contractor's Test Results*
At the beginning and throughout the contract, the Engineer and the Contractor shall compare each other's test procedures and results. The comparison should be based on the methods described in Appendix-F of the "AASHTO Implementation Manual for Quality Assurance, 1995".

Should the Engineer determine from the comparison or for any other reason that any of the acceptance test results are incorrect, such results will be discarded. In this case, additional acceptance sampling and testing will be performed to supplement the remaining, valid test results.

If the Engineer and the Contractor are unable to resolve their differences, the Contractor may request referee testing by an independent testing laboratory accredited by AASHTO. Such laboratory must be acceptable to both the Engineer and the Contractor.

The request for referee testing shall be submitted in writing by the Contractor to the Engineer within thirty (30) days after completion of the lot. Referee testing will be independent from any previous testing by either the Engineer or the Contractor and the results of such referee testing shall be considered final. Should the referee testing results in higher pay factors for the lot(s) in question, the Authority will pay the cost of referee testing. Otherwise, the entire cost of testing shall be borne by the Contractor.

3. *Extreme Values (Outliners)*
Test results apparently inconsistent with the results of the majority of tests will also be closely examined by the Engineer in order to determine their validity. The examination will cover the procedures used in sampling and testing and, if necessary, a mathematical analysis will be performed in accordance with ASTM E-178 using the upper 2.5% significance level. Test results thus determined by the Engineer to be non-representative of the material being evaluated will be

discarded. In this case, additional acceptance sampling and testing will be performed to supplement the remaining, valid test results.

(Add **Subsection 414.04.(x) "CONSTRUCTION METHODS. Contractor's Quality Control Testing and Inspection. *Plant Startup Requirements."*** *as follows):*

(x) *Plant Startup Requirements.* At or prior to the Pre-Work Conference, the Contractor shall provide a quality control system. The system shall include the fully equipped laboratory and the full complement of quality control personnel that are to perform the quality control functions for the remainder of the project.

Plant startup production shall be limited to that necessary to calibrate the plant, testing equipment and procedures using the approved mix design(s). The concrete thus produced shall be sampled and tested by both the Contractor and Engineer for air content and strength (3-day). The Contractor's test results shall then be reconciled with those from the Engineer.

No concrete from the startup operation that does not meet the requirements of Acceptable Quality Level as defined in Subsection (v)-1.4 shall be placed in the pavement. Instead adjustments to the process shall continue to be made until all of the requirements are met. Concrete not meeting the requirements shall become the property of the Contractor and will not be paid for. Costs associated with the startup operations will not be measured separately for payment but will be included in the payment for Contractor's Quality Control.

Appendix B: OTA's Percent-Within-Limits (PWL) Special Provisions for AC Pavements

OKLAHOMA TRANSPORTATION AUTHORITY
SPECIAL PROVISIONS
FOR
QUALITY CONTROL AND ACCEPTANCE PROCEDURES
FOR
PLANT MIX ASPHALT CONCRETE PAVEMENT

These Special Provisions revise, amend and, where in conflict, supersede applicable sections of the *Standard Specifications for Turnpike Construction*, Edition of 1999. These Special Provisions apply to all types of Asphalt Concrete Pavement.

The Contractor is responsible for the quality of the materials and construction, whereas the Engineer will be responsible for determining the acceptability of such materials and construction.

It is the intent of these Special Provisions that materials and construction of acceptable quality shall receive an average pay factor of one hundred (100) percent; that materials and construction of truly superior quality shall be awarded a bonus payment; and that materials and construction of deficient quality will receive a reduced payment or be removed and replaced.

(Add the following to **Subsection 411.01. "DESCRIPTION."***):*
Contractor's Quality Control and Acceptance Procedures will apply to this work in accordance with the applicable requirements of Section 106 and as herein specified.

(Revise **Subsection 411.04.(m)2 "CONSTRUCTION METHODS. Contractor's Quality Control Testing and Inspection. *Contractor's Testing.*"** *to read as follows):*

2. *Contractor's Testing.* As a minimum, the Contractor's sampling and testing shall comply with the following schedule:

Property	Sampling and Testing Frequency
Gradation	1 per 5000 SQ YDS per lift
Asphalt Content	1 per 5000 SQ YDS per lift
Air Voids	1 per 5000 SQ YDS per lift
Roadway Density	1 per 5000 SQ YDS per lift
Thickness	As needed to control operations
Smoothness	As needed to control operations

Additional sampling and testing to ensure compliance with Standard Specifications and other Special Provision requirements shall be in accordance with the Contractor's Quality Control Plan.

(Add the following to **Subsection 411.04.(m) "CONSTRUCTION METHODS. Contractor's Quality Control Testing and Inspection."***):*

4. *Control Charts.* Control charts covering as a minimum the characteristics of gradation, asphalt content, air voids, roadway density and thickness shall be maintained by the Contractor and displayed at the plant or job site. The charts shall identify the project number, the contract item number, the characteristics, the date, the lot and sublot numbers, the applicable upper and/or lower specification limits, the Contractor's test results and any other data needed to facilitate control of the process and identify problems before they become serious. Copies of the Contractor's quality control tests shall be provided to the Engineer at time intervals acceptable to the Engineer.

(Revise **Subsection 411.04.(n) "CONSTRUCTION METHODS. Contractor's Quality Control Testing and Inspection. *Acceptance.*"** *to read as follows):*

(n) **Acceptance.** While the Contractor shall be fully and exclusively responsible for producing an acceptable product, acceptance responsibility rests with the Engineer. The entire lot of asphalt, as defined in Subsection (n)-1.1, will be accepted or rejected and paid for on the basis of acceptance test results. The Engineer may choose to use the Contractor's tests for acceptance after the Contractor's test results have been demonstrated to be consistent with tests taken by the Engineer and that they adequately represent the material being evaluated.

1. *Basis for Acceptance and Payment.* The following characteristics will be considered in evaluating materials and construction for acceptance and payment:

- Asphalt Cement Content
- Gradation
- Air Voids
- Roadway Density
- Pavement Structure Thickness (as provided in Section 431)
- Pavement Surface Smoothness (as provided in Section 430)

1.1 Lot and Sublot Definition

Except for surface smoothness and thickness, acceptance and pay adjustments will be made on a lot-by-lot basis. Each lot of Asphalt Concrete will be sampled at random and tested for all the quality characteristics described in (n)-1.1, in accordance with the following requirements.

The standard lot size shall consist of five equal sublots of 5000 square yards each per lift. Each sublot will be sampled at random to obtain one or more test specimens as follows:

a) Gradation and asphalt cement content determination: one specimen and one test for each characteristic per sublot.
b) Air voids determination: three specimens per sublot averaged and considered as one individual test.
c) Roadway density determination: three cores and/or three nuclear gauge test determinations per sublot with additional sampling and testing made as necessary. The average of the density measurements is considered as one individual test.

In the event that operational conditions cause work to be interrupted before the standard lost size has been achieved, the lot size may be redefined by the Engineer. However, the number of test determinations required to evaluate each lot will be at least four. Each partial lot will be divided into at least four equal sublots, and each sublot will be sampled at random to obtain the required number of test specimens.

Quantities of mixture less than 250 tons may be accepted by the Engineer upon visual inspection by which the Engineer has reason to believe that the materials and construction are of acceptable quality. At the Engineer's option, this quantity may be treated as a separate lot, combined with the previous lot or combined with the following lot, as the case may warrant. On a multiple project contract, the lots of concrete will carry over from project to project.

1.2 *Smoothness and Thickness Acceptance and Pay Adjustments*
For smoothness and thickness determination and pay adjustment purposes, the pavement will be tested on an extent-to-extent basis in accordance with Sections 430 and 431. Acceptance and pay adjustment determinations made under Sections 430 and 431 will be completely independent of those made under this Special Provision.

1.3 *Criteria for Lot Acceptance and Payment*
Except for surface smoothness and thickness, conformance with the specifications will be judged on the basis of the following two criteria:

a) The estimated lot percent defective with respect to gradation, asphalt cement content, air voids and roadway density. The lot percent defective with respect to a particular quality characteristic is the amount of materials and construction which falls outside the specified limit(s) listed in the following table:

Quality Characteristic	Lower Limit (L)	Upper Limit (U)
Gradation:		
Sieves #4 and larger	(Target – 6.0)%	(Target + 6.0)%
Sieves #10 through #80	(Target – 4.5)%	(Target + 4.5)%

Sieves #200	(Target – 2.0)%	(Target + 2.0)%
Asphalt Cement Content	(Target – 0.6)%	(Target + 0.6)%
Air Voids (LMS)	(Target – 2.5)%	(Target + 2.5)%
Roadway Density	93%	98%

b) Any load of asphalt mixture that is excessively segregated or having aggregate improperly coated will be rejected for use in the work. Excessively high or low temperature will also be cause for rejection. Furthermore, sections of completed pavement which from visual observation or known deficiencies appear to be seriously inadequate will be extensively tested. The results of such tests will not be used for pay adjustment purposes, but will be used to determine whether the section is totally unacceptable and must be removed. In the event that a section is determined to be unacceptable, its removal and replacement shall be at no additional cost to the Authority.

1.4 Acceptable Quality Level

A lot shall be considered of acceptable quality with respect to a particular characteristic if the percent defective, as defined in Subsection (n)-1.3(a), is no more than ten (10) percent. The contractor shall perform the necessary quality control sampling and testing to ensure that acceptable quality level requirements are met.

1.5 Determination of Lot Percent Defective

The lot percent defective with respect to each of the characteristics of gradation, asphalt content, air voids and roadway density will be determined as follows:

1. Compute the sample mean ($\overline{X}$) and the standard deviation (S) of the N=5 (or N=4 if fewer than 5 sublots) test results (X_i):

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

2. Compute the upper quality index (Q_U) and/or the lower quality index (Q_L) corresponding to the upper and/or lower limits listed in Subsection (n)-1.3(a):

$$Q_U = \frac{U - \overline{X}}{S} \qquad Q_L = \frac{\overline{X} - L}{S}$$

3. Using the appropriate portion of Table 1 (for N = actual sample size), determine the percentage of materials and construction falling outside the specification limits PD_U and/or PD_L associated with Q_U and/or Q_L, respectively. Add these two values to obtain the lot percent defective:

$PD = PD_U + PD_L$

1.6 *Pay Factors for Lot Quality Characteristics*

Except for pavement smoothness and thickness, the pay factor (PF) for each quality characteristic will be determined as follows:

1. If PD is less than fifty (50) percent proceed to Step 4.

2. If PD is greater than or equal to fifty (50) percent but less than sixty (60) percent, the Engineer may elect to reevaluate the lot with additional test specimens as described in Step 2b.

   a) If no additional test specimens are taken, proceed to Step 4.
   b) If the Engineer elects to reevaluate the lot, five additional test specimens will be taken at new random locations. Using the five new test results, estimate the total percent defective PD as explained in Subsection (n)-1.6. The final PD value for the lot will be the average of the PD values determined using the two sets of test specimens.

3. If PD is greater than or equal to sixty (60) percent, the Engineer may require removal and replacement of the defective lot at the Contractor's expense. If this option is not exercised, the Contractor may elect to replace the lot or leave it in place subject to a pay factor of PF=0%.

4. Compute the pay factor percentage for the lot using the following equation:

   $$PF = 102 - 0.04\ (PD) - 0.016\ (PD)^2$$

1.7 *Pay Adjustment For Lots*

Once a lot has been defined, its identity will be maintained throughout the mixing and placement process. When the lot is completed, the individual pay factors determined in Subsection (n)-1.6 for gradation, asphalt content, air voids and roadway density will be used to calculate a composite pay factor percentage (CPF) and a pay adjustment (PA) for the subject lot as follows:

$$CPF = \frac{PF_G + 3PF_A + 3PF_V + 3PF_D}{10}$$

Where:

PF_A = Pay factor percentage for asphalt content,
PF_V = Pay factor percentage for air voids,
PF_D = Pay factor percentage for roadway density, and

PF_G = Pay factor percentage for gradation -- the smallest of the individual pay factors for the sieves listed in Subsection (n)-1.3(a)

The pay adjustment for the completed lot will be determined in accordance with the following formula:

$$PA_{Lot} = (CPF - 1) \cdot (CUP) \cdot (Q_{Lot}) \cdot \frac{T_{Lift}}{T_{Struct}}$$

Where:

PA_{Lot} = Pay adjustment for the lot
CPF = Composite pay factor percentage
CUP = Contract unit price ($/SQ YD), and
Q_{Lot} = Quantity of concrete in the lot (SQ YDS)
T_{Lift} = Nominal lift thickness (as shown on the plans)
T_{Struct} = Nominal structure thickness (as shown on the plans) to which CUP applies

1.8 Pay Adjustments Not Covered in Special Provisions

Adjustments in pay, for deviations from specified standards of characteristics other than those described in these Special Provisions (if any) will be made in accordance with Subsection 105.03.

1.9 Total Pay Adjustment for Entire Project

The total adjustment in pay for the entire project is the sum of (1) the pay adjustments for individual lots per Subsection (n)-1.7; plus (2) the pay adjustments for smoothness per Section 430; plus (3) the pay adjustments for thickness per Section 431; plus (4) other pay adjustments, if appropriate, per Subsection (n)-1.8.

2. Conflicts Between Engineer's and Contractor's Test Results

At the beginning and throughout the contract, the Engineer and the Contractor shall compare each other's test procedures and results. The comparison should be based on the methods described in Appendix-F of the "AASHTO Implementation Manual for Quality Assurance, 1995".

Should the Engineer determine from the comparison or for any other reason that any of the acceptance test results are incorrect, such results will be discarded. In this case, additional acceptance sampling and testing will be performed to supplement the remaining, valid test results.

If the Engineer and the Contractor are unable to resolve their differences, the Contractor may request referee testing by an independent testing laboratory accredited by AASHTO. Such laboratory must be acceptable to both the Engineer and the Contractor.

The request for referee testing shall be submitted in writing by the Contractor to the Engineer within thirty (30) days after completion of the lot. Referee testing will be independent from any previous testing by either the Engineer or the Contractor and the results of such referee testing shall be considered final. Should the referee testing results in higher pay factors for the lot(s) in question, the Authority will pay the cost of referee testing. Otherwise, the entire cost of testing shall be borne by the Contractor.

3. *Extreme Values (Outliners)*
Test results apparently inconsistent with the results of the majority of tests will also be closely examined by the Engineer in order to determine their validity. The examination will cover the procedures used in sampling and testing and, if necessary, a mathematical analysis will be performed in accordance with ASTM E-178-80 using the upper 2.5% significance level. Test results thus determined by the Engineer to be non-representative of the material being evaluated will be discarded. In this case, additional acceptance sampling and testing will be performed to supplement the remaining, valid test results.

(Add **Subsection 411.04.(o) "CONSTRUCTION METHODS. Contractor's Quality Control Testing and Inspection. *Plant Startup Requirements.*"** *as follows):*

(o) *Plant Startup Requirements.* Prior to beginning production of asphalt for the mainline, the Contractor shall provide a quality control system. The system shall include the fully equipped laboratory and the full complement of quality control personnel that are to perform the quality control functions for the remainder of the project.

Plant startup production shall be limited to that necessary to calibrate the plant and the testing equipment and procedures using the mix design approved for mainline construction. The asphalt concrete thus produced shall be sampled and tested by both the Contractor and Engineer for VMA, Hveem Stability, gradation and asphalt content. The Contractor's test results shall then be reconciled with those from the Engineer.

No asphalt concrete from the startup operation shall be placed on the mainline or the control strip. Instead adjustments shall continue to be made until all of the requirements are met. Asphalt concrete from the plant startup operation may be utilized and paid for in the construction temporary facilities or if no temporary facilities are available they shall remain the property of the Contractor and will not be paid for. Costs associated with startup operations will not be measured separately for payment but will be included in the payment for Contractor's Quality Control.

(Add **Subsection 411.04.(p) "CONSTRUCTION METHODS. Contractor's Quality Control Testing and Inspection.** ***Control Strip Requirements."*** *as follows):*

(p) *Control Strip Requirements.* After fulfilling the plant startup requirements, one or more control strips shall be constructed on the shoulder or detour for the purpose of verifying the required production mix characteristics and establishing rolling patterns to obtain target requirements. The initial placement of asphalt shall be limited to approximately 500 tons. This material shall then be sampled and tested by the Contractor and the Engineer for VMA, Hveem Stability and all of the characteristics in Subsection (n)1.1. No additional asphalt shall be placed until all the results are evaluated and necessary adjustments in production and placement procedures are made. No pay adjustments will be made on the approximately 500 ton placement.

After necessary adjustments are made, the above process shall be repeated for the next approximately 500 tons of asphalt placed. Pay adjustments for deviations from target on this second placement will be made at the rate of one half those specified. If required, additional control strips shall be made on the shoulder until an acceptable product (i.e., percent defective of no more than ten percent) is produced. Pay adjustments will be applied to all asphalt mixture in excess of the first approximately 1000 tons as described in Subsection 411.04.(n)1.7. Control strips will not be measured separately for payment. Work and materials associated with control strips will be paid for at the contract unit price (as adjusted) for the appropriate type of asphalt concrete.

Appendix C: Normality Test Results

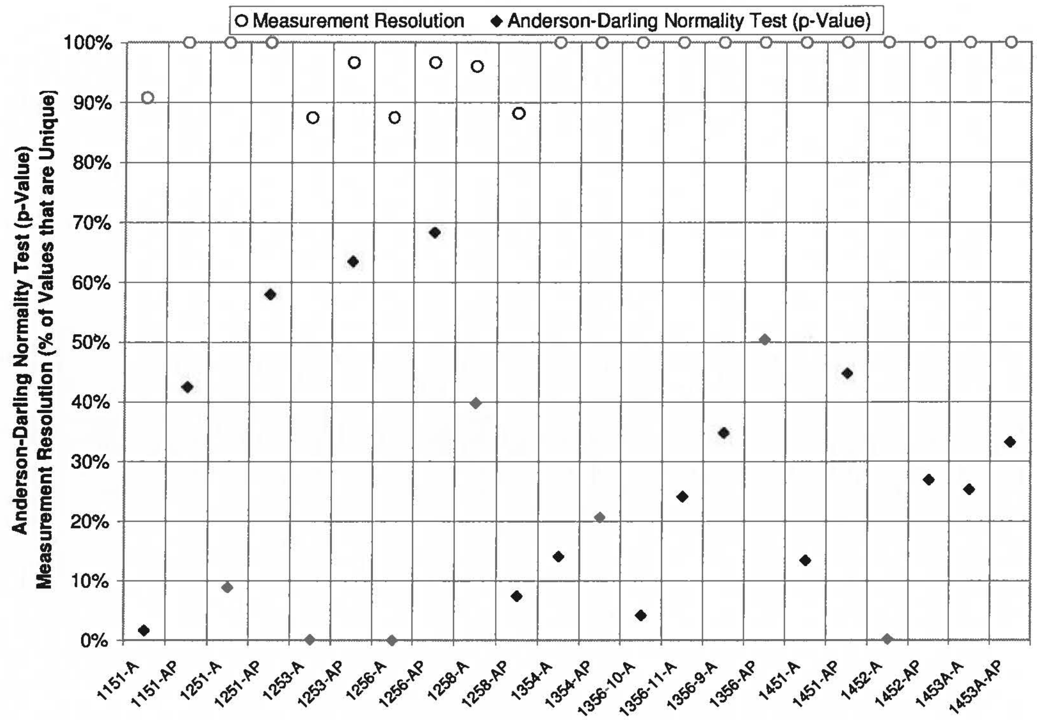
The charts in this appendix provide the results of the non-normality analysis of the data from the ten PCC and seven AC projects included in the study. The results are presented as "p-values" based on the Anderson-Darling normality test. The p-value represents the likelihood that the observed data came from a normally-distributed population. Statistical hypothesis-testing uses p-values to make inferences about a given population based on samples taken from the population. In this instance, the statistical hypothesis test is as follows:

H_o: the data are from a normal population

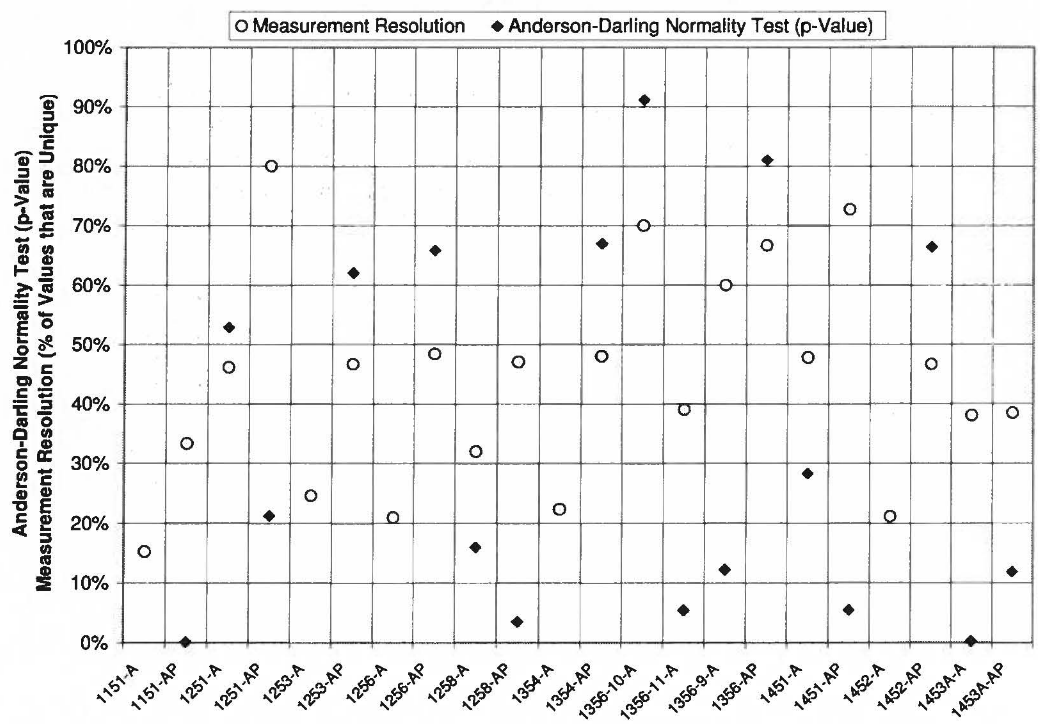
H_a: the data are not from a normal population

Hypothesis testing involves assuming the null hypothesis (H_o) is true unless the data provide *strong evidence* to the contrary. How strong the evidence must be depends upon the chosen level of Type-I decision risk (α), where α is the risk of wrongly rejecting the null hypothesis even if it is, in fact, true. In practice, if the p-value is less than or equal to α, we reject the null hypothesis and conclude that the alternate hypothesis (H_a) is true and the null hypothesis (H_o) is not true. Typical values of α are 0.05 and 0.10.

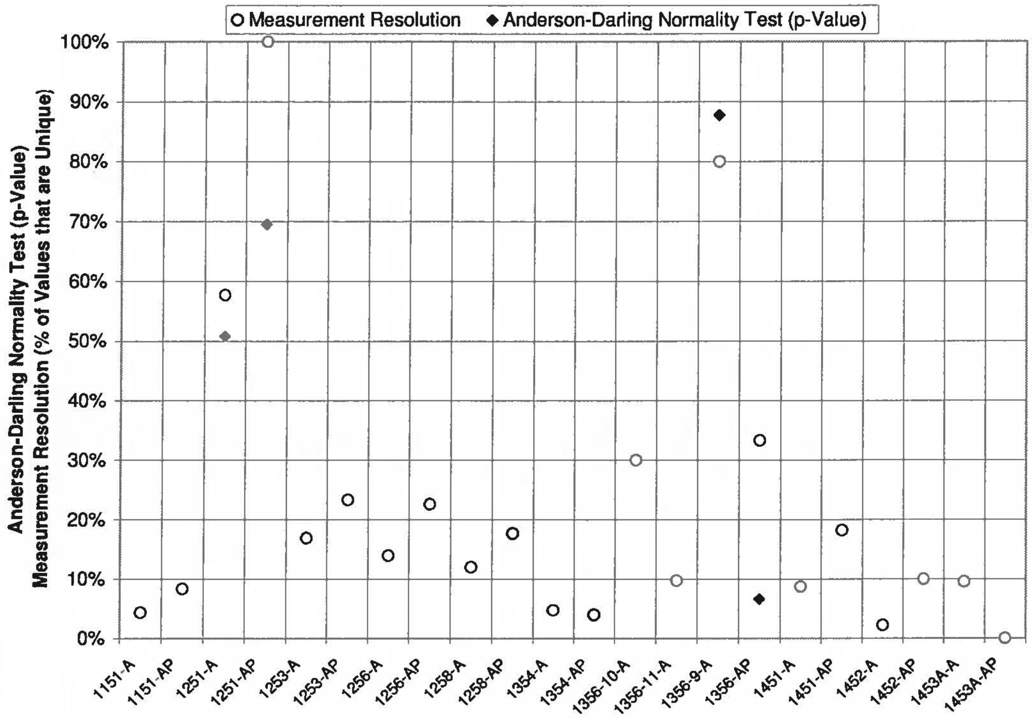
Much of the non-normality observed with the aggregate gradation data was due to the presence of rounding for the "percent-passing" values. In an attempt to quantify the instances where the observed non-normality was due to measurement-system resolution rather than distribution shape, the non-normality charts also provide the "measurement resolution" for each sample, where the measurement resolution is reported as the percentage of observed values that are truly unique. As an example, if the sample includes the following four observations: (3, 3, 4 and 4), only 50% of the observed values are unique (3 and 4 – two out of the four values). Similarly, a sample including the following ten observations: (2, 2, 2, 3, 3, 3, 3, 4, 5, and 6) would also have 50% unique values (2, 3, 4, 5 and 6 – five out of the ten values). Whenever the measurement resolution is below about 30%, the data are very likely to fail the normality test even if the true shape of the distribution is essentially normal. As such, whenever the measurement resolution is below 30%, the normality p-values are not shown.



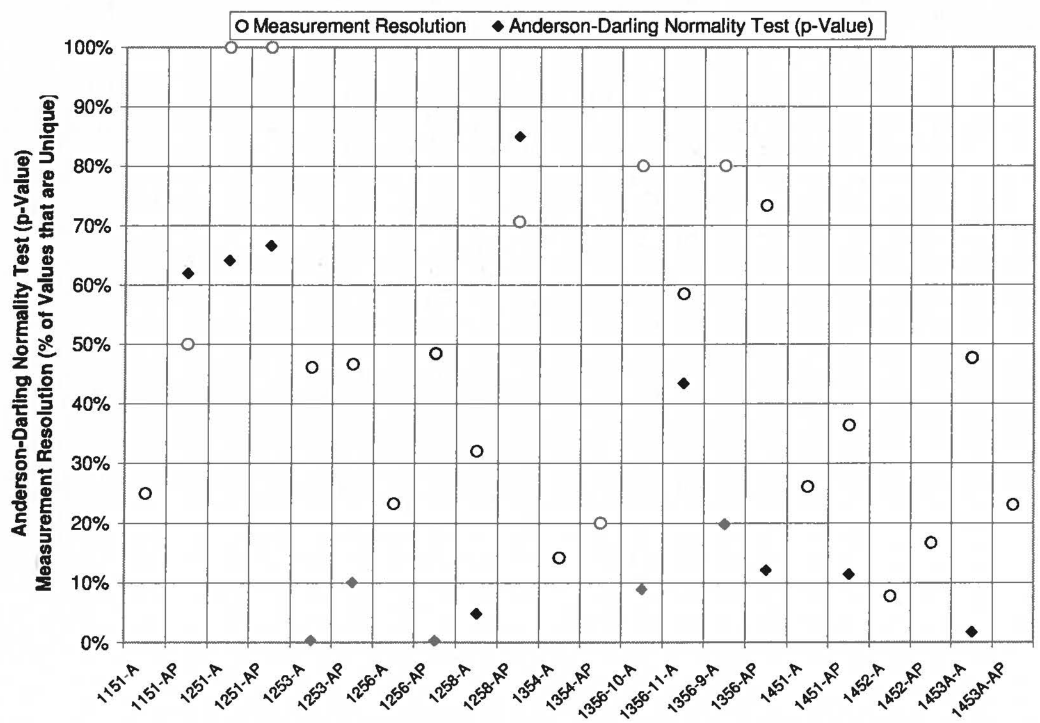
PCC Strength (Normality Test and Measurement Resolution)



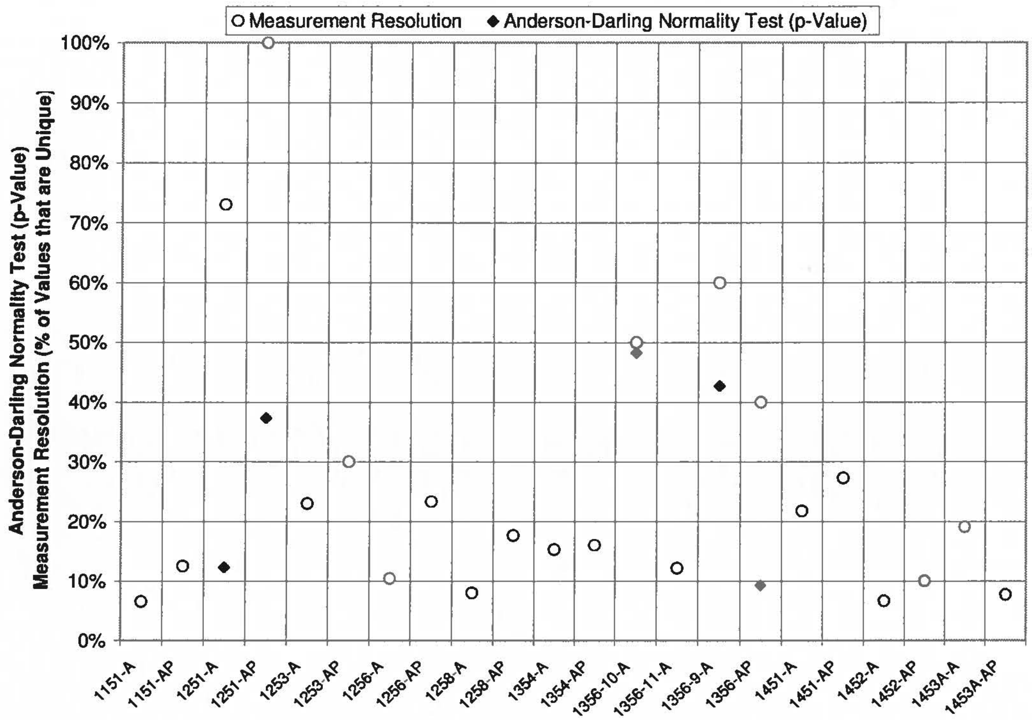
PCC Air Content (Normality Test and Measurement Resolution)



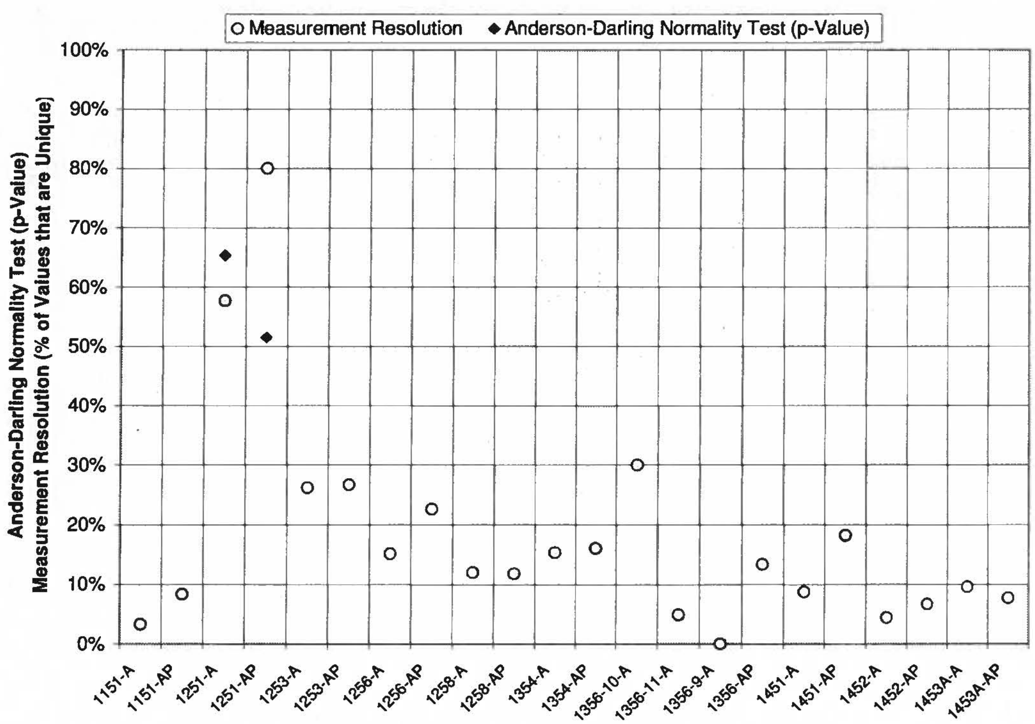
PCC Coarse Aggregate Gradation 1-inch Sieve (Normality Test and Measurement Resolution)



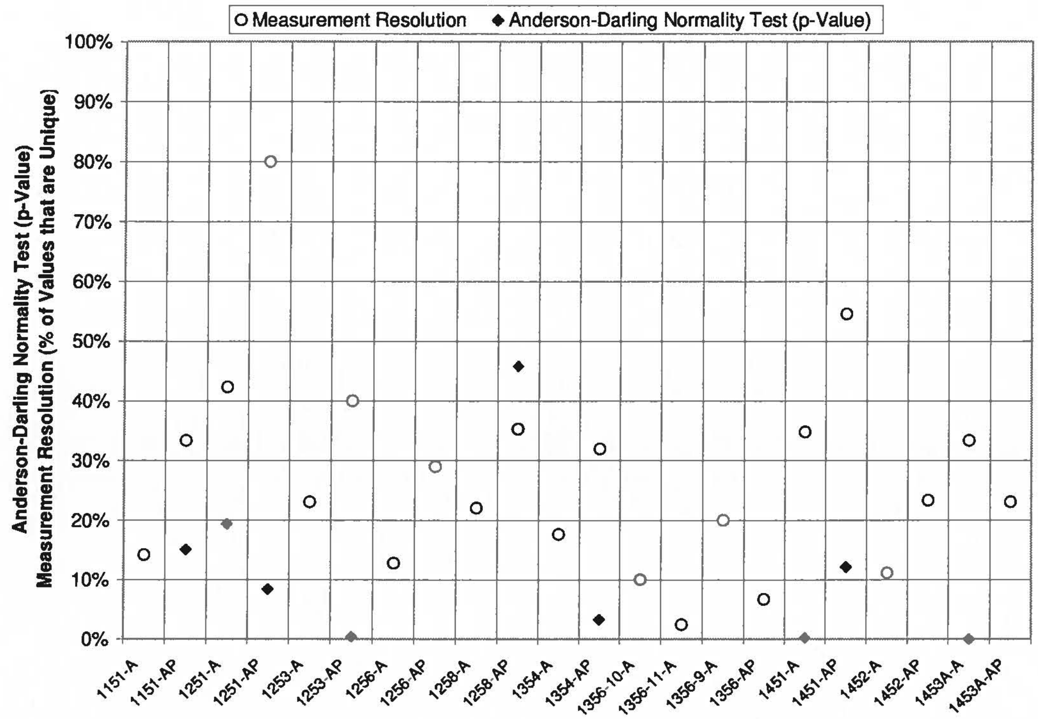
PCC Coarse Aggregate Gradation ½-inch Sieve (Normality Test and Measurement Resolution)



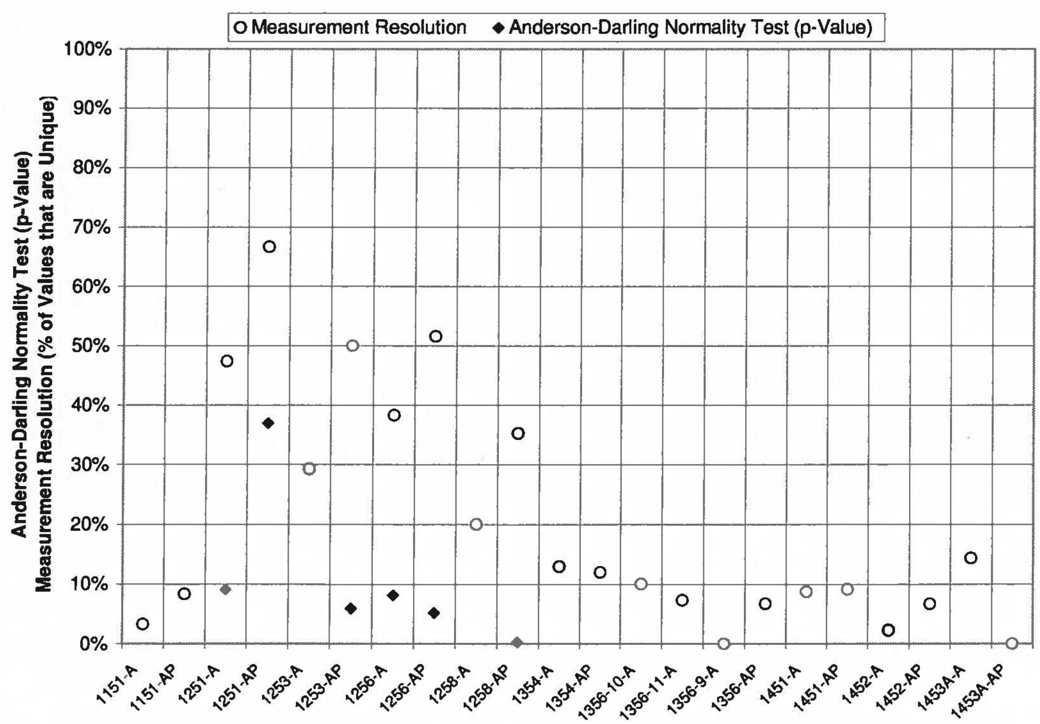
PCC Coarse Aggregate Gradation No. 4 Sieve (Normality Test and Measurement Resolution)



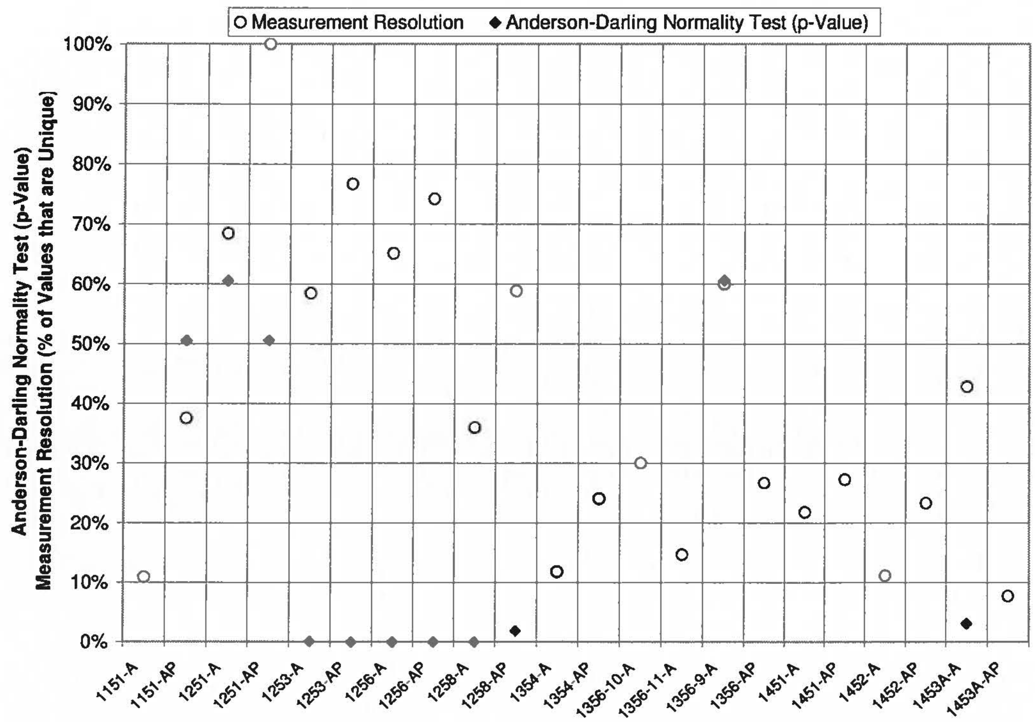
PCC Coarse Aggregate Gradation No. 8 Sieve (Normality Test and Measurement Resolution)



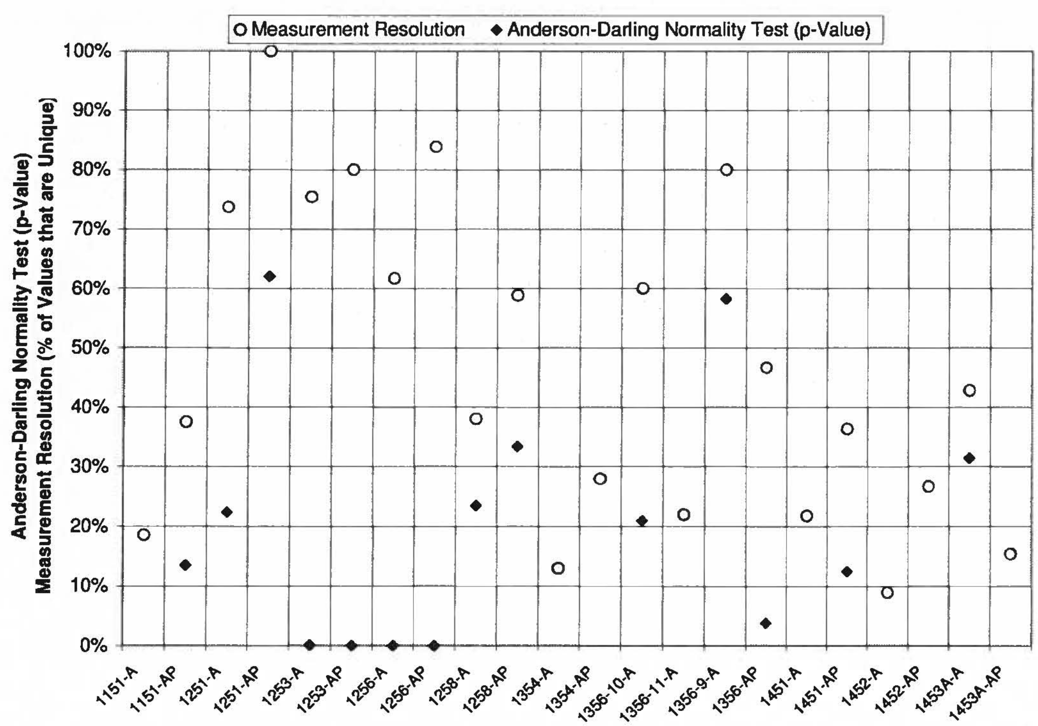
PCC Coarse Aggregate Gradation No. 200 Sieve (Normality Test and Measurement Resolution)



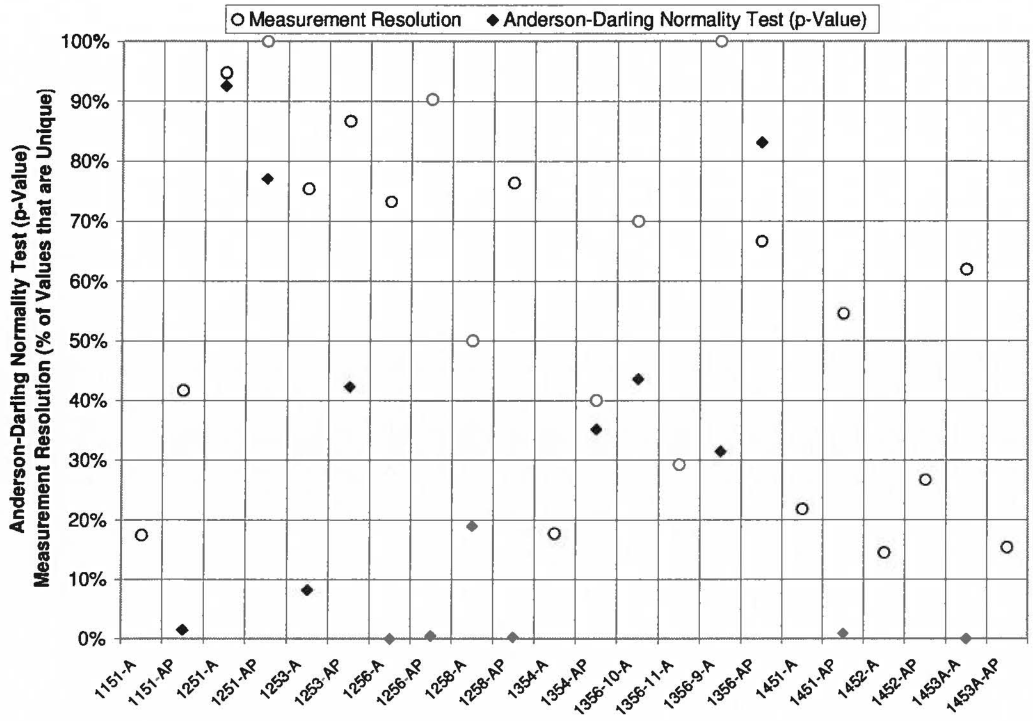
PCC Fine Aggregate Gradation No. 4 Sieve (Normality Test and Measurement Resolution)



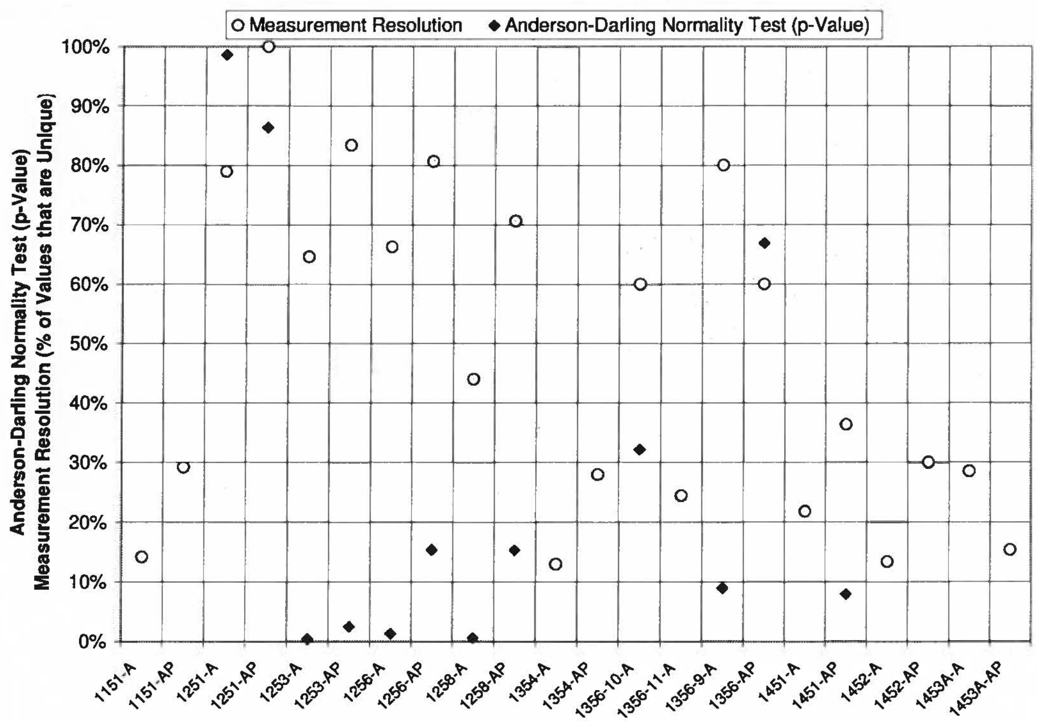
PCC Fine Aggregate Gradation No. 8 Sieve (Normality Test and Measurement Resolution)



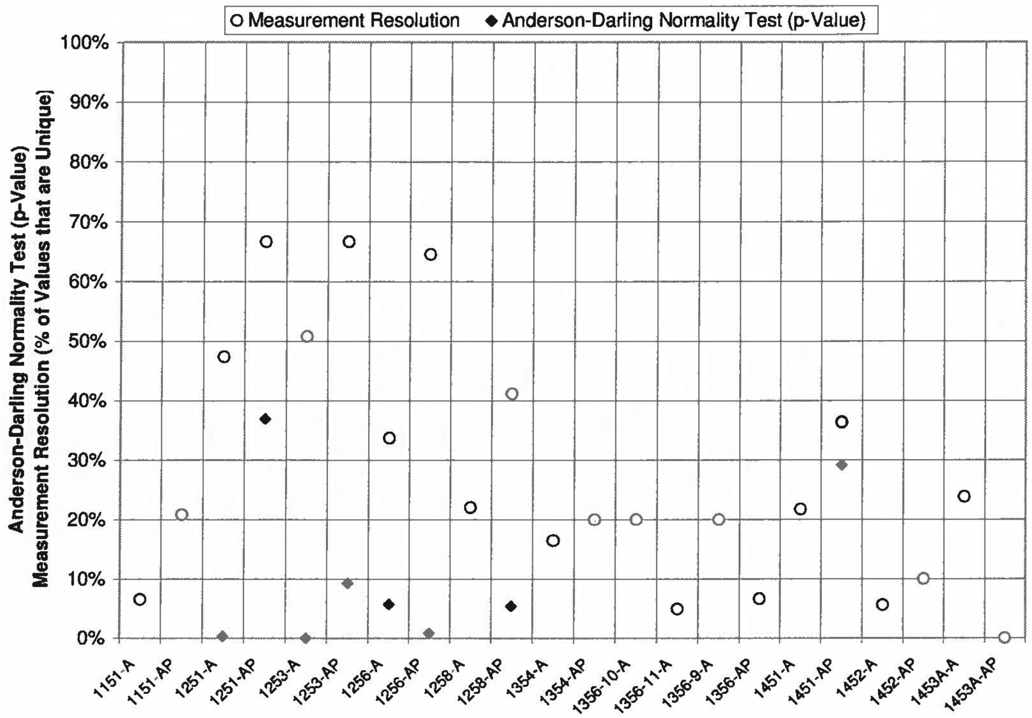
PCC Fine Aggregate Gradation No. 16 Sieve (Normality Test and Measurement Resolution)



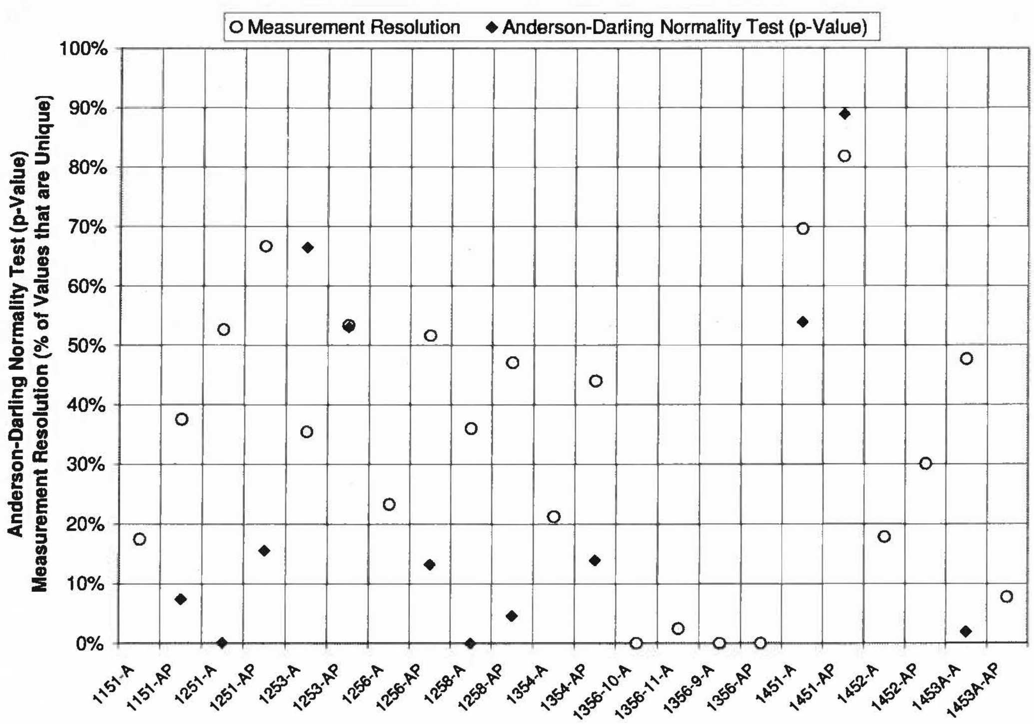
PCC Fine Aggregate Gradation No. 30 Sieve (Normality Test and Measurement Resolution)



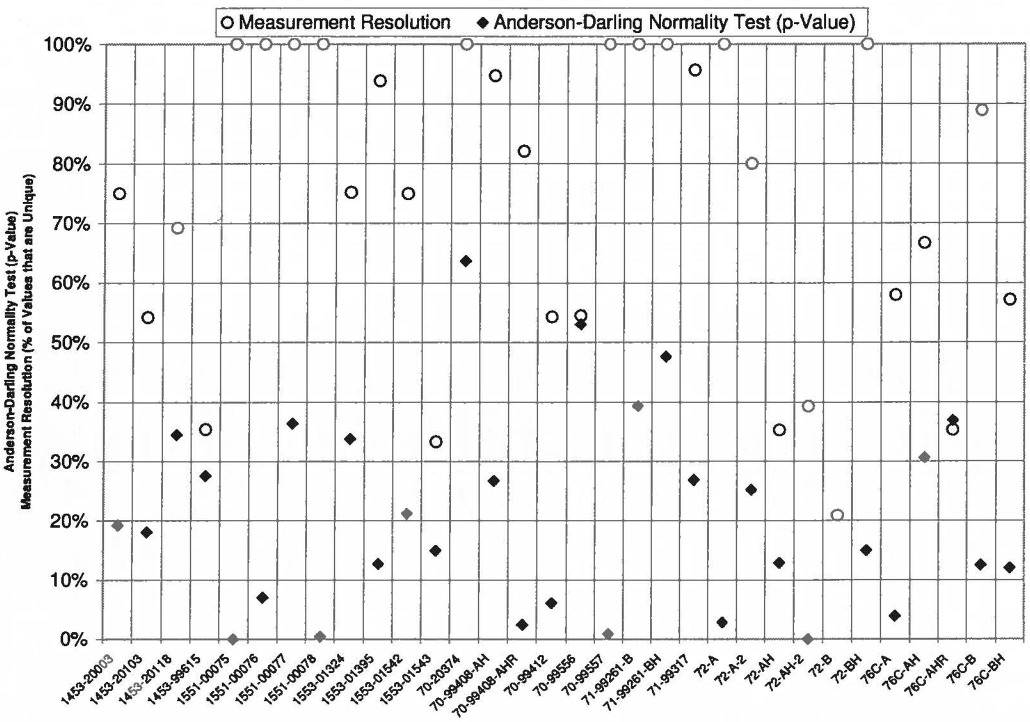
PCC Fine Aggregate Gradation No. 50 Sieve (Normality Test and Measurement Resolution)



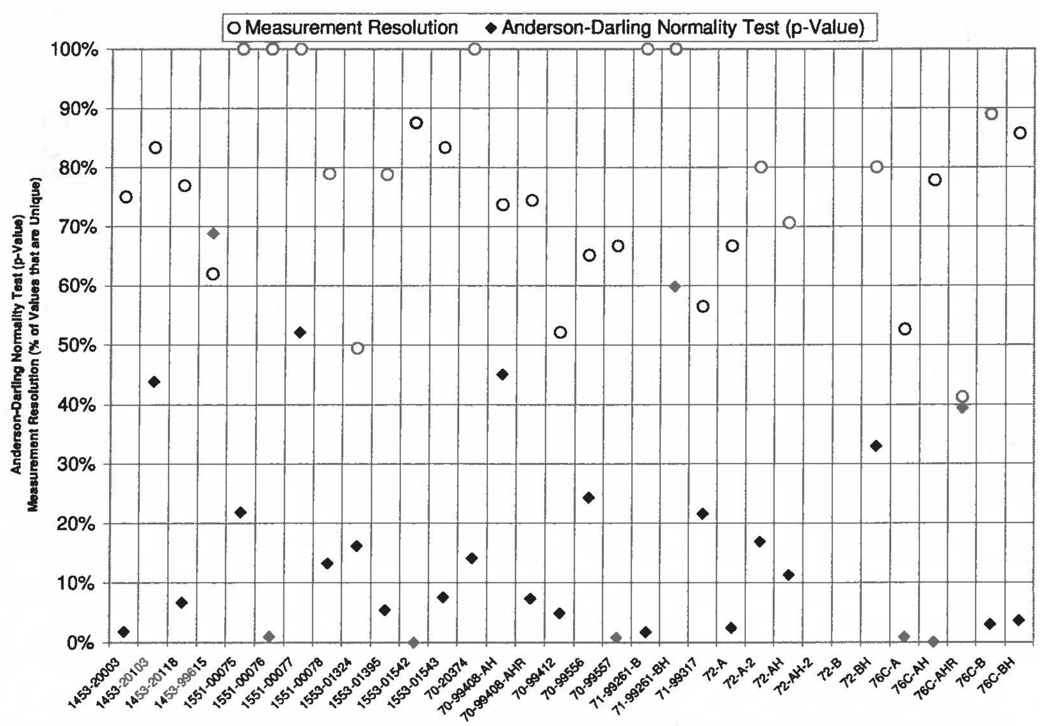
PCC Fine Aggregate Gradation No. 100 Sieve (Normality Test and Measurement Resolution)



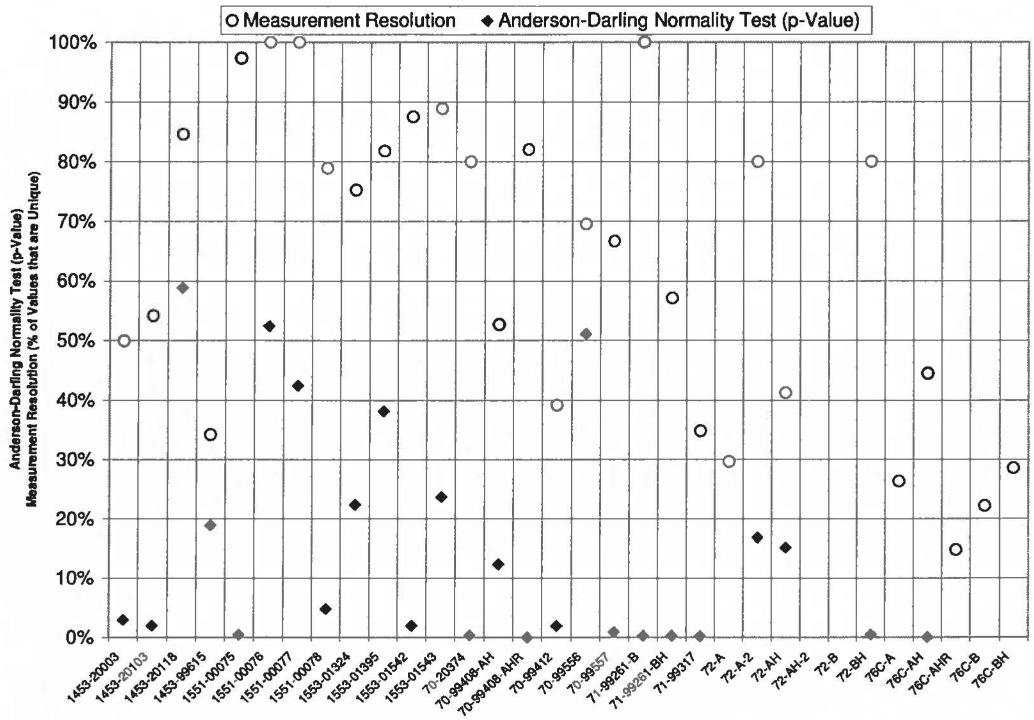
PCC Fine Aggregate Gradation No. 200 Sieve (Normality Test and Measurement Resolution)



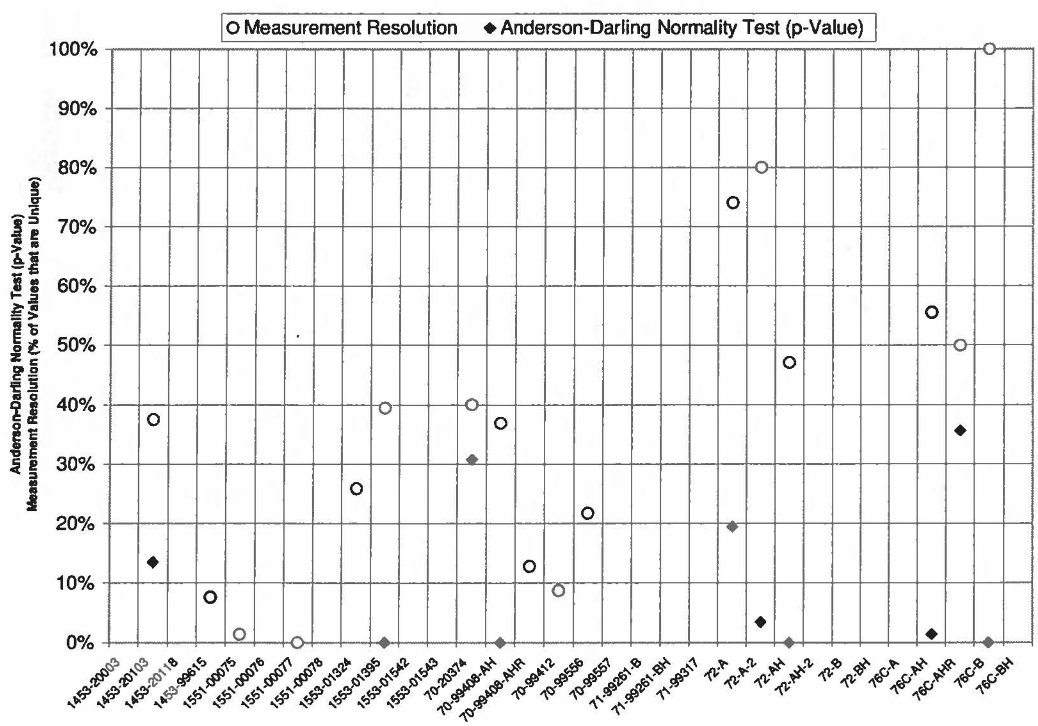
AC Roadway Density (Normality Test and Measurement Resolution)



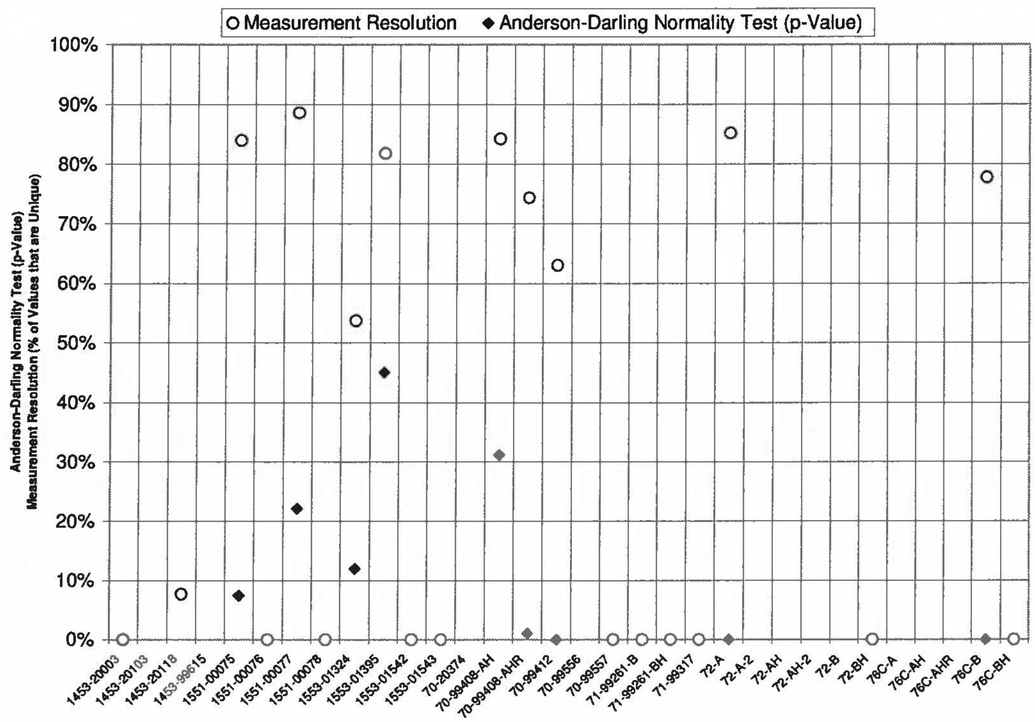
AC Asphalt Cement Content (Normality Test and Measurement Resolution)



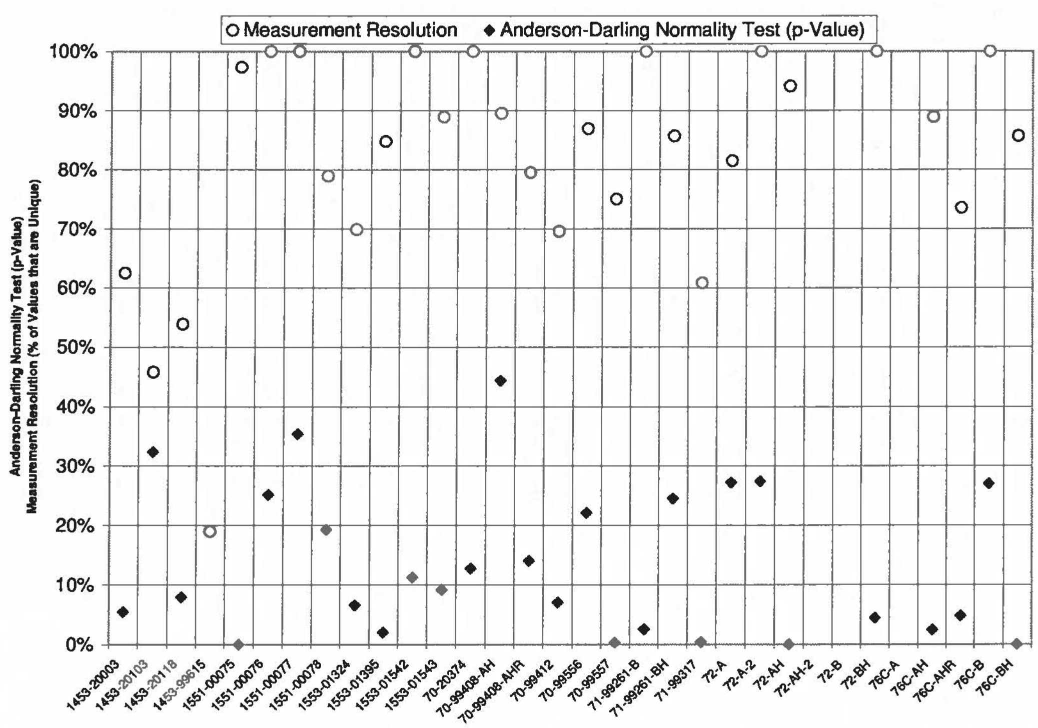
AC Air Voids (Normality Test and Measurement Resolution)



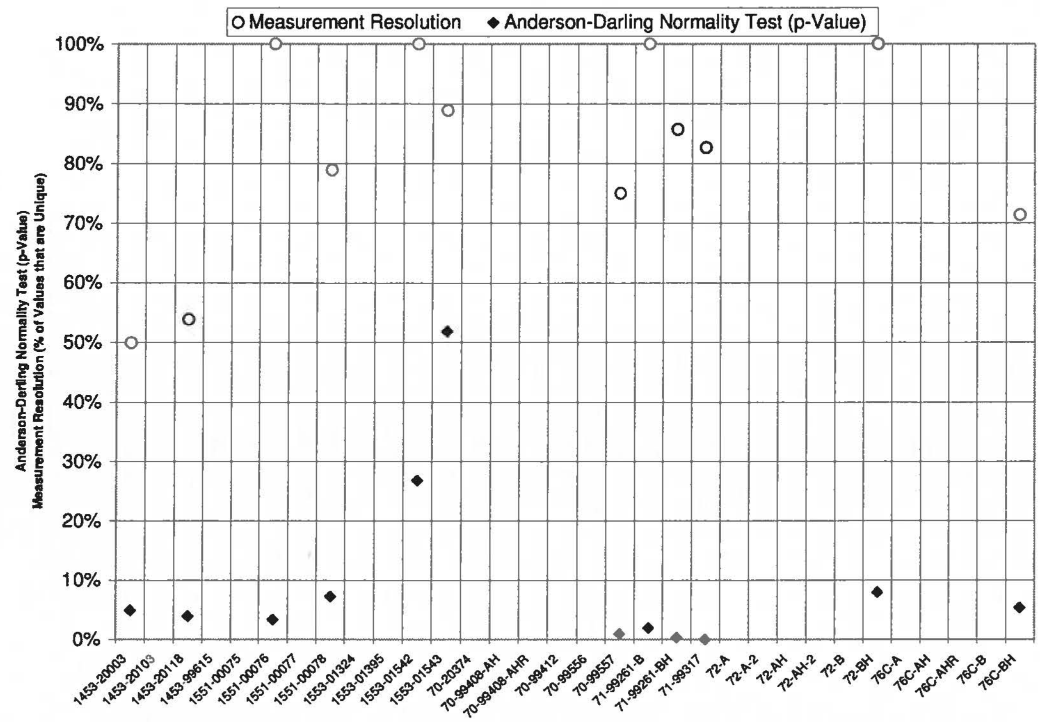
AC Gradation 1-inch Sieve (Normality Test and Measurement Resolution)



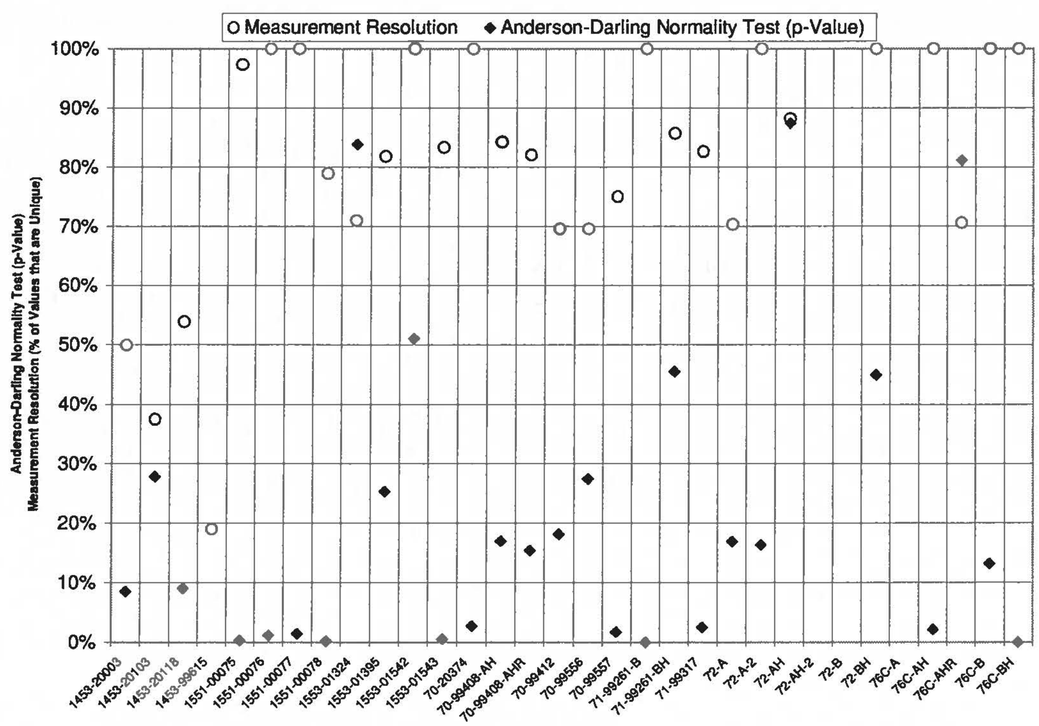
AC Gradation ¾-inch Sieve (Normality Test and Measurement Resolution)



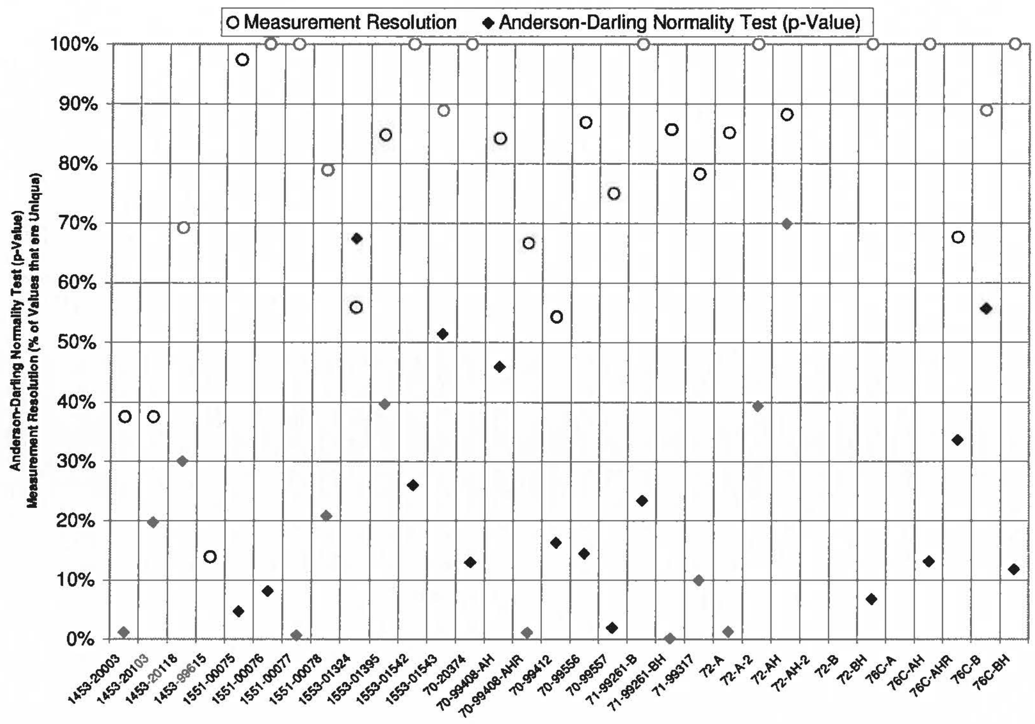
AC Gradation ½-inch Sieve (Normality Test and Measurement Resolution)



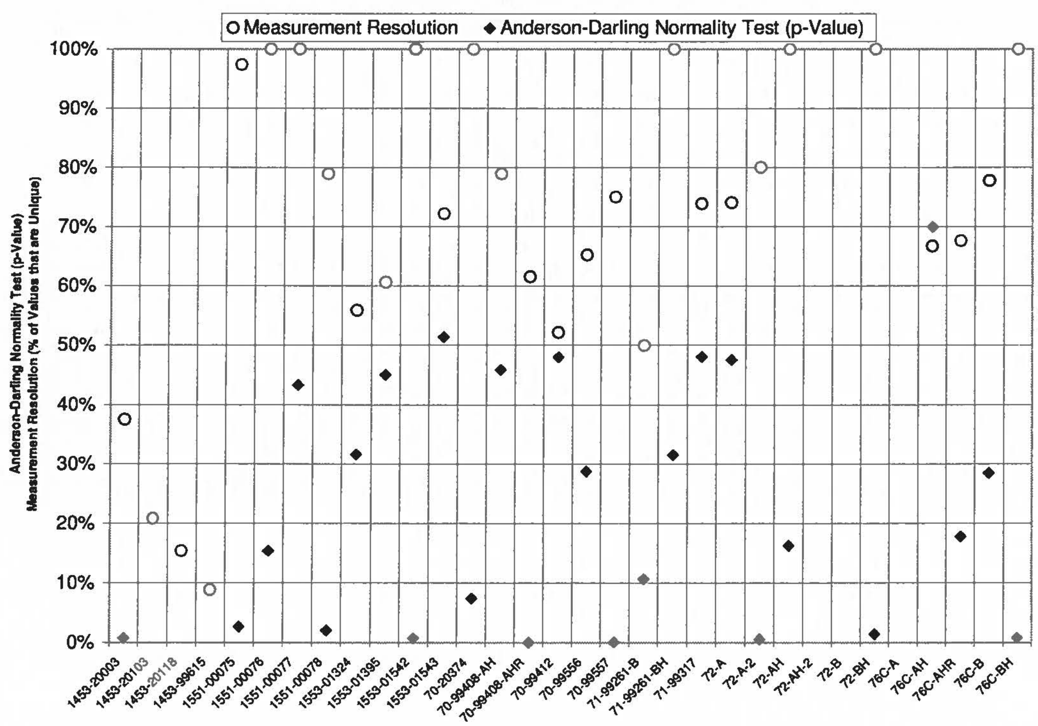
AC Gradation 3/8-inch Sieve (Normality Test and Measurement Resolution)



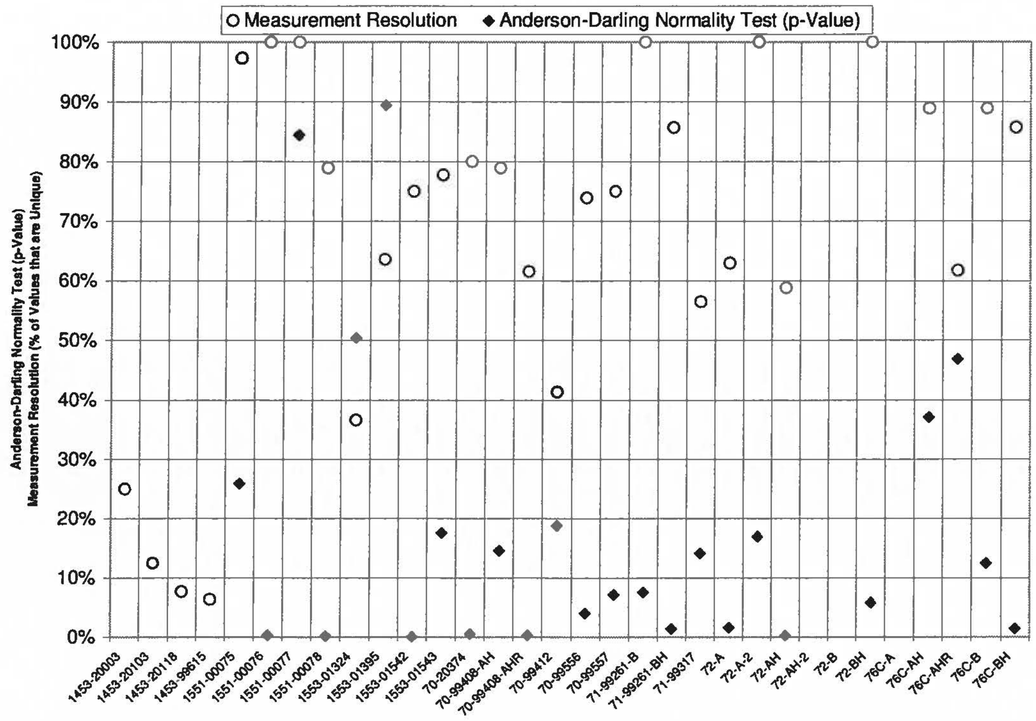
AC Gradation No. 4 Sieve (Normality Test and Measurement Resolution)



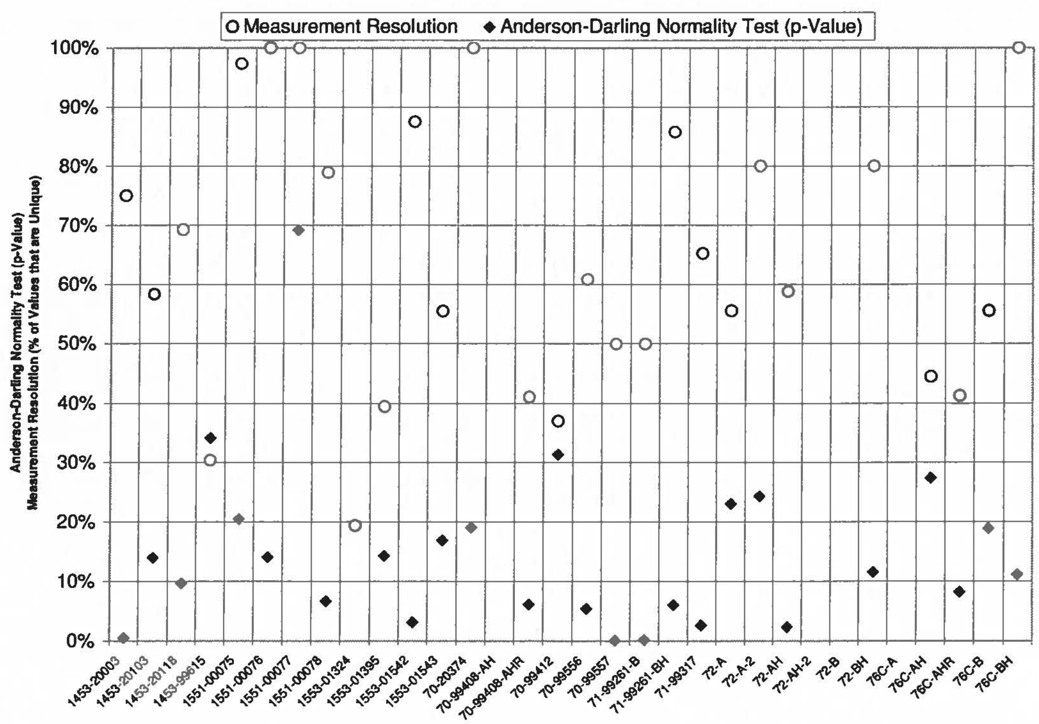
AC Gradation No. 10 Sieve (Normality Test and Measurement Resolution)



AC Gradation No. 40 Sieve (Normality Test and Measurement Resolution)



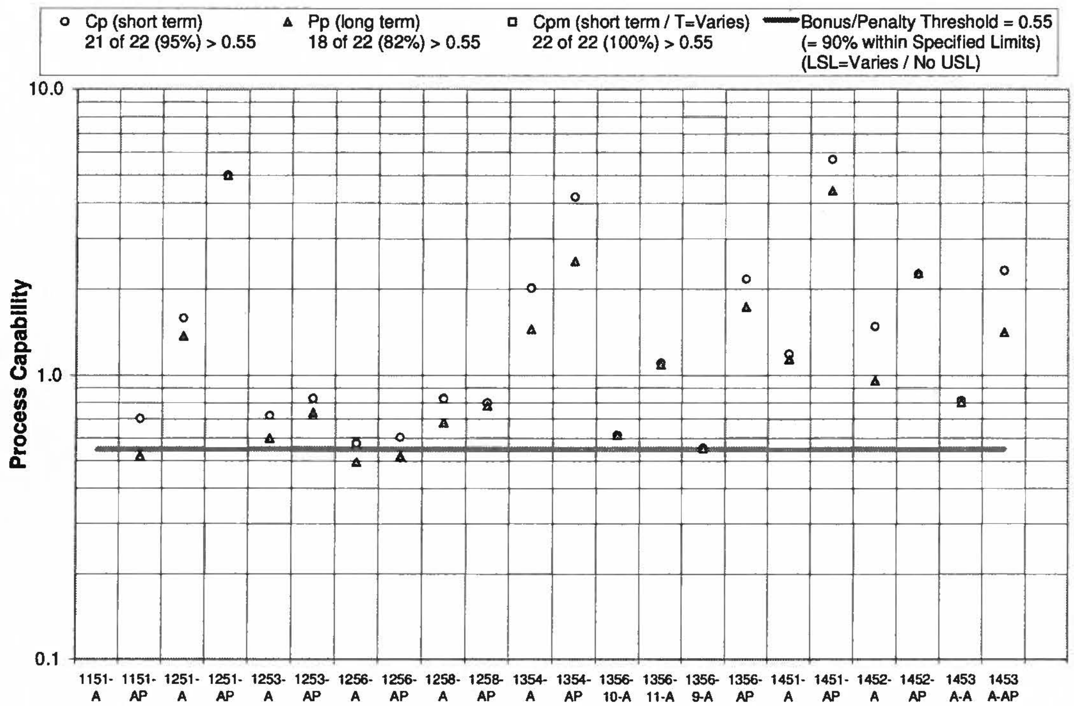
AC Gradation No. 80 Sieve (Normality Test and Measurement Resolution)



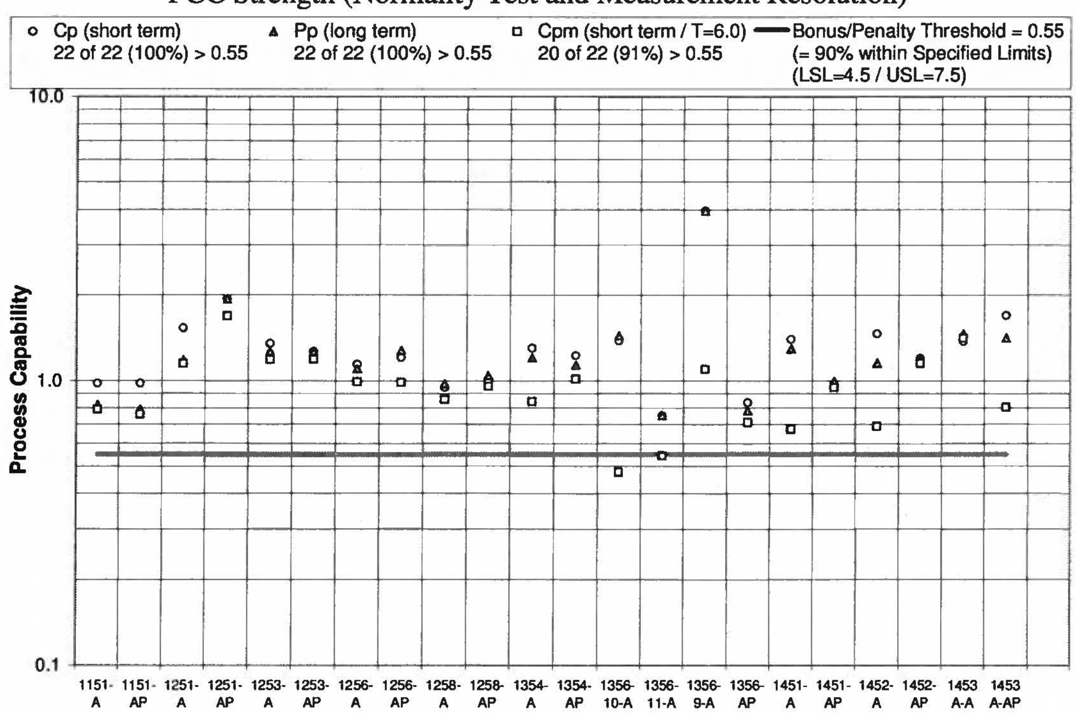
AC Gradation No. 200 Sieve (Normality Test and Measurement Resolution)

Appendix D: Process Capability Results (OTA Projects using OTA PWL Specifications)

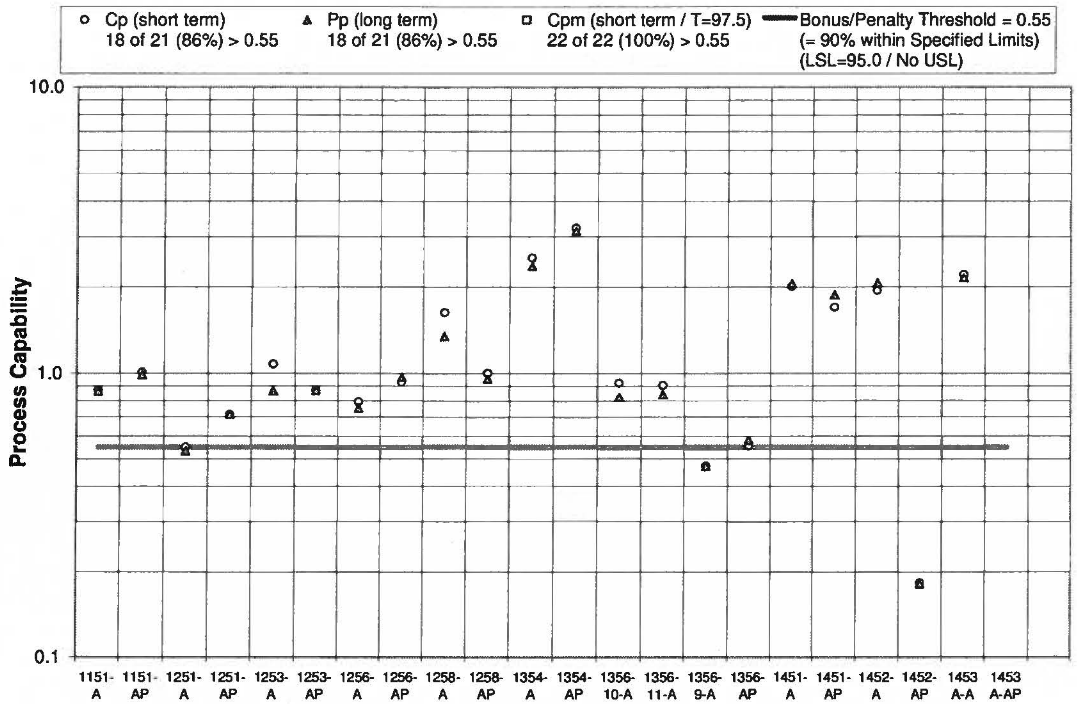
[This space intentionally left blank.]



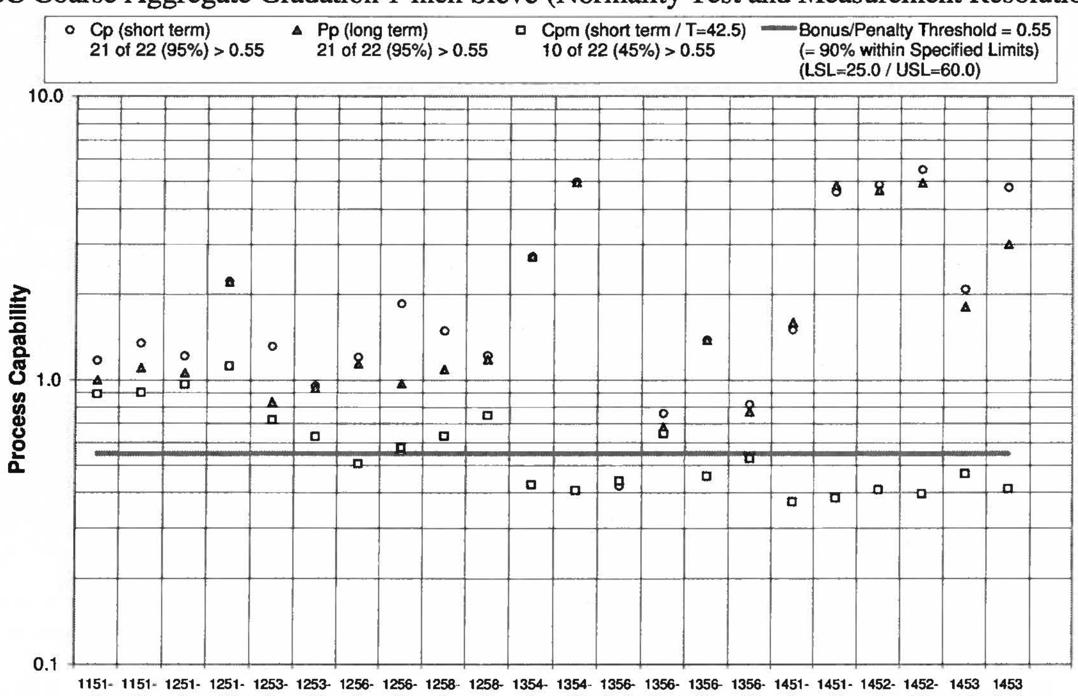
PCC Strength (Normality Test and Measurement Resolution)



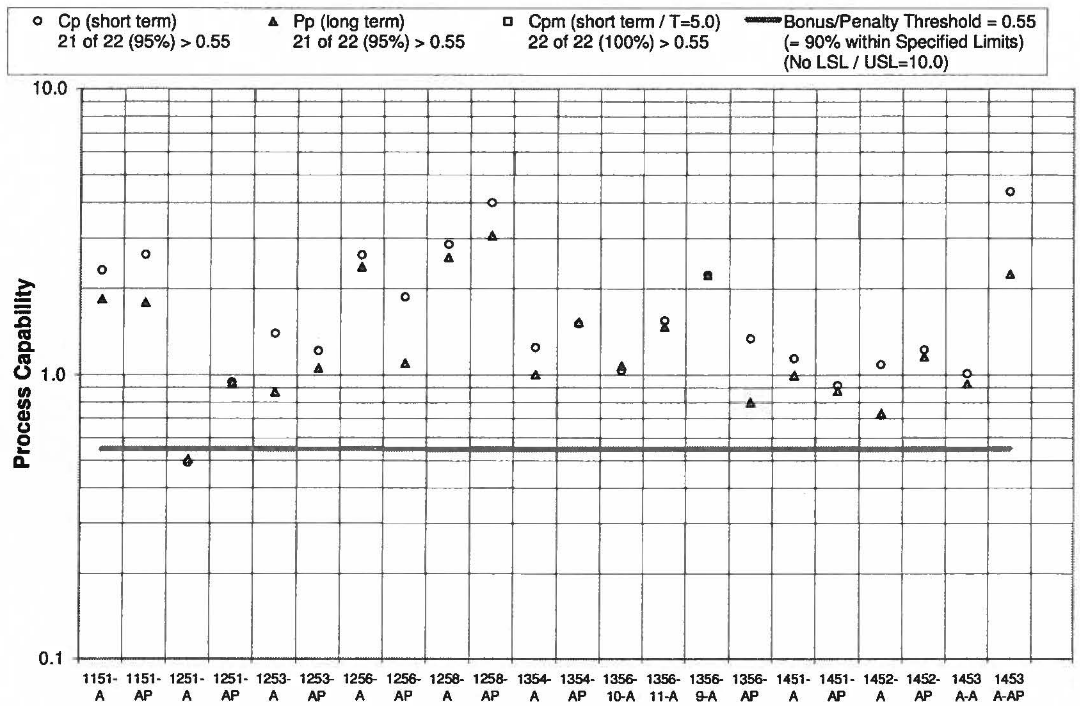
PCC Air Content (Normality Test and Measurement Resolution)



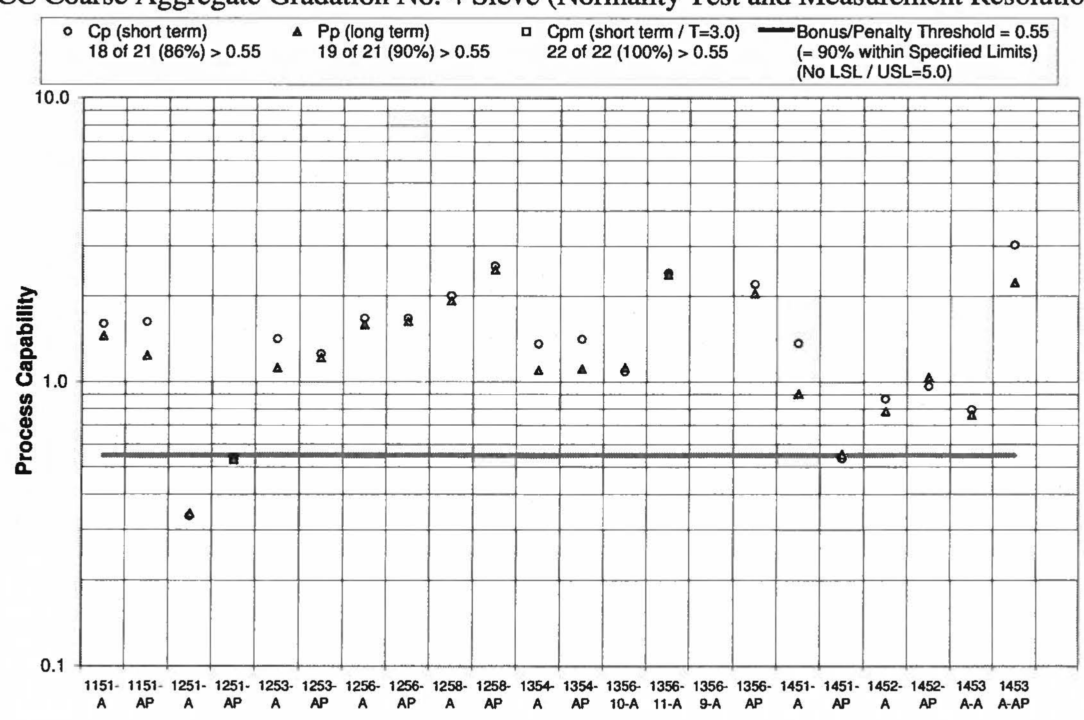
PCC Coarse Aggregate Gradation 1-inch Sieve (Normality Test and Measurement Resolution)



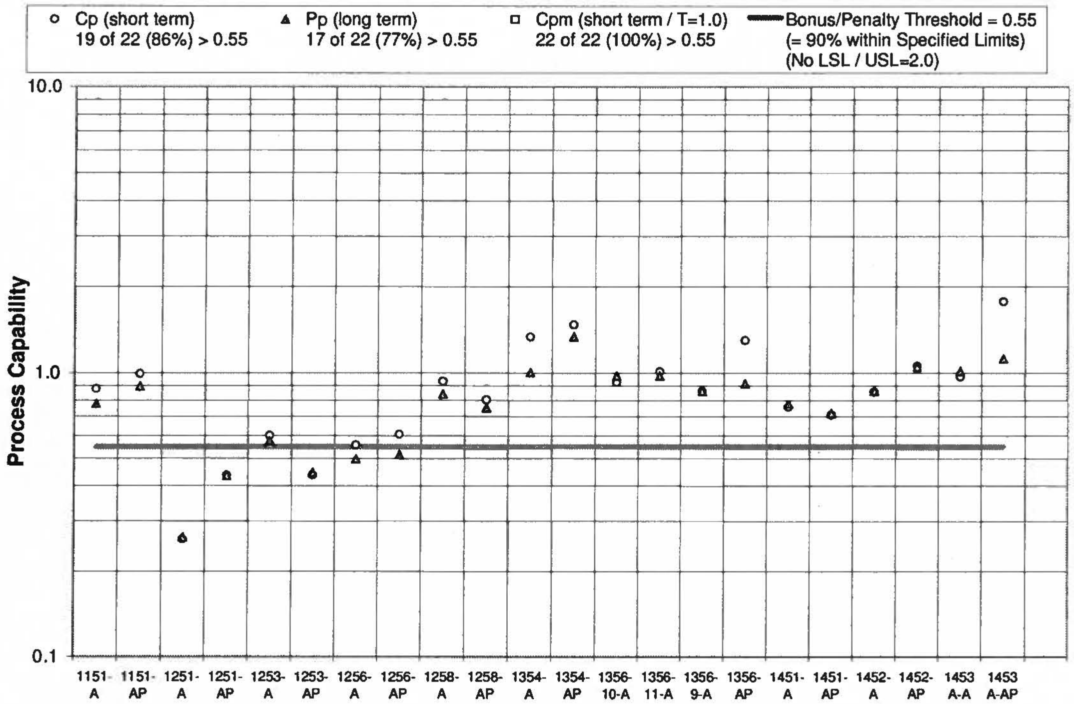
PCC Coarse Aggregate Gradation ½-inch Sieve (Normality Test and Measurement Resolution)



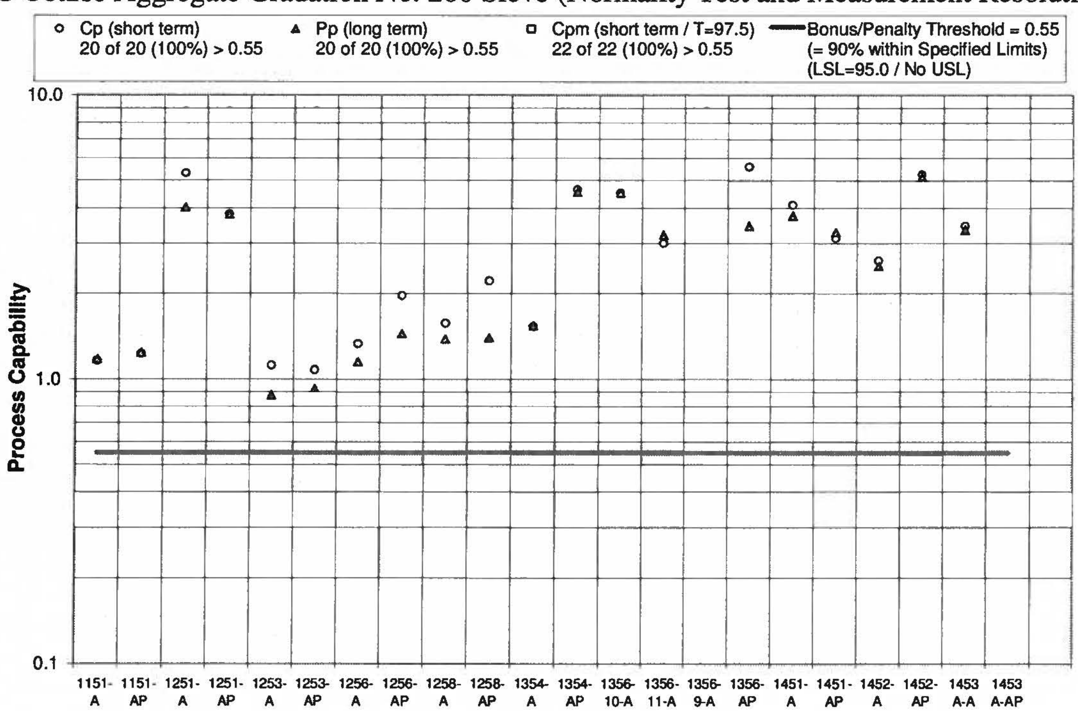
PCC Coarse Aggregate Gradation No. 4 Sieve (Normality Test and Measurement Resolution)



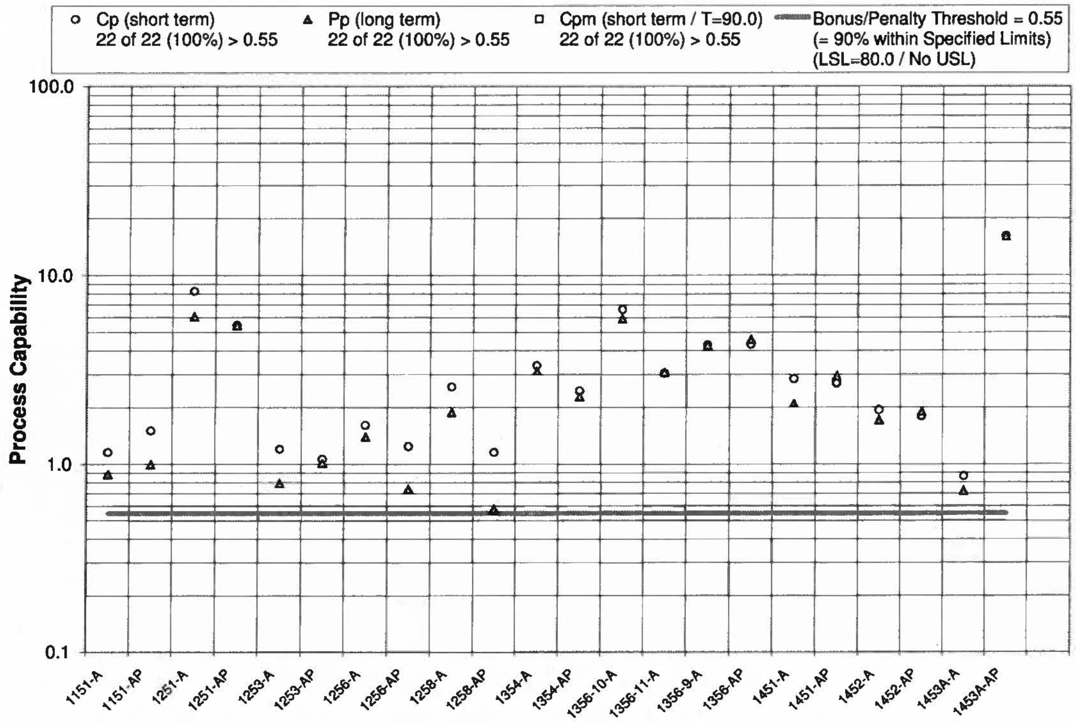
PCC Coarse Aggregate Gradation No. 8 Sieve (Normality Test and Measurement Resolution)



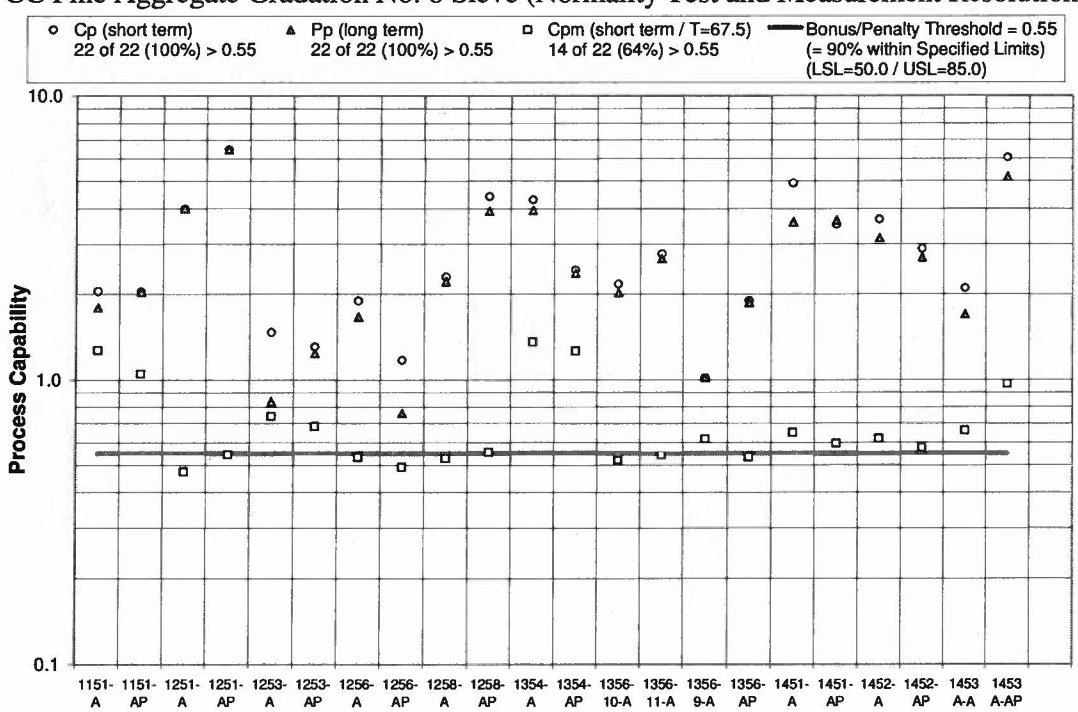
PCC Coarse Aggregate Gradation No. 200 Sieve (Normality Test and Measurement Resolution)



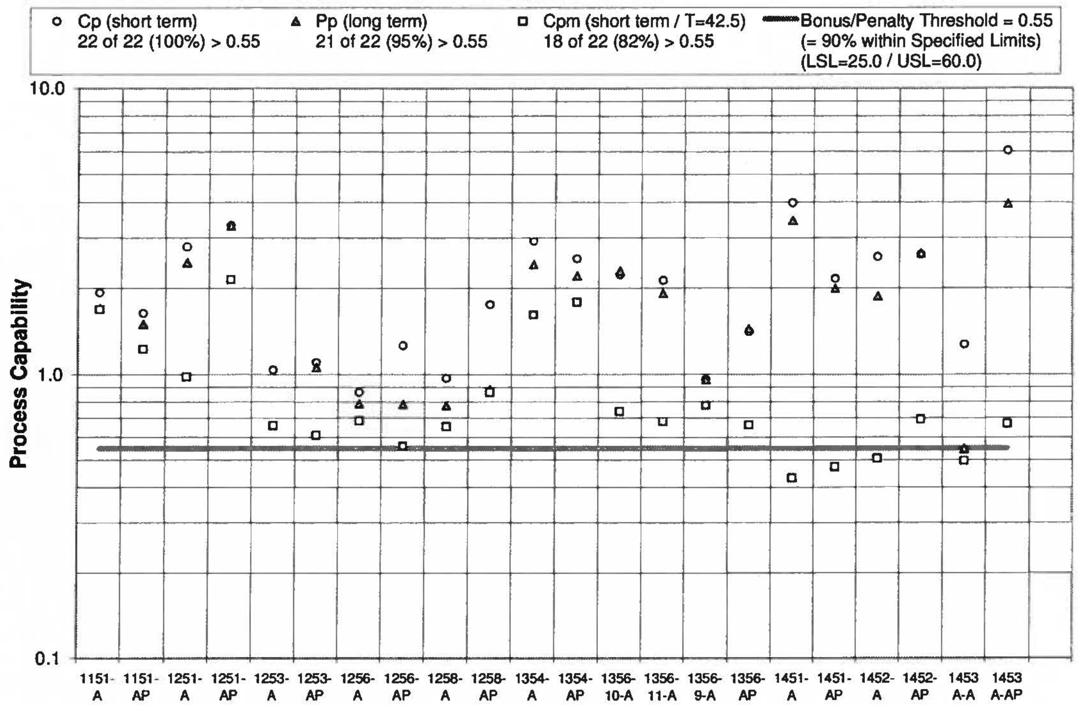
PCC Fine Aggregate Gradation No. 4 Sieve (Normality Test and Measurement Resolution)



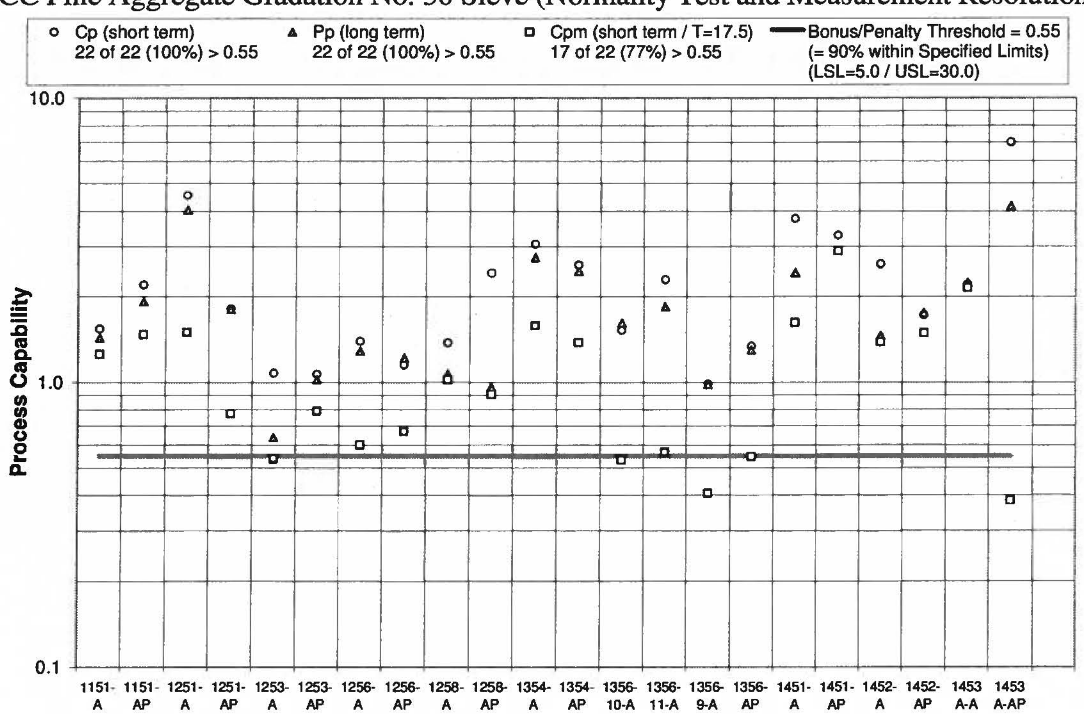
PCC Fine Aggregate Gradation No. 8 Sieve (Normality Test and Measurement Resolution)



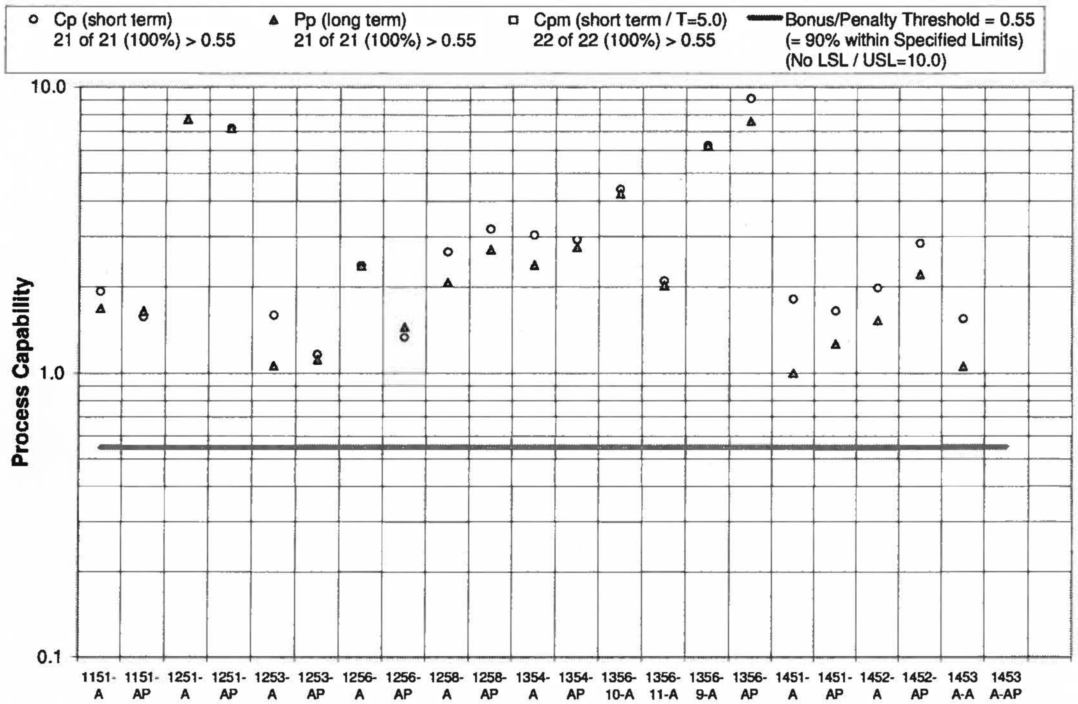
PCC Fine Aggregate Gradation No. 16 Sieve (Normality Test and Measurement Resolution)



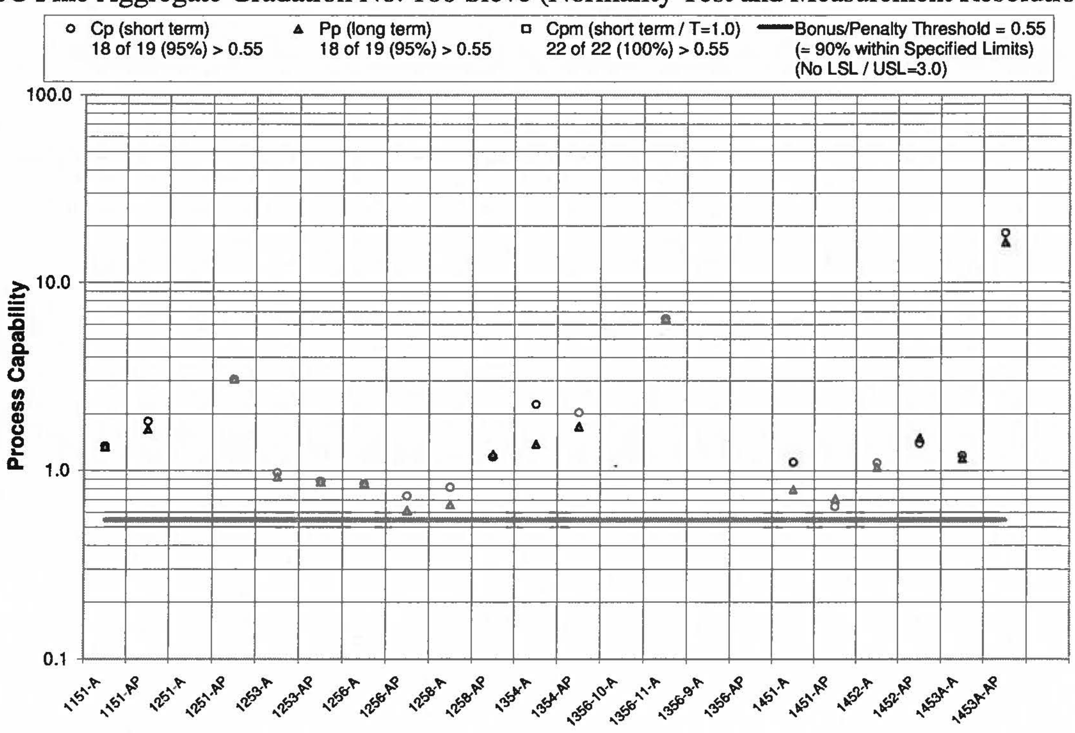
PCC Fine Aggregate Gradation No. 30 Sieve (Normality Test and Measurement Resolution)



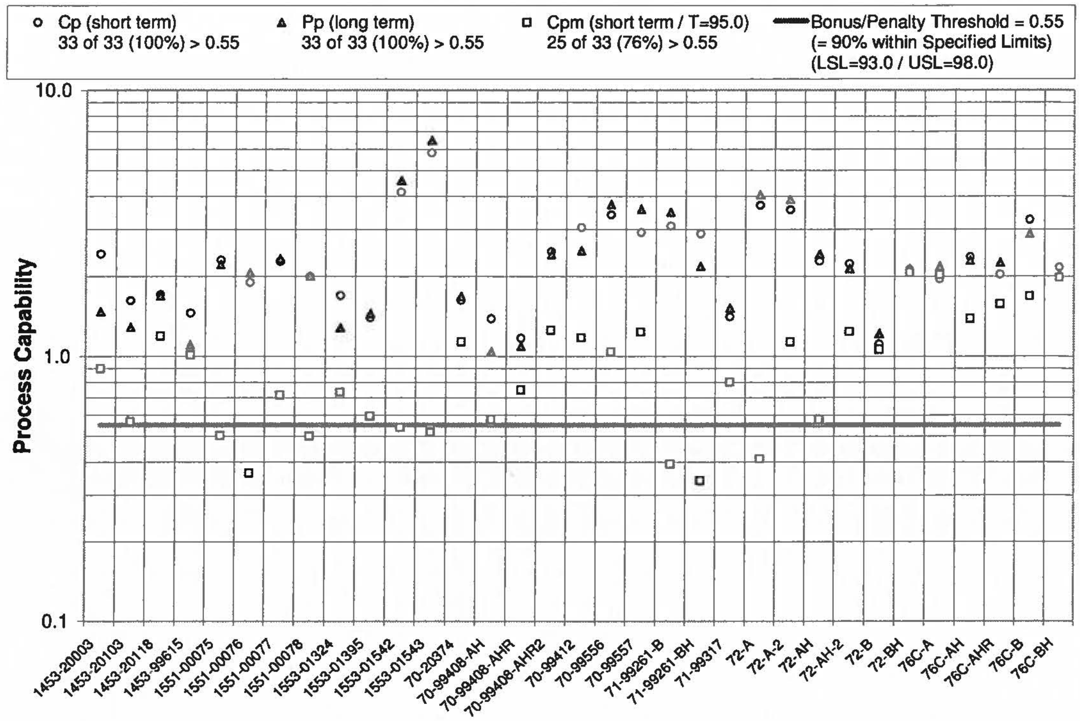
PCC Fine Aggregate Gradation No. 50 Sieve (Normality Test and Measurement Resolution)



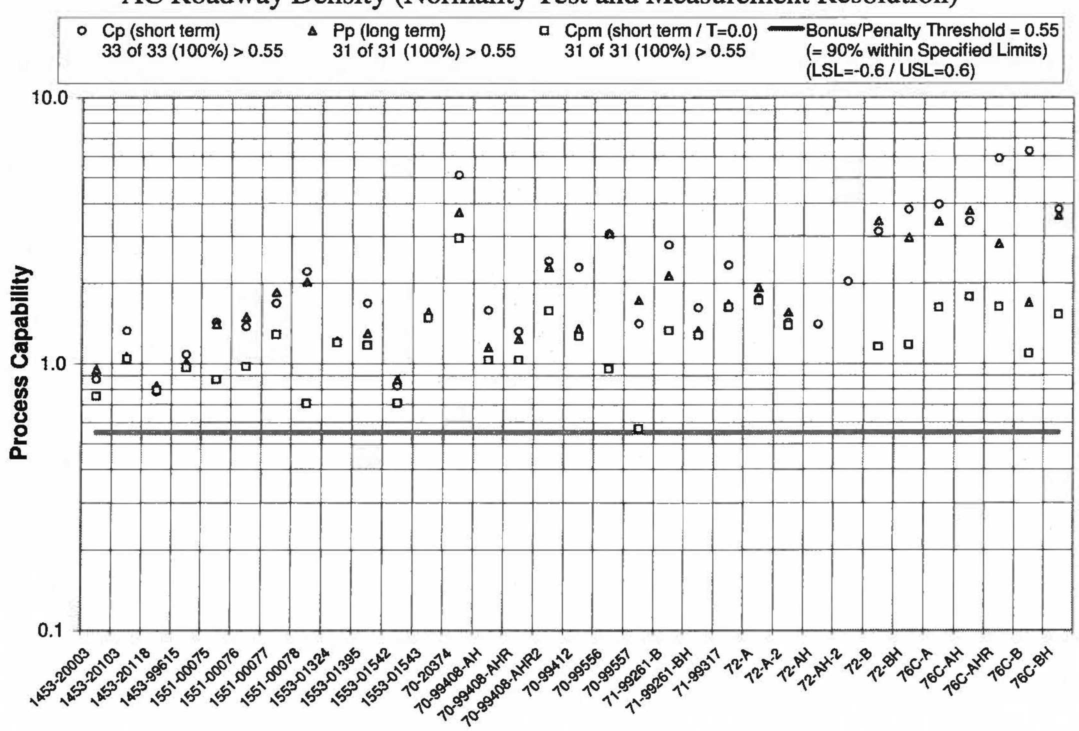
PCC Fine Aggregate Gradation No. 100 Sieve (Normality Test and Measurement Resolution)



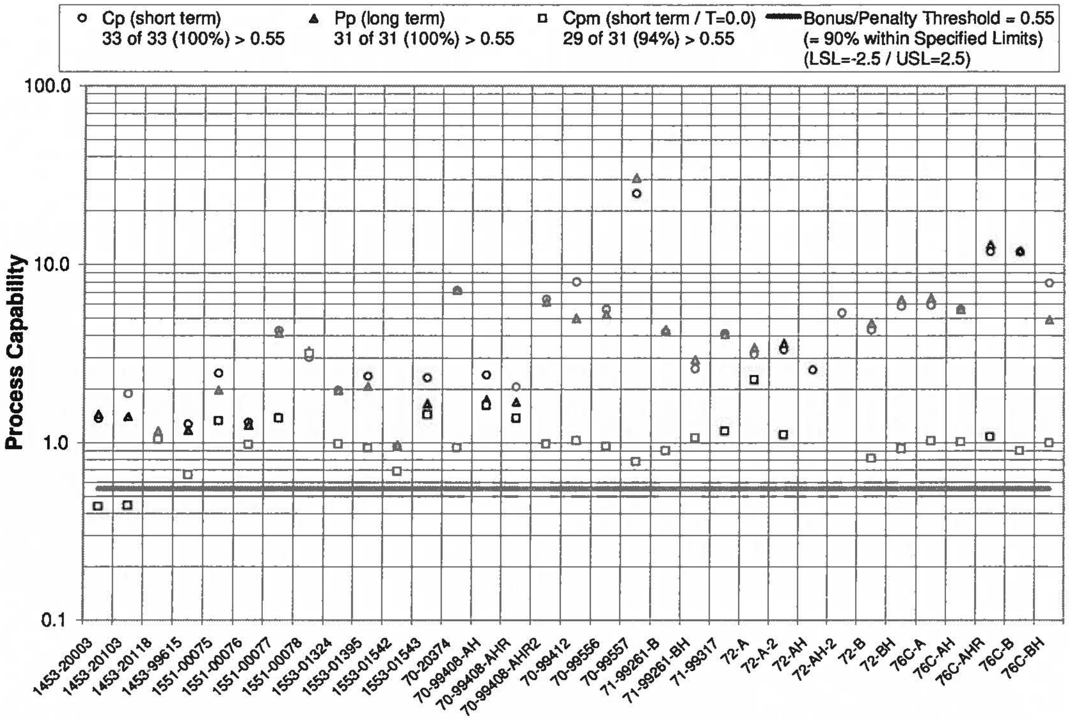
PCC Fine Aggregate Gradation No. 200 Sieve (Normality Test and Measurement Resolution)



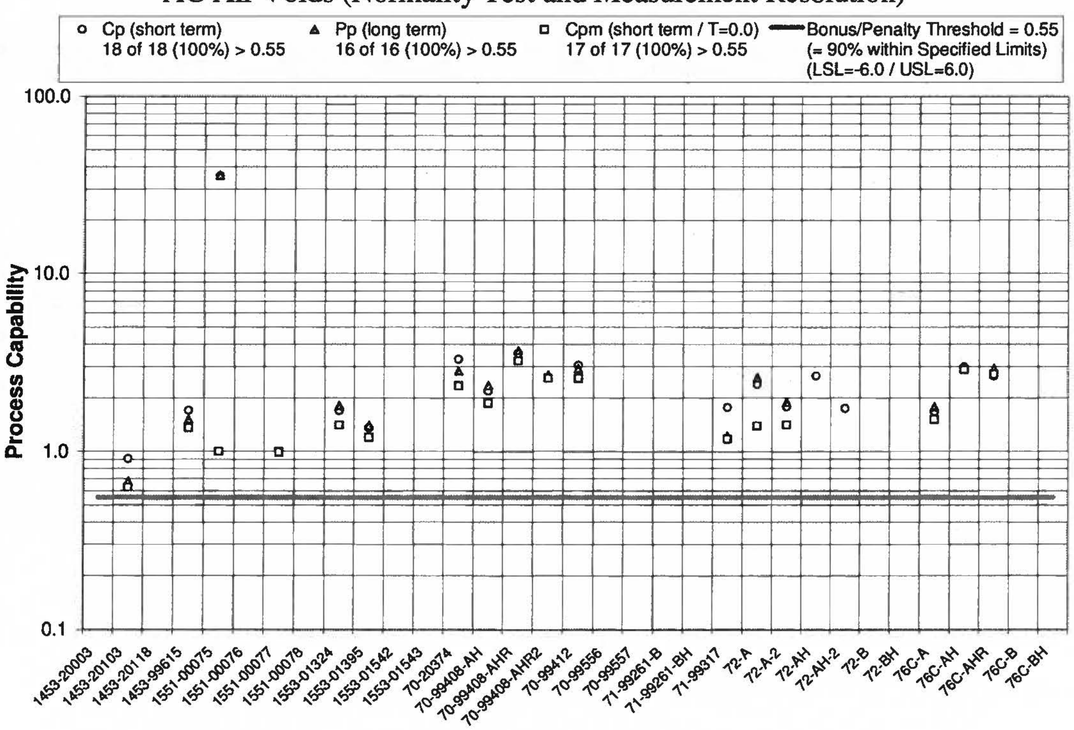
AC Roadway Density (Normality Test and Measurement Resolution)



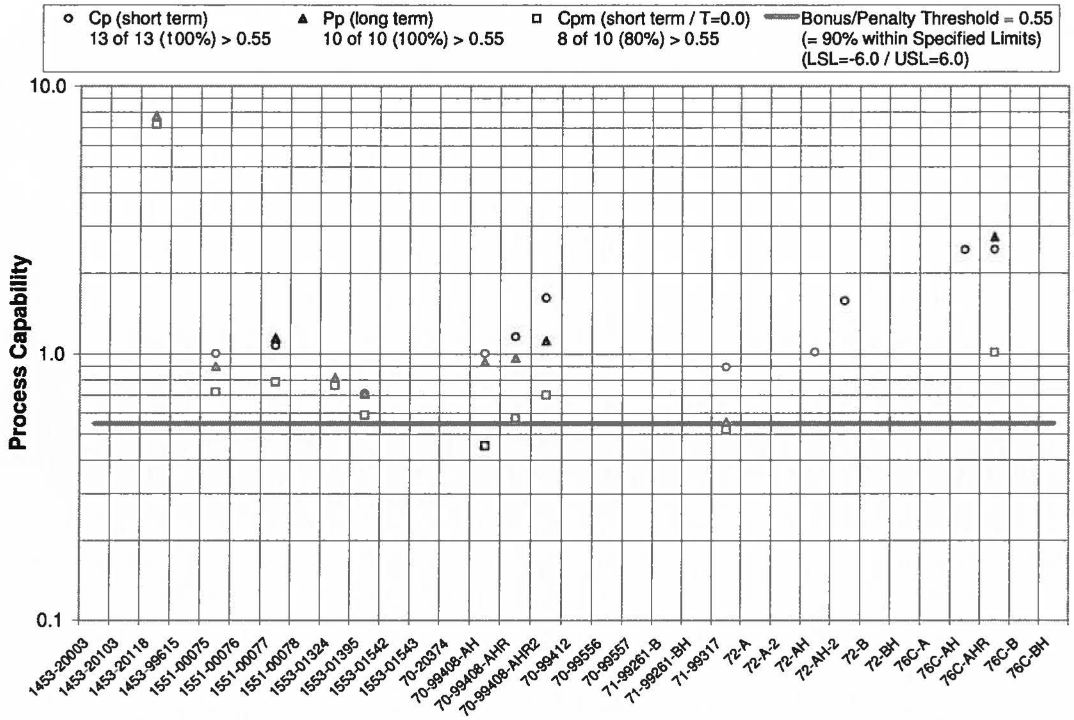
AC Asphalt Cement Content (Normality Test and Measurement Resolution)



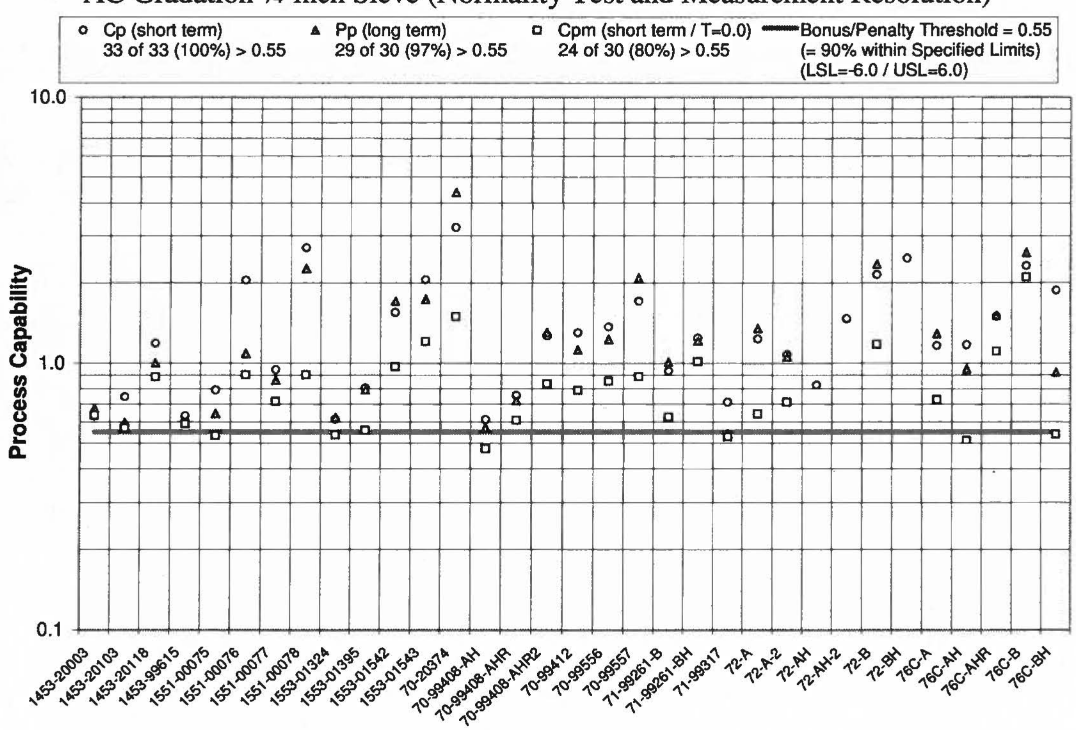
AC Air Voids (Normality Test and Measurement Resolution)



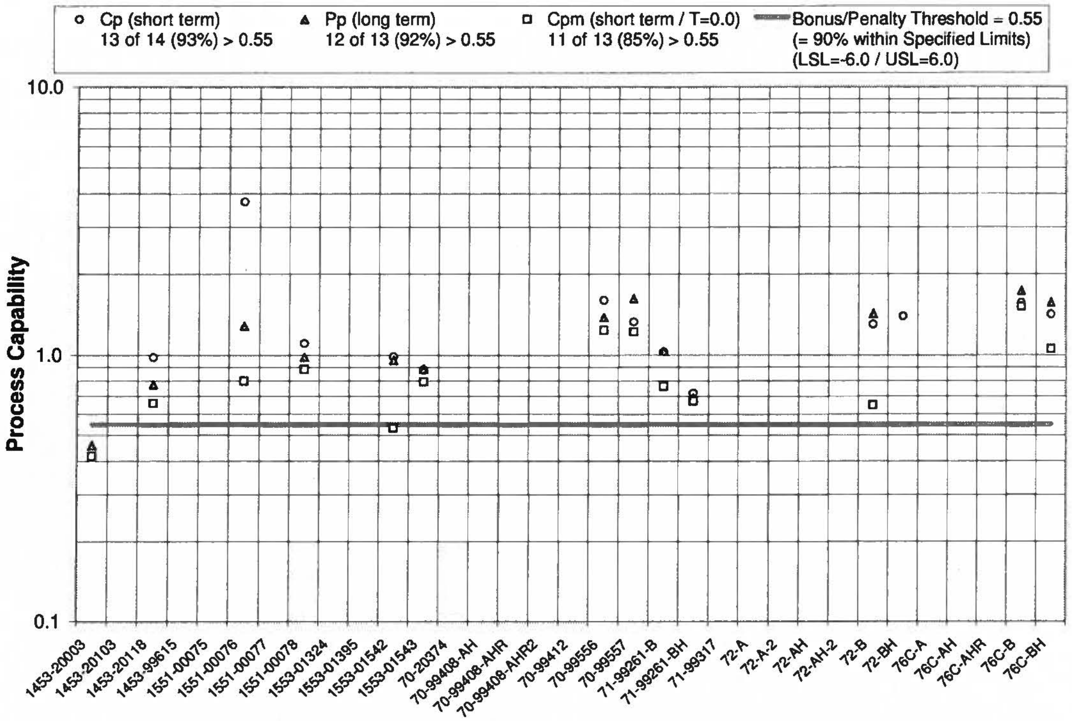
AC Gradation 1-inch Sieve (Normality Test and Measurement Resolution)



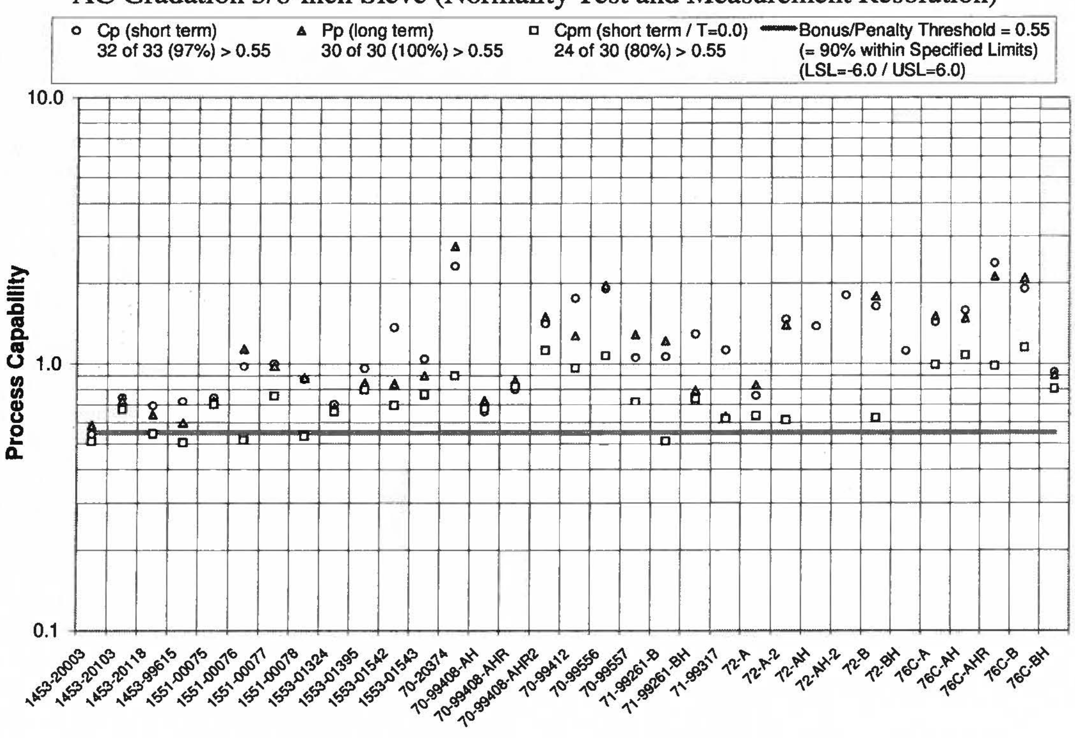
AC Gradation ¾-inch Sieve (Normality Test and Measurement Resolution)



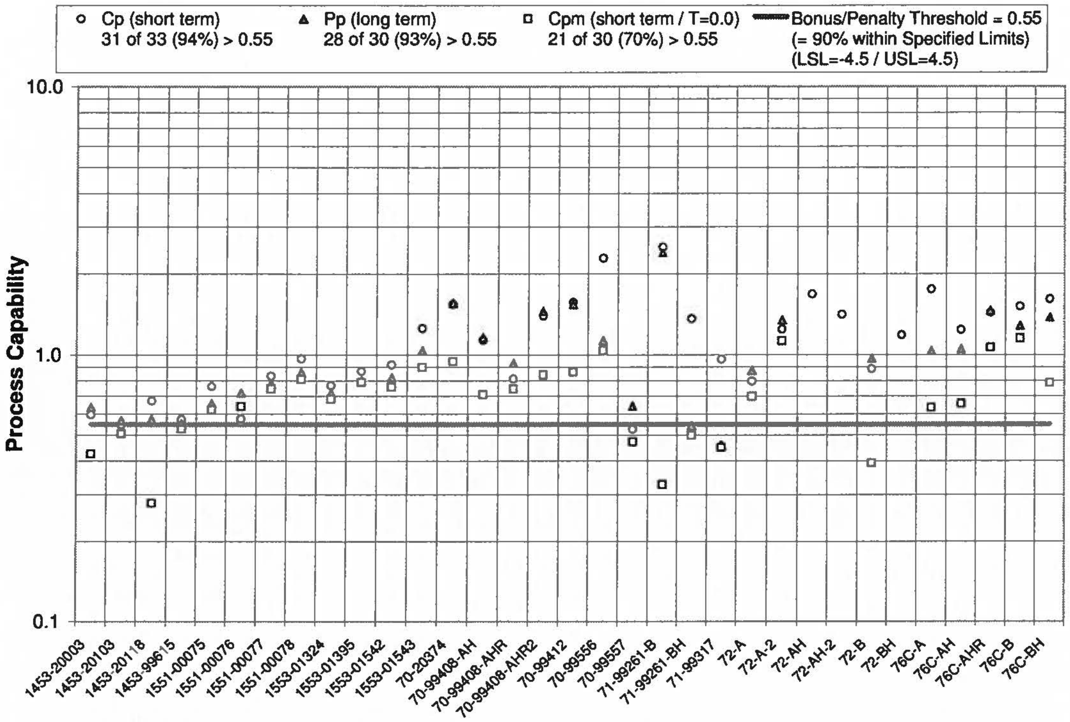
AC Gradation ½-inch Sieve (Normality Test and Measurement Resolution)



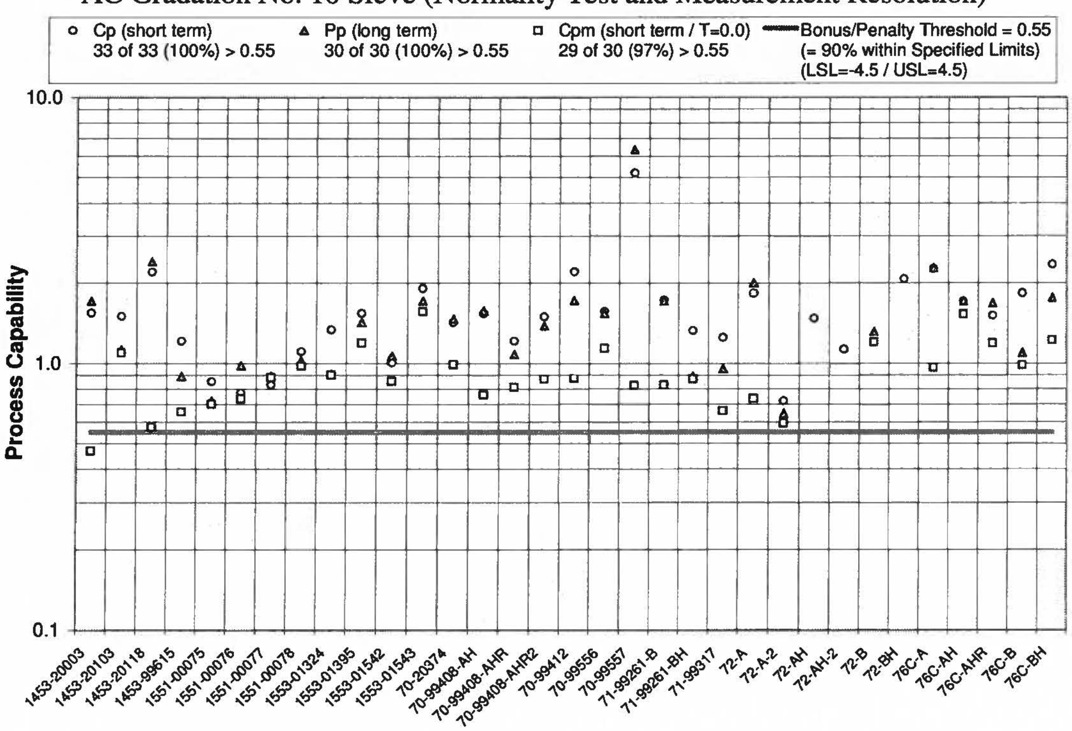
AC Gradation 3/8-inch Sieve (Normality Test and Measurement Resolution)



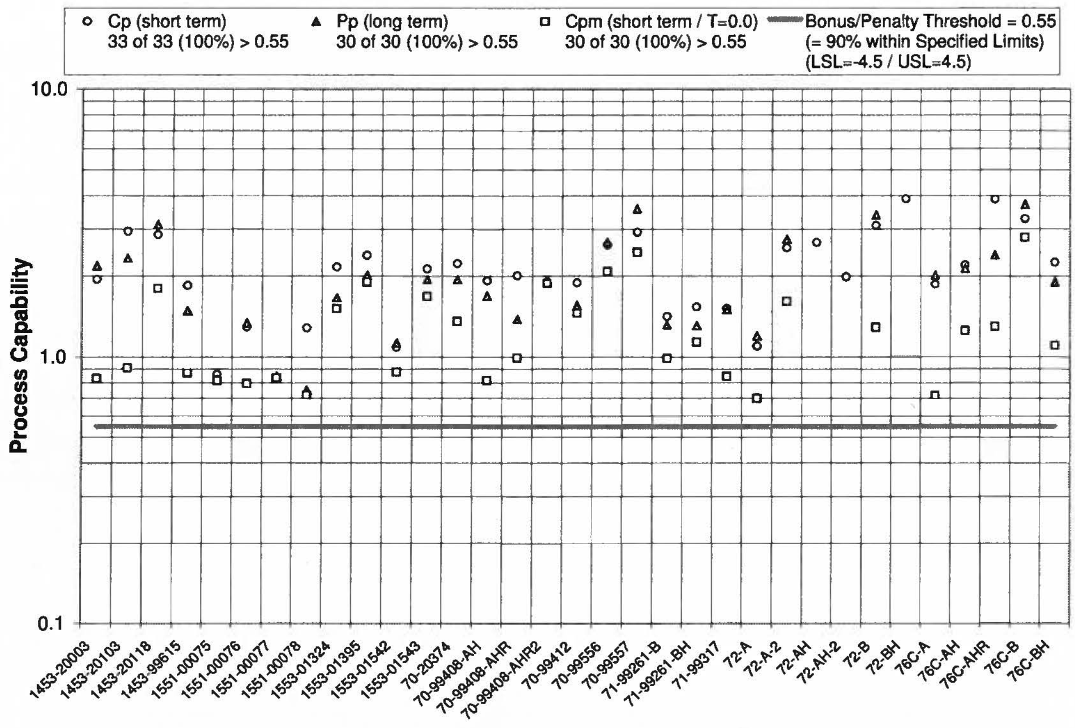
AC Gradation No. 4 Sieve (Normality Test and Measurement Resolution)



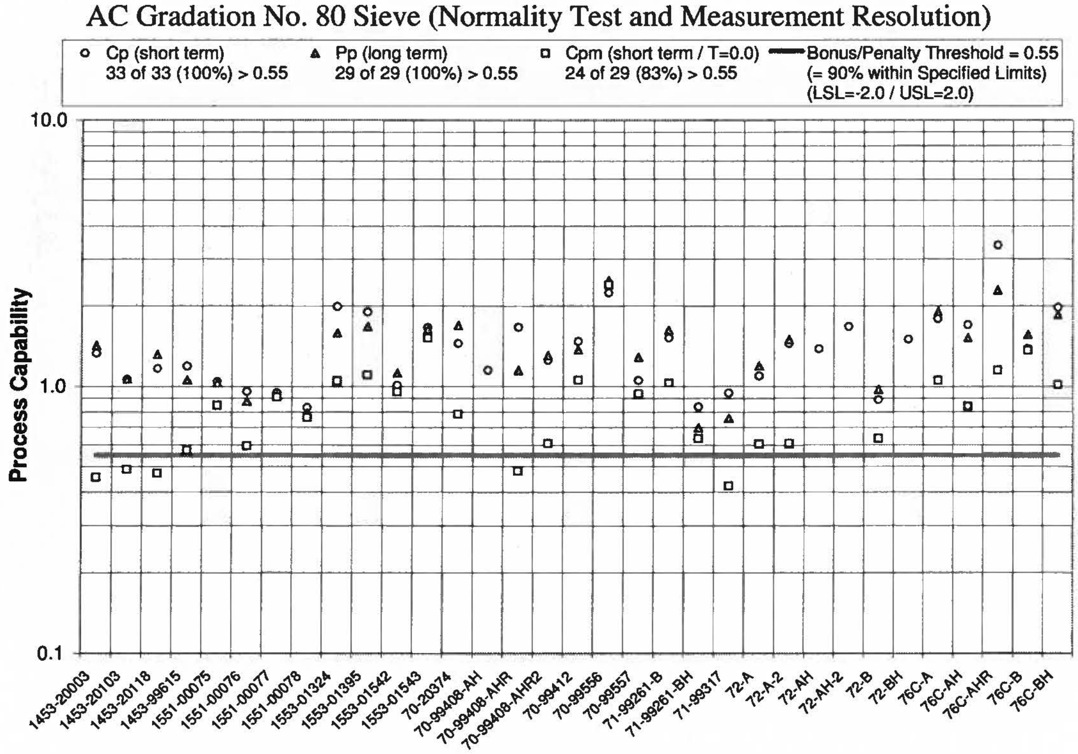
AC Gradation No. 10 Sieve (Normality Test and Measurement Resolution)



AC Gradation No. 40 Sieve (Normality Test and Measurement Resolution)



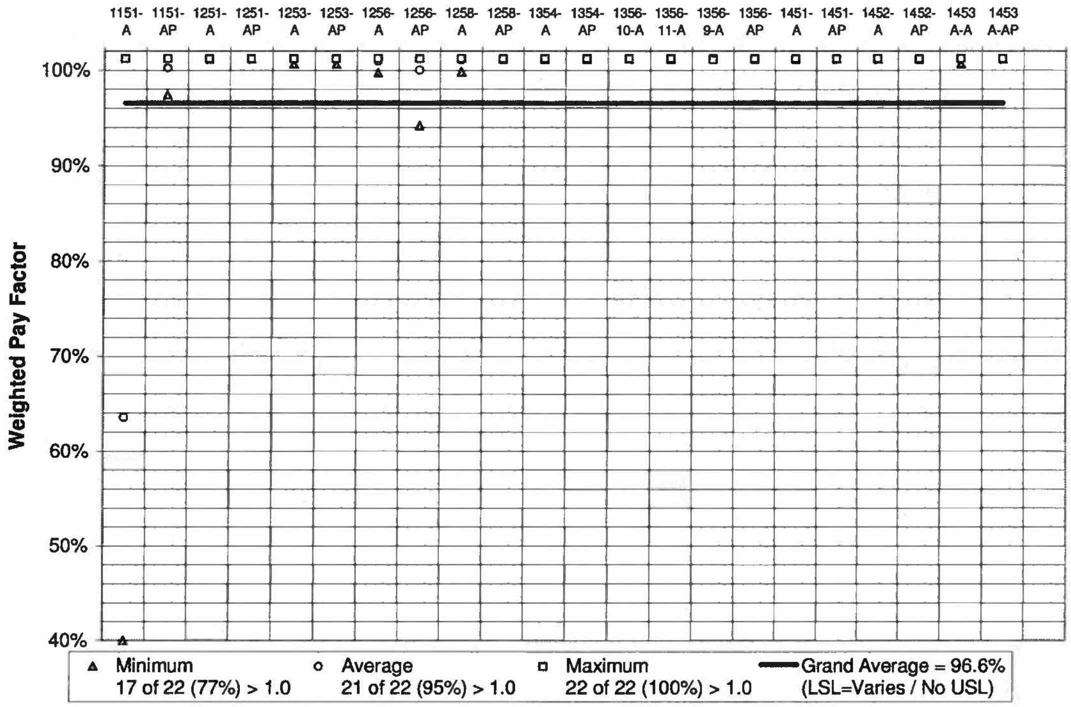
AC Gradation No. 80 Sieve (Normality Test and Measurement Resolution)



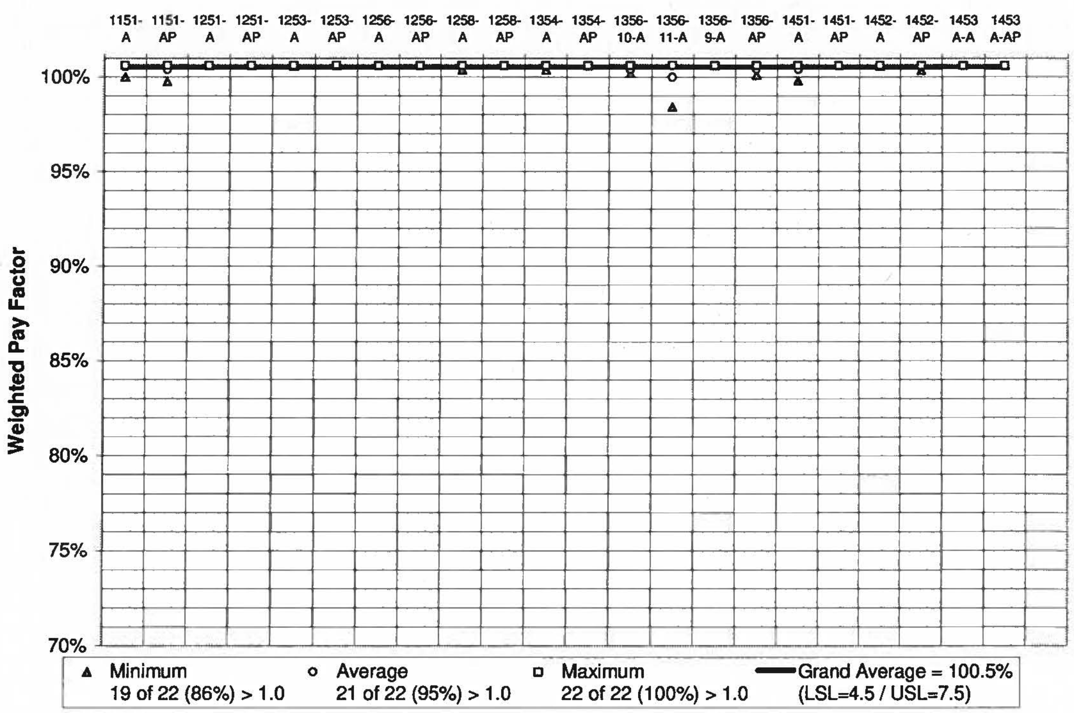
AC Gradation No. 200 Sieve (Normality Test and Measurement Resolution)

Appendix E: Weighted Pay Factors (OTA Projects using OTA PWL Specifications)

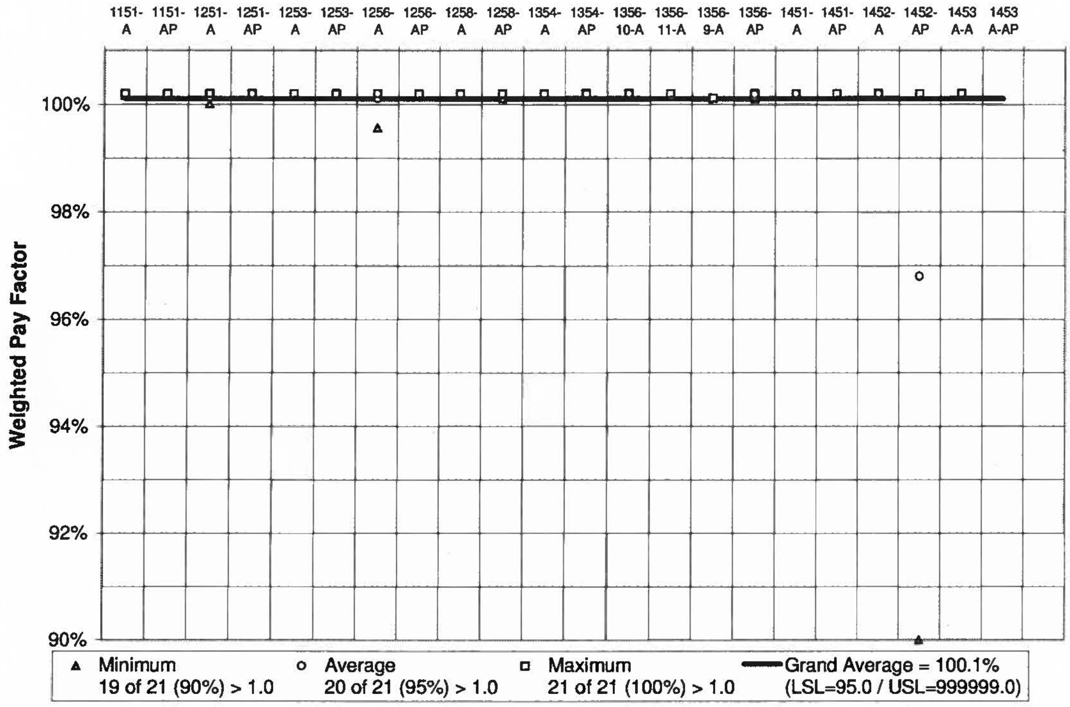
[This space intentionally left blank.]



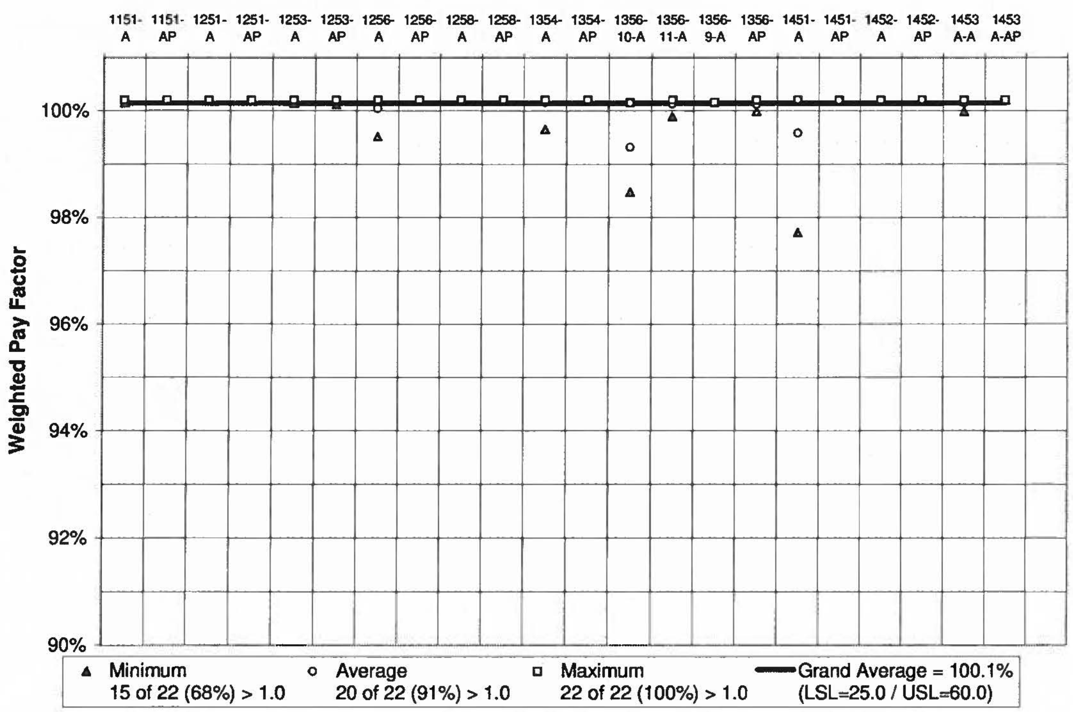
PCC Strength (Weighted Pay Factors)



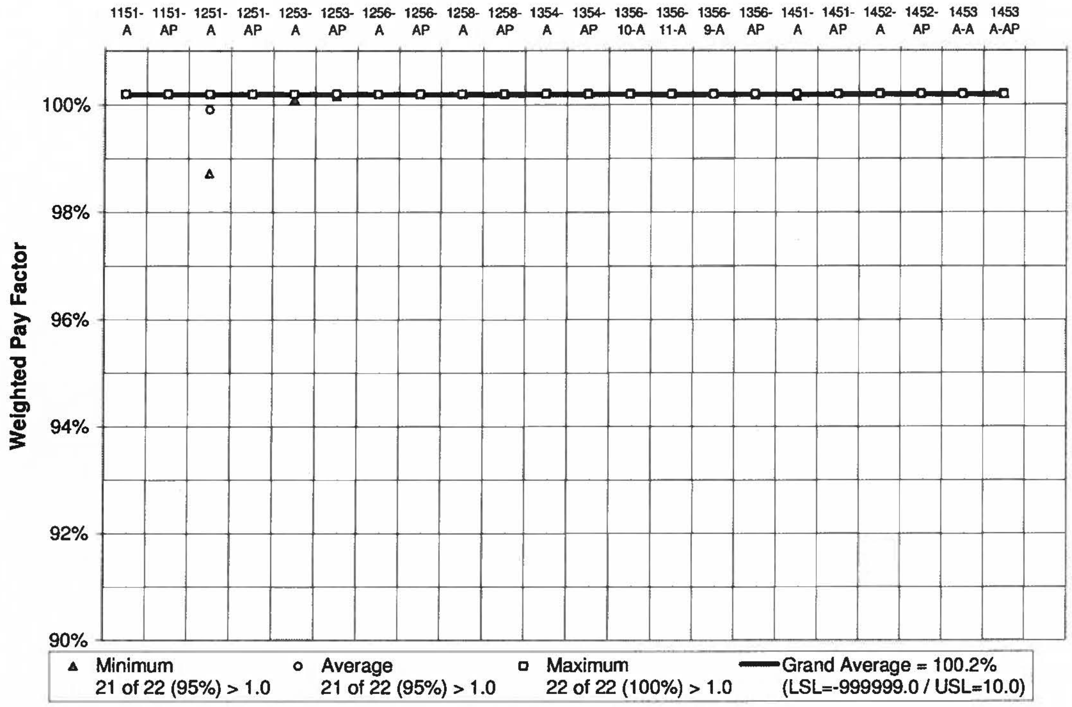
PCC Air Content (Weighted Pay Factors)



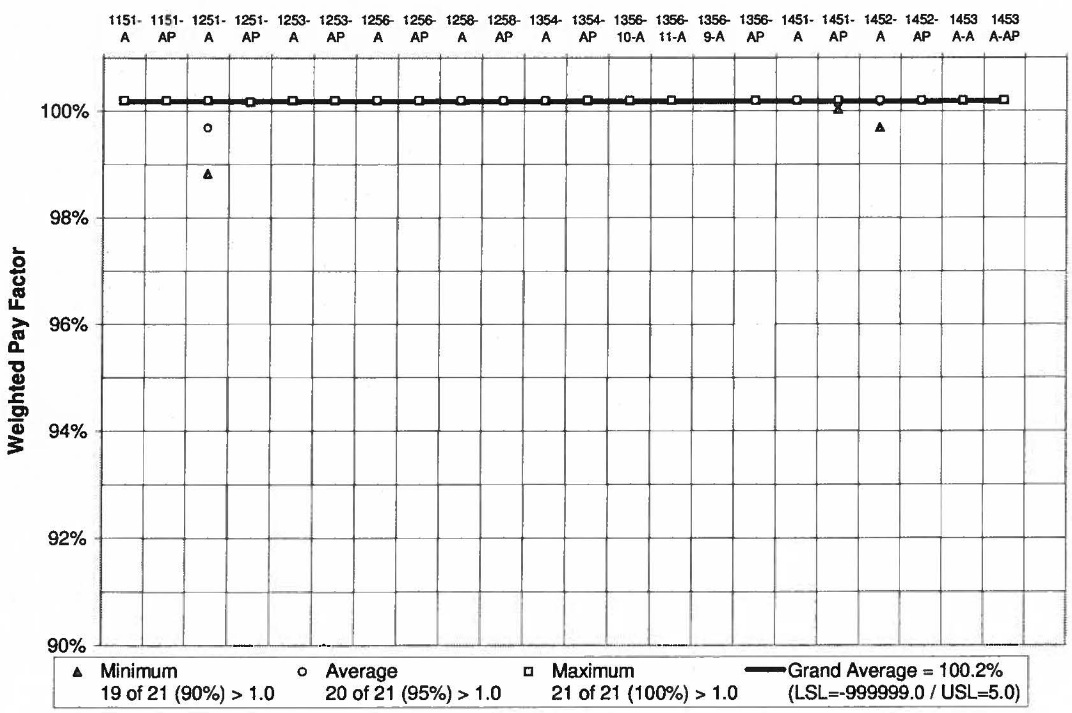
PCC Coarse Aggregate Gradation 1-inch Sieve (Weighted Pay Factors)



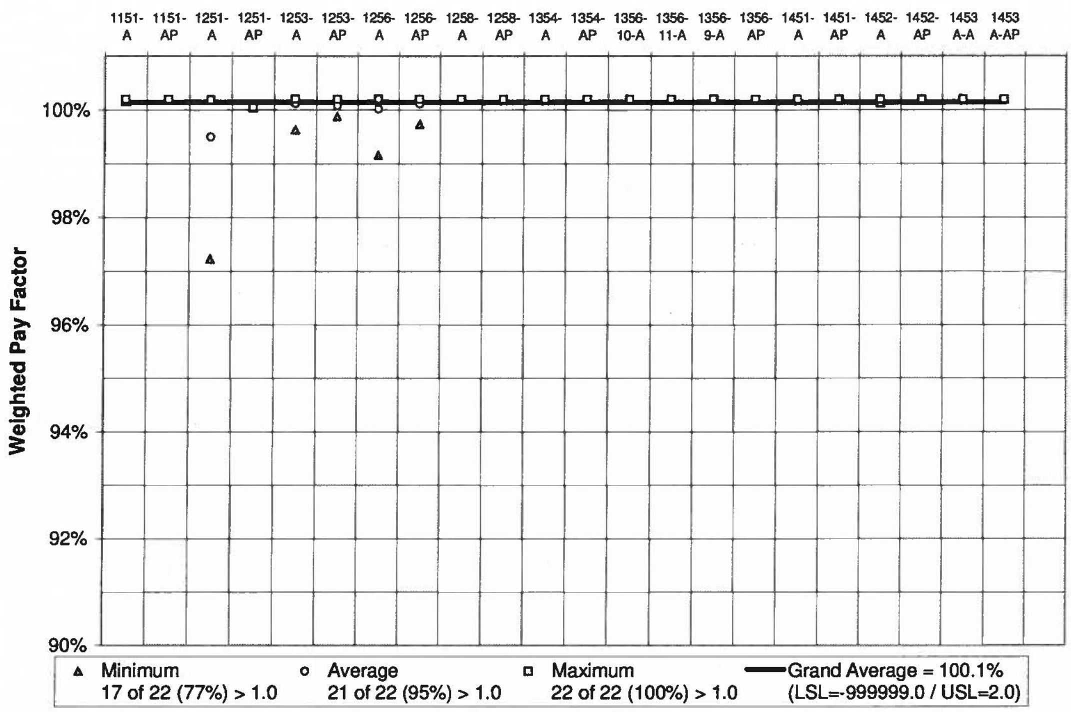
PCC Coarse Aggregate Gradation ½-inch Sieve (Weighted Pay Factors)



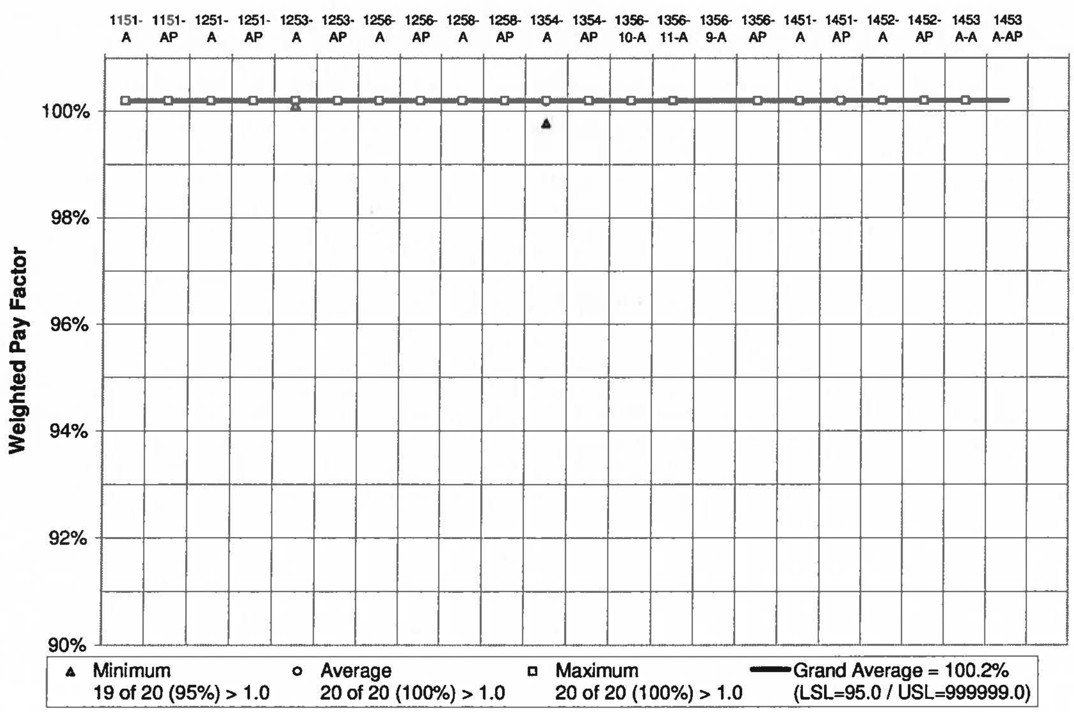
PCC Coarse Aggregate Gradation No. 4 Sieve (Weighted Pay Factors)



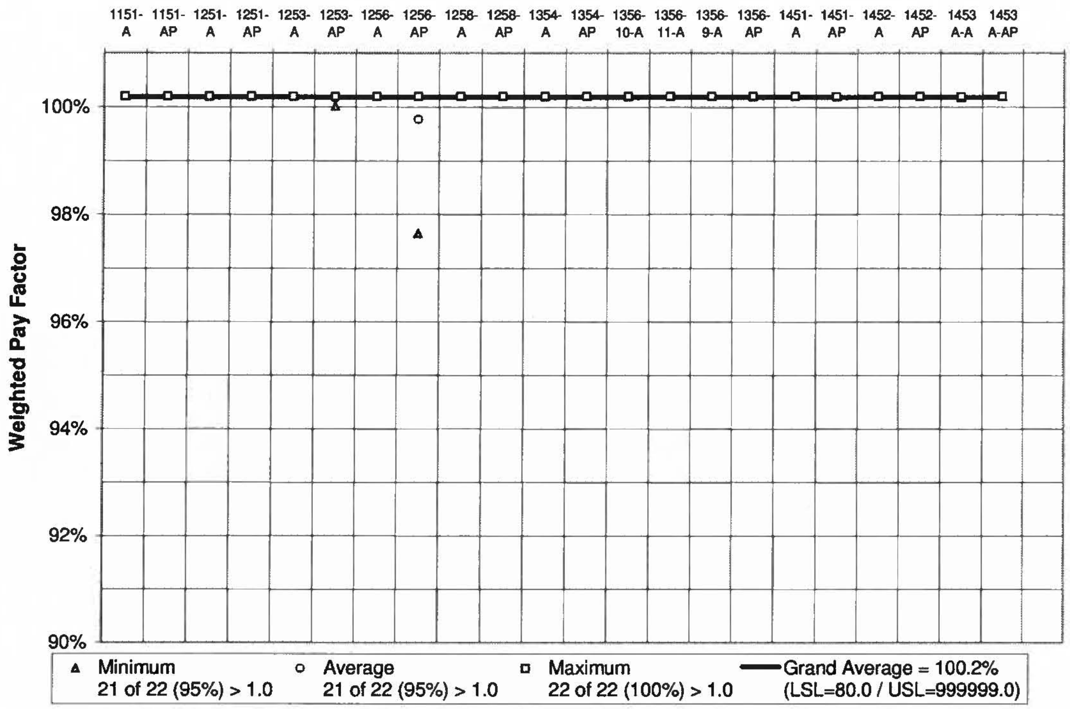
PCC Coarse Aggregate Gradation No. 8 Sieve (Weighted Pay Factors)



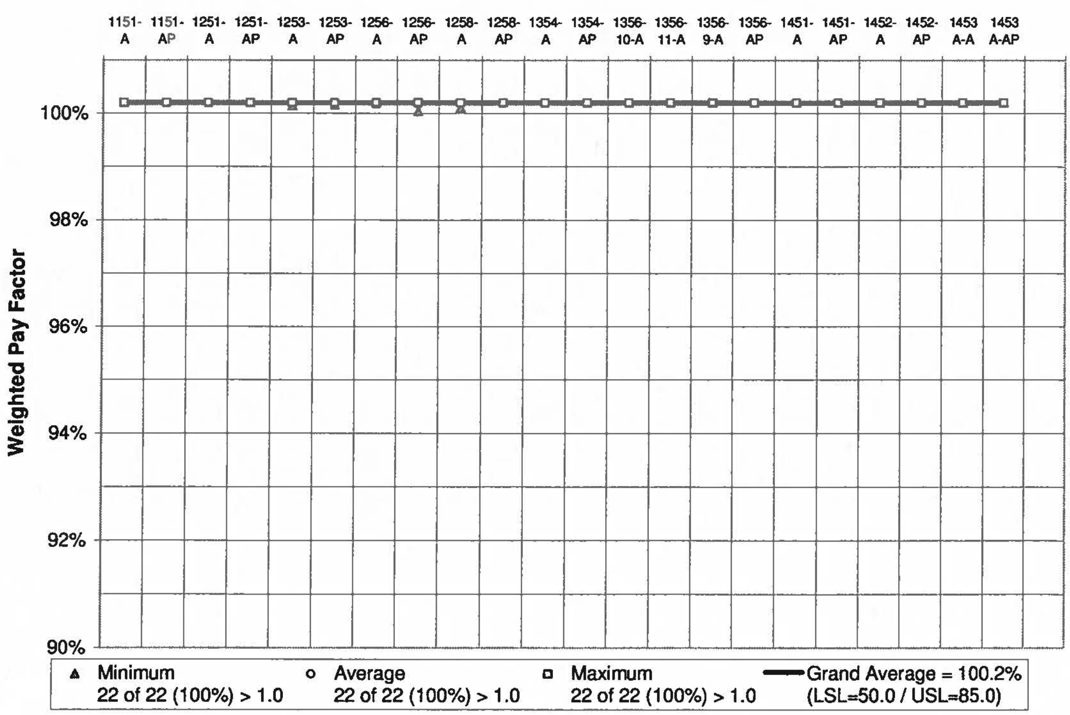
PCC Coarse Aggregate Gradation No. 200 Sieve (Weighted Pay Factors)



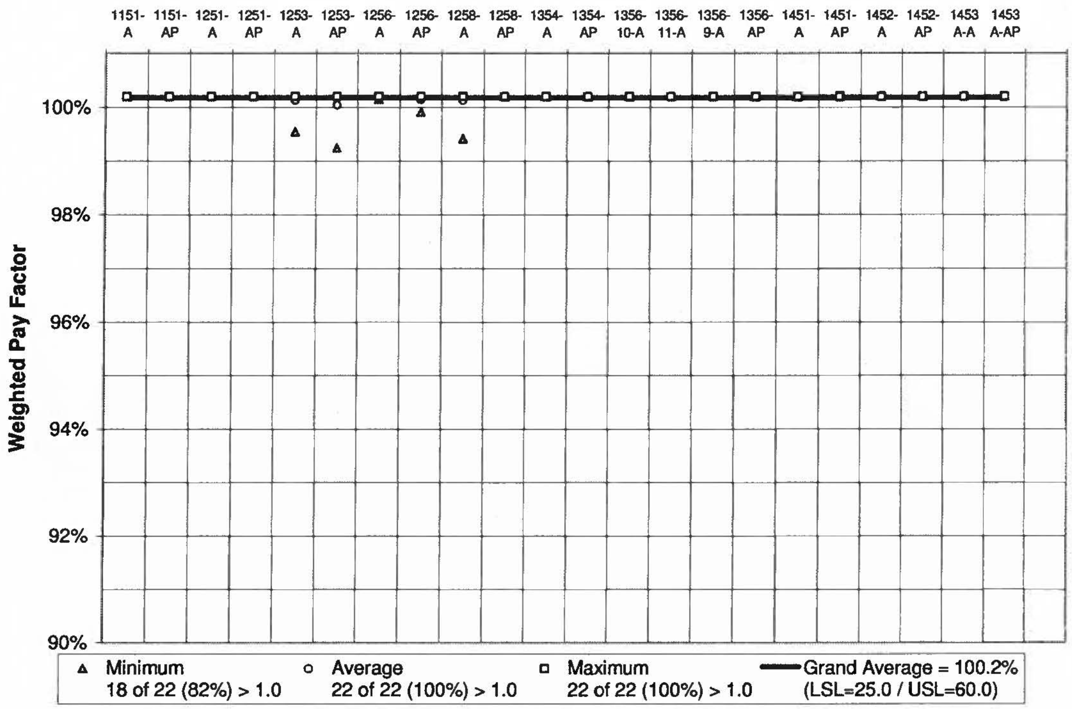
PCC Fine Aggregate Gradation No. 4 Sieve (Weighted Pay Factors)



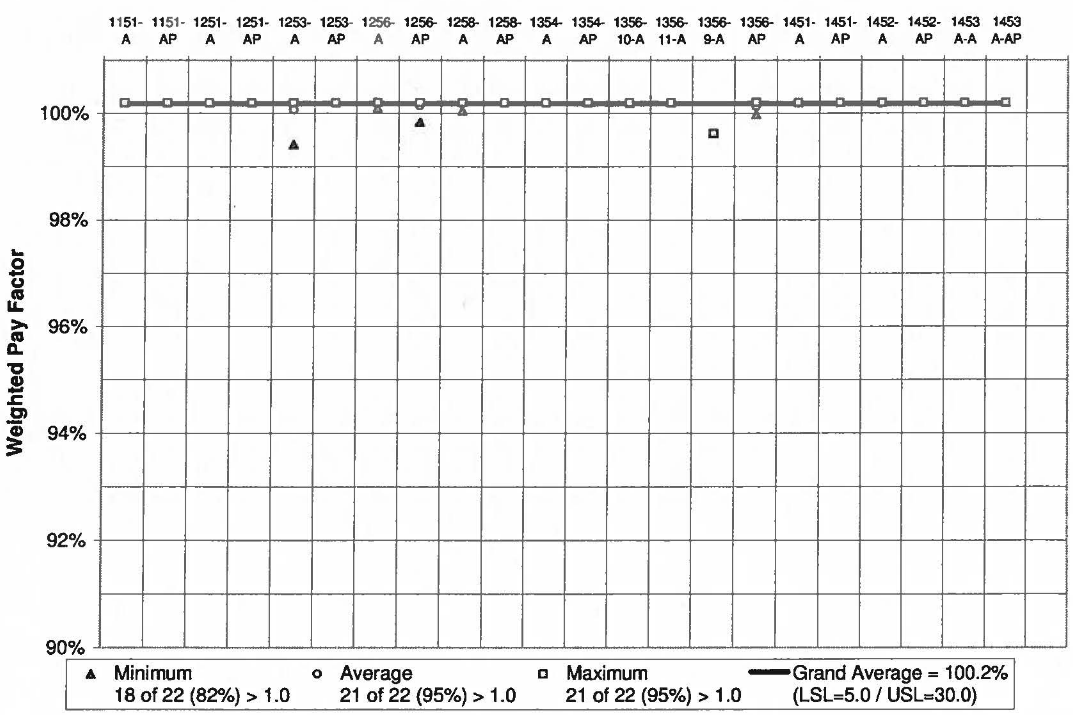
PCC Fine Aggregate Gradation No. 8 Sieve (Weighted Pay Factors)



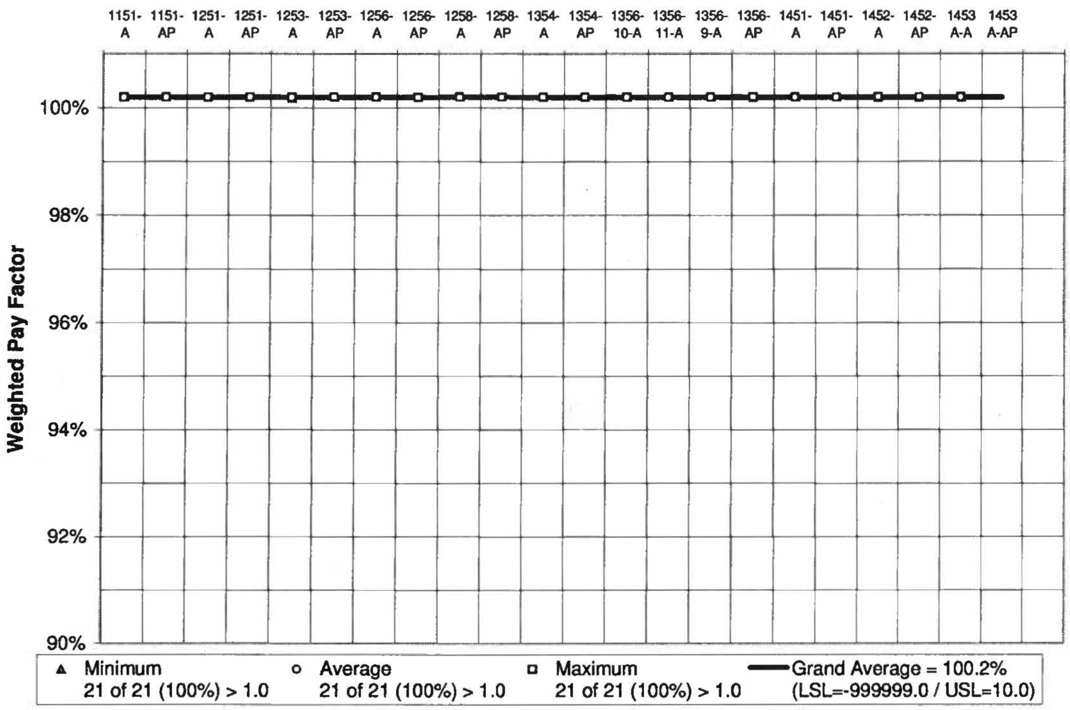
PCC Fine Aggregate Gradation No. 16 Sieve (Weighted Pay Factors)



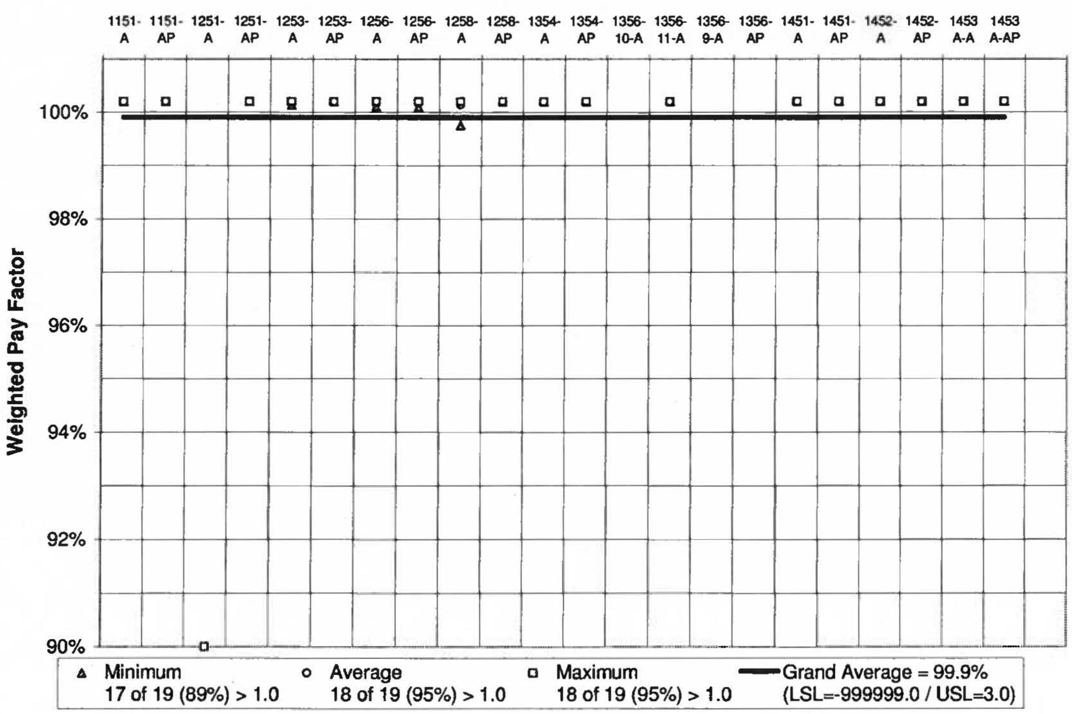
PCC Fine Aggregate Gradation No. 30 Sieve (Weighted Pay Factors)



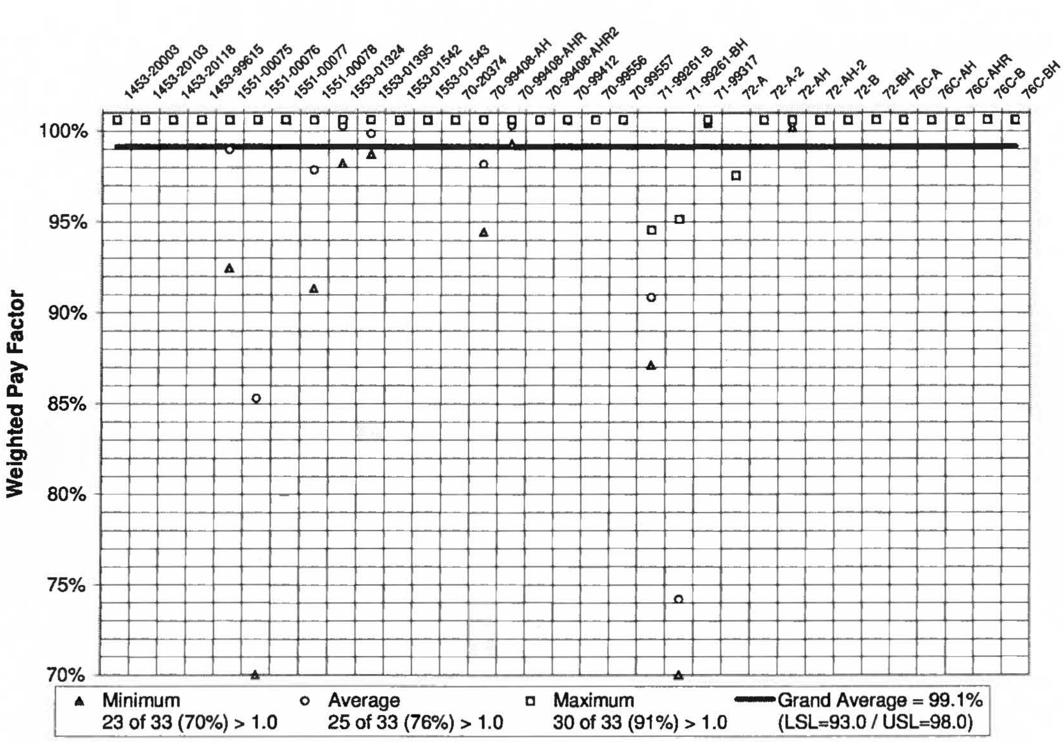
PCC Fine Aggregate Gradation No. 50 Sieve (Weighted Pay Factors)



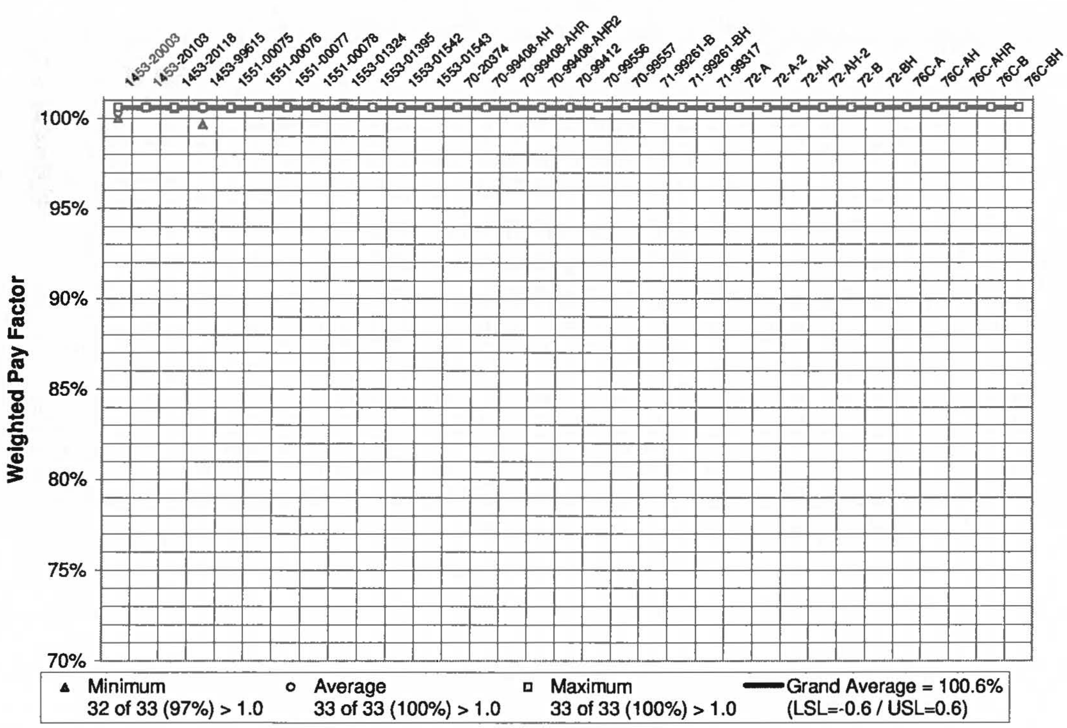
PCC Fine Aggregate Gradation No. 100 Sieve (Weighted Pay Factors)



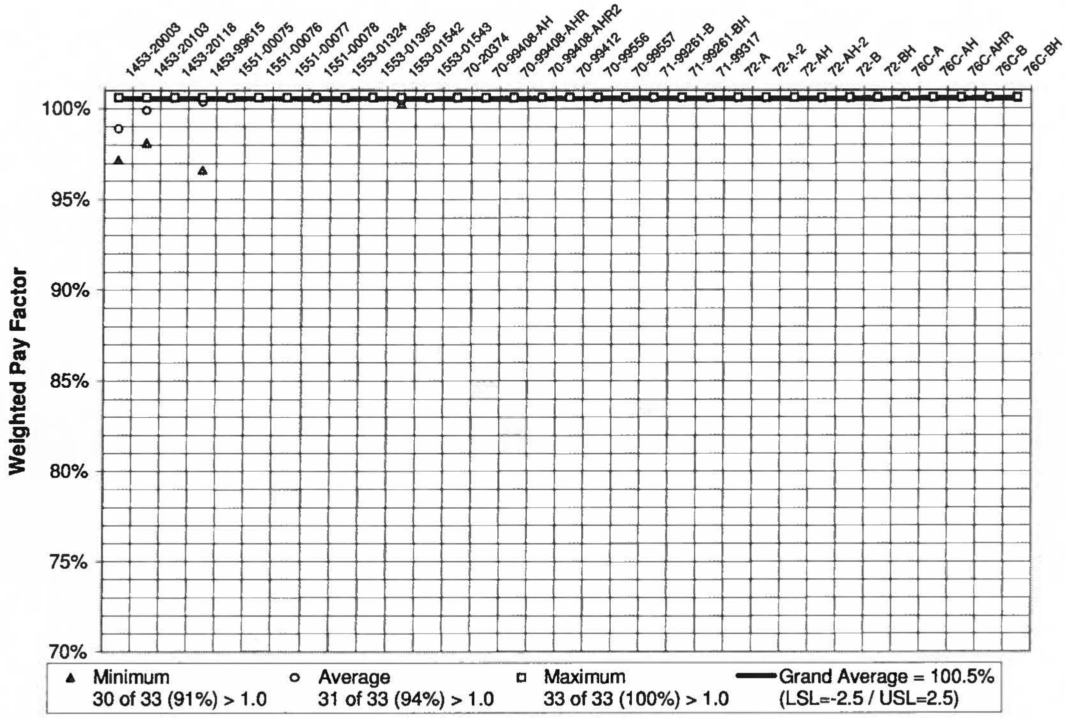
PCC Fine Aggregate Gradation No. 200 Sieve (Weighted Pay Factors)



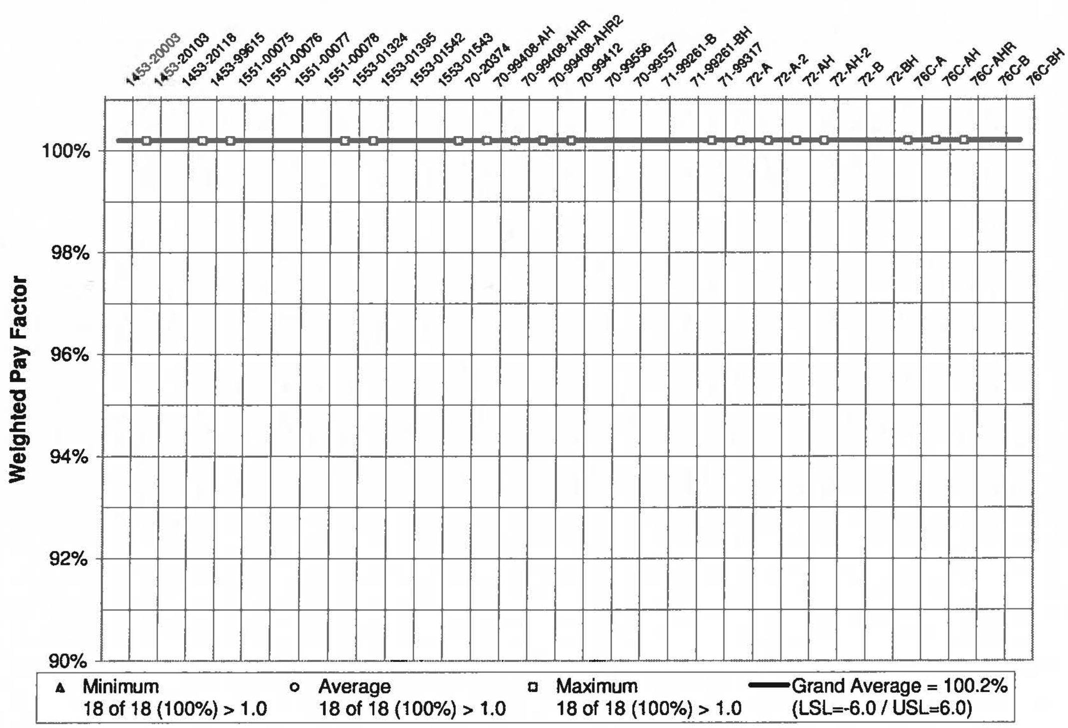
AC Roadway Density (Weighted Pay Factors)



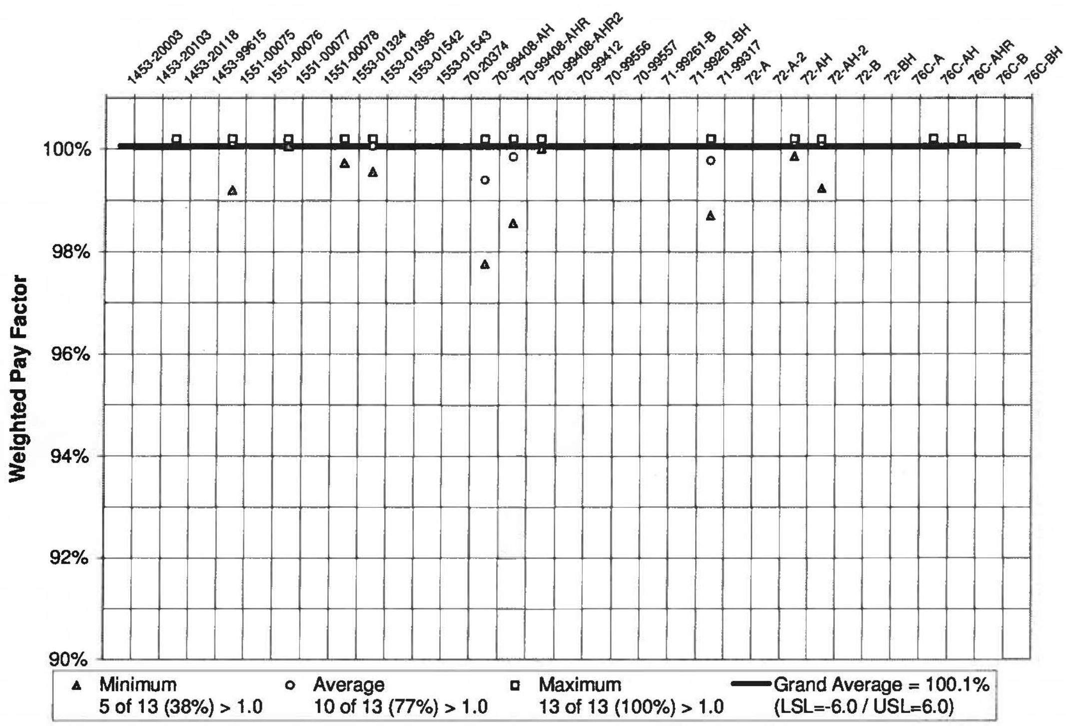
AC Asphalt Cement Content (Weighted Pay Factors)



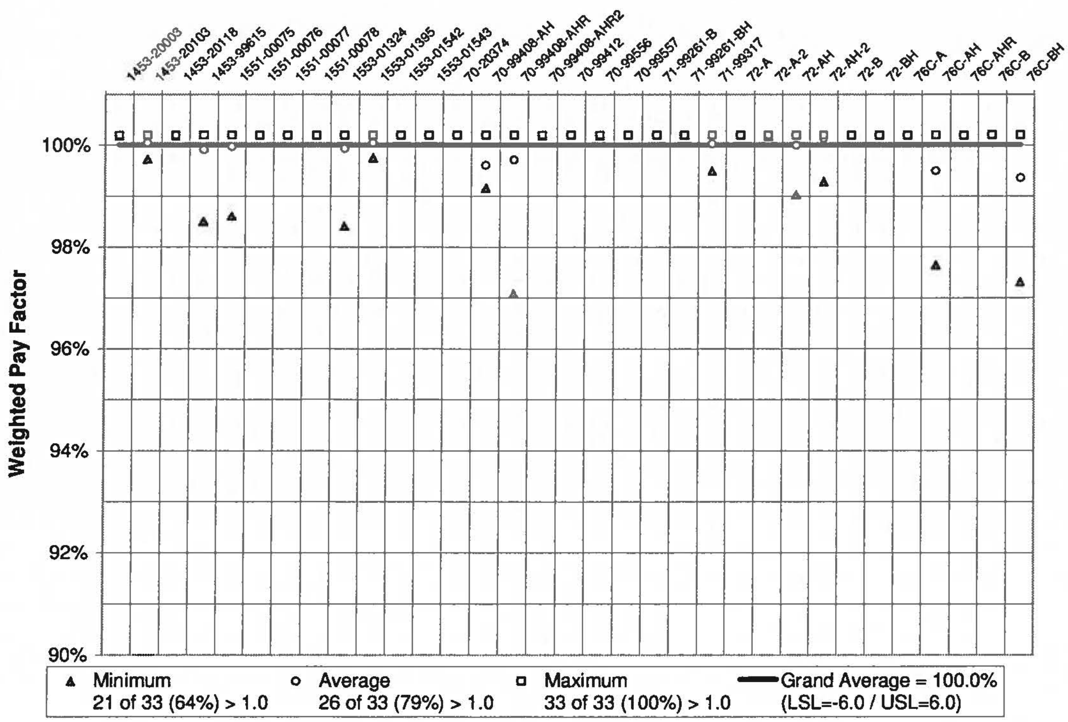
AC Air Voids (Weighted Pay Factors)



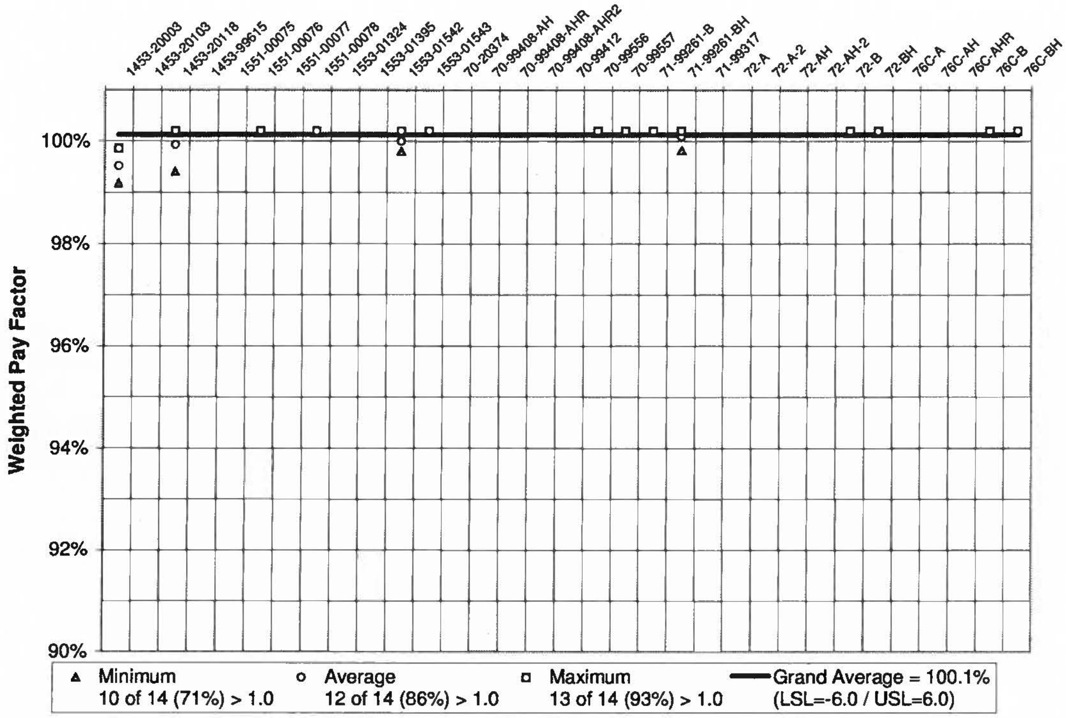
AC Gradation 1-inch Sieve (Weighted Pay Factors)



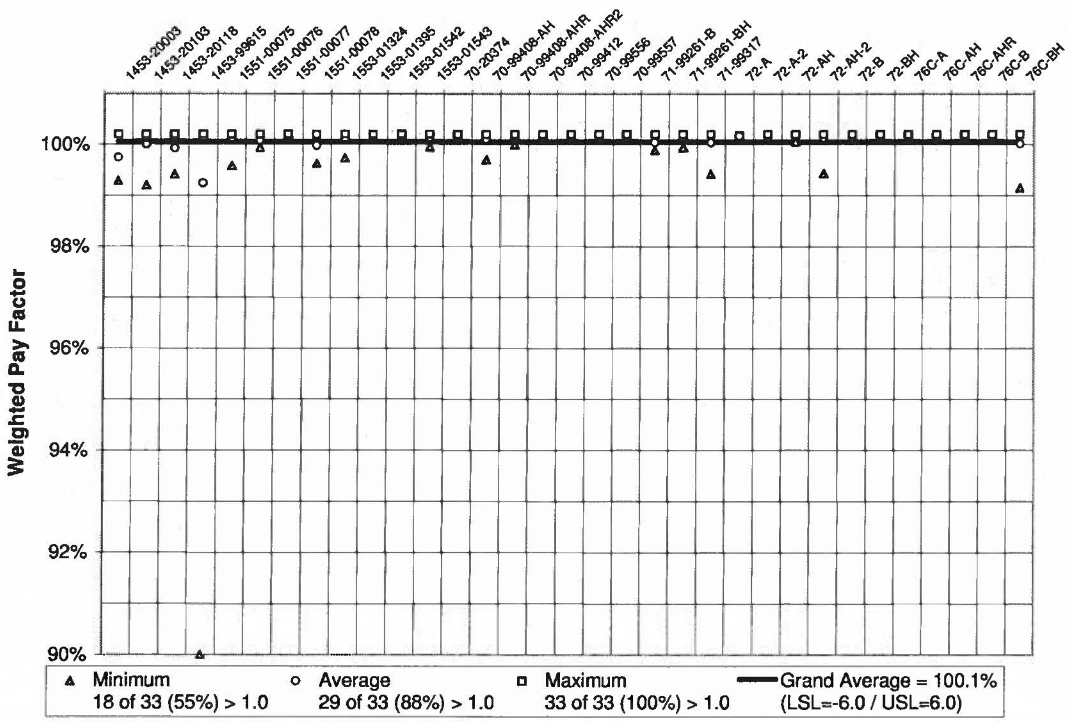
AC Gradation ¾-inch Sieve (Weighted Pay Factors)



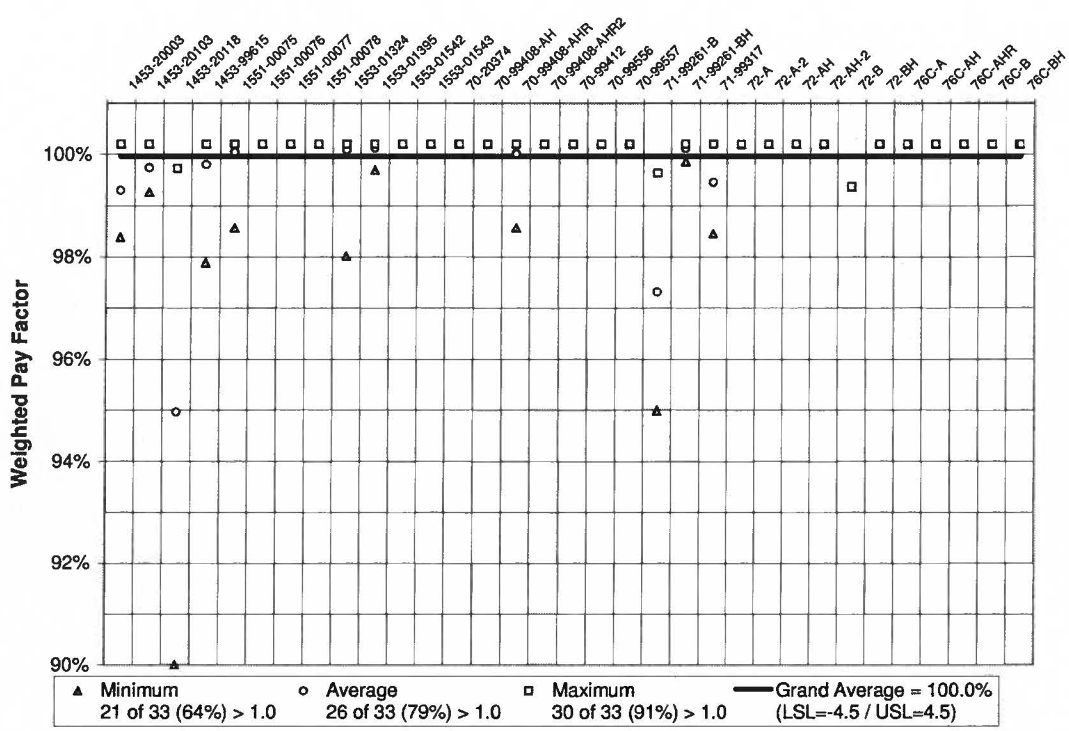
AC Gradation ½-inch Sieve (Weighted Pay Factors)



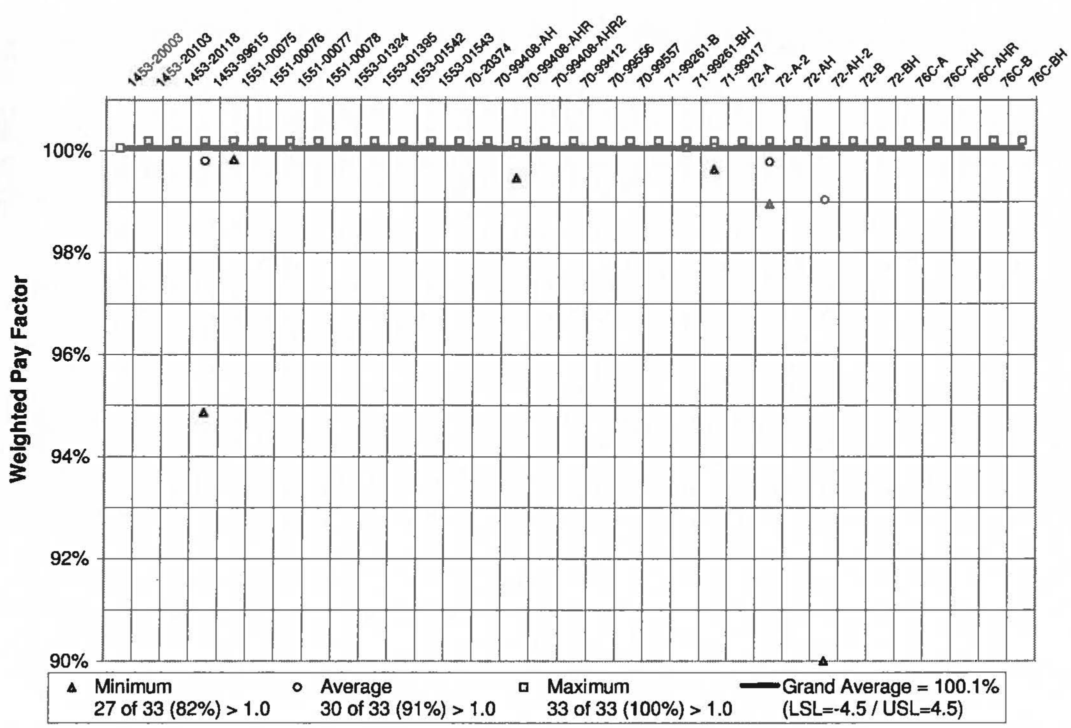
AC Gradation 3/8-inch Sieve (Weighted Pay Factors)



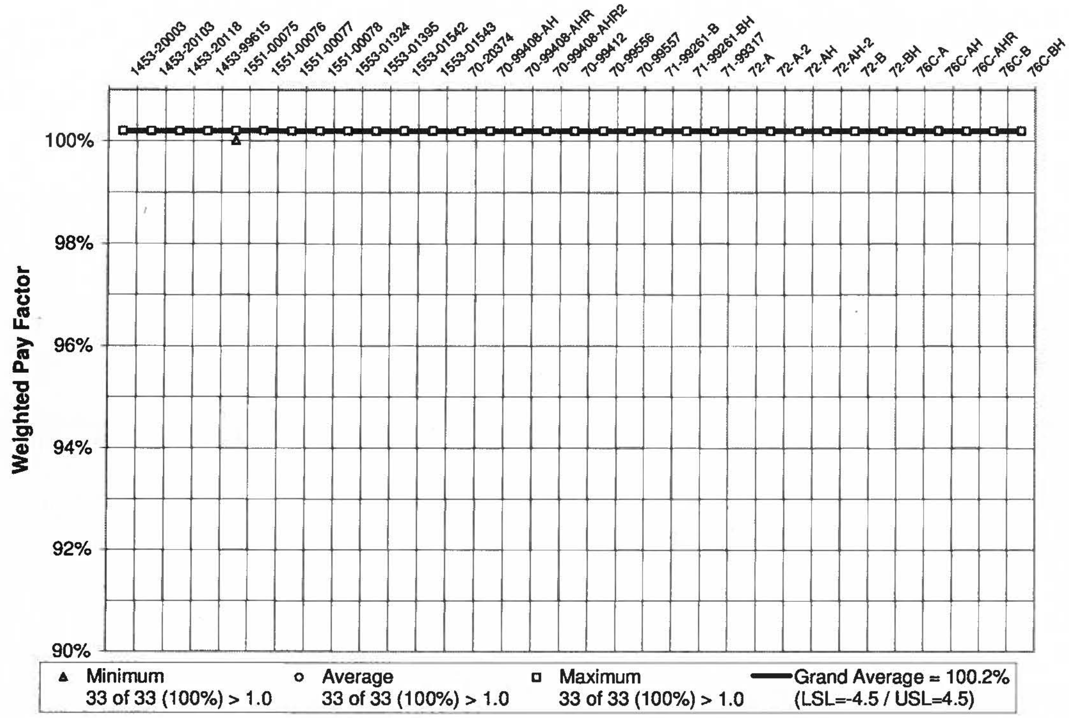
AC Gradation No. 4 Sieve (Weighted Pay Factors)



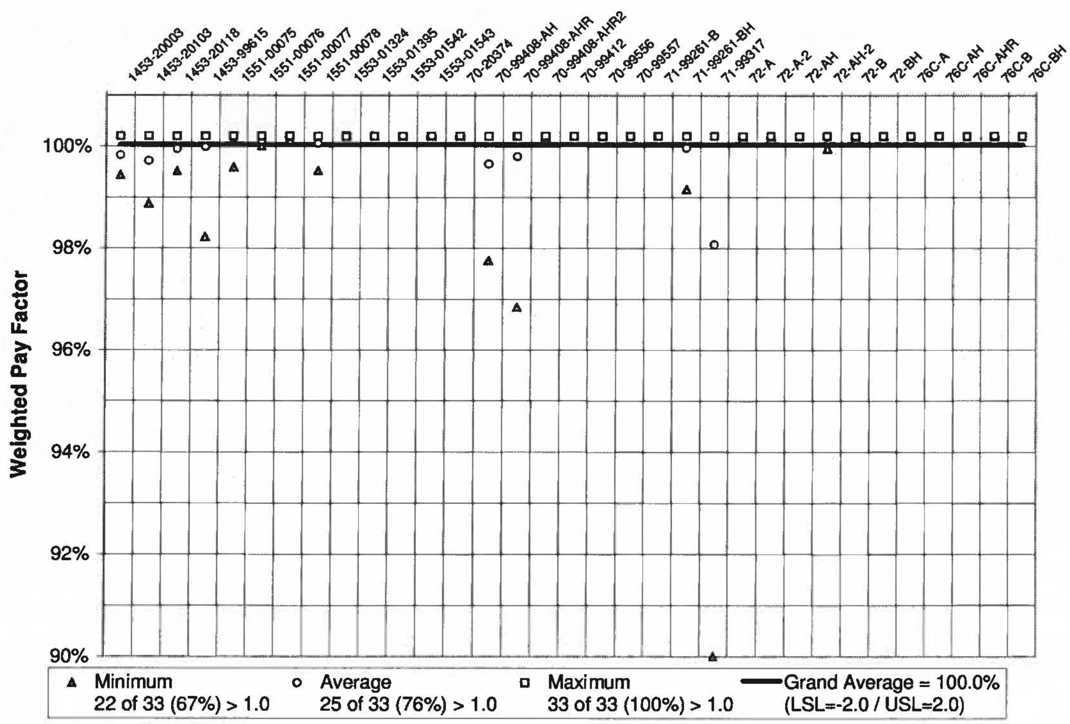
AC Gradation No. 10 Sieve (Weighted Pay Factors)



AC Gradation No. 40 Sieve (Weighted Pay Factors)



AC Gradation No. 80 Sieve (Weighted Pay Factors)



AC Gradation No. 200 Sieve (Weighted Pay Factors)

Appendix F: Guidelines for Quality Control Testing for PWL

Contractors Quality Control Testing and Inspection. The Contractor shall provide quality control personnel as necessary to assure the production of quality products as specified. Such personnel shall include one or more Quality Control Technicians who either individually or collectively are fully qualified in the production, placement and testing of [plant mix asphalt concrete] [concrete and reinforcing steel placed in portland cement concrete pavement]. Sampling and/or testing of construction materials for either control or acceptance purposes shall be accomplished by persons certified in the appropriate area(s) by the Oklahoma Highway Construction Materials Technician Certification Board.

The Contractor shall be responsible for the formulation of all mix designs. The mix design shall be prepared by an approved [asphalt] [concrete] mix design laboratory of the Contractor's choice. All mix designs and changes to the mix designs shall be submitted to the Materials Engineer for review. The Contractor shall perform or have performed all field sampling and testing necessary to ensure that materials and products are within the specified acceptable range. Control charts displaying results of these tests shall be maintained by the Contractor and displayed at the plant site. Copies of the Contractor's quality control tests shall be provided to the Engineer within 24 hours or at time intervals acceptable to the Engineer. Certification by the manufacturers may be used in lieu of field tests when such tests in the field are impracticable. [Asphalt] [Portland] cement and additives are examples of materials in this category.

(A) *Contractor's Process Control.* The Contractor shall be responsible for the process control of all materials during handling, blending, mixing and placing operations to produce an acceptable [asphalt] [portland cement] concrete pavement.

At no time will the Engineer issue instructions to the Contractor or producer as to the setting of dials, gauges, scales and meters. However, he/she may advise the Contractor against the continuance of any operations or sequence of operations which will result in non-compliance with Specification requirements.

(B) *Contractor's Testing.* For the characteristics subject to pay adjustments in this Special Provision, the Contractor's sampling and testing shall, as a minimum, comply with the schedule in Subsection (M) (*Contractor's Testing and Engineer's Acceptance Procedures*) found in the *Guidelines for Acceptance Testing for PWL* [Appendix G]. Additional sampling and testing to ensure compliance with Standard Specifications and other Special Provision requirements shall be in accordance with the Contractor's Quality Control Plan.

(C) *Contractor's Laboratory.* The Contractor shall provide a fully-equipped laboratory at a location no more than 50 road miles from the production site. The laboratory shall be subject to approval of the Engineer.

(D) *Contractor's Quality Control Plan.* Prior to initiation of work, the Contractor shall prepare a plan to ensure that acceptable quality can and will be obtained. The Plan which is to be submitted to the Engineer at least one week prior to the prework conference shall comply with SP 643-6QA and cover all of the items discussed in Section [411] [414] and [708] [701] of the Standard Specifications. However, the Contractor must tailor the plan to meet specific needs of the project. Once accepted by the Engineer, the plan becomes a part of the Contract and shall be enforced accordingly. Subsequent changes to the plan may be required by the Engineer in order to adjust to changes in the process or to correct problems in meeting Specification requirements

(E) *Control Charts for Quality Control.* The Contractor shall maintain and keep current control charts covering, as a minimum, the characteristics of [air voids, asphalt cement content, roadway density] [unit weight, slump, air content, thickness, compressive strength], and gradation (for each sieve). In all instances, technician- and equipment-identification numbers shall be recorded with each test to assist in troubleshooting potential testing discrepancies.

1. The control charts shall be "individuals" or "individuals and moving range" (I&MR) type control charts or as otherwise approved by the Engineer.
2. The charts shall identify the project number; the contract item number; the characteristic being measured; the date, time, lot #, sublot #, technician identification # and equipment identification # for each measurement; the applicable upper and lower control limits (but NOT the specification limits); the Contractor's test results; and any other data needed to facilitate control of the process and identify out-of-control conditions for the process in a timely manner.
3. The centerline, standard deviation, and upper and lower control limits for each control chart shall initially be calculated based on the initialization test results for each characteristic and contract item, excluding any initialization test results clearly due to explainable special-cause variation. Written approval by the Engineer will be required prior to any such exclusion of test data.
4. An "out of control" condition is defined as the condition resulting from any one of the following eight (8) "alarm" conditions occurring on a single control chart:
   (1) Any one point is more than 3 standard deviations from the centerline.
   (2) Nine points in a row are on the same side of the centerline.
   (3) Six points in a row are all increasing or all decreasing.
   (4) Fourteen points in a row are alternating up and down.
   (5) Two out of three points are more than 2 standard deviations from the centerline (and on the same side of the centerline).
   (6) Four out of five points are more than 1 standard deviation from the centerline (and on the same side of the centerline).
   (7) Fifteen points in a row are all within 1 standard deviation of the centerline.
   (8) Eight points in a row are all more than 1 standard deviation from the centerline (on either side of the centerline).
5. Whenever an out-of-control condition corresponding to alarm criteria 1, 2, 3, 5, 6 or 8 is observed for any of the control charts, the Contractor shall provide written notification to the Engineer concerning said out-of-control condition within 18 hours of the time the alarm-generating test was performed. In addition, the Contractor shall provide written notification to the Engineer (within 36 hours of the time the alarm-generating test was performed) concerning the investigative

and/or corrective actions taken or to be taken. **Failure to provide written notification to the Engineer within the time periods specified shall result in an automatic one-half-percent reduction in the composite pay factor for the affected lot for each failure to comply with the specified notification procedures.** After the probable-cause investigation for the out-of-control condition has been completed, written notification shall be provided to the Engineer stating the probable cause and corrective actions taken or to be taken to reduce the likelihood of reoccurrence (or, stating that the probable cause could not be determined, if such is the case). The Engineer may, at his discretion, exclude from pay-adjustment calculations any out-of-control test data for which the special-cause variation has been adequately identified and explained, provided the cause of said variation has been corrected so as to be unlikely to reoccur. (NOTE: Not all out-of-control conditions will require corrective action. For example, [a change in roadway densities due to a change in rolling pattern] [a steady drift in strength values due to temperature-induced changes in water demand] may not require corrective action if the [densities being obtained remain comfortably within the specified limits] [strengths being obtained remain comfortably above specified levels]. However, all out-of-control conditions shall be investigated for probable cause, regardless of the proximity of the measured values to the specification limits.)

6. In the event a significant systemic change to the process occurs or is observed (e.g. change in raw material sources, change in mix proportions, steady drift in [roadway densities] [compressive strength], etc.), the control limits (including centerline and standard deviation) for the affected characteristics shall be recalculated using the available test data that best represent the new process. Any such changes to control limits must be approved in writing by the Engineer.

Appendix G: Guidelines for Acceptance Testing for PWL

Acceptance. While the Contractor shall be fully and exclusively responsible for producing an acceptable product, acceptance responsibility rests with the Engineer. The entire lot of [asphalt] concrete as defined in Subsection (L) "*Lot and Sublot Selection*" will be accepted or rejected and paid for on the basis of acceptance test results.

(A) *General.* The following characteristics will be considered when determining the acceptability and pay factors for [Plant Mix Asphalt] [Portland Cement] Concrete Pavement. However, all of the requirements of the Standard Specifications on materials and workmanship except those superseded by Special Provisions in this Contract, shall remain in effect.

(a) Gradation
(b) [Asphalt Cement Content] [Strength]
(c) [Air Voids] [Air Content]
[(d) Roadway Density]

Several methods are available to test for the above characteristics. While only one method will be used, several tests may be made to measure each characteristic. The pay factors that relate to gradation will be considered in a group with only the lowest pay factor for the individual sieves to be considered in determining payment. All sieves specified in [section 708.04] [sections 701.05(e) and 701.06(c)] of the Standard Specifications or as modified by Special Provisions in this Contract shall be run. [Only the percentage passing the No. 200 sieves for both coarse and fine aggregate will be included in the *Percent-Within-Limits* pay-factor calculations. However, the standard acceptance requirements for all sieves will remain in force as specified in section 701]. The remaining applicable pay factors will be considered individually in determining payment. Pay factors for [asphalt cement content, air voids, roadway density] [strength, air content] and gradation (lowest), will apply to all [asphalt] [portland cement] concrete pavement placed.

(B) *Identifying Testing Precision and Bias.* For each test characteristic of each contract item for which control charting is required, the following initialization procedure shall be performed:

1. Initial testing shall be performed to identify any testing biases between the Contractor's and the Department's testing equipment and procedures. This testing will be referred to as "initialization" testing and will include the plant startup testing as well as the first lot for each contract item for which control charting is required. In all instances, technician- and equipment-identification numbers shall be recorded with each test result to assist in troubleshooting potential testing discrepancies.
2. The frequency of testing for initialization lots shall be double the normally-specified testing frequency. As such, each initialization lot shall be broken down into [ten (10)] [twelve (12)] equal sublots with the sampling within each sublot to

be performed at a random location or time interval in accordance with ASTM D 3665 or other acceptable means for ensuring random sampling of the materials.

3. [During the initialization testing, for each sublot, the Contractor shall obtain fresh asphalt concrete samples in accordance with AASHTO T 2 under direct observation by the Department (for subsequent air voids, asphalt cement content, and gradation testing). In so doing, the Contractor shall obtain the equivalent of "split" samples in accordance with applicable Department guidelines and procedures. The Contractor and the Department shall each perform the following tests (all from the same "split" sample of asphalt concrete): one (1) air voids test; one (1) asphalt cement content test; one (1) set of gradation sieve analyses.
   ]
   [During the initialization testing, for each sublot, the Contractor shall obtain gradation samples for coarse and fine aggregate in accordance with AASHTO T 2 under direct observation by the Department. The Contractor shall split each sample in accordance with AASHTO T 248. The Contractor and the Department shall each perform two (2) sets of gradation sieve analyses on the split samples in accordance with AASHTO T 27.
   ]

4. [During the initialization testing, for each sublot, the Contractor shall obtain roadway density cores from two (2) randomly-selected locations. At each location, two (2) cores shall be obtained, both within five-feet of each other and at the same distance from the longitudinal edge. Furthermore, if nuclear-methods for determining roadway density are to be used on the project, four (4) two-minute nuclear density measurements shall be taken directly over each coring location prior to coring and without moving the gauge between readings. Each core shall be uniquely identified to enable direct comparisons with each other and, if applicable, with the nuclear density measurements. The Contractor and the Department shall each perform two (2) density measurements on each core.
   ]
   [During the initialization testing, for each sublot, fresh concrete samples shall be obtained simultaneously by the Contractor and the Department from the same batch and the same proximity within the batch. For each sublot, the Contractor and the Department shall each perform the following tests (all from the same batch of concrete): two (2) slump tests; two (2) unit weight tests; two (2) air-content tests; cast three (3) cylinder specimens. The duplicate tests for slump, unit weight, and air content shall be recorded according to the order performed. The six (6) cylinder specimens (three cast by the Contractor and three by the Department) shall be tested for compressive strength at the same age (with said age being between 3- and 7-days as determined by the Engineer). At the Engineer's discretion, the cylinders may be cured in high-temperature water tanks to accelerate the strength-development of the cylinder specimens. In this instance, all cylinders cast from a single batch (i.e. the three cast by the Contractor **and** the three by the Department) shall be cured in the *same* curing tank and thus at the same elevated temperatures. [NOTE: Additional cylinder specimens will be cast by the Department and tested for compressive strength at 28-days to be used for acceptance purposes until such time as the Contractor's test

methods for compressive strength are validated.]
]

5. During initialization testing, the Contractor and the Department shall not divulge their respective test results until the conclusion of all sampling and associated testing for a given lot and quality characteristic.
6. At the conclusion of the sampling and associated testing for a given initialization lot and quality characteristic, a statistical "paired-t test" will be performed by the Engineer (in accordance with the *Guidelines for Initial Validation of Contractor's Test Methods*) using the pairs of initialization test data. The Engineer may, at his discretion, exclude from this and subsequent analyses any initialization test data that are clearly due to explainable special-cause variation provided the cause of said variation has been corrected so as to be unlikely to reoccur.
7. For those characteristics showing a *statistically-* **and** *practically*-significant bias between the Contractor's and the Department's test methods, the following shall govern:
    (1) The Department's test results shall be relied upon for acceptance and pay adjustment until such time as:
        (a) the source of the bias has been identified and eliminated and the lack of bias subsequently validated in accordance with the *Guidelines for Initial Validation of Contractor's Test Methods*, or
        (b) as provided in Subsection B.7.(2) below.
    (2) The Contractor may request evaluation of the testing bias via side-by-side three-way testing with an independent-assurance laboratory. In this instance, the following shall govern:
        (a) The Department will select the independent-assurance laboratory, which may be the Department's own independent-assurance laboratory or a third-party laboratory of the Department's choosing.
        (b) The steps outlined in Subsection B.2. through B.5. shall be performed utilizing three-way split samples, with the Contractor, Department, and independent-assurance laboratory each performing tests on the split-sample specimens.
        (c) At the conclusion of the three-way split-sample testing, three statistical "paired-t tests" will be performed by the Engineer (in accordance with the *Guidelines for Initial Validation of Contractor's Test Methods*) using the three pairs of evaluation test data (e.g. Contractor's versus Department's, Contractor's versus independent-assurance laboratory's, and Department's versus independent-assurance laboratory's test results).
        (d) If either of the following conditions are identified as a result of the three-way split-sample testing and analysis, Subsection B.7.(2)(f) shall govern, otherwise Subsection B.7.(2)(e) shall govern:
            (i) There is a *statistically-* **and** *practically*-significant bias between the Contractor's test methods and the Department's test methods **and** between the Contractor's test methods and the independent-assurance laboratory's test methods, or

(ii) the test methods employed by the Contractor during the three-way split-sample testing are not representative of the test methods to be employed by the Contractor during the normal execution of the project.

(e) The following shall govern subject to Subsection B.7.(2)(d):

(i) The Contractor shall not be held responsible for any additional costs incurred by the Department in conjunction with the three-way split-sample testing.

(ii) The Contractor's test results shall be relied upon for acceptance and pay adjustment subject to the provisions of Subsection B.8.

(f) The following shall govern subject to Subsection B.7.(2)(d):

(i) The Contractor shall reimburse the Department for all additional costs incurred by the Department as a result of the three-way split-sample testing.

(ii) The Department's test results shall be relied upon for acceptance and pay adjustment in accordance with Subsection B.7.(1)(a).

8. For those characteristics where the statistical paired-t test validates the Contractor's test methods, the Contractor's test results will be used for acceptance and pay adjustment subject to the following:

(1) The Department will perform ongoing paired testing at the approximate frequency of one (1) paired test per lot, but not less than one (1) paired test for every ten (10) sublots.

(2) The Engineer will keep and maintain testing-bias control charts in accordance with the *Guidelines for Ongoing Validation of Contractor's Test Methods*.

(3) In the event the Engineer's testing-bias control charts demonstrate an out-of-control condition for testing bias, the following procedures shall be followed:

(a) The Contractor and Engineer shall immediately investigate the probable cause.

(b) If the probable cause is identified and corrected or if the resulting bias is (according to the judgment of the Engineer) not likely to exceed the limits for Allowable Testing Bias (ATB) set forth in the *Guidelines for Initial Validation of Contractor's Test Methods*,

(i) Acceptance and pay adjustments for that quality characteristic will continue to be based on the Contractor's test results.

(ii) At the Engineer's discretion, the errant test results may be discarded from the acceptance and pay-adjustment calculations. If replacement test results are available (i.e. Department tests were performed on the same sublots as the discarded Contractor test results) the Department's test results may, at the Engineer's discretion, be substituted for the discarded test results.

(c) If the probable cause is not corrected and the resulting bias is (according to the judgment of the Engineer) likely to exceed the limits for Allowable Testing Bias (ATB) set forth in the *Guidelines for Initial Validation of Contractor's Test Methods*,

(i) All subsequent testing for that quality characteristic will be paired testing until such time as the source of the unacceptable testing bias is clearly identified and corrected and the validity of the Contractor's test methods has been re-established.

(ii) During this period of 100% paired testing, acceptance and pay adjustments for that quality characteristic will be based on the Department's test results rather than the Contractor's

(C) *Criteria for Lot Acceptance and Payment.* Except for surface smoothness, conformance with the specifications will be judged on the basis of the following criteria:

1. The estimated Percent-within-Limits (PWL) with respect to gradation, [asphalt cement content, air voids, and roadway density] [air content, and strength]. The PWL with respect to a particular quality characteristic is the amount of materials and construction which falls within the specified limits listed in the following tables [(where "JMF" refers to the corresponding values from the Job Mix Formula)]:

[Quality Characteristic	Lower Specification Limit (LSL)	Upper Specification Limit (USL)
Roadway Density	93 %	97 %
Air Voids (LMS)	JMF – 2.0 %	JMF + 2.0 %
Asphalt Cement Content	JMF – 0.4 %	JMF + 0.4 %
Gradation:		
Sieves #4 and larger	JMF – 6.0 %	JMF + 6.0 %
Sieves #8 through #100	JMF – 4.5 %	JMF + 4.5 %
Sieve # 200	JMF – 2.0 %	JMF + 2.0 %

Quality Characteristic	Lower Target Limit (LTL)	Upper Target Limit (UTL)
Roadway Density	94 %	96 %
Air Voids (LMS)	JMF – 0.8 %	JMF + 0.8 %
Asphalt Cement Content	JMF – 0.16 %	JMF + 0.16 %
Gradation:		
Sieves #4 and larger	JMF – 2.4 %	JMF + 2.4 %
Sieves #8 through #100	JMF – 1.8 %	JMF + 1.8 %
Sieve # 200	JMF – 0.8 %	JMF + 0.8 %

]

[Quality Characteristic	Lower Specification Limit (LSL)	Upper Specification Limit (USL)
28-Day Compressive Strength		

Class A Concrete	3,800 psi	No Upper Limit
Class AP Concrete	3,000 psi	No Upper Limit
Air Content	4.5 %	7.5 %
Gradation:		
Sieve # 200 (Coarse)	No Lower Limit	2.0 %
Sieve # 200 (Fine)	No Lower Limit	3.0 %

Quality Characteristic	Lower Target Limit (LTL)	Upper Target Limit (UTL)
28-Day Compressive Strength		
Class A Concrete	4,500 psi	No Upper Limit
Class AP Concrete	3,750 psi	No Upper Limit
Air Content	5.5 %	6.5 %
Gradation:		
Sieve # 200 (Coarse)	No Lower Limit	1.0 %
Sieve # 200 (Fine)	No Lower Limit	1.0 %

]

(D) *Acceptable and Rejectable Quality Levels.* A lot shall be considered of acceptable quality with respect to a particular characteristic if the PWL, as defined in Subsection (E) is no less than 90 percent. A lot shall be considered of rejectable quality with respect to a particular characteristic if the PWL is less than 50 percent. Lots exceeding the Acceptable Quality Level shall be subject to positive pay factors as defined in Subsection (F). Lots failing to achieve the Acceptable Quality Level but exceeding the Rejectable Quality Level shall be subject to negative pay factors as defined in Subsection (F). Lots failing to achieve the Rejectable Quality Level shall be subject to removal and replacement or a "zero" pay factors (for that quality characteristic) as defined in Subsection (F).

The contractor shall perform the necessary quality-control sampling and testing to ensure that acceptable quality level requirements are consistently met.

(E) *Determination of Percent-within-Limits (PWL).* The PWL with respect to each of the characteristics of gradation, [asphalt cement content, air voids, and roadway density] [air content, and strength], will be determined as follows:

1. Compute the sample mean ($\overline{X}$) and the sample standard deviation (S') as follows, where N = the number of individual test results (i.e. the number of X_i's):

$$\overline{X} = \frac{\sum X_i}{N} \qquad S' = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N-1}}$$

2. If $\overline{X}$ falls *outside* the target limits (LTL and UTL) **and** *inside* the specification limits (LSL and USL), compute the target-adjusted standard deviation (S'') as follows:

$$S'' = \sqrt{S'^2 + \left(X_{target} - \overline{X}\right)^2}$$

where X_{target} = the nearest target limit (LTL or UTL)

3. If $\overline{X}$ falls *inside* the target limits (LTL and UTL) **or** *outside* the specification limits (LSL and USL), compute the target-adjusted standard deviation (S'') as follows:

$$S'' = S'$$

4. Using the target-adjusted standard deviation (S''), compute the upper quality index (Q_u) and the lower quality index (Q_L) corresponding to the upper and lower specification limits listed in Subsection (C):

$$Q_U = \frac{USL - \overline{X}}{S''} \qquad Q_L = \frac{\overline{X} - LSL}{S''}$$

5. Using Table 1 (for sample size N = [5] [6] or the appropriate table for other values of N), determine the percentage of materials and construction falling *outside* the specification limits PD_U and PD_L associated with Q_U and Q_L, respectively. Add these two values to obtain the total percent defective (for that quality characteristic) for the lot (*PD*):

$$PD = PD_U + PD_L$$

6. Determine the percentage of materials and construction falling *within* the specification limits (*PWL*) as follows:

$$PWL = 100 - PD$$

(F) *Pay Factors for Lot Quality Characteristics.* Except for pavement smoothness, the pay factor (*PF*) for each quality characteristic will be determined as follows:

1. If *PWL* is greater than or equal to 50 percent, compute the pay factor using the equation:

$$PF = 4.3(PWL) - 0.0215(PWL)^2 - 111$$

2. If *PWL* is less than 50 percent, the Engineer may require removal and replacement of the defective lot at the Contractor's expense. If this option is not exercised, the Contractor may elect to replace the lot or leave it in place subject to a pay factor (for that quality characteristic) of *PF* = 0%.

(G) *Pay Adjustment for Lots.* Once a lot has been defined, it's identity will be maintained throughout the mixing and placement process. When the lot is completed, the individual pay factors determined in Subsection (F) for gradation, [asphalt content, air voids, and roadway density] [air content, and strength] will be used to calculate a composite pay factor (*CPF*) and a pay adjustment (*PA*) for the subject lot as follows:

[

$$CPF = \frac{4PF_D + 3PF_V + 2PF_{AC} + PF_G}{10}$$

where:

CPF = Composite pay factor
PF_D = Pay factor for roadway density
PF_V = Pay factor for air voids
PF_{AC} = Pay factor for asphalt cement content
PF_G = Pay factor for gradation – the smallest of the individual pay factors for the sieves listed in Subsection (C).

]
[

$$CPF = \frac{6PF_S + 3PF_A + PF_G}{10}$$

where:

CPF = Composite pay factor
PF_S = Pay factor for strength
PF_A = Pay factor for air content
PF_G = Pay factor for gradation – the smallest of the individual pay factors for the sieves listed in Subsection (C).

]

The pay adjustment for the completed lot will be determined in accordance with the following formula:

$$PA_{Lot} = (CPF - 1)(CUP)(Q_{Lot})$$

where:

PA_{Lot} = Pay adjustment for the lot,
CPF = Composite pay factor,
CUP = Contract unit price [($/Ton ($/Metric Ton))] [($/Yd^2 ($/m^2))], and
Q_{Lot} = Quantity of [asphalt] concrete [pavement] in the lot [(Tons (Metric Tons))] [(Yd^2 (m^2))]

(H) *Smoothness Acceptance and Pay Adjustment.* For smoothness determination and pay adjustment purposes, the pavement surface will be tested on an extent-to-extent basis in accordance with Special Provisions 430-1QA. Acceptance and pay adjustment determinations made under Special Provisions 430-1QA will be completely independent of those made under this Special Provision.

(I) *Pay Adjustments Not Covered in Special Provisions [411-9QA] [414-10QA] or 430-1QA.* Adjustments in pay, for deviations from specified standards for characteristics other than those described in these Special Provisions (if any) will be made in accordance with General Provision 105.03.

(J) *Total Pay Adjustment for Entire Project.* The total adjustment in pay for the entire project is the sum of: (1) the pay adjustments for individual lots per Subsection (G); plus (2) the pay adjustments for smoothness per Special Provision 430-1QA; plus (3) other pay adjustments, if appropriate, per Subsection (I).

(K) *Extreme Value (Outliers).* Test results apparently inconsistent with the results of the majority of tests will also be closely examined by the Engineer in order to determine their validity. The examination will cover the procedures used in sampling and testing and, if necessary, a mathematical analysis performed in accordance with ASTM E178-02 (upper 2.5% significance level). Test results thus determined by the Engineer to be non-representative of the material being evaluated will be discarded. The remaining test results will then be supplemented, if necessary, and treated in a manner consistent with the *Guidelines for Initial Validation of Contractor's Test Methods* and the *Guidelines for Ongoing Validation of Contractor's Test Methods.*

(L) *Lot and Sublot Section.* The asphalt concrete will be randomly sampled and tested for all control test characteristics on a lot to lot basis in accordance with the following requirements. However, any load of mixture which is visually unacceptable for reasons of being excessively segregated or aggregate improperly coated will be rejected for use in the work. Excessively high or low temperature will be cause for rejection. Furthermore, sections of completed pavement which form visual observation or known deficiencies that appear to be seriously inadequate will be tested. The results of such tests will not be used for pay adjustment purposes but will be used to determine whether the section is totally unacceptable and must be removed. In the event that it is determined to be unacceptable, its removal and replacement shall be at no additional cost to the Department. A standard size lot [at the asphalt plant] shall consist of [five (5)] [six (6)] equal sublots of [1,000 tons (metric tons)] [2500 Yd^2 (m^2)] each. Any partial lot (one with less than [5,000 tons (metric tons)] [15,000 Yd^2 (m^2)]) shall be treated as a separate lot when four (4) or more sublots exist. When a lot contains three (3) or less sublots, it shall be combined with the previous lot. On multiple-project contracts, the lots of [asphalt] concrete will carry over from project to project within that contract. All acceptance testing shall be performed at a random location or time interval within each sublot in accordance with ASTM D 3665 or other acceptable means for ensuring random sampling of the materials

(M) *Contractor's Testing and Engineer's Acceptance Procedures.* Once a lot has been defined, its identity will be maintained throughout the mixing and placement process. Pay factors, determined from random sampling and testing of the lot at appropriate locations, will be used in computing its pay adjustment.

1. The Contractor is require as a minimum to comply with the following schedule for sampling and testing. Depending upon the available time and his confidence in the Contractor's Process Control, the Engineer may elect to perform more or less sampling and testing.

   [

   (1) Asphalt Cement Content and Gradation – one (1) test for each characteristic sampled at a *single* randomly-selected location per sublot.

   (2) Roadway Density – three (3) specimens sampled at *three* randomly-selected locations per sublot with the three resulting test values considered as three (3) separate tests. [NOTE: For the calculation of Percent-Defective (*PD*) from the Quality Index (*Q*) for Roadway Density, the number of "sublots" (N) will thusly be fifteen, instead of five.]

(3) Air voids (except NMS mixes) – three (3) specimens per sublot sampled at a *single* randomly-selected location per sublot, with the three resulting test values averaged and considered as one (1) test.

(4) Air voids (NMS mixes) – two (2) specimens per sublot sampled at a *single* randomly-selected location per sublot, with the two resulting test values averaged and considered as one (1) test.

]

[

(1) Gradation , slump, unit weight, and air content – one (1) test for each characteristic sampled at a *single* randomly-selected location per sublot.

(2) Strength – three (3) cylinders sampled at a *single* randomly-selected location per sublot, with the three resulting test values averaged and considered as one (1) test.

]

Appendix H: Guidelines for *Initial* Validation of Contractor's Test Methods

In order to utilize the Contractor's material test results for acceptance and payment, the Department must ensure that the Contractor's results compare favorably with the Department's test results for the same material and same quality characteristic. The following procedure can be used to determine the initial validity of the Contractor's test methods:

1. A "paired test" (as referred to herein) will be any time two (2) separate tests are conducted by the Department and the Contractor (one test each) where the material tested has been collected as a true split sample or was sampled at the same time, from the same lot, sublot and batch, and from the same proximity within the batch. For asphalt air voids, asphalt cement content, and gradation testing, "paired" tests are those tests performed on split samples or samples taken from the same batch and from the same proximity within the batch. For roadway density and thickness measurements from cores, "paired" tests are those measurements performed on the exact same cores. For concrete unit weight, slump, and air content, "paired" tests are those conducted on samples taken from the same batch and from the same proximity within the batch. For concrete compressive and flexural strength each "paired" test will generally be the average strength of two or three specimens each (by the Department and the Contractor) where the specimens are cast from concrete sampled from the same batch and from the same proximity within the batch.

2. For each quality characteristic for which the Contractor's test methods are to be validated, the Department will conduct a minimum of ten (10) initial paired tests with the Contractor.

3. The Engineer may, at his discretion, exclude from this and subsequent analyses any paired test data that clearly exhibit explainable special-cause variation provided the cause of said variation has been identified and corrected so as to be unlikely to reoccur.

4. For each quality characteristic, calculate the paired-t test statistic (t_P), the average of the paired differences ($\overline{X}_P$), and the standard deviation of the paired differences (S_P) as follows:

$$t_P = \left| \sqrt{N_P}\,\frac{\overline{X}_P}{S_P} \right| \qquad \overline{X}_P = \frac{\sum_{i=1}^{N_P}(C_i - D_i)}{N_P} \qquad S_P = \sqrt{\frac{\sum_{i=1}^{N_P}(C_i - D_i - \overline{X}_P)^2}{N_P - 1}}$$

where

t_P = Paired-t test statistic. This value will be compared to a critical t-value (t_{crit}) to be obtained from Table 1.

N_P = The number of paired tests conducted. Each "paired test" consists of two individual tests (one by the Contractor and one by the Department) performed on a single split sample.

$\overline{X}_P$ = The average of the differences between the paired tests. This should not be confused with the difference between the averages. The correct order for computing $\overline{X}_P$ is to calculate all the differences between the paired tests, then take the average of those differences.

S_P = The standard deviation of the differences between the paired tests. As with $\overline{X}_P$, the correct order for computing S_P is to calculate all the differences between the paired tests, then take the standard deviation of those differences.

C_i = The Contractor's individual test result for a given split sample *i*.

D_i = The Department's individual test result for a given split sample *i*.

$N_P - 1$ = The number of degrees of freedom for use with Table 1.

5. Obtain the critical t-value t_{crit}, from Table 1, using $N_P - 1$ as the number of degrees of freedom.

6. Compare the calculated paired-t statistic (t_P) to the corresponding critical value from Table 1 (t_{crit}).

7. If t_P is less than t_{crit}, the Contractor's test methods for that quality characteristic are without significant bias. The Contractor's test methods for that quality characteristic can be considered valid at the present time. Proceed to Step 12.

8. If t_P is greater than or equal to t_{crit}, there exists a *statistically*-significant bias (= $\overline{X}_P$) between the Contractor's and the Department's test methods.

9. Compare the Contractor's testing bias ($\overline{X}_P$) to the relevant Allowable Systemic Testing Bias (ATB) from the following table:

PCC Quality Characteristic	Allowable Testing Bias (ATB)	Units
Gradation:		
Sieve # 200 (Coarse)	± 0.4	%-passing
Sieve # 200 (Fine)	± 0.4	%-passing
Unit Weight	± 1.0	pcf
Slump	± 0.3	inch
Air Content	± 0.3	% (by volume)
Compressive Strength	± 125	psi
Flexural Strength	± 50	psi

AC Quality Characteristic	Allowable Testing Bias (ATB)	Units
Gradation:		
Sieves #4 and larger	± 1.2	%-passing
Sieves #8 through #100	± 0.9	%-passing

Sieve # 200	± 0.4	%-passing
Asphalt Cement Content	± 0.2	% (by weight)
Air Voids (LMS)	± 1.0	% (by volume)
Roadway Density	± 0.4	% max. theor.

10. If the magnitude of $\overline{X}_P$ is less than the magnitude of the corresponding ATB, the Contractor's bias for that quality characteristic, though *statistically*-significant, is not *practically* significant. The Contractor's test methods for that quality characteristic can be considered valid at the present time. Proceed to Step 12.

11. If the magnitude of $\overline{X}_P$ is greater than or equal to the magnitude of the corresponding ATB, the Contractor's bias for that quality characteristic is *practically* significant and unacceptable. The Contractor's test results for that quality characteristic will not be used for acceptance unless and until the Contractor's test methods are validated as follows:
    (1) The respective test equipment used during the paired testing (both the Contractor's and the Department's) shall be inspected and tested for calibration by a qualified independent calibration specialist.
    (2) A calibration test certificate shall be prepared by the calibration specialist identifying the pre-calibration errors at various measurement levels such that the paired test data are fully bracketed by the pre-calibration error estimates.
    (3) The paired test results will then be adjusted by the Engineer so as to effectively nullify the errors identified on each calibration test certificate.
    (4) Return to Step 4, using the adjusted paired test data in lieu of the original paired test data.

12. Once the Contractor's test methods for a given quality characteristic have been validated, the Engineer will provide ongoing validation of the Contractor's test method in accordance with the *Guidelines for Ongoing Validation of Contractor's Test Methods.*

Table 1 – Critical Values for the Paired-t Test Statistic (based on $\alpha = 0.01$)

Degrees of Freedom (N_P - 1)	Critical t-value (t_{crit})	Degrees of Freedom (N_P - 1)	Critical t-value (t_{crit})	Degrees of Freedom (N_P - 1)	Critical t-value (t_{crit})
2	9.925	25	2.787	48	2.682
3	5.841	26	2.779	49	2.680
4	4.604	27	2.771	50	2.678
5	4.032	28	2.763	60	2.660
6	3.707	29	2.756	70	2.648
7	3.499	30	2.750	80	2.639
8	3.355	31	2.744	90	2.632
9	3.250	32	2.738	100	2.626
10	3.169	33	2.733	110	2.621
11	3.106	34	2.728	120	2.617
12	3.055	35	2.724	130	2.614
13	3.012	36	2.719	140	2.611
14	2.977	37	2.715	150	2.609
15	2.947	38	2.712	160	2.607
16	2.921	39	2.708	170	2.605
17	2.898	40	2.704	180	2.603
18	2.878	41	2.701	190	2.602
19	2.861	42	2.698	200	2.601
20	2.845	43	2.695	300	2.592
21	2.831	44	2.692	400	2.588
22	2.819	45	2.690	500	2.586
23	2.807	46	2.687	1000	2.581
24	2.797	47	2.685	10000	2.576

Appendix I: Guidelines for *Ongoing* Validation of Contractor's Test Methods

After the Department has initially validated the Contractor's test method for a particular quality characteristic as outlined in the *Guidelines for Initial Validation of Contractor's Test Methods*, the following procedure will be used for ongoing validation of the Contractor's test methods for that quality characteristic. The ongoing validation will be performed by the Engineer through the creation and maintenance of a "testing-bias" control chart for that quality characteristic. The purpose of the control chart will be to document the bias (over time) between the Contractor's and the Department's testing procedures and to identify any changes that occur with the Contractor's (or Department's) test methods during the prosecution of the project. The following procedures are meant to guide the Engineer in the use of testing-bias control charts:

1. The values to be reported on the testing-bias control chart will be the differences (for that quality characteristic) between the Contractor's and the Department's test results for all paired tests conducted for that quality characteristic for the project. In general, "paired" tests are those tests performed on materials sampled at the same time, from the same batch and from the same proximity within the batch. The differences will be recorded and charted as $C_i - D_i$, where C_i = the Contractor's individual test result for a given split sample i, and D_i = the Department's individual test result for that same split sample.

2. In addition to the initial paired tests (performed during the initial validation of the Contractor's test methods), the Engineer will periodically conduct additional ongoing paired tests with the Contractor. The frequency of ongoing paired testing will, in general, be one (1) paired test per lot, but should not be less than one (1) paired test for every ten (10) sublots.

3. The control charts will be maintained and kept current and will follow the same guidelines and requirements as those specified for the Contractor's control charts in the *Guidelines for Quality Control Testing for PWL.*

4. In the event an out-of-control condition is observed for testing bias,
   a) The Contractor and Engineer shall immediately investigate the probable cause.
   b) If the probable cause is identified and corrected or if the resulting bias is (according to the judgment of the Engineer) not likely to exceed the limits for Allowable Testing Bias (ATB) set forth in the *Guidelines for Initial Validation of Contractor's Test Methods*,
      i) Acceptance and pay adjustments for that quality characteristic will continue to be based on the Contractor's test results.
      ii) At the Engineer's discretion, the errant test results may be discarded from the acceptance and pay-adjustment calculations. If replacement test results are available (i.e. Department tests were performed on the same sublots as the discarded Contractor test results) the Department's test results may, at the Engineer's discretion, be substituted for the discarded test results.
   c) If the probable cause is not corrected and the resulting bias is (according to the judgment of the Engineer) likely to exceed the limits for Allowable Testing Bias (ATB) set forth in the *Guidelines for Initial Validation of Contractor's Test Methods*,

i) All subsequent testing for that quality characteristic will be paired testing until such time as the source of the unacceptable testing bias is clearly identified and corrected and the validity of the Contractor's test methods has been re-established.

ii) During this period of 100% paired testing, acceptance and pay adjustments for that quality characteristic will be based on the Department's test results rather than the Contractor's.