



TRC0001

# **Development of Quality Control Procedures for Hot-Mix Asphalt**

Kevin D. Hall

Final Report

2012

**FINAL REPORT**

**TRC-0001**

**DEVELOPMENT OF QUALITY CONTROL PROCEDURES FOR HOT-MIX  
ASPHALT**

by

Kevin D. Hall

Conducted by

Department of Civil Engineering  
University of Arkansas

In cooperation with

Arkansas State Highway and Transportation Department

U.S. Department of Transportation  
Federal Highway Administration

University of Arkansas  
Fayetteville, Arkansas 72701

**MAY 2012**

## **TRC-0001**

### **Development of Quality Control Procedures for Hot-Mix Asphalt**

#### **EXECUTIVE SUMMARY**

The Arkansas quality control / quality assurance (QC/QA) specification for hot mix asphalt (HMA) construction was initially developed from data and experience gained with the Marshall method of mix design. The full implementation of Superpave resulted in questions relating to the suitability of the QC/QA specification. Research was undertaken to revise the existing program or create a new specification.

Six projects were randomly sampled and tested by three operators (the contractor, the agency, and the research team) to establish typical HMA construction variability, in terms of the pay factors used in Arkansas: air voids, voids in the mineral aggregate (VMA), binder content, and field density. The testing data was normally distributed, validating a typical assumption regarding construction. Analyses showed testing variability among the three operators could be absorbed into the overall variability of test properties, rather than being considered a distinct factor. Three levels of HMA quality were identified by the mean and standard deviation of test data.

A specification for QA/QC was proposed, which features acceptance criteria based on both the mean value of a given property (similar to the existing specification) and an acceptable range of the property. The acceptance criteria for both the mean and range are based on the standard deviation of the property expected in the field; for initial implementation, it is recommended that the 'medium' quality level as defined by the testing program be used.

A limited validation study was performed to gauge the effect of implementation. The study suggested that implementation of the proposed system would not result in drastic cases of materials/pavement rejection. Further, the study indicated that field compaction would be the current acceptance property most affected by the proposed system. It is strongly recommended a full-scale field validation study be conducted; details of such a study are included.

## TABLE OF CONTENTS

INTRODUCTION .....	2
BACKGROUND .....	3
Quality Control/Quality Assurance.....	4
Types of Specifications.....	4
Superpave and NCHRP 409.....	7
WesTrack .....	10
RESEARCH OVERVIEW .....	11
ESTABLISHMENT OF CONSTRUCTION VARIABILITY.....	12
Testing Program.....	14
Data Analysis .....	15
Data Normality.....	16
Operator Variability .....	17
HMA Properties .....	21
DEVELOPMENT AND IMPLEMENTATION OF A QA/QC SPECIFICATION.....	26
Mean-and-Range Approach .....	26
Validation.....	27
RECOMMENDATIONS.....	28

## TRC-0001

### Development of Quality Control Procedures for Hot-Mix Asphalt

#### INTRODUCTION

Acceptance specifications for hot-mix asphalt materials and construction shifted away from a ‘materials and methods’ approach to a statistically-based end-result specification in the 1980’s. This required a set of specifications which defined the desired product (e.g. hot-mix asphalt pavement) quality, and specifications concerning a quality control/quality assurance (QC/QA) sampling and testing program. The Arkansas State Highway and Transportation Department (AHTD) fully completed implementation of a QC/QA program with the publication of the 1993 Edition of the *Standard Specifications for Highway Construction*.

The current QC/QA system used in Arkansas for flexible pavement construction bases acceptance of hot-mix asphalt materials and the completed pavement on the average of quality control (QC) tests performed by the contractor and quality assurance (QA) tests performed by the agency. A potential shortcoming of the existing system relates to the lack of control on the *variability* of materials and construction, i.e. QC and QA tests may vary widely – indicating a lack of production/construction process control – and yet the *average* of these tests generates full acceptance or even payment incentives. Project TRC-0001 sought to develop a QC/QA testing and acceptance program which provides requisite assurance that the final product quality is acceptable, and that the variability of the materials and pavement qualities is minimized.

## **BACKGROUND**

Prior to the 1970s, quality control and quality assurance plans for highway construction in most states, including Arkansas, consisted of a “recipe” or method for construction specified by the appropriate agency. It was believed that as long as the methods conformed to the specifications, then the finished product would be acceptable. In many cases, this was true. Past experience and engineering judgment proved successful for the most part. However, in the 1960s, several highway failures attracted the attention of Congress. The result of this attention was the formation of the House Committee on Oversight and Investigations. The findings of this committee revealed poor record keeping, as well as large amounts of inadequate materials being used for construction. Additionally, the AASHO Road Test increased awareness concerning improvements in highway construction. It was evident that a more capable method was necessary for ensuring the quality of highways.

In the 1970s the mood concerning quality control and quality assurance began to change. The idea began to shift from “materials and methods” specifications (MRS) towards statistically based end-result specifications (ERS). This meant that a contractor would have more flexibility in how the highway was constructed, as long as the desired quality was achieved. This required a set of specifications that define the desired product quality. For example, rather than specifying a particular size and type of roller for the compaction of hot-mix asphalt (HMA), a required level of density was specified. The method used to produce an acceptable density became the responsibility of the contractor. The transition from method to end-result specifications has been slow to occur, partially due to the reluctance to change by both contractors and agencies.

## **Quality Control/Quality Assurance**

Quality assurance specifications also referred to as quality control/quality assurance (QC/QA) specifications usually contain a combination of end-result specifications and method-related specifications. The contractor is responsible for quality control while the highway agency is responsible for acceptance testing of the product. These specifications are typically statistically based, utilizing methods such as random sampling and lot-by-lot testing, allowing contractors to ensure that their operations are producing an acceptable product (Hughes, et al, 1996). A 1992 AASHTO survey indicated that all but eight of the fifty states either used or had made plans to use quality control/quality assurance (QC/QA) specifications (Smith, et al, 1998).

In 1996, AASHTO Highway Subcommittee on Construction released the Implementation Manual for Quality Assurance. This manual deviates from the National Research Council's definition of quality assurance specifications (given above) as it recommends moving completely away from the use of "method type" specifications. The report refers exclusively to the use of end-result specifications and advocates the development and use of performance-related specifications. Among other things, the importance of personnel training, laboratory accreditation, and defining the barrier between quality control and quality assurance are discussed in detail. It could be strongly argued that this document should be used as the blueprint of any new specification system and for evaluating the effectiveness of any in-place system (AASHTO, 1996a).

### Types of Specifications

Quality control and quality assurance (QC/QA) is important in ensuring quality construction of pavements. In general, there are four basic forms of quality control. To more thoroughly

understand the evolution of QC/QA, these forms will be explained in greater detail. The first is a materials and methods specification, or “recipe” specification. This type of specification was used for many years prior to the 1970’s. Method-related specifications (MRS) are primarily based on experience and engineering judgment. Within MRS, the equipment, material, and procedure the contractor is to use for construction are all laid out in detail. Using this form of specification forces the supervising agency to monitor every facet of the contractor’s operations closely to ensure adherence to the process as well as making the agency ultimately responsible for the outcome. While proponents of MRS felt that this form of specification ensured that the work was done right the first time, faults have become apparent in many aspects of MRS. In most cases, MRS lacked a section describing methods for payment, which can sometimes be a hotly contested issue. Additionally, a complete description of the work process reduced innovation, and method-related specifications removed all responsibility for quality from the contractor. As long as the contractor followed the outlined procedures for construction, it could not be clearly faulted if the HMA did not perform properly. Also, requiring agency personnel to constantly watch the contractor set the foundation for an adversarial relationship between the contractors and inspectors (Smith, et al, 1998).

Given these problems, HMA researchers searched for a new way of ensuring the quality of pavement construction. A more current type of specification, known as the end-result specification (ERS), has become the most popular quality control method. In an end-result specification, desired properties and attributes of the finished product are defined, which are to be met by the contractor. Though more flexibility is allowed during construction, it becomes the contractor’s responsibility to make sure that the desired properties are attained. Fewer problems with acceptance and rejection were identified with end-result specifications. Typical properties

that are tested include yield (thickness), compaction, and smoothness. During the last twenty years, the overall concept of quality control has shifted from the traditional materials-and-methods specification to an end-result specification. The transition has been slow, however, due to lack of support and resistance to change by contractors and agencies.

The third type of specification is a statistically based quality control specification usually coming in the form of an extension of an end-result specification. It uses specifications and acceptance procedures based on random sampling and statistical probability. It also establishes allowable variations that are typical of the various components of the construction process. Technical weaknesses of the end-result and statistical end-result specifications, however, began to reveal holes in the new methods. The major criticism of the end-result approaches was that they did not necessarily measure characteristics related to the performance of the pavement, as they contain no definitive criteria for identification of the performance characteristics. ERS are unable to quantify substantial compliance or to determine price adjustment factors that relate to reduced or enhanced quality (Smith, et al, 1998). While a pavement may meet statistically determined requirements, the core of the specifications was still dependent upon what had worked well in the past. Thus, the push towards performance specifications began.

The final type of specification is performance based. This “specification of the future” is not yet fully implemented in the United States. In this type of specification, the agency is concerned only with the performance of the final product, while giving less emphasis to methods of construction and the materials used. Test methods, which are based on estimating the actual performance of the in-service pavement, would determine the acceptability of the product and possibly determine a level of pay. To finalize these test methods, the relationships between certain hot-mix asphalt properties and the performance of the pavement must first be established.

The criteria for performance-related specifications (PRS) are currently being developed from data collected on long-term pavement performance projects, and accelerated loading facilities such as WesTrack in Nevada that will be discussed later. The goal of PRS is to identify the level of quality providing the best balance between cost and performance (Smith, et al, 1998). An appropriate PRS will include a statistical quality assurance program based on predicted pavement performance, will be enforced by a pay adjustment penalty for work that does not meet agency minimum limits, and will include a bonus for work exceeding some limit. Performance in the system will be defined by a numerical value, such as equivalent single-axle loads (ESALs) to some level of present serviceability index (PSI), or repetitions to other measures of performance like smoothness (Shook, et al, 1993). However, it may be some time before performance based test methods are used. Fortunately, the foundation is already under way for PRS on pavement performance projects such as WesTrack and statistically based research investigating possible relationships for predicting pavement stress, pavement distress, and pavement performance from particular combinations of predictors that represent traffic, environment, roadbed, and structural conditions (Shook, et al, 1993).

### **Superpave and NCHRP 409**

For years, the Arkansas Highway and Transportation Department (AHTD) has used a “statistically based” QC/QA end-result type of specification. Currently for hot mix asphalt construction, the contractor tests a random sample in each subplot for quality control, and the state tests a random sample in each lot for quality assurance. A lot consists of 3,000 tons, and a subplot consists of 750 tons. Thus, there are four sublots within each lot. This comprises what is known as a stratified random sampling procedure, such that a random sample is obtained within a specified section of material.

Though the current QC/QA system is statistically based, it was developed for the Marshall method of mix design. Bruce Marshall, a former bituminous engineer with the Mississippi Department of Transportation, formulated the concepts of the Marshall mix-design method. Through additional research and correlation studies, the U.S. Army Corps of Engineers improved upon and added to Marshall's test procedure eventually formalizing the process into a set of mix design criteria (MS-2, 1995). Empirical in nature, the Marshall mix-design method has served the asphalt industry well; however, its lack of flexibility to innovation and the current push towards PRS probably will mean its downfall.

With PRS development in mind, the Strategic Highway Research Program (SHRP) produced a new asphalt mixture design procedure called Superpave (Superior Performing Asphalt Pavements). The development of Superpave stemmed from a federally funded research effort to improve the performance and durability of roadways across the United States. The Superpave system incorporates performance-based asphalt materials characterizations with specified environmental conditions to improve performance by controlling rutting, low temperature cracking, and fatigue cracking (SP-2, 1996). Superpave does not represent the final evolution into a performance-based specification, but it does illustrate a step in the right direction. Nationally, there was concern that the then-current quality control procedures might not be applicable to Superpave designs as they had been originally developed under the Marshall system.

In order to investigate this potential problem, the National Cooperative Highway Research Program (NCHRP) performed a research project designated NCHRP 9-7. The results of this study were published in NCHRP Report 409. This publication outlines a quality control plan for Superpave mixtures. This plan is currently being evaluated by several states. The

proposed plan for QC/QA involves a rigorous testing effort, particularly for field verification of a mix. For example, during field verification of a mix, the contractor and the state take random samples from each of five 100-ton sublots. Tests are then performed for gradation, asphalt binder content, maximum specific gravity, gyratory compaction, bulk specific gravity, air void content, voids in the mineral aggregate, and the slope of the gyratory compaction curve. At least ninety percent of the results must fall within the limits specified for the test based on the original mixture design.

During actual construction, the contractor is required to continue testing a random sample from each subplot, while the state tests a random sample from each lot. For density testing, NCHRP defines a lot as a pavement section 5,000 feet long and 12 feet wide. This lot is then divided into a minimum of five sublots. For most other tests, a lot size of 1,000 tons with a minimum of five sublots per lot is recommended. In a recent survey of state agencies, it was noted that current typical lot sizes range from 500 tons to 4,000 tons. Also, most states specify five sublots per lot, while some specify only 4 sublots per lot. The NCHRP Report 409 also recommends that the contractor utilize control charts for the monitoring of the construction process. Control charts are simply to be used as a tool for graphing trends and alerting the contractor when adjustments to the process may be necessary.

Overall conclusions of the study were that quality control and quality assurance limits should be based on test variances determined during field verification. While this appears to be statistically valid, a contractor with greater initial variation could be allowed a greater variation during actual construction than a contractor with minimal initial variation. This issue should be resolved. The benefits of a good quality control specification are many. An adequate, yet not excessive, amount of testing should be required in order to ensure maximum quality of our

state's highways. In particular, the need for and use of quality control cannot be overemphasized for Superpave. Quality must be built into Superpave; it cannot be tested or inspected into the mix (Cominsky, et al, 1998).

### WesTrack

Striving towards the development of performance-related specifications for hot-mix asphalt (HMA), the Federal Highway Administration (FHWA) sponsored the WesTrack project. A team made up of consulting firms, universities, and road contractors ran the experiment. As defined by the research contract, the WesTrack project had two primary objectives. The first objective was to promote the advancement of performance-related specifications (PRS) for HMA pavement construction by assessing the impact on performance of variations in materials and construction properties such as asphalt contents, air voids, and aggregate gradation from design values in a large-scale, accelerated field loading test. The second objective of the project was to provide an early field verification of the Strategic Highway Research Program (SHRP) Superpave mixture design procedures (Seeds, et al, 1997).

In the past, HMA mix design and construction procedures have tended to be more of an art form than scientifically based, relying greatly on the expertise of engineers, mix designers, and contractors. The move from method-related specifications (MRS) to end-result specifications (ERS) and now to performance-related specifications (PRS) has been a tedious and demanding process on all parties involved in the production of HMA. The current push towards the development and implementation of performance-related specifications has taken place for a number of reasons. A PRS system provides a way to equitably reward or penalize contractors for the quality of the constructed pavement. A valid set of PRS focuses on the actual material properties and construction practices that have the most profound effect on the long-term

performance of the pavement. At WesTrack, researchers attempted to acquire the performance data necessary to develop the connections between these properties/practices and performance. After the completion of the loading period, the WesTrack team focused on utilizing the collected performance data to either verify or calibrate existing performance prediction relationships (Seeds, et al, 1997). While some progress has been made, research remains ongoing.

## **RESEARCH OVERVIEW**

For years, AHTD has used a “statistically based” QC/QA end-result type of specification. Currently for hot mix asphalt construction, the contractor tests a random sample in each subplot for quality control, and the state tests a random sample in each lot for quality assurance. A lot consists of 3000 tons, and a subplot consists of 750 tons. Thus, there are four sublots within each lot. This actually comprises what is known as a stratified random sampling procedure, such that a random sample is obtained within a specified section of material.

Using the current stratified random sampling procedure, a new specification was sought, which would include criteria for materials and construction acceptance based on both the average (mean) of measured properties – identical to the current AHTD system – and the range of those properties. In this manner, the proposed system establishes controls on the acceptable variability of the materials and construction process. In developing such a specification, one key issue must be fully addressed: the establishment of an initial estimate of the variability of hot-mix asphalt production/construction in Arkansas. Any statistically-robust acceptance specification must be based on some estimate of variability.

A field testing effort was completed to provide data necessary for making this initial estimate. A description of this effort follows in subsequent sections of this report. A new

specification was developed based on ‘quality levels’ estimated through the testing effort. The new specification contains estimates of both the material properties and of production and construction processes.

## **ESTABLISHMENT OF CONSTRUCTION VARIABILITY**

The primary objective of TRC-0001 is to develop a new QC/QA system for Arkansas (or refine the existing system) if necessary. An initial requirement to accomplish this objective was to establish or estimate typical HMA construction variability within the state, for the four primary “pay factors” that are used to control HMA quality: air voids (VTM), voids in the mineral aggregate (VMA), binder content ( $P_b$ ), and field density, expressed as percent compaction ( $\%G_{mm}$ ). This section describes the testing program and data analysis undertaken to develop an estimate of typical HMA construction variability in Arkansas. However, prior to discussing the details of the Arkansas study, it is useful to briefly review the current Arkansas system.

In Arkansas, an HMA project is divided for quality control purposes into lots, each equaling 3,000 tons of mix. These lots are further subdivided into four equal sublots containing 750 tons each. The current specification requires the contractor to randomly sample each of the four sublots within a given lot while the AHTD inspector randomly samples the lot. The contractor’s results represent quality control, while the agency results represent quality assurance. Four mixture/construction properties are used to determine pay: binder content, air voids, VMA, and field density. The bulk specific gravity, maximum theoretical specific gravity, and binder content of the mix are determined from mix sampled at the HMA plant. Field density is measured after compaction and prior to traffic placement.

The acceptance and pay determination for HMA is based on the average of the five tests performed within the lot – four contractor tests and one agency test. In addition, the single test performed in a subplot by the contractor is used for acceptance of the subplot. A lot average within compliance limits results in full payment for the contractor. A lot average within the reduction limits results in a reduced payment for the contractor, while a lot average outside the rejection limits results in no payment to the contractor. Sublots outside the specification limits are also rejected. When lots or sublots are rejected, they must be removed at the cost of the contractor and replaced with a new batch of acceptable mix.

The adjustment of the contract price of a lot is reduced by a given percentage, which depends on the property being measured, for each deviation outside the compliance limits. A deviation represents 0.1 percentage point. This reduction in payment is cumulative for all four of the pay factors measured for a given lot. If a reduction in the contract price for a lot is warranted, adjustments to the lot price are made in the following manner:

- Binder Content: lot price reduced by 12 percent per deviation outside the specification limits (up to a maximum of 3 deviations)
- Air Voids: lot price reduced by 10 percent per deviation outside the specification limits (up to a maximum of 5 deviations)
- VMA: lot price reduced by 10 percent per deviation outside the specification limits (up to a maximum of 5 deviations)
- Field Density: lot price reduced by 4 percent per deviation outside the specification limits (up to a maximum of 10 deviations)

## **Testing Program**

The first major objective of the research involved the determination or estimation of typical HMA construction variability in Arkansas. To this end, a sampling and testing plan was developed and executed to generate data relating to the variability of the HMA factors used for payment – Air voids, VMA, binder content, and field density. Projects were chosen for the sampling plan carefully, in hopes of obtaining a relatively broad range of perceived “quality” in construction.

A total of six asphalt projects were examined. The projects ranged in size from approximately 10,000 tons (just over three lots) to 35,000 tons (just under twelve lots) of HMA surface mix. The projects will be referred to as: Black Rock (BLA), Harrison (HAR), Little Rock (LIT), Prescott (PRE), Russellville (RUS), and Texarkana (TEX). Each of these projects were overlay projects – the most common type of HMA construction currently performed in Arkansas. For this study, only 12.5 mm Superpave coarse-graded surface courses were sampled and tested. A total of four sampling times were used in each project. Two consecutive sublots were chosen from within each of two randomly chosen lots for sampling. Specific locations for sampling within the sublots were determined using the random number technique contained in the Arkansas specification.

Hot-mix asphalt specimens were sampled from truckbeds immediately following loading. Three operators – the contractor, the agency (AHTD), and the research team (UA) – collected materials from the same truck. Contractor and AHTD technicians tested their material on site immediately following each sampling event. Due to logistical difficulties UA personnel were forced to transport samples to the testing laboratory at the University in Fayetteville, which required reheating specimens for testing. Field density was determined using cores cut from the

compacted asphalt mat. Core locations were determined by tagging the truck sampled at the plant and, insofar as possible, marking the location on the roadway corresponding to that particular load.

Each operator performed tests required by Arkansas QA/QC specifications, including the bulk specific gravity ( $G_{mb}$ ), the maximum theoretical specific gravity ( $G_{mm}$ ), and the asphalt binder content ( $P_b$ ) of the mix. All tests were performed according to applicable AASHTO and AHTD specifications, e.g. AASHTO T-166 (bulk specific gravity), AASHTO T-209 (maximum specific gravity), and AHTD 449 (binder content by nuclear methods). Arkansas specifications require only two HMA specimens to be compacted for determining air voids and VMA. For this research effort each operator compacted six specimens. Each operator, using two cores each, determined field density – which typically only requires one field core per subplot.

### **Data Analysis**

The ultimate goal of the sampling and testing program was to determine construction variability. Most agencies assume that construction data follows a normal distribution. The first analysis of the project data, then, sought to establish the normality of the data obtained from the sampling and testing program to evaluate the suitability of the data for its intended purpose.

Unfortunately, testing variability is one component of the overall variability exhibited by a field data set. The second analysis performed on the testing data sought to establish the effect of testing operator on measured HMA properties. It was hoped that, statistically, no significant differences in testing data would be found among the three operators – so that this potential source of variability could be dismissed or at least accounted for in subsequent analyses. A final analysis performed on the data related to the actual variability of each of the “pay factor” properties, as estimated by the standard deviation of the property tested (or calculated from test

data). This measure of variability will potentially form the basis for a revised pay factor schedule for Arkansas HMA construction.

### Data Normality

The first analysis involved the determination of whether the data exhibited characteristics consistent with a normal distribution. This determination is actually quite important in terms of the suitability of the data for providing the basis of a QA/QC system. Most, if not all, agencies using statistically-based QA/QC systems assume construction data and mixture characteristics are normally distributed. This assumption is certainly not required to create a control system, but greatly simplifies the calculations involved and the understanding of the system by those charged with the system's implementation.

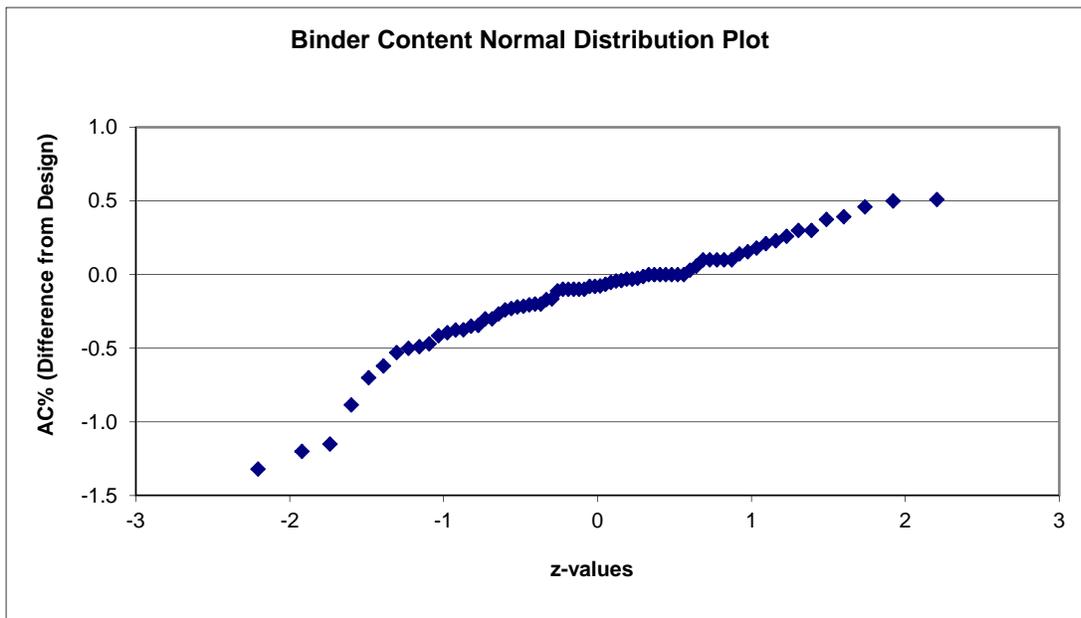
A descriptive technique for checking normality was used. Normal probability plots were created for each of the four critical properties (binder content, air voids, VMA, and field density). In a normal probability plot, the observations in the data set are ordered from smallest to largest and then plotted against the expected z scores, or z-values, of the observations calculated under the *assumption* that the data are from a normal distribution. When the data are, in fact, normally distributed, a linear trend will result. A nonlinear trend in the normal probability plot suggests that the data are nonnormal.

For this analysis, the results from the UA, the AHTD, and the contractor were combined into one data set for each property. Figures 1-4 are the normality plots for binder content, air voids, VMA, and compaction, respectively. A general linear trend is evident in the data, suggesting the data is normally distributed. Based on these analyses, the data generated by the

sampling and testing program executed for this project was found to represent a population of results that follow a normal distribution.

### Operator Variability

An analysis of variance (ANOVA) was used to determine if significant differences existed among operators. If significant differences were indicated, a Duncan's Multiple Range Test (DMRT) was used to establish which data set(s) caused the difference. Table 1 summarizes the results of the ANOVA/DMRT analyses. No significant difference was indicated in the data for cases in which operators are shown as having the same "ranking letter" (A or B). In those cases in which a significant difference was noted, the ranking letter is different.



**Figure 1. Binder Content Normal Distribution Plot**

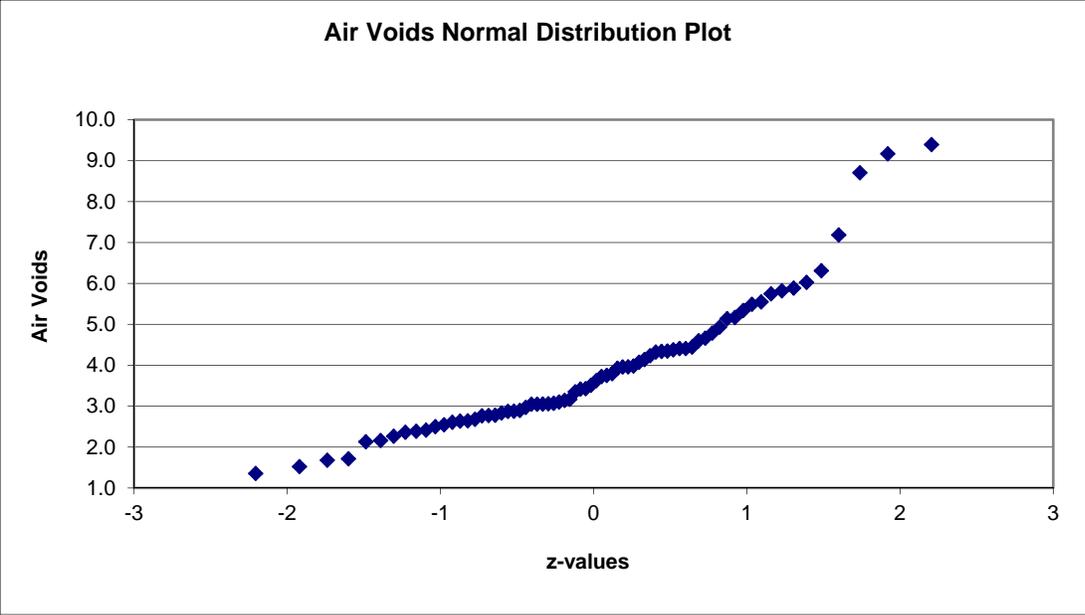


Figure 2. Air Voids Normal Distribution Plot

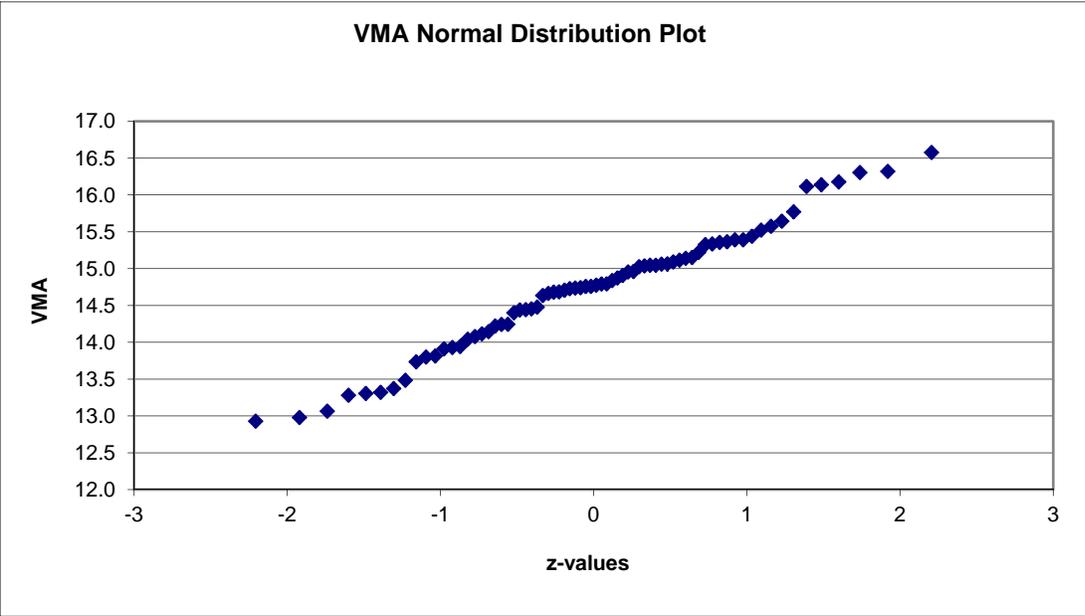
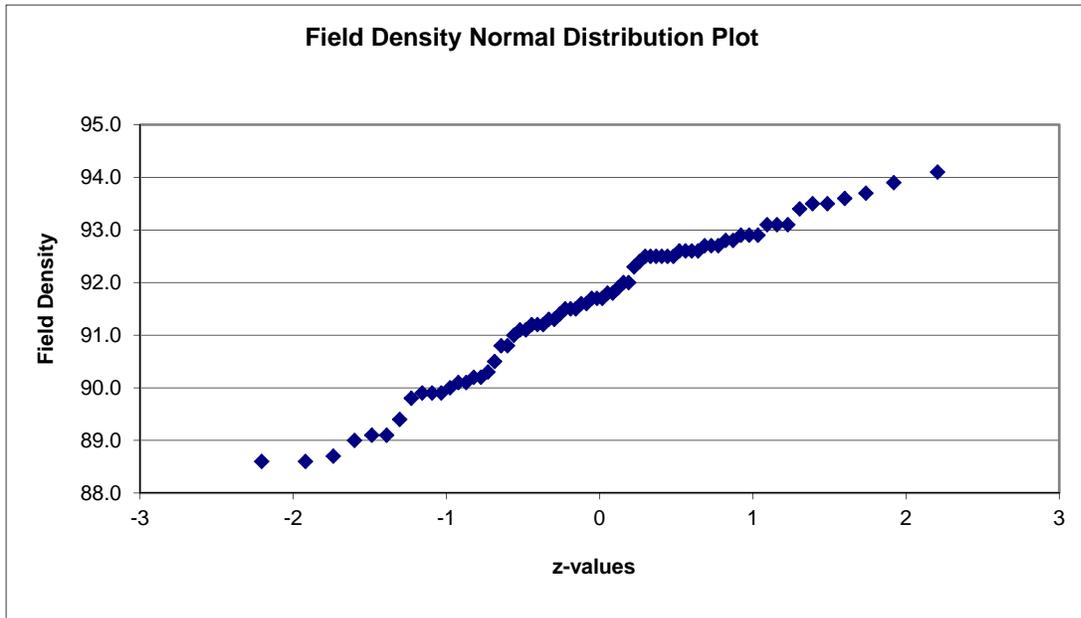


Figure 3. VMA Normal Distribution Plot



**Figure 4. Compaction Normal Distribution Plot**

Overall, a reasonable conclusion based on the results shown in Table 1 is that there is no consistent difference among operators for these test results. In only eight of twenty-four cases (six projects times four properties) a significant difference was exhibited among at least two of the operators. It is noted that in six of these eight instances, both binder content and air voids were found to be significantly different; it could be argued that, if all other mix characteristics remain relatively constant, changes in air voids will closely track changes in binder content. Therefore, it is not surprising that such a pattern exists in Table 1. Based on the overall results of the analyses, the project team concluded that for practical purposes all data from all operators could be considered to be similar, and could be combined for further analyses.

**Table 1. Summary of Results from Statistical Analyses of Project Data**

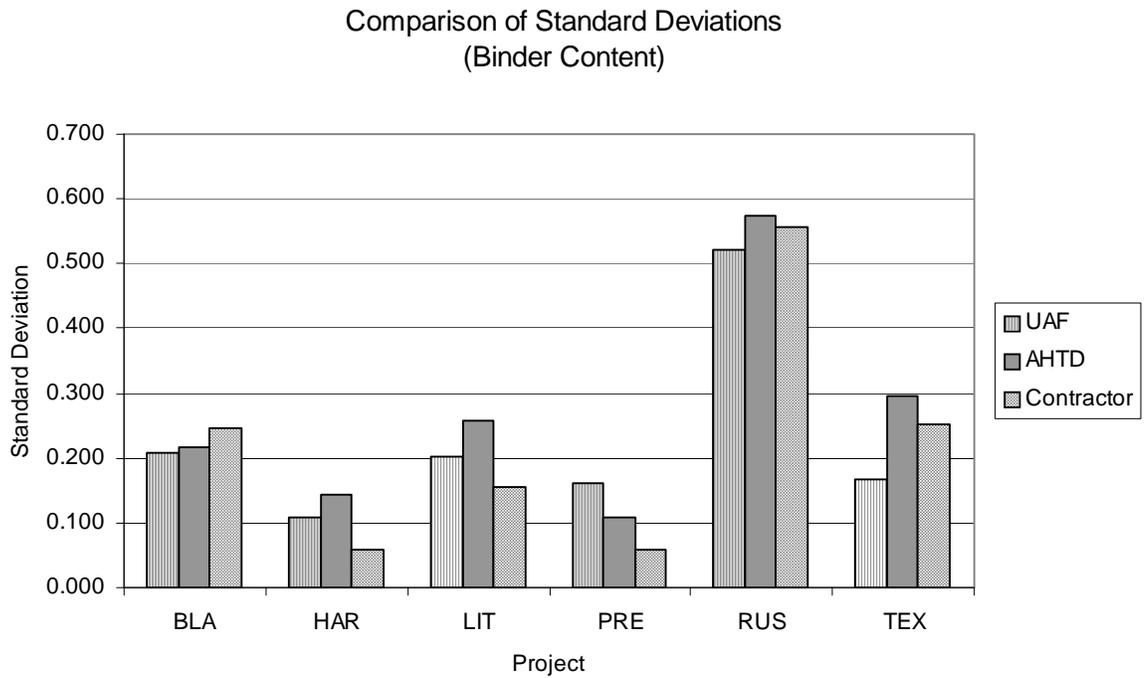
Property	Project	Significant Difference?	Statistical Ranking <sup>1</sup>		
			UA	AHTD	Contractor
<b>Binder Content</b>	<b>BLA</b>	YES	A	B	B
	<b>HAR</b>	YES	A	B	B
	<b>LIT</b>	No	A	A	A
	<b>PRE</b>	No	A	A	A
	<b>RUS</b>	No	A	A	A
	<b>TEX</b>	YES	A	B	A
<b>Air Voids</b>	<b>BLA</b>	YES	A	B	B
	<b>HAR</b>	YES	A	B	A
	<b>LIT</b>	No	A	A	A
	<b>PRE</b>	No	A	A	A
	<b>RUS</b>	No	A	A	A
	<b>TEX</b>	YES	AB	A	B
<b>VMA</b>	<b>BLA</b>	No	A	A	A
	<b>HAR</b>	YES	A	B	A
	<b>LIT</b>	YES	A	B	AB
	<b>PRE</b>	No	A	A	A
	<b>RUS</b>	No	A	A	A
	<b>TEX</b>	No	A	A	A
<b>Field Density</b>	<b>BLA</b>	No	A	A	A
	<b>HAR</b>	No	A	A	A
	<b>LIT</b>	No	A	A	A
	<b>PRE</b>	No	A	A	A
	<b>RUS</b>	No	A	A	A
	<b>TEX</b>	No	A	A	A

<sup>1</sup>Ranking of “A” or “B” – identical rankings indicate no statistical difference in data set

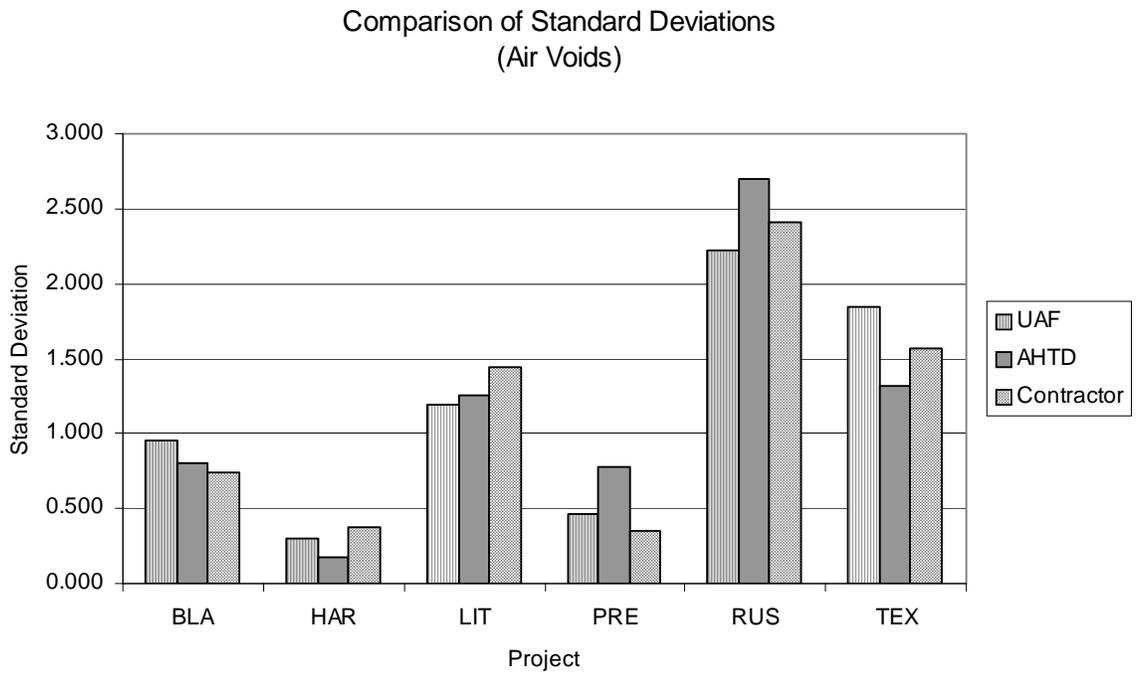
## HMA Properties

Having established that the research data could be approximated by a normal distribution, the final analyses related to an examination of the mean and standard deviation of each property. The mean and standard deviation for each property at each of the four sampling times within the six projects were calculated. For each of the critical properties, the standard deviations were plotted to compare the relative variability of the different operators and the various projects. Figures 5 through 8 illustrate the data. It should be noted that for binder content, the data is expressed in terms of “percent away from the target (design) binder content”. In this manner, binder content data can be appropriately compared across projects having dissimilar design binder content.

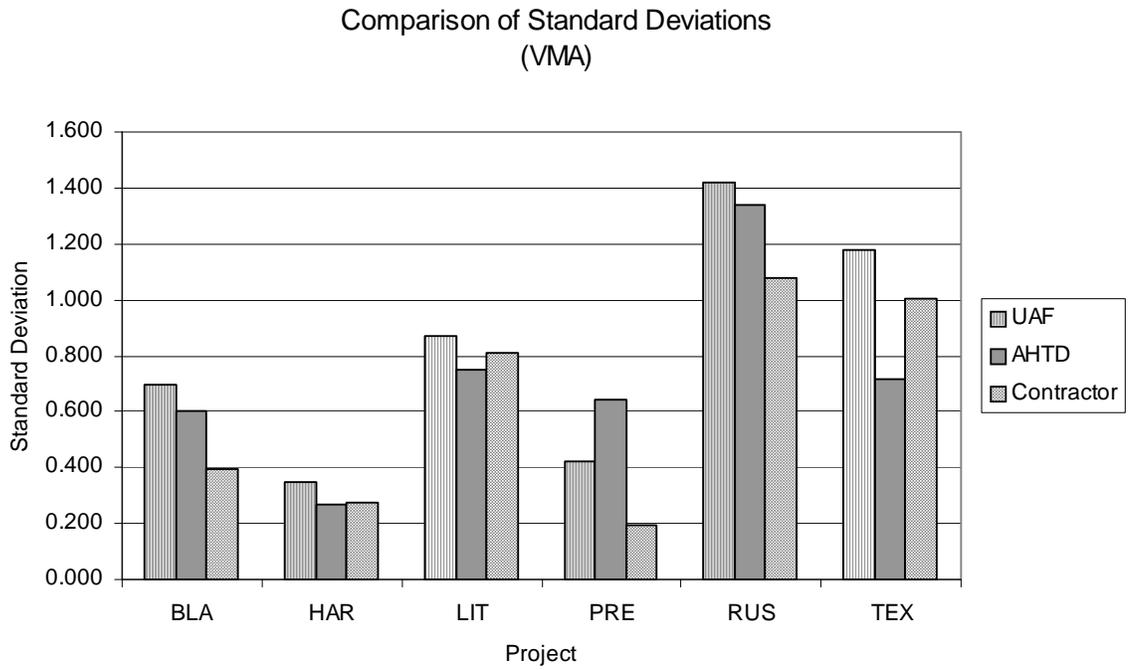
Figures 5 through 8 graphically illustrate the results of the ANOVA/DMRT analysis – namely, there is no significant effect of operator on test results. In addition, no apparent bias is present in the data – that is, no single operator was consistently high or low compared to the other two. The agency (AHTD) testing technician and the contractor technician changed with each project, while the UA operator was consistent throughout all projects. Given this fact, the relative consistency of standard deviations recorded around the state among operators on a given project lends support to the assumption that testing variability is relatively constant for any given location. This assumption allows the item of “testing variability” to be basically ignored in formulating overall variability. The range of variability shown in each of Figures 5 through 8 suggest that a goal of the testing program was achieved, namely, a good cross-section of “quality” was sampled during the program.



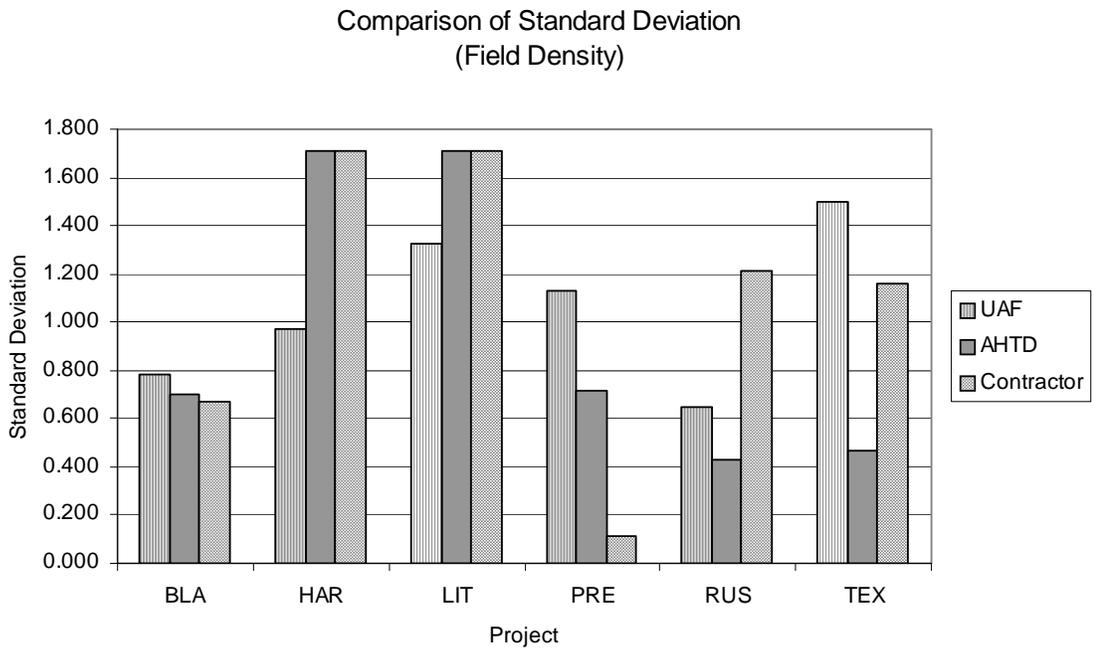
**Figure 5. Standard Deviation for Binder Content**



**Figure 6. Standard Deviation for Air Voids**



**Figure 7. Standard Deviation for VMA**



**Figure 8. Standard Deviation for Field Density**

The results shown in Figures 5 through 8 also reveal a pattern of certain projects having lower standard deviations than others. As a case in point, PRE consistently had a lower standard deviation than the other projects for nearly all of the properties. This concept was expanded, and the data was grouped into three qualitative descriptions of apparent “quality” as defined by the standard deviation of test data: high quality (projects HAR and PRE), medium quality (projects LIT and BLA), and low quality (projects TEX and RUS). It should be noted that while the selection of individual projects into various “quality” categories was based primarily on the standard deviation of test results, the mean value of test results also was considered. In other words, a project not only had to show precise (or consistent) data, but also show accurate data to be considered for higher quality status.

All data recorded for each of the two projects in a given quality category were grouped to calculate an overall mean and standard deviation for each pay factor property. It had been previously established that individual operator data for a given project could be considered statistically similar to allow for grouping; it was assumed that for a given subjective quality “category”, project data could be likewise grouped. Table 2 summarizes the mean and standard deviation value calculated for each quality level.

As expected, moving from a high quality to a low quality category for each property increases the standard deviation for that property. It is interesting to note that for most properties, moving from a higher to lower quality category also moves the mean value farther away from the “target” mean, i.e. design air voids, design binder content, etc. In other words, not only does testing data become less precise at lower quality levels, but also less accurate.

**Table 2. Summary of Material Properties by Quality Level**

	Quality Level					
	High		Medium		Low	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
<b>Binder Content (%)<sup>1</sup></b>	0.06	0.184	0.21	0.251	0.33	0.413
<b>Air Voids (%)</b>	3.58	0.649	3.09	0.768	5.02	2.097
<b>VMA (%)</b>	14.92	0.346	14.32	0.589	14.94	1.136
<b>Field Density (%G<sub>mm</sub>)</b>	92.57	0.790	91.82	0.959	90.43	1.313

<sup>1</sup>Binder content expressed as percent away from design binder content

From a purely qualitative perspective, there are some disturbing trends in the data presented in Table 2. Standard deviations recorded for both the medium and low quality projects appear quite large. In many cases, the standard deviation for a given HMA property is greater than the allowable field tolerance for that property. High standard deviations should be a cause for concern for both the contractor (from a payment viewpoint) and the agency (from a quality viewpoint). An interesting note is that the Arkansas specification does not explicitly consider the variability (sublot-to-sublot or lot-to-lot) of test properties except for cases of extreme “swings” in the data. Payment decisions are based overwhelmingly on the property value averaged from five individual tests. The data generated by this project suggest that some check on the

variability of HMA property data might be warranted in a revised or new construction specification.

Another disturbing, but unfortunately not surprising, data trend involves field density. Variability numbers related to field density seem to be reasonable; however, the mean values shown in Table 3 indicate that even “high quality” projects experience difficulty in obtaining proper field compaction (Arkansas specifies a minimum of 92 percent compaction). It is apparent from the data generated on this project that measures to help ensure proper compaction are warranted in a revised or new construction specification.

## **DEVELOPMENT AND IMPLEMENTATION OF A QA/QC SPECIFICATION**

As stated, the primary objective of Project TRC-0001 is to develop a new (or refine the existing) QA/QC specification which allows some control on materials and construction variability while maintaining an acceptance system which has produced good quality pavements in Arkansas. In developing this specification, one key issue was fully addressed: the establishment of an initial estimate of hot-mix asphalt production/construction in Arkansas – as described in previous sections of this report. For the purposes of developing and implementing a proposed QC/QA specification, it is recommended the ‘medium’ quality level be used. As hot-mix asphalt production / construction data continue to be collected, the quality level – as expressed by the standard deviation of the property tests – may be re-evaluated.

### **Mean-and-Range Approach**

In an attempt to ‘balance’ the use of the existing AHTD specification against the implementation of new procedures, it was decided to develop a ‘mean-and-range’ system for QA/QC purposes. In such a system, acceptance is governed by the mean of a given property (in keeping with the

current system) and an appropriate range for the property – which sets limits on the variability of the property.

For the mean, it is proposed to set acceptance limits for the mean at 1.5 times the standard deviation; in the initial implementation of the proposed system, the standard deviation is a function of the ‘medium quality level’ (see Table 2) standard deviation for each property, modified by the sample size used in Arkansas (five):

$$s_n = \frac{s}{\sqrt{n}}$$

Where:  $s_n$  = standard deviation based on a sample size of n  
s = historic standard deviation  
n = sample size (e.g. 5)

The range of values is calculated as the largest value in the lot minus the smallest value in the lot. Acceptance limits for the range are proposed as two times the historic standard deviation of the property. The initial estimate of standard deviation is proposed as the ‘medium quality level’ as shown in Table 2.

## **Validation**

A key issue was identified for implementation efforts: validation of the proposed system. In order to gauge the effect of implementation, a limited simulation study was performed, using field QC/QA data generated from five projects constructed under the Interstate Rehabilitation Program (IRP, commonly referred to as “bond jobs”) conducted in Arkansas between 2000 and 2005. Overall, the proposed specification produced more lots in which a negative price adjustment was indicated – however, the total number of such adjustments was not significant, nor were the magnitude of the adjustments. In addition, it appeared that field compaction would be the property most affected by the proposed specification. The major value of this small effort

is to suggest that implementation of the proposed specification should not result in catastrophic negative price adjustments to contractors performing quality work (note that, in general, the ‘bond jobs’ on the IRP would typically feature relatively higher-quality production/construction).

## RECOMMENDATIONS

It is recommended to implement the mean-and-range system described. However, a major feature of the implementation effort related to the proposed specification is large-scale field validation. To fully understand the total effect of implementation, large production/construction data sets – representing a wide variety of projects, materials, contractors, etc. – should be analyzed using current QC/QA acceptance methods and the proposed specification. From such an effort, refinements to the proposed system might be made to balance production and construction consistency with necessary levels of compliance.

A recommended implementation plan for the proposed specification is detailed in the listing with follows:

- For a given construction season, collect field QC/QA testing data (via Site Manager) and apply **current** AHTD acceptance specifications. Concurrently, apply the proposed QC/QA specification to the data set.
  - Establish initial acceptance limits for the mean and range terms using the historic standard deviation values representing “medium” quality level (Table 1).
  - Track subplot and lot acceptance (100% pay) for both the current and proposed specifications.
  - Calculate payment reductions for both the current and proposed specifications.
    - For the Mean (proposed specification), calculate payment reductions using reduction factors in the current specification (Section 4.10 in the *Standard Specifications...*).
    - For the Range (proposed specification) calculate payment reductions using reduction factors in the current specification (Section 4.10 in the *Standard Specifications...*).

- Analyze all data and comparisons between the current and proposed specifications during the period October to December. Specific items of interest in the data and in the data comparisons include:
  - Standard deviation of all hot-mix asphalt related properties – to be used in potential subsequent refinements of the proposed specification
  - Number of lots judged out-of-compliance for each pay factor
    - For the proposed specification, note whether the determination of out-of-compliance was caused by Mean and/or Range
  - Magnitude of payment adjustment for each pay factor
  - Trends in differences between the current and proposed specifications related to:
    - Job size (tonnage)
    - Geographic region
    - Type of mix (surface, base, binder)
    - Overlay versus new construction
- Host a QC/QA summit in December to present the data analysis. Identify necessary refinements to the proposed specification.
- Contract with the Center for Training Transportation Professionals (CTTP) to develop and present QC/QA training to AHTD and industry personnel. Initial training courses should be targeted for a January time frame. Refine course as necessary.

## REFERENCES

AASHTO Highway Subcommittee on Construction. Manual for Quality Assurance. Issued February 1996a.

AASHTO Highway Subcommittee on Construction. Quality Assurance Guide Specification. Issued February 1996b.

Abdulshafi, O., Fitch, M., Kedzierski, B., Powers, D. "Laboratory Optimization of Asphalt Concrete Intermediate Course Mixes to Improve Flexible Pavement Performance." Transportation Research Record. Issue 1681, pg. 69-75, 1999.

Arkansas State Highway and Transportation Department. Standard Specification for Highway Construction, Section 410. 1996.

Arkansas State Highway and Transportation Department. Supplemental Specification: Asphalt Concrete Hot Mix Base, Binder, and Surface Courses (SS-400-1). Pg. 1-15. Rev. March 30, 2000.

Buttler, W., Harrell, M. "Development of End-Result and Performance-Related Specifications for Asphalt Pavement Construction in Illinois." Crossroads 2000: Transportation Conference Proceedings (pg. 195-202). Center for Transportation Research and Education: Ames, Iowa, 1998.

Cominsky, R., Killingsworth, B., Anderson, R., Anderson, D., and Crockford, W. NCHRP Report 409: Quality Control and Acceptance of Superpave-Designed Hot Mix Asphalt. Transportation Research Board. National Academy Press: Washington, D.C., 1998.

Coree, B., Hislop, W. "Difficult Nature of Minimum Voids in the Mineral Aggregate Historical Perspective." Transportation Research Record. Issue 1681, pg. 148-156. 1999.

Coree, B., Hislop, W. A Laboratory Investigation into the Effects of Aggregate-Related Factors of Critical VMA in Asphalt Paving Mixtures. Center for Transportation and Education Iowa State University. June 2000.

D'Angelo, J., Harman, T., Paugh, C. "Evaluation of Volumetric Properties and Gyrotory Compaction Slope for the Quality Control of Hot-Mix Asphalt Production." Annual Meeting of the Association of Asphalt Paving Technologist, 2001.

Epps, A., and Hand, A. "A Comparison of HMA Field Performance and Laboratory Volumetric Sensitivities." Hot Mix Asphalt Conference. Pg. 1-26. 2000.

Epps, J., Committee, Staff, and Consultant Scherocman, J. Hot-Mix Asphalt Paving Handbook. U.S. Army Corps of Engineers. Issued July 1991.

Epps, J., Leahy, R., Mitchell, T., Ashmore, C., Seeds, S., Alavi, S., Monismith, C. “Westrack—The Road to Performance-Related Specifications.” International Conference on Accelerated Pavement Testing. Pg. 1-28. Reno, Nevada, October 18-20, 1999.

Epps, J., Seeds, S., Alavi, S., Monismith, C., Ashmore, S., Mitchell, T. “Westrack Full-Scale Test Track: Interim Findings.” Eighth International Conference on Asphalt Pavements. Pg. 147-164. Seattle, Washington, 1997.

Harvey, J., Popescu, L. “Accelerated Pavement Testing of Rutting Performance of Two Caltrans Overlay Strategies.” Transportation Research Record. Issue 1716, pg. 116-125. 2000.

Hughes, C. and Committee, and Heijl, F. “Glossary of Highway Quality Assurance Terms.” Transportation Research Circular. Number 457, April 1996.

Hughes, C., Simpson, A., Cominsky, R., Pendleton, O., Weed, R., Wilson, T. Measurement and Specification of Construction Quality, Volume 1: Final Report. (Report No. FHWA-RD-98-077). U.S. Department of Transportation: McLean, VA, May 1998).

Jay, M.L. “Achieving Rut-Resistant Pavements.” Hot Mix Asphalt Technology. Vol. 5, no. 5, pg. 43-47. 2000.

Kandhal, P., Foo, K., Mallick, R. A Critical Review of VMA Requirements in Superpave. NCAT Report no. 98-1. January 1998.

Killingsworth, B., and Cominsky, R. Recommendations and Products from NCHRP Project 9-7. Presented at Transportation Research Board 77<sup>th</sup> Annual Meeting, January 11-15, 1998, Washington, D.C.

Mix Design Methods for Asphalt Concrete and Other Hot-Mix Types. The Asphalt Institute. Manual Series no. 2 (MS-2). May 1984.

Mix Design Methods for Asphalt Concrete and Other Hot-Mix Types. The Asphalt Institute. Manual Series no. 2 (MS-2). Sixth Edition. March 1995.

Parker, Jr., F., Hossain, M., Song, J. “Analysis of Quality Control and Quality Assurance Data for Superpave Mixes.” Transportation Research Record. Issue 1712, pg. 25-34. 2000.

Principles of Construction of Hot-Mix Asphalt Pavements. The Asphalt Institute. Manual Series no. 22 (MS-22). January 1983.

Quality Assurance in Pavement Construction. American Society for Testing and Materials. (ASTM Special Technical Publication 709). December 1978.

Rillett, L.R. “Effect of Variability on End Product Specification.” Journal of Construction Engineering and Management. Pg. 139-145. March/April 1998.

Roberts, F., Kandhal, P., Brown, E., Lee, D., Kennedy, T. Hot Mix Asphalt Materials, Mixture Design, and Construction. NAPA Education Foundation: Lanham, Maryland, 1991.

Schmitt, R., Hanna, A., Russell, J., Nordheim, E. “Analysis of Bias in HMA Field Split-Sample Testing.” Hot Mix Asphalt Conference. Pg. 1-30. 2000.

Shook, J., Diaz, M., Stroup-Gardiner, M., Seeds, S. Performance-Related Specifications for Asphalt Concrete—Phase II. (Publication No. FHWA-RD-91-070). U.S. Department of Transportation: McLean, VA, December 1993.

Sinich, T., Levine, D., and Stephan, D. Practical Statistics by Example Using Microsoft Excel. Prentice-Hall: Upper Saddle River, NJ, 1999.

Smith, G. and Committee. NCHRP Synthesis 263: State DOT Management Techniques for Materials and Construction Acceptance. Transportation Research Board. National Academy Press: Washington D.C., 1998.

Superpave Mix Design. Asphalt Institute. Superpave Series No. 2 (SP-2). 1996.

“Superpave: Quality Control is Essential.” Better Roads Magazine. February 1998.

U.S. Department of Transportation, Federal Highway Administration. Performance of Coarse-Graded Mixes at Westrack—Premature Rutting Final Report. Federal Highway Administration: Washington D.C., June 1998.

“A Winning Performance.” Hot Mix Asphalt Technology. Vol. 5, no. 1, pg. 8-10. 1999.