Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-02/4047-1		

4. Title and Subtitle	5. Report Date
INITIAL ASSESSMENT OF TxDOT BINDER QUALITY ASSURANCE PROGRAM	December 2001
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Amy Epps, Clifford Spiegelman, Eun-Sug Park, Edith Arambula, Tazeen Ahmed, and Tatiyana Apanasovich	Report 4047-1
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Texas Transportation Institute The Texas A&M University System College Station, Texas 77843-3135	11. Contract or Grant No. Project No. 0-4047
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Texas Department of Transportation Research and Technology Implementation Office P. O. Box 5080 Austin Texas 78763-5080	Research: August 2000-August 2001
	14. Sponsoring Agency Code

15. Supplementary Notes

Research performed in cooperation with the Texas Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration.
Research Project Title: Analysis and Development of Asphalt Quality Assurance Procedures

16. Abstract

Most state departments of transportation (DOT) maintain asphalt binder quality assurance (QA) programs to ensure that asphalt binders used in the construction of their road system meet specifications stipulated for each project. TxDOT samples and approves asphalt materials at the source and utilizes these materials in highway projects without consideration of possible changes in properties that may occur between production and use during construction. Historic concern and limited recent data indicate that binder properties do change, contributing to difficulties during construction operations and poor performance. The primary goal of this study is to evaluate the current TxDOT binder QA program and recommend revisions as necessary toward improving quality. This interim report documents an initial assessment of the current program based on (1) an extensive information search and review that included two detailed surveys of TxDOT districts and nine other state DOTs and (2) partial results from a comprehensive laboratory testing program that simulated factors that may affect changes in binder properties prior to use. Preliminary recommendations that will be evaluated as the project continues include performing a more detailed and comprehensive evaluation of the current binder QA program and implementing revisions in terms of a total binder quality system under the direction of a quality manager. In this system, the binder QA program would be one tool to improve material quality used during construction. Other recommended tools include mandated quality control (QC) plans for both suppliers and contractors that require special handling for binders, training programs for sampling and testing, general education seminars presenting the goals of the system to all involved, and a round-robin program to establish testing variability. Researchers also recommend development of a user-friendly database for storing binder QA data that includes storage times, storage temperatures, and pavement performance data. They suggest frequent analysis of this data, possibly to set frequency rates for field sampling by binder and supplier.

17. Key Words	18. Distribution Statement
Asphalt, Binder, Quality Assurance, Evaluation	No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	130	

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

INITIAL ASSESMENT OF TxDOT BINDER QUALITY ASSURANCE PROGRAM

by

Amy Epps
Assistant Research Scientist

Clifford Spiegelman
Research Scientist

Eun-Sug Park
Assistant Research Scientist

Edith Arambula
Graduate Assistant Researcher

Tazeen Ahmed
Graduate Assistant Researcher

and

Tatiyana Apanasovich
Graduate Assistant Researcher

Texas Transportation Institute

Report 4047-1
Project Number 0-4047
Research Project Title: Analysis and Development of Asphalt Quality Assurance Procedures

Sponsored by the
Texas Department of Transportation
In Cooperation with the
U.S. Department of Transportation
Federal Highway Administration

December 2001

TEXAS TRANSPORTATION INSTITUTE
The Texas A&M University System
College Station, Texas 77843-3135

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration (FHWA) and the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation, nor it is intended for construction, bidding, or permit purposes. Trade names were used solely for information and not for product endorsement. The engineer in charge was Joe Button, P.E. (Texas No. 40874).

ACKNOWLEDGMENTS

The authors thank the Texas Department of Transportation (TxDOT) and the Federal Highway Administration (FHWA) for their support in funding this research project. Special thanks goes to Darren Hazlett and Jerry Peterson of TxDOT for their efforts in providing technical guidance, support, and direction to this project. Special thanks are also due to Jacob Bell of the Texas Transportation Institute (TTI) for his help in the laboratory and in the field. Thanks also to Amanda Hughes and Sara Khan from TTI for revising TxDOT data into a usable format. Thanks to the many state Departments of Transportation (DOTs) who provided the authors with an understanding of their binder quality assurance (QA) programs and, in some cases, a data set for analysis. The suppliers who donated materials for this project and provided interest and support are also thanked. Finally, thanks to the Center for Asphalt and Materials Chemistry in the Department of Chemical Engineering at Texas A&M University for its help during the laboratory testing program.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

CHAPTER 1. INTRODUCTION

Most state departments of transportation (DOTs) maintain asphalt binder quality assurance (QA) programs to ensure that asphalt binders used in the construction of their road system meet specifications stipulated for each project. These specifications include the binder grade, selected based on environmental and traffic conditions expected over the design life of the project. Therefore, production of a quality asphalt pavement requires that the binder used during construction meet the specifications for the selected grade.

Binder QA programs may require sampling at the production source, during construction, or both. Possible sampling points are shown in Figure 1 as boxes and include the following:

- a storage tank at the production site or refinery,
- a storage tank at a supplier terminal,
- a transfer line to load transports at the production site or refinery,
- a blending line to load transports without intermediate tank storage,
- a transfer line from a transport to a storage tank at the contractor site,
- a storage tank at the contractor site,
- a transfer line from the contractor storage tank to the hot-mix asphalt (HMA) plant for asphalt cements, and
- a spray bar on a distributor truck for liquid asphalts (asphalt emulsions or cutback asphalts).

Programs that do not require sampling during construction, either at the HMA plant or at the project site for spray applications of liquid asphalts, do not consider possible changes in material properties that may have occurred between production and use during construction. Some of these changes may be detrimental in terms of performance or create difficulties during construction operations. Performance problems may surface if changes in material properties render a binder used during construction to have insufficient capacity to resist the primary forms of distress under the environmental and traffic conditions for a specific project. Construction difficulties may arise if, for example, contractors select compaction temperatures based on the specified grade and these temperatures are inadequate in terms of consistency for the actual material used.

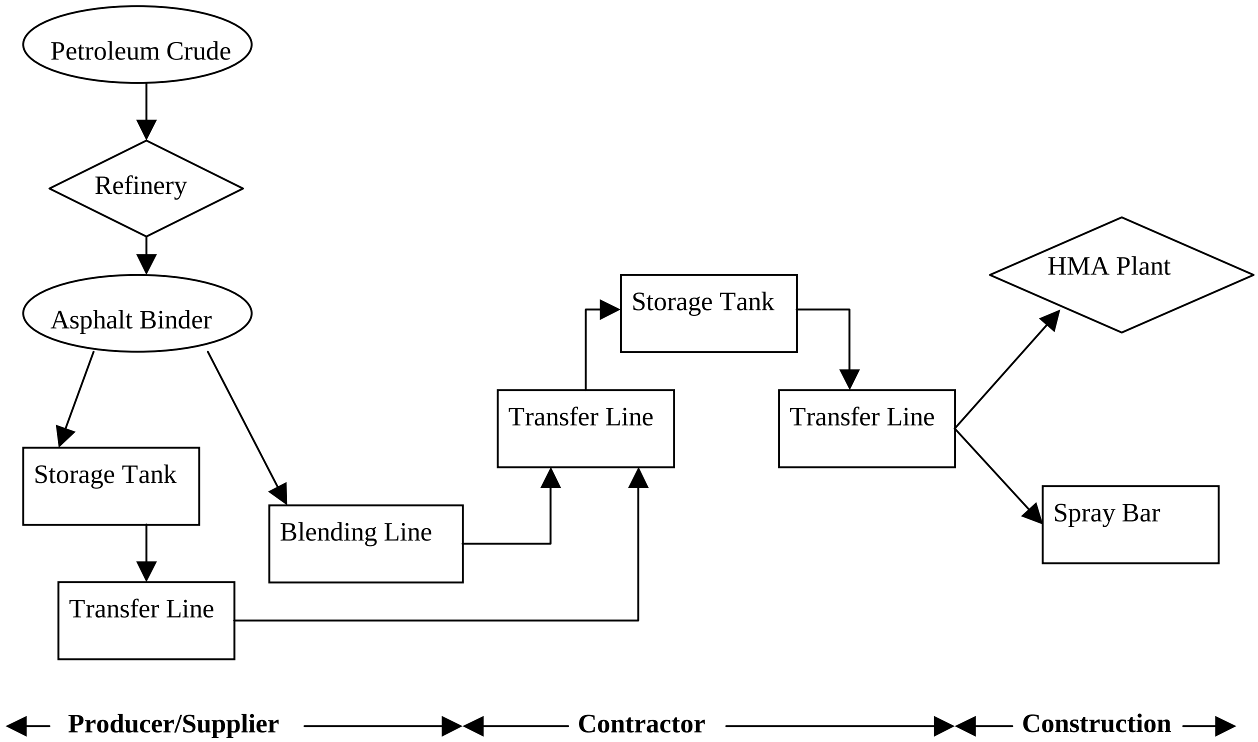


Figure 1. Asphalt Binder from Production to Construction.

TxDOT samples and approves asphalt materials at the source, and these materials are then utilized in highway projects without consideration of possible changes in properties that may occur between production and use during construction. Historic concern and limited recent data indicate that binder properties do change, contributing to difficulties during construction operations and poor performance. The primary goal of this study is to evaluate the current TxDOT QA program for binders and recommend revisions as necessary toward improving quality. This interim report documents an initial assessment of the current TxDOT binder QA program. The final report for this project will contain additional evaluation and recommendations based on an extensive ongoing experimental testing program that includes laboratory simulation of factors that may affect binder properties and actual field sample results.

To evaluate the TxDOT binder QA program, researchers needed an understanding of factors that may cause changes in binder properties between production and use during construction, the effect of these changes on performance, current TxDOT QA practices, and other state DOT QA programs. This report documents the results of an extensive information search and review and the design and partial results from a comprehensive laboratory testing program toward gaining this understanding. The report concludes with preliminary recommended changes to the TxDOT binder QA program that will be evaluated as the project continues.

CHAPTER 2. INFORMATION SEARCH AND REVIEW

Researchers conducted a literature search and review with the assistance of the TTI library staff and completed an extensive survey to accomplish the following goals:

- obtain general definitions of and recommendations for QA programs with an emphasis on binder QA programs,
- identify prospective binder properties directly related to performance that can be measured in a timely manner for use in a QA system,
- identify any performance models that relate off-target values of binder properties to loss of field performance and associated costs,
- identify factors that may cause changes in properties of binders sampled from the source to those sampled just prior to use,
- define the current binder QA program in Texas and its impact on TxDOT districts, and
- define the state-of-the-practice in binder QA programs in Texas and other selected states.

This chapter provides descriptions of the results of each part of the information search and review, including summaries of the relatively small body of literature found and general comparative descriptions of the binder QA programs in Texas and selected states.

QUALITY CONTROL (QC) AND QUALITY ASSURANCE (QA)

General references by A. Mitra, D. Summers, R. Aguayo, and A. Gabor define quality control (QC) and QA and describe the use of statistics to enhance quality and aid in decision making (*1*, *2*, *3*, *4*). QC is generally defined as a system used to maintain a desired level of quality in a product or service. This goal may be achieved through different measures such as planning, design, use of proper equipment and procedures, inspection, and corrective action when a deviation is observed between the product, service, or process output and a specified standard. QA is generally defined as all planned or systematic actions necessary to provide confidence that a product or service will satisfy given needs.

Several people have made significant contributions in the field of QC/QA. W. Edwards Deming may be the most recognized (*5*). He conducted a thriving worldwide consulting practice for more than 40 years with clients that included manufacturing companies, telephone companies, railways, carriers of motor freight, consumer researchers, census methodologists, hospitals, legal firms, government agencies, and research organizations in universities and in industry. He suggested the following 14 points for management that are fundamental to the implementation of any quality program:

- Create and publish to all employees a statement of the aim and the purposes of the company or other organization. The management must consistently demonstrate their commitment to this statement.
- Everyone, including top management, must learn the new philosophy.
- Understand the purpose of inspection, for process improvement and cost reduction.
- End the practice of awarding business on the basis of the price tag alone.
- Constantly and continuously improve the system of production and service, to improve quality and productivity and, thus, constantly decrease costs.
- Institute training on the job.
- Institute leadership.
- Drive out fear, so that everyone may work effectively for the company.
- Break down barriers between departments.
- Eliminate slogans, exhortations, and targets for the work force asking for zero defects and new levels of productivity.
- Eliminate work standards (quotas) on the factory floor. Substitute leadership.
- Create pride in the job being done.
- Institute a vigorous program of education and self-improvement.
- Put everybody in the company to work to accomplish the transformation.

Researchers integrated these 14 points, integral to a successful QC/QA program, into the preliminary recommendations presented in this interim report.

Statistics can be utilized in both QC and QA environments to aid in decision making. Process control charts are used in QC to compare material properties during production with required test values and to determine when a change in the process is required to consistently produce material that meets specifications. Statistics can also be used in this setting to determine if a particular process can produce material that meets specific requirements. Confidence intervals are used in QA to account for material, sampling, and testing variability and to determine when a material fails a single property or multiple properties required in a specification. In this report, results obtained through statistical analysis techniques demonstrate the potential for establishing a binder QA program with continuous improvement and availability of information relevant to decision making toward improving quality.

With regard to binder quality, suppliers and contractors are responsible for maintaining their own QC system. The owner, generally a DOT, defines and maintains the QA system to ensure a binder has all properties required by the specification and related to adequate performance to guard against premature failure. Many states utilize the American Association of State Highway and Transportation Officials (AASHTO) PP26 Standard Practice for Certifying Suppliers of Performance-Graded (PG) Asphalt Binders as a guideline for establishing their QC/QA systems (*6*). This standard defines PG suppliers and their responsibilities in terms of assuring specification compliance. The supplier must submit a QC plan to the agency that details the testing procedures and frequency to assure compliance.

AASHTO PP26 provides guidance for minimum QC plan components and a standard form for reporting data. QC plan requirements include transport inspection guidelines and initial, reduced, and minimum testing frequencies. This standard also provides sampling and laboratory accreditation requirements. If historical compliance is demonstrated, the standard defines an approved supplier certification program that agencies may use to minimize disruption in the construction process. Agency responsibilities outlined in AASHTO PP26 include acceptance of the QC plan, administration of the certification program, and inspection of supplier facilities. The standard also describes provisions for split sample and QA sampling and testing, but it does not specify guidelines for sampling and testing frequencies or specific acceptable tolerances for specification parameters. For reduced testing frequencies in supplier QC plans, the variability of each test is suggested for the tolerance level.

The Northeast Center of Excellence for Pavement Technology (NECEPT) is currently addressing deficiencies in AASHTO PP26 through a pooled-funds study (*7*, *8*, *9*, *10*). These deficiencies include failure to specify sampling and testing frequencies for QA samples, sampling locations for QA samples to account for changes in binder properties that may occur subsequent to production, acceptable tolerances for specification compliance that consider all possible sources of variability, and corrective action for noncompliance. Their goals include development of a QC/QA system that includes multiple components to address these inadequacies. They have developed a binder technician and laboratory certification program, a split sampling program to establish expected testing variability, a QC program for suppliers, a QA program that includes conflict resolution guidelines and payment schedules incorporated in a simulation program that ensures a balance between agency and supplier risk, and a regional database with common specification certificates of analysis to support these programs. Implementation of this entire system is expected in 2002. Researchers will monitor this process to ascertain if further improvements to the TxDOT binder QA program are possible.

BINDER PROPERTIES RELATED TO PERFORMANCE

The recently implemented specification system for binders used in HMA was developed during the Strategic Highway Research Program (SHRP) and utilizes laboratory tests that measure fundamental physical properties that can be directly tied to field performance of asphalt-aggregate mixtures. This system specifies binder properties for unmodified or modified asphalt cements used in HMA to ensure safety, provide for ease in pumping and handling, guard against excessive aging, and mitigate the three major forms of distress in asphalt concrete pavements: permanent deformation, fatigue cracking, and thermal cracking (AASHTO MP1) (*11*). The PG binder specification system was developed based on unmodified asphalt cements, but the equipment and form of the specification is expected to be applicable to modified binders. Ongoing research under National Cooperative Highway Research Program (NCHRP) Project 9-10 is exploring the applicability of the PG specification to modified binders and assessing what changes are needed to support evaluation of these materials (*12*).

The properties specified in the PG system are consistent for all binders, only the temperatures at which these properties must be met vary. Each property specified is measured using a characterization test described in this section. For a specific project, predicted pavement temperatures and traffic conditions determine the binder grade needed for satisfactory performance.

The characterization tests required to specify a binder measure physical properties related to pavement performance directly through engineering principles. A historical database of past performance is not needed to use test results as a prediction tool, although validation is required and has been completed in terms of laboratory mixture performance tests (*13*). A characterization test related to rutting performance is conducted on binder that has been short-term aged in the Rolling Thin Film Oven (RTFO) (American Society for Testing and Materials (ASTM) D2872), to simulate the critical state for this type of distress after mixture production and construction (*6*, *14*). Tests related to cracking performance are conducted on binder that has been short-term aged in the RTFO and long-term aged in the Pressure Aging Vessel (PAV) (AASHTO PP1) to simulate the critical state for both fatigue and thermal cracking (*6*, *14*).

A dynamic shear test (AASHTO TP5) characterizes binder resistance to rutting and fatigue cracking (*6*). This test is used to evaluate the time- and temperature-dependent behavior of binders at intermediate and high temperatures. A controlled stress dynamic shear rheometer (DSR) measures the viscoelastic behavior of the material in terms of complex shear modulus (G^*) and phase angle (δ). The DSR applies a sinusoidal variation in shear stress (τ) to a thin film of binder at a frequency of 10 rad/s, and the resulting sinusoidal variation in shear strain (γ) is measured (*11*). The rheologic parameters are computed as follows (*11*):

$$G^* = \frac{\tau_{\max} - \tau_{\min}}{\gamma_{\max} - \gamma_{\min}}, \text{and}$$
$$\delta = \Delta t = \text{time lag of strain response,}$$

where:

τ_{max}	=	maximum value of applied sinusoidal shear stress,
τ_{min}	=	minimum value of applied sinusoidal shear stress,
γ_{max}	=	maximum value of shear strain response, and
γ_{min}	=	minimum value of shear strain response.

The complex shear modulus (G^*) provides a measurement of the total material resistance to repeated shear stress, including the elastic or recoverable deformation and the viscous or nonrecoverable deformation. The phase angle (δ) provides an indication of the relative amount of elastic response as compared to viscous response, with $G^*\cos\delta$ or the component in phase with the stress measuring the elastic response and $G^*\sin\delta$ or the loss modulus relating the viscous response. Phase angles vary from 0 to 90°, with a zero angle representing a purely elastic material and a right angle representing a purely viscous material. At low temperatures, binders behave more like elastic solids, with δ approaching zero. To completely characterize a binder, both properties are needed as functions of temperature and time of loading, as two binders may have equivalent G^* values but behave differently due to the relative amount of elastic versus viscous response to applied shear stress, indicated by the phase angle (δ). The specification combines both rheologic properties by specifying a minimum value of $G^*/\sin\delta$ for short-term aged binders. This parameter controls permanent deformation by limiting the dissipated energy in a controlled stress repetitive shear loading test. The minimum $G^*/\sin\delta$ is set at 2.20 kPa in the specification for a loading frequency of 10 rad/s. In the development of the specification, this limit was determined based on measured $G^*/\sin\delta$ for unaged and commonly used AC-10 binders and an average measured value of aging index (ratio of absolute viscosity after RTFO to viscosity before RTFO) for these materials that historically has shown adequate performance in terms of resistance to permanent deformation in moderate climates (represented by the conventional 60 °C viscosity measurements) (*15*). The specification for long-term aged binders requires a maximum $G^*/\sin\delta$ value of 5000 kPa for a loading rate of 10 rad/s as measured in the DSR. This parameter is assumed to control fatigue cracking in thin pavement structures by limiting the dissipated energy in a controlled strain repetitive loading test. The maximum value for $G^*/\sin\delta$ was selected based on a large study of 42 binders, with 15 percent failing to meet the specified maximum value (*15*). The effects of pavement structure and mixture stiffness in terms of HMA resistance to fatigue cracking are not currently included in the PG binder specification.

The bending beam rheometer (BBR) and the direct tension tester (DTT) are used to determine the low-temperature behavior of binders. The BBR characterizes binder stiffness at temperatures too low for accurate measurement with the DSR. With both pieces of equipment, binder stiffness is evaluated over a wide range of temperatures critical to performance in the field. The BBR subjects a small beam of binder to a constant creep load and measures the resulting deflection at a temperature related to the lowest service temperature encountered by a pavement (AASHTO TP1) (*6*, *11*). Using beam theory, the binder stiffness is calculated as follows:

$$S(t) = \frac{PL^3}{4bh^3\delta(t)},$$

where (*11*):

$S(t)$	=	creep stiffness (MPa) at time t,
P	=	constant applied load (N),
L	=	distance between beam supports (102 mm),

b	=	beam width (12.5 mm),
h	=	beam thickness (6.25 mm), and
$\delta(t)$	=	deflection (mm) at time t.

This stiffness provides a measure of the binder resistance to creep loading at low temperatures, simulating thermal stresses incurred in pavements as temperatures decrease. The creep rate (m) is also determined from test results as the change in stiffness with time as measured on a log-log plot. The BBR testing temperature is 10 °C higher than the low pavement temperature expected in the field to reduce testing time to 240 s using the principle of time-temperature superposition (*11*). This provides results equivalent to the creep stiffness and creep rate after a 2-hour loading time at the minimum pavement temperature. The binder specification sets limits on the stiffness and *m*-value at a 60-s loading time. These parameters represent critical properties of the binder that directly relate to HMA resistance to thermal cracking. For adequate resistance to this form of distress, the binder plays a predominant role. For a given change in temperature, binders with more resistance to thermal cracking will exhibit smaller induced tensile stresses (controlled by stiffness) and relax these induced stresses at a faster rate (controlled by the *m*-value). The specification requires a creep stiffness at 60 s to be less than 300 MPa and an *m*-value at this same time of loading to be at least 0.30. If the stiffness is between 300 and 600 MPa, the requirement for direct tension failure strain may be used to pass the specification.

The DTT provides an indication of the strain that can be sustained by a binder prior to failure. Although relationships exist to relate the creep stiffness measured with the BBR to the strain at break for unmodified binders, these relationships do not apply to all binders, especially modified ones. The DTT pulls a dog-bone-shaped sample of binder at a slow constant rate until failure (*11*). This test is performed at low temperatures on PAV residue of binders with creep stiffnesses between 300 and 600 MPa. The failure strain (ε_f) is calculated as follows (*11*):

$$\varepsilon_f = \frac{\Delta L}{L_e}$$

where:

ΔL	=	change in length, and
L_e	=	effective gauge length

This failure strain is defined at the load where the failure stress reaches a maximum. Failure stress is defined as the ratio of failure load and original cross-sectional area (36 mm^2). The SHRP specification requires that the failure strain be at least 1 percent.

The recently completed NCHRP Project 9-10 recommended significant changes to the Superpave binder specification for modified binders (*12*). These changes addressed deficiencies in the original specification that included a lack of consideration for the following:

- storage stability,
- additives used in modification,
- the effect of non-Newtonian behavior on mixing and compaction temperatures,
- damage accumulation from repeated traffic loading,
- pavement structure effects,
- traffic speed (other than grade shifting), and
- the effect of cooling rate and variable glass-transition temperatures on low-temperature behavior.

As part of NCHRP Project 9-10, researchers developed screening tests to evaluate storage stability and additives. Based on an extensive laboratory study involving binder and mixture testing, they also recommended new binder parameters to improve characterization of the binder contribution to the three primary forms of asphalt concrete distress. These new parameters are the viscous component of creep stiffness (G_v) measured in a repeated shear creep test at high temperatures, the number of cycles to crack propagation (N_p) measured in a repeated shear controlled stress test at intermediate temperatures, and the critical thermal cracking temperature based on both failure stress and failure strain criteria at representative cooling rates. Researchers also developed new procedures for determining glass-transition temperature and mixing and compaction temperatures for modified mixtures.

NCHRP 9-10 researchers recommended a three-level grading system to accommodate different levels of reliability and available data. Level 1 is based only on environmental conditions, with Level 2 also incorporating traffic conditions. Environmental conditions, traffic speed and volume, and pavement structure are all considered in Level 3. Other recommendations included changes to mixture testing procedures. For binder QA purposes, measurement of the new binder parameters after short-term aging in the RTFO was suggested.

As implementation of these results following a field validation experiment proceeds, further changes to the TxDOT binder QA program may become necessary.

MODELS RELATING BINDER PROPERTIES TO PERFORMANCE

The literature on models relating binder properties to performance is extremely limited (*8*). Most researchers recognize the need for these types of models for a number of different applications, but robust models are not available at this time. One limited study conducted at the University of Nevada, Reno, produced a report by Stephane Charmot titled, "Pay Adjustment Factors for Superpave Performance Graded Asphalt Binders," that provides the following (*16*):

- recent models that relate Superpave binder properties to mixture performance, and
- pay factors associated with inadequate performance for each type of distress (rutting, fatigue cracking, and low temperature cracking) due to off-target Superpave binder properties.

The Nevada Department of Transportation (NDOT) developed a pay factor system based on Charmot's results (*16*). Key economic factors in developing such a system include inflation, discount rate, and analysis period. In Charmot's life-cycle cost analysis, a discount rate of 4 percent with no inflation was used over an analysis period of 30 years for rutting and fatigue or 22 years for low-temperature cracking.

Charmot analyzed mixture performance test results and binder test results gathered during the SHRP validation studies. He then developed pay factors due to an inadequate binder based on a methodology that incorporates the following two alternatives, one when an adequate binder is used and one when an inadequate binder is used:

- calculation of total present worth,
- transformation to an equivalent uniform annual cost, and
- conversion to a total cost over the expected performance life.

The difference in total costs as a percent of binder cost is then subtracted from 100 to determine the pay factor. Maintenance costs, user costs, and nonuser costs were not considered in the life-cycle cost analysis because they were considered equivalent for both the adequate and inadequate binder scenarios. Only rehabilitation costs were considered affected by a reduction in performance life. A brief discussion of the data used for each primary form of distress follows:

- <u>Rutting</u>: Mixture resistance to rutting was defined as the number of Repeated Simple Shear

Test at Constant Height (RSST-CH) cycles to 2 percent permanent shear strain after short-term oven aging. Binder rutting performance was assessed by $G^*/\sin\delta$ values after RTFO. The RSST-CH cycles were converted to Equivalent Single Axial Loads (ESALs) using the SHRP relationship. The sensitivity analysis showed the rutting pay factor model to be stable, with the most significant effect from HMA thickness.

- Fatigue Cracking: Mixture resistance to fatigue cracking was defined as the number of cycles in the flexural beam fatigue test (20 °C, 10 Hz) to reduce the flexural stiffness by 50 percent after short-term oven aging. Binder fatigue performance was assessed by $G^*\sin\delta$ values after RTFO and after RTFO and PAV. The sensitivity analysis showed the fatigue cracking pay factor model is also stable, with the most significant effect from HMA thickness as expected for this type of distress.
- Low-Temperature Cracking: Mixture resistance to thermal cracking was measured in terms of a transverse cracking index after 7 years for six test pavements in Pennsylvania. Binder low-temperature cracking performance was assessed by *S* values and *m*-values at -34 °C after RTFO and PAV. Maintenance costs had to be considered for this type of distress. Two different sets of pay factors were developed based on the two different binder properties. The sensitivity analysis showed the low-temperature cracking pay factor model is very stable, with the most significant effect from HMA specific gravity.

In the absence of identifying other viable models, researchers will utilize the resulting models from this study to the extent possible to evaluate the benefits of recommended changes to the TxDOT binder QA program in the second year of the project.

FACTORS AFFECTING BINDER PROPERTIES PRIOR TO CONSTRUCTION

A possible limitation of the current TxDOT binder QA program is the inability to account for binder properties that may change between production and use during construction. A number of factors may affect or cause these changes. Based on the literature review, Table 1 provides a preliminary list of these factors that can be separated into three categories based on the location of the binder (Figure 1) during its journey from production to use during construction. Researchers selected the highlighted factors for inclusion in a laboratory testing program to identify factors that have the most impact on measured binder properties that may change between production and use during construction.

Table 1. Factors that May Affect Binder Properties prior to Construction.

Category	Factors
Supplier Location	Storage Time
	Storage Temperature (Overheating)
	Blending
	Changing Crude Source
	Refinery Process (Temperature and/or Pressure)
	Contamination in Tanks
Transportation	Contamination in Tanks
	Overheating
Contractor Location	Storage Time
	Storage Temperature (Overheating)
	Contamination/Mixing Different Binders
	Separation
	Dilution
	Presence of Modifier

Aging is one critical effect caused by extended storage time at elevated temperatures. This effect is generally the result of one or more of the following six processes, rendering an increase in the binder stiffness and resulting in a brittle material with reduced resistance to cracking (*17*):

- oxidation,
- volatization,
- thixotropy,
- polymerization,
- separation, and
- syneresis.

The most important processes in terms of the factors suggested in Table 1 are steric hardening (thixotropy), volatization, and oxidation. Researchers anticipate that the effect of aging resulting from these processes will be one of the primary mechanisms causing changes in binder properties from production to use during construction. Researchers expect other primary effects to be related to contamination or mixing of different materials either in the blending or modification process.

Physical and/or chemical changes in properties are a particular problem with polymer-modified asphalts. Most researchers believe excessive heating will cause certain polymers to depolymerize (partially) into monomers that have very low viscosities. The result may be that an expensive modified asphalt required because of its superior properties may be placed in construction with properties commensurate with a lower grade that will result in poor performance. Increased storage temperature is one of the factors explored in the laboratory testing program described in the following chapter.

BINDER QA PROGRAM IN TEXAS

Currently TxDOT samples and approves asphalt materials at the source based on procedures set forth in October 1998 (*18*). The source is defined as either the production site (refinery) or the supplier terminal, and the TxDOT procedures use the terms supplier and producer interchangeably. Prior to the approval process by TxDOT, the supplier must provide test results that indicate specification compliance. In addition, TxDOT samples materials for QA testing according to Test Method Tex-500-C with the supplier present (*19*). TxDOT obtains samples from tanks if batched or as transports are being loaded if blended. TxDOT may also sample transports on a random basis prior to departure from the production site or the supplier terminal. The TxDOT Asphalt Branch of the Materials Section, Construction Division, subsequently referred to as the TxDOT laboratory in Austin, conducts as many tests on these supplier samples as deemed necessary to verify specification compliance. These verification tests constitute the current TxDOT binder QA program. Costs are covered by TxDOT for all materials that meet the specification and by the supplier if a material fails to meet the requirements. If transport samples fail, TxDOT cancels shipment rights for the originating tank. TxDOT approves asphalt cements for up to 60 days and liquid asphalts (asphalt emulsions and cutback asphalts) for a maximum of 30 days. Advance acceptance prior to verification or QA testing is also possible if the supplier has established a QC plan and a good record of compliance, defined as test results for three consecutive samples verified by TxDOT through QA testing and provision of acceptable test results by the supplier. TxDOT can withdraw this privilege if a sample does not meet the specification.

In addition to the established QA program that relies on monitoring the quality of binders at the supplier source, a program of random sampling in the field by TxDOT districts has also been suggested to increase overall binder quality (*20*). Guidelines for taking samples as close to the point of use as possible, making the contracting community aware of the program in advance, detecting any problems early in the project, and giving priority for completing the QA testing were presented in a May 1999 memo from Mr. Michael Behrens to all district engineers (*20*). Testing may take place at either the TxDOT laboratory in Austin or in a district laboratory that has the capability to conduct the required tests. In addition, the May 1999 memo states that all remedial actions for noncompliance with specifications are available, including pay-factor adjustments.

TxDOT does not require the field sampling QA program at this time, but suggestions made to the district engineers stem from recent attempts to revise the asphalt binder specification for PG asphalts to include QC/QA testing of samples taken as close to the point of use during construction as possible. Provisions for bonus/penalty pay-factor adjustments were also explored. Three draft versions that include these types of revisions were proposed over a 2-year period from 1996 to 1998 (*21*, *22*, *23*). Figure 2 highlights the similarities and differences of the three draft versions.

The first version requires obtaining four samples per day during construction and includes both a bonus and penalty pay structure for compliance over the entire project and noncompliance within specific limits for part of the project, respectively. For preconstruction, the contractor is required to provide a complete set of test results indicating specification compliance. The TxDOT laboratory in Austin then conducts verification testing and bears the cost of this process. If the specification compliance is not confirmed, the contractor supplies a second sample and complete set of test results to TxDOT. For the second round of confirmation testing, the contractor bears the costs.

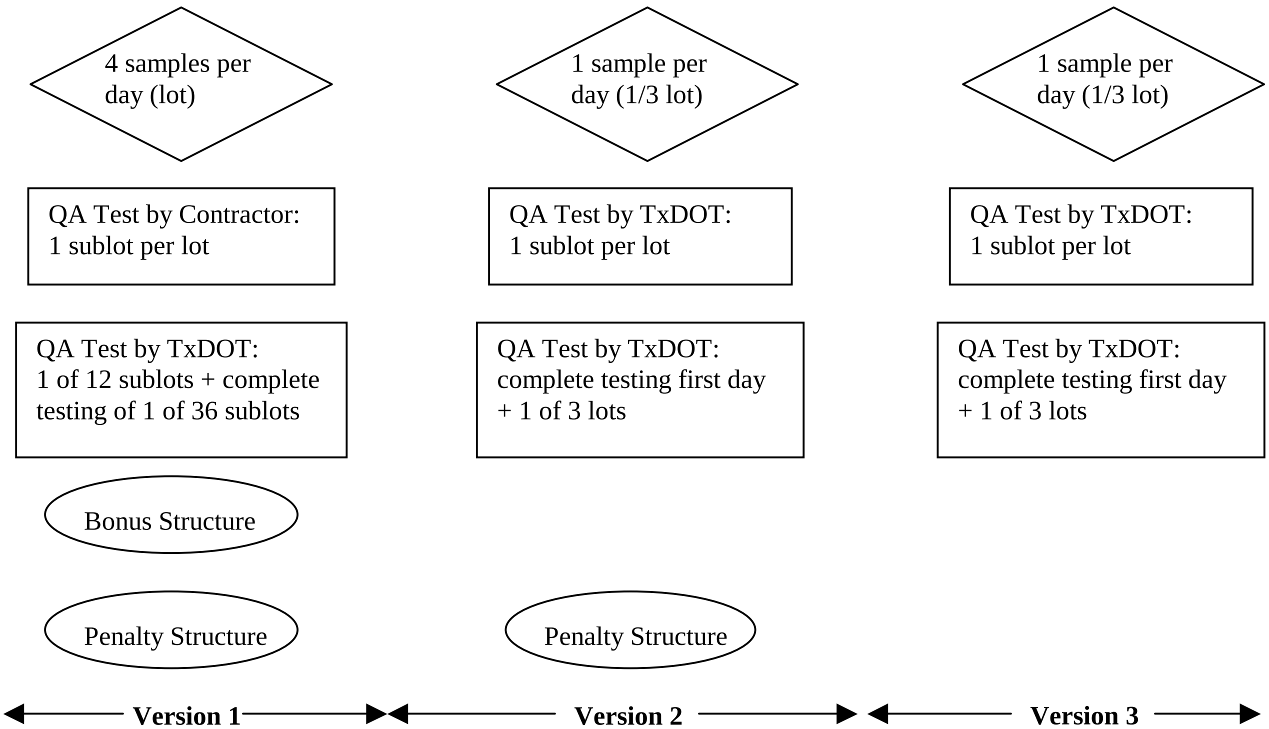


Figure 2. Previously Proposed Binder QC/QA Programs for TxDOT.

During construction, the specification requires that samples be taken and labeled as specific lots and sublots. A lot in the sampling plan is defined as the amount of binder used during one day's production of HMA for a specific project. Each lot contains four sublots. The contractor samples materials with TxDOT personnel present. In this version, QC testing by the contractor is optional and QA testing by the contractor is required. The QA testing requires the contractor to determine the rutting parameter ($G^*/\sin\delta$) from DSR results after short-term aging in the RTFO for one sublot per lot selected at random (*11*). TxDOT district laboratories conduct verification testing for this high-temperature rutting parameter on a minimum of one out of every twelve sublots. For one out of every 36 sublots, the TxDOT laboratory in Austin conducts complete specification verification. Pay-factor adjustments are then determined based on the high-temperature properties as measured in QA testing if the contractor QA results and the verification results are consistent according to a specified maximum difference. If the results differ by more than this maximum, the remaining sublots in the lot in question are tested and either an agreement is made to use all of the QA tests or all of the verification tests to characterize the lot or referee testing is undertaken by the TxDOT laboratory in Austin. A schedule is also provided to allow accumulation of penalty pay factors based on the DSR rutting parameter after RTFO.

The second version of the proposed QC/QA specifications reduced the number of samples per day to one and eliminated the bonus pay-factor adjustment. The only change made in the third version was to eliminate pay-factor adjustments altogether. Other changes from the first version in both subsequent versions (2 and 3) included a definition of a lot in the sampling plan as three consecutive sublots with one sublot sampled each day and required QA testing to be conducted by TxDOT instead of the contractor. Required QA testing includes determination of the DSR rutting parameter after RTFO for one sublot per lot selected at random. TxDOT also conducts confirmation testing on the first day of production and for a minimum of one for every three lots thereafter. This testing includes all tests to ensure complete specification compliance. Penalty pay factors in the second version are adjusted based on QA testing by lot unless QC testing conducted by the contractor can isolate a particular sublot in the lot classified as noncompliant. The maximum allowable difference in the QC and QA test results is 0.5 kPa in this version. Other than the changes noted, the second and third versions replicate the first version.

After evaluation of each of these versions of possible QC/QA specifications, TxDOT decided that this type of specification required excessive administration and that district personnel were not available at the time (*20*). As a result, the decision to implement field sampling in a QA program was left to the individual districts and was not required. In the second year of this project, researchers will evaluate changes to the current QA program such as those presented in the three draft versions described (*21*, *22*, *23*).

Survey of TxDOT Districts and Suppliers

Researchers developed two evaluation surveys for TxDOT district personnel and binder suppliers (Appendix A). The survey questions addressed satisfaction with the current TxDOT binder QA program, suppliers and contractors for each district, and sampling and testing of binders including resources and commonly failed tests. Researchers faxed these surveys to all

TxDOT districts and suppliers that serve Texas after contact was made by phone. They did not receive any surveys from suppliers, but 14 out of 25 TxDOT districts responded.

Appendix B contains a summary of the TxDOT district survey responses in tabular form in a common format for ease of comparison with survey results from state DOT personnel responsible for the overall binder QA program. Tables 2 through 12 further highlight the similarities and differences between the perceptions of the 14 TxDOT districts.

Table 2. TxDOT District Satisfaction.

Answer	Count
yes	6 AMA, CRP, DAL, LFK, PAR, WFS
no	7 ATL, BMT, BRY, CHS, ELP, HOU, LBB

Table 3. TxDOT District Fairness.

Answer	Count
fair • in-line testing of field samples	4 AMA, CRP, LFK, WFS 1 LFK
not fair • infrequent testing • lack of contractor QC	5 BMT, BRY, CHS, ELP, HOU 1 BRY 1 HOU

Table 4. TxDOT District Achievement of Goal.

Answer	Count
yes	3 AMA, CRP, LFK
no • infrequent testing • no guidelines for failing	7 ATL, BRY, CHS, DAL, ELP, LBB, WFS 1 BRY 1 ELP

Table 5. TxDOT District Shortcomings.

Answer	Count
ineffective – material specified not on road	4 ATL, CHS, DAL, LBB
lack of contractor QC	3 BMT, ELP, HOU

Table 6. TxDOT District Responsibility.

Answer	Count
DOT	4 BRY, CHS, ELP, WFS,
contractor	7 ATL, BMT, CRP, DAL, HOU, PAR, WAC
contractors and suppliers	2 AMA, LFK
contractor and DOT	1 BRY

Table 7. TxDOT District Size.

District	# of Major Suppliers	# of Major Contractors	# of Laboratories	# of Technicians
AMA	2	7	1	5
ATL	1	4	1	1
BMT	4	4	0	0
BRY	3	5	1	3
CHS	4	3	1	1
CRP	2	2	1	2
DAL	2	2	1	2
ELP	1	1	1	1
HOU	5	5	1	3
LBB	1	2	0	0
LFK	3	1	1	3
PAR	1	2	0	0
WAC	4	1	1	1
WFS	1	2	0	0

Table 8. TxDOT District Sampling.

Answer	Count
DOT contract employee and contractor	3 AMA, ATL, LFK
DOT	7 CHS, CRP, ELP, HOU, LBB, PAR, WFS
other	1 DAL
some training	8 ATL, CHS, CRP, ELP, HOU, LBB, WFS
asphalt cement (ac): in-line at HMA plant ac: in-line or contractor tank ac: contractor tank	6 ATL, CRP, DAL, ELP, HOU, PAR 2 AMA, LBB 3 CHS, LBB, WFS
daily weekly biweekly monthly by truckload by project as requested	2 CRP, WFS 1 ELP 2 ATL, LFK 1 HOU 2 CHS, LBB 2 DAL, PAR 1 AMA

Table 9. TxDOT District Testing.

Answer	Count
DSR after RTFO • and DSR-unaged • and penetration • and Abson recovery • and Brookfield	9 AMA, ATL, CHS, CRP, DAL, ELP, HOU, LFK, WFS 3 CRP, ELP, HOU 3 AMA, HOU, LFK 1 AMA 1 HOU
daily weekly 1:5 samples	1 CHS 1 ELP 1 WFS
multiple replicates	1 HOU

Table 10. TxDOT District Sampling and Testing.

Sample Type	Answer	Count
field samples	all suppliers	8 ATL, CHS, DAL, ELP, LBB, LFK, PAR, WFS
	some suppliers, by request	3 AMA, CRP, HOU
	no suppliers • program in development	2 BMT, BRY 1 BRY

Table 11. TxDOT District Equipment.

Answer	Count

AASHTO
• none
• DSR
  • and RTFO
  • and Brookfield
  • and penetration
  • and Abson recovery
	
	1 LBB
	11 AMA, ATL, CHS, CRP, DAL, ELP, HOU, LFK, PAR, WAC, WFS
	10 AMA, ATL, CHS, CRP, DAL, ELP, HOU, LFK, WAC, WFS
	4 CHS, ELP, HOU, WFS
	4 AMA, ATL, HOU, LFK
	3 AMA, ATL, HOU

QA officer	3 ATL, LFK, PAR
calibration • yearly • every 6 months • prior to use	 6 ATL, DAL, ELP, HOU, LFK, PAR 1 HOU (RTFO) 2 HOU (DSR), WAC

Table 12. TxDOT District Specification Compliance: Failure.

Answer	Count
rates for supplier samples • 0-3% rates for field samples • 0-3% • >10% agreement with other results • 70-80% • 90-100%	 5 CRP, ELP, HOU, LFK, WAC 7 ATL, CRP, DAL, ELP, HOU, LFK, WAC 2 ATL (1 supplier), CHS 1 ATL 7 CHS, CRP, ELP, DAL, HOU, LFK, WAC
testing failure • retest	 3 ATL, DAL, LFK

Approximately half of the districts are satisfied with the current TxDOT binder QA program, and half are not. Districts that at least take field samples from some suppliers believe the program is fair and achieves a stated goal of obtaining the material as specified on the road in order to produce asphalt concrete that lasts its intended design life. The districts were not asked specifically to identify the goal of the current TxDOT binder QA program, so an assessment of district understanding of the primary motivation behind the program could not made. Five districts including two that do not currently take any field samples (Beaumont (BMT) and Bryan

(BRY)) think that the program is not fair, and a total of seven districts feel that the current program does not achieve its goal. Witchita (WFS) is an anamoly in assessesing the program as fair but unable to achieve its goal. BRY cites infrequent testing as a reason for its assessment, and this district has a field sampling program in development. El Paso (ELP) suggested that guidelines need to be developed for materials that fail the specification. Four districts identified the current program as ineffective, and three other districts cited the lack of contractor QC as a shortcoming of the existing program.

The survey indicated confusion among the districts in term of responsibility for a quality product following construction. Half of the responses indicate the contractor is responsible, while four districts accept the responsibility as the DOT. Two districts spread responsibility between the contractors and the suppliers, and BRY splits responsibility between the contractor and the DOT. According to the survey of TxDOT personnel who oversee the binder QA program, responsibility transfers from the contractor following construction and acceptance by the DOT. To improve the program, the primary goals and responsibilities should be clear to all involved.

Table 7 shows a few statistics that indicate the size or magnitude and resources used in the TX binder QA program in the 14 districts that responded to the survey. Some districts that take field samples do not have a laboratory, and they send their samples to Austin for testing. Amarillo (AMA) and Lufkin (LFK) have the largest laboratory testing capabilities, testing field samples from some or all suppliers, respectively. Most of the other districts with a laboratory utilize one or two technicians for binder testing.

Eight of the 14 districts surveyed collect field samples from all suppliers, and three districts collect these samples from some suppliers. Eleven of the districts indicate that DOT personnel or a contract employee hired by the DOT take the sample, and three districts specify that the contractor is also present. Eight districts respond that these personnel undergo some training. Most samples are taken from either the contractor storage tank or closest to the point of use, in-line at the HMA plant. Sampling frequencies vary by district from daily to monthly and from once per truckload to once per project or as requested.

All TxDOT districts with laboratories utilize AASHTO equipment and test standards when testing binder field samples. Three districts indicate that a QA officer is in charge of calibrating the equipment at least on a yearly basis. Six districts cite an annual calibration frequency, and two other districts calibrate more frequently. Eleven district laboratories have DSR equipment, with 10 also having a RTFO. One district also has the Brookfield viscometer and penetration equipment. Three districts also have Brookfield viscometers, and three different district laboratories contain penetration equipment. Nine districts use the DSR and RTFO equipment for obtaining high-temperature properties before and after short-term aging, and a few other districts utilize penetration equipment and Brookfield viscometers. Testing usually includes an abbreviated program based on available equipment, and frequencies vary by district from daily to weekly or once for every five samples. Only Houston (HOU) conducts multiple replicate tests, while the other districts utilize single replicates to check for specification compliance.

Half of the districts surveyed have relatively low failure rates (0-3 percent) for field samples, with 90 to 100 percent of district test results in agreement with the supplier results contained in the current binder QA program. Three districts indicate that retesting of the same sample is the prescribed procedure if a test result does not satisfy the specification. Only two districts (DAL and ELP) specified a test (DSR) for PG asphalt cements where the material fails to meet the specification most often. No tests were cited for asphalt emulsions. Researchers offer these results taking into account the fact that these districts only conduct limited testing of field samples.

Contractor Visit and Interview

A visit with Bill Thomas of Young Brothers in Bryan focused on the concerns and responsibilities of HMA plant owners in relation to the binder QA program in Texas. As the binder QA program in Texas is now formulated, HMA plant owners are not involved in binder acceptance testing. They assume that the binder purchased from the supplier meets the required specifications. Young Brothers has three binder tanks, and they generally use the material in a single tank over a 24-hour period. Generally they only use one grade of binder in HMA production at a rate of 220 tons per hour. They only utilize one supplier, and they do not conduct any binder tests. They report tracking numbers for the binder printed on the work orders obtained from the suppliers to the TxDOT district.

BINDER QA PROGRAMS IN TEXAS AND OTHER STATES

In addition to the evaluation surveys of TxDOT districts and Texas (TX) binder suppliers, researchers gathered additional information through a two-part phone survey of state DOTs, including TxDOT. The goal of this additional information search was to collect general and then detailed information from binder QA programs in both Texas and nine other selected states. Researchers selected states based on contacts or others suggested by these contacts. The more general survey involved collecting general information, any documentation including specifications, and a sample data set (over a 1-year period). Information gathered in the more detailed survey (Appendix A) through multiple phone conversations and e-mail included the following:

- contact information;
- general satisfaction, goals, and shortcomings;
- responsibility for premature failures;
- size of the program (number of major suppliers, major contractors, laboratories, technicians);
- impact on suppliers and contractors;
- general sampling, testing, and handling requirements and output;
- DOT sampling and testing of both supplier and field samples;
- equipment;
- specification compliance requirements;
- pay factor / penalty systems;
- cost estimates; and
- analysis of results.

Appendix C contains a summary of the state DOT survey responses in tabular form in a common format for ease of comparison with survey results from TxDOT districts. Tables 13 through 29 further highlight the similarities and differences between the 10 state binder QA programs.

Researchers did not complete a detailed review and analysis of the statistical validity of each state binder QA program as proposed because of time and resource limitations. TxDOT may pursue this type of analysis through an ongoing statistical support contract or a multi-year project focused specifically toward achieving this goal.

Table 13. Satisfaction with Binder QA Program.

Answer	Count
yes	9 AZ, CA, CO, MD, MN, NV, OR, UT, WA
no	1 TX

Table 14. Fairness of Binder QA Program.

Answer	Count
fair to contractors • price reduction perceived as fair	8 AZ, CA, CO, MD, MN, NV, TX, WA 1 CO
fair to suppliers	6 CA, CO, MD, MN, TX, WA
may not be fair to suppliers • contractors pass on penalty	4 AZ, NV, OR, UT 6 AZ, CO, MD, NV, UT WA

Table 15. Goal of Binder QA Program.

Answer	Count
material specified on road • without delays • fair with minimum resources	10 1 CA 1 TX
save time and effort through shared certification and inspection	1 MN

Table 16. Shortcomings of Binder QA Program.

Answer	Count
ineffective – material specified not on road	1 TX
lack of contractor QC • required /expanded in near future • requires too many resources to check	5 AZ, NV, OR, UT, WA 3 CO, NV, UT 1 OR

Table 17. Responsibility in Binder QA Program.

Answer	Count
DOT-penalty system in place • before acceptance, can shut down construction or revoke supplier certification • can leave unpaid for up to 2 years	6 AZ, CO, MN, NV, OR, UT 1 CA 1 WA
contractor-first yr, DOT-after first yr	1 MD
DOT-approved based on supplier samples assumed OK in field	1 TX

Table 18. Size of Binder QA Program.

State	# of Major Suppliers	# of Major Contractors	# of Laboratories	# of Technicians
AZ	4 PG 4 emulsion	6 Asphalt Concrete (AC) 6 spray	1 central 3 regional	3 full time 2 summer at central 1-2 at regional
CA	11	40-50	1 central	3 full time
CO	5	84 total	1 central	2.5 full time 0.5 summer
MD	6	6	1 central 2 regional (western, eastern)	3 full time @ central 2 full time @ western 1 full time @ eastern
MN	10 PG 7 emulsion	52 AC	1 central	4 full time (in summer)
NV	5	3 AC (10 minor)	1 N 1 S	5 full time 1 summer/lab
OR	5	3-5 AC 3-5 spray	1 central	3 full time 1 summer
TX	18 total	40-50 90 total	1 central 25 district	4 full time 2 @ 25% in central
UT	5-6	5-6	1 central	2 full time 1 summer
WA	9 PG 9 spray	~ 15	1 central	2 full time 1 prep

Table 19. Impact of Binder QA Program.

Answer	Count
some supplier disputes with contractors • penalty passed to suppliers	1 AZ 1 NV
none on contractors • unless shutdown job	4 AZ, MN, TX, WA 5 CO, MD, MN, NV, OR
no delays unless trend of failing results	1 CA
required contractor QC plan • required/expanded in near future	1 UT (no testing) 2 CO, UT

required supplier QC plan
• only for some emulsion certification
• advance acceptance
• certification
    • part of combined states group
• annual inspection
    • if necessary
	7 CA, CO, MD, MN, TX, UT, WA
	1 AZ
	2 TX, UT
	9 AZ, CA, CO, MD, MN, NV, TX, UT, WA
	2 MD, MN
	2 MN, WA
	1 CO

some delay to suppliers w/out advance acceptance	1 TX

Table 20. Sampling in Binder QA Program.

Answer	Count
contractor w/DOT witness	6 AZ, CO, MD, MN, UT, WA
contractor, DOT not required • 10% witnessed by DOT	2 NV, OR 1 OR
DOT contract employee	1 TX
DOT	1 CA
AASHTO	10
ac: in-line at HMA plant ac: in-line or contractor tank ac: contractor tank ac: contractor truck	4 AZ, CA, NV, UT 4 CO, MD, TX, WA 1 OR 1 MN
emulsion: distributor truck emulsion: distributor truck or contractor tank emulsion: supplier emulsion: none	4 AZ, NV, OR, WA 4 CA, CO, MD, TX 1 MN 1 UT
some training	5 CA, MN, NV, OR, WA
adjustable frequency	4 MD, MN, NV, UT

Table 21. Testing in Binder QA Program.

Answer	Count
single replicate to check compliance	10
round-robin testing	3 CO, MN, OR
AASHTO accreditation	10
in-house training	3 AZ, CO, NV
formal technician training	4 CA, MD, MN, TX (central)
adjustable frequency	4 MD, MN, NV, UT

Table 22. Supplier Testing in Binder QA Program.

Answer	Count
yes	9 AZ, CA, CO, MD, MN, NV, TX, UT, WA
no	1 OR

Table 23. DOT Sampling and Testing in Binder QA Program.

Sample Type	Answer	Count
supplier samples	no samples taken	4 AZ, MD, NV, OR
	all suppliers of new binder-only at beginning of season	1 CA
	all suppliers -only at beginning of season	1 CO
	option in special situations	3 MN, UT, WA
	only during mix design	1 OR
	primary basis for acceptance	1 TX
field samples	all suppliers	9 AZ, CA, CO, MD, MN, NV, OR, UT, WA
	suggested but resources not available; some districts on regular basis or if problem suspected	1 TX

Table 24. Equipment in Binder QA Program.

Answer	Count
AASHTO	10
• complete	3 central labs AZ, MD, TX
• DSR, RTFO	1 regional lab MD
• Brookfield	1 regional lab MD
QA personnel	
• 1	4 CO, OR, TX, WA
• 2	2 AZ, UT
• >2	2 MN, NV

Table 25. Specification Compliance: Tolerance Intervals in Binder QA Program.

Answer	Count
yes	5 MD, MN, NV, UT, WA
• AMRL proficiency tests	1 MD (PG)
• ASTM precision and bias	1 MD (emulsion)
• round-robin testing	1 MN
• AASHTO repeatability	1 NV
no	5 AZ, CA, CO, OR, TX

Table 26. Specification Compliance: Failure in Binder QA Program.

Answer	Count
rates	
• 0-1%	3 MN, OR (ac), UT
• 2-5%	5 CA, CO (PG), MD, TX (supplier), WA
• 6-10%	1 NV
• >10%	3 AZ, CO (emulsion), OR (emulsion)
testing failure	
• retest	7 CA, CO, MD, MN, NV, OR, TX
• test adjoining samples	8 AZ, CA, CO, MD, MN, OR, UT, WA
• complete resample and retest	2 MD, TX
• compare w/other results	4 MD, NV, TX, UT
• referee testing	1 UT

Table 27. Specification Compliance: Tests that Fail Most Often in Binder QA Program.

Material	Test	State
PG	DSR ORIG	4 CO, MD, TX, UT
	DSR-PAV	1 WA
	DSR-RTFO	5 AZ, CO, MD, MN, TX
	m-value	2 AZ, MD
	toughness & tenacity	3 CO, NV, UT
	ductility	2 CO, NV
emulsion	Saybolt viscosity	4 AZ, OR, TX, WA
	Sieve	2 AZ, CA
	residue by evaporation	2 CA, MN
	penetration of residue	3 CO, MN, OR
ac	absolute viscosity	2 CA, OR
	penetration	1 OR

Table 28. Penalties and Pay Factors in Binder QA Program.

Answer	Count
yes • based on one property • based on multiple properties	7 AZ, CO, MN, NV, OR, UT, WA 4 AZ, MN, OR, WA 3 CO, NV, UT
no	2 CA, TX
no formal system	1 MD

Table 29. Analysis of Benefits and Costs in Binder QA Program.

Answer	Count
qualitative confidence	9 AZ, CA, CO, MD, MN, NV, OR, UT, WA
no quantitative confidence estimate	10
acceptance	10
database • old/paper • in near future • forensic • compare/track performance • track suppliers • track use, costs • laboratory assessment • communication • research	6 AZ, CO, MD, MN, NV, OR 2 CA, TX 2 UT, WA 8 AZ, CA, CO, MN, NV, TX, UT, WA 6 CA, MD, MN, NV, UT, WA 9 AZ, CA, CO, MD, MN, NV, OR, UT, WA 5 MD, MN, NV, TX, UT 2 MN, OR 3 AZ, MD, UT 3 MN, NV, WA
cost estimate not available	4 CA, NV, OR, WA
rough cost estimate	6 AZ, CO, MD, MN, TX, UT
benefit/cost (B/C) ratio not available	10
balance resources and quality assessment	4 CO, MN, UT, WA

All of the states except TX are satisfied with their current binder QA program. Following completion of this project, researchers expect TxDOT's satisfaction with their program to improve. The goal of all of the states' programs is to obtain the material that was specified on the road. Minnesota (MN) also cited a secondary goal of saving time and effort through a coordinated program where multiple states share certification and inspection of suppliers. California (CA) indicated that there must also be no delay in construction caused by the binder QA program. A specific goal of the TX program is to promote fairness to all parties through a program that requires minimum resources. In most of the states, responsibility for a quality product transfers from the contractor following construction and acceptance by the DOT. The DOT is then responsible for premature failures, usually after the first year in service. This system works well in many of the states where penalties are assessed to the contractors based on an estimate of the difference in performance of the as-constructed and as-designed or as-specified pavement. Most of the states felt that their program was fair to contractors, but many questioned the issue of fairness with respect to suppliers because any penalties assessed the contractors are usually passed on to the suppliers even if the there is a lack of QC during transportation or at the contractor location, a problem cited by half of the states. Three states plan to introduce or expand a required contractor QC plan in the near future, and one state recognizes that resources are not available to maintain this type of system.

Table 18 shows statistics that indicate the size or magnitude of binder QA programs in the 10 states surveyed. TX and CA have the largest number of major binder suppliers, but many of the other states have larger laboratory testing programs in terms of the number of laboratories and the number of technicians assigned to the binder QA program. The workload in terms of number of tests per year varies from state to state and is difficult to compare because of differences in sampling and testing frequencies and abbreviated testing requirements. Most states with large testing programs require testing of field samples for acceptance by the DOT. Testing of supplier samples is left to the suppliers themselves and is required in almost all of the states, although each state differs in terms of the frequency of complete and abbreviated specification compliance testing. In some of the states, the DOT tests supplier samples at the beginning of the season, for new binders, or in special situations. Currently, the TX system is opposite, requiring DOT testing of supplier samples and no regular system of testing field samples.

In terms of sampling either supplier or field samples for testing by the DOT, most states allow the contractor to take the sample according to AASHTO guidelines with a DOT witness present. In Nevada (NV) and Oregon (OR), this witness is not required but is present some of the time in OR. In CA, DOT employees take samples, and these personnel are trained, as they are in half of the states surveyed. In TX, neither the contractor nor the DOT is present; TxDOT hires a contract employee with no formal training to take supplier samples. Most states take field samples from either the contractor storage tank or closest to the point of use, in-line at the HMA plant or from the emulsion distributor truck. Sampling frequencies are also adjustable in some states to account for a continued record of compliance or noncompliance or to adjust the laboratory workload.

Most of the states, including TX, require supplier QC plans, but currently only Utah (UT) requires some form of a contractor QC plan. Thus, the impact on contractors is minimal in most states unless construction is shut down for a serious problem that may be related to binder quality. Different states have different supplier requirements that may include an annual inspection, certification, or an advance acceptance program to reduce delays. Certification in two states is good for a combined group of states, reducing the number of resources required for each individual state.

All states surveyed utilize AASHTO equipment and test standards when testing binders in their QA programs. Laboratories are AASHTO accredited through the efforts of one or two people in the majority of states. A complete set of testing equipment is found in the central laboratory in each state, while regional laboratories may only have a limited set of equipment. Less than half of the states have a formal technician training program. Three states participate in round-robin testing programs, and four states allow for adjustment of testing frequencies. Testing frequencies vary by state, with some samples remaining untested, some undergoing an abbreviated specification compliance testing program, and others subjected to a complete testing sequence.

Single replicate test results are compared to specification limits that include tolerance intervals in half of the states evaluated. The basis for these tolerances is different for each state, ranging from proficiency or round-robin test results to ASTM or AASHTO precision and repeatability statements. Each state defines compliance and rejection limits in a schedule. The other half of the states, including TX, do not allow for any variability in the result from the specification limit. In these states, the supplier is expected to account for any variability and ensure that the specified value can be met.

Most states are satisfied with their binder QA program, as illustrated by their relatively low failure rates (less than 5 percent), especially for PG asphalt cements. Each state prescribes a different procedure following failure of a material to meet a specified test, but the majority require retesting the same sample and testing of samples immediately surrounding the failed sample. These results are used to estimate the quantity of material out of specification for calculation of penalties through pay factors. A few states compare failed test results with other results from the supplier, round-robin testing programs, or AASHTO repeatability limits. Complete resampling and retesting or testing by a third party is another less-common option in a few states, with the supplier or the contractor paying for testing of noncompliant material in TX and UT, respectively. Tests for PG asphalt cements where the material fails to meet the specification most often according to the survey results include the DSR on unaged or short-term aged (in the RTFO) and a toughness and tenacity test the intermountain west states (Colorado (CO), NV, and UT) include in their PG+ specification. For asphalt emulsions, a number of states cited Saybolt viscosity and penetration of the residue as tests where the material most often fails the specification.

When materials fail the specification, pay factors are calculated in seven of the 10 states. Pay factors are also determined in Maryland (MD), although there is no formal system. The two states without pay factors (CA and TX) are also the largest states that probably use the largest volume of binders in asphalt construction per year. Issues associated with these large states may partially explain the lack of a formal pay-factor system. Penalties are assessed based on only one binder property in four states and on an accumulation of failing binder properties in three states. Often dependent on the materials involved, properties measured, environmental conditions, and facility type, each state uses different schedules and equations to determine the penalty assessed of the contractor.

The final category analyzed through the detailed survey of state binder QA programs was the analysis of benefits and costs. In all states, the main use of the data is to allow the DOT to accept the material and responsibility for use in asphalt pavement construction. The majority of states use an electronic database for a variety of purposes, including forensic analyses and historical analysis of the quality of materials from each supplier and the performance of different binders. Other benefits cited include improved communication with suppliers, laboratory assessment, research, and the ability to track binder use and costs to the state. The larger states of CA and TX currently have inadequate databases that do not allow for some of these benefits. Again, increased resources are required, but creation of electronic databases is forthcoming in all states surveyed. No detailed cost information and therefore benefit to cost (B/C) ratios were available from any of the 10 states. Only a qualitative sense of confidence is obtained in all states except TX. Unfortunately, none of the 10 states has any quantitative confidence estimate of the quality of material utilized in asphalt pavement construction due to limited resources and the lack of a need to quantitatively justify the program. Four of the states highlighted the fact that their binder QA programs attempt to balance resources while at the same time assessing the quality of materials used in asphalt pavement construction and qualitatively obtaining a sense of confidence in these materials.

CHAPTER 3. EVALUATION

In addition to the qualitative comparison of binder QA programs documented in the previous chapter, researchers quantitatively evaluated existing binder data from three states and partially completed an extensive laboratory testing program. This chapter provides a description of and results from the analysis of existing data sets and the partially completed laboratory testing program. Researchers will document a complete set of results from the laboratory testing program and their implications in the final research report for this project.

ANALYSIS OF EXISTING DATA

Researchers statistically analyzed data received through the information search from Colorado and Oregon using cluster analysis to compare test results required by specification to their corresponding specified values. They used a different classification tool called classification and regression trees (CART) to statistically analyze existing data from Texas toward the same goal. They pursued this second type of analysis with the Texas data because the cluster analysis did not produce meaningful results useful to TxDOT for decision making. This section provides descriptions of these two analysis methods, followed by the resulting classifications and their implications.

Cluster Analysis

Researchers used two approaches to examine the Colorado and Oregon data through cluster analysis. One approach compared each individual test result with its required value in the specification. The second approach compared all test results to their specified values simultaneously. The analysis also focused on statistically describing results from each test and the collection of tests for each type of binder material. For each test, researchers examined central tendency, variation, and shape and type of the distribution of results through graphical and mathematical techniques. The focus of this analysis was to show, using data from the other states, what information can be obtained if the Texas data included results from field samples stored in an organized, easily accessible manner. One goal was to understand the variability to facilitate establishment of a rational basis for pay factors and determination of the confidence level that the material used meets the specification.

As a first step with the Colorado PG data set, researchers explored correlation of different binder test parameters to aid in selecting those most relevant for use in a QA program. Then they examined statistical distributions of the selected parameters using kernel estimation, a nonparametric smoothing method. This initial analysis showed bimodal distributions, with one group of measurements that generally exceeded the specification in one mode and a second group of measurements clustered around the specified value. As a result of the multimodality of the data, researchers chose cluster analysis as a more appropriate tool.

Cluster analysis is an exploratory data analysis tool for solving classification problems. Its objective is to sort cases into groups, or clusters, so that the degree of association is strong between members of the same cluster and weak between members of different clusters. Each cluster thus describes, in terms of the data collected, the class to which its members belong. As a result, cluster analysis can reveal similarities in data that may have been otherwise impossible to find.

The results from cluster analysis can be used in several ways. Cluster analysis aids in the identification of outliers (observations lying very far from the main body of the data) by assigning them to one cluster. These outliers may be the result, for instance, of measurement errors or typing errors made while entering the observations into a database. Outliers can be discarded so as not to affect the result of the analysis. When future QA tests are conducted, they can be assigned to clusters, enabling prediction of tests that might cause problems and whose results should therefore be examined more closely. This assignment can be done using different statistical procedures to find a cluster where observations have relationships between variables similar to the one under investigation. For experimental design purposes, clusters can be used as blocks. Thus, it would be important to pick an equal number of samples from each cluster to make the analysis less biased and to reduce supplier-to-supplier variability. Other anticipated advantages of this type of analysis include identification of materials (and corresponding suppliers) that consistently fail specific property requirements.

Cluster analysis groups observations so that the observations in each group are similar with respect to the clustering variables. The various clustering techniques fall into two categories, hierarchical and nonhierarchical. Hierarchical cluster analysis is an iterative procedure. Initially, each data point is a cluster. In each succeeding step, the two "closest" clusters are merged, reducing the total number of clusters by one. This continues until there is only one cluster, or the desired predetermined number of clusters is reached.

Determining which clusters are "closest" requires a measure of the distance between clusters. The various hierarchical clustering algorithms differ mainly in the way they compute distance. Sharma (*24*) gives a summary of the various clustering algorithms together with the empirical studies comparing the performance of different clustering algorithms. From the survey, it appears that single-linkage, average-linkage, and Ward's method perform best.

For this analysis, Ward's method was chosen. Ward's method does not compute distances between clusters, but rather forms clusters by maximizing within-cluster homogeneity.

The main problem with all hierarchical clustering methods is that only observations with complete data can be used. In this study, 62 percent of the observations have missing data, so these methods are of limited use.

In nonhierarchical clustering, the data are partitioned into *g* predetermined clusters. This requires that the researcher have some *a priori* knowledge of how the data will cluster. This is usually obtained by clustering the data using one or more hierarchical techniques. Observations with missing data can also be handled since once the cluster centroids or seeds are identified, clusters are formed by assigning observations to the seed to which the observation is closest, based on available information.

CART Analysis

Researchers statistically analyzed existing data from Texas using CART, with the majority of records labeled Pass, Fail, and For Information Only. Researchers used this type of analysis to develop simple rules that produce classification trees and corresponding classes with these three labels. For each type of material, they identified several critical properties ($x_1, \ldots, x_p$) and used them in the CART analysis to decompose the data using binary (two way) splitting rules. In each of the resulting subsets of data, a majority-voting rule determined the class label (Pass, Fail, or For Information Only). For example, a splitting rule of ($x_1 \leq 150$) meant that all data with x_1 values less than or equal to 150 were assigned into one class and the remaining data were assigned to another class. The overall class label was determined by the most common label in the subset. For example, the 38 Pass labels in 50 cases with ($x_1 \leq 150$) identified this class as Pass. CART recursively split and resplit the properties until a simple tree was produced that accurately reflected the classifications in the existing database, if possible. An example output tree for CRS2 materials from this analysis is shown in Figure 3, with Saybolt2 indicating the Saybolt viscosity measured at 122 °F (50 °C).

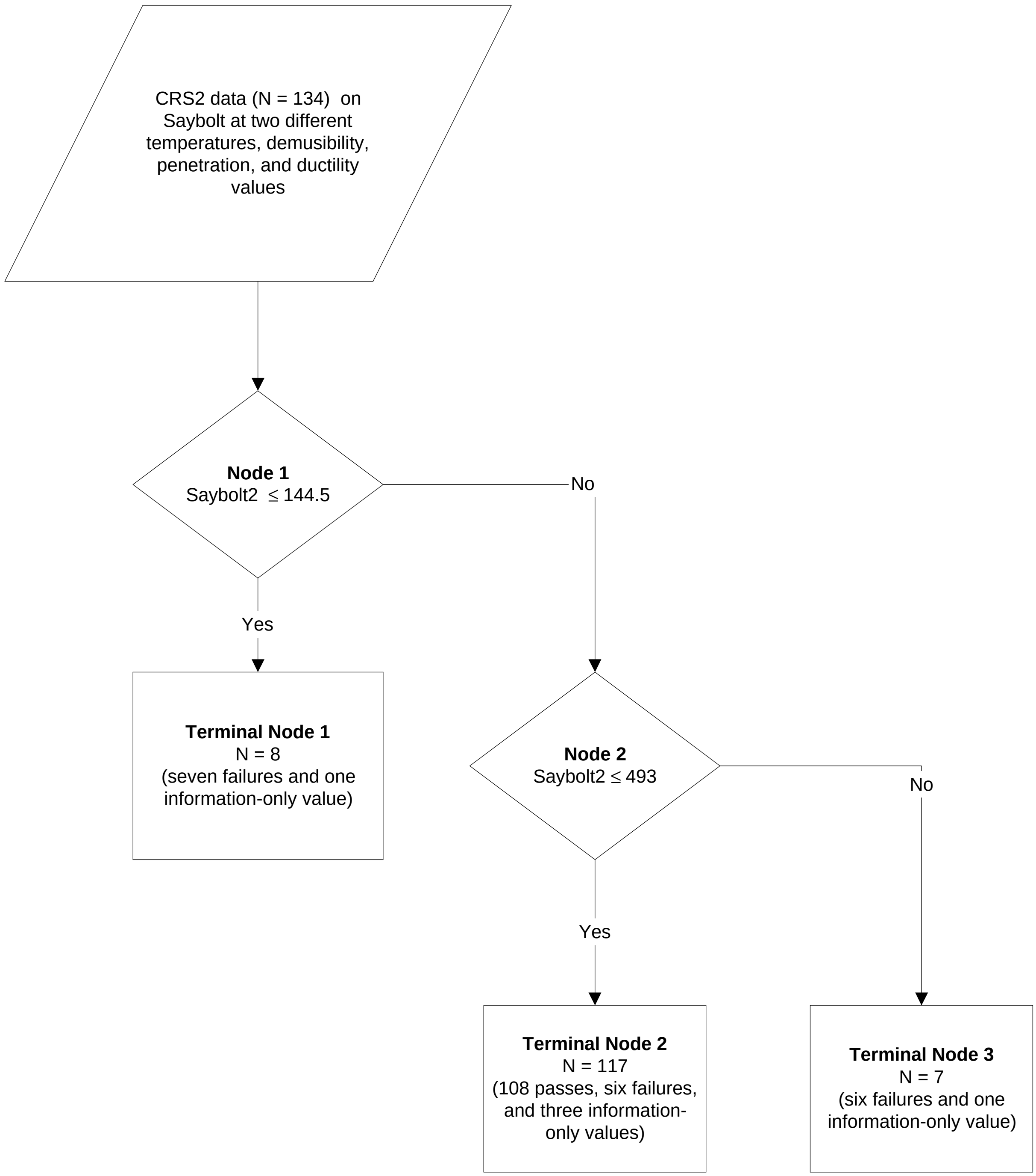


Figure 3. CRS2 CART Tree.

COLORADO

The Colorado PG dataset covered a 1-year (2000) time period. The data consist of the results from three QA tests (DSR, RTFO-DSR, PAV-DSR) for eight different binders (PG binder grades labeled Binders 1-8) produced by twelve suppliers (Supplier A-M, without Supplier I, to avoid confusion with J). DSR represents the $G^*/\sin\delta$ value measured on an unaged binder. RTFO-DSR is used for the $G^*/\sin\delta$ value for a short-term aged binder, and PAV-DSR indicates the $G^*\sin\delta$ value measured on a binder that has been both short-term and long-term aged. Of the 577 observations, only 217 had complete data. The DSR data were missing from some observations, but this test was performed more frequently than either the RTFO-DSR or PAV-DSR tests.

To standardize the data, researchers transformed each property in the following manner:

$$std.value = \frac{value - spec}{spec}$$

where:

spec	=	the specified value for the test,
value	=	a test result, and
std.value	=	the standardized test result for further analysis

They then relabeled the standardized QA test results for DSR, RTFO-DSR, and PAV-DSR as STDSR, STRTFO, and STPAV, respectively. These standardized results must all be greater than zero to meet the specification.

The goal of this analysis was to separate suppliers based on the quality of their binder. Researchers used hierarchical clustering with Ward's method to identify the number of clusters and cluster seeds. Then, they obtained a nonhierarchical cluster solution for the data.

Researchers identified and deleted one obvious outlier for RTFO-DSR. This outlier might be due to an error when results were typed into the database.

Table 30 shows descriptive statistics for the entire data set: number of observations (N), mean, standard deviation (Std. Dev.), minimum, and maximum for each variable.

Table 30. Descriptive Statistics for All Data.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	556	0.587	0.725	0.998	5.000
STRTFO	239	0.448	0.618	0.854	2.690
STPAV	222	0.490	0.258	0.630	0.999

Researchers chose the number of clusters to be four based on several statistics that measure cluster homogeneity. Table 31 presents descriptive statistics for each variable by cluster.

Table 31. Descriptive Statistics for All Data by Cluster.

Cluster #	Variable	N	Mean	Std. Dev.	Minimum	Maximum
1	STDSR	203	0.480	0.146	0.150	1.160
	STRTFO	99	0.310	0.231	-0.095	0.995
	STPAV	98	0.653	0.161	0.330	0.999
2	STDSR	219	0.164	0.148	-0.998	0.550
	STRTFO	109	0.158	0.213	-0.854	0.518
	STPAV	90	0.334	0.205	-0.630	0.808
3	STDSR	52	2.374	1.123	1.390	5.000
	STRTFO	17	1.820	0.407	0.845	2.614
	STPAV	20	0.220	0.256	-0.366	0.696
4	STDSR	82	0.847	0.246	0.480	1.470
	STRTFO	14	2.002	0.327	1.586	2.690
	STPAV	14	0.750	0.047	0.672	0.842

Figure 4 shows box-plots for each variable by cluster.

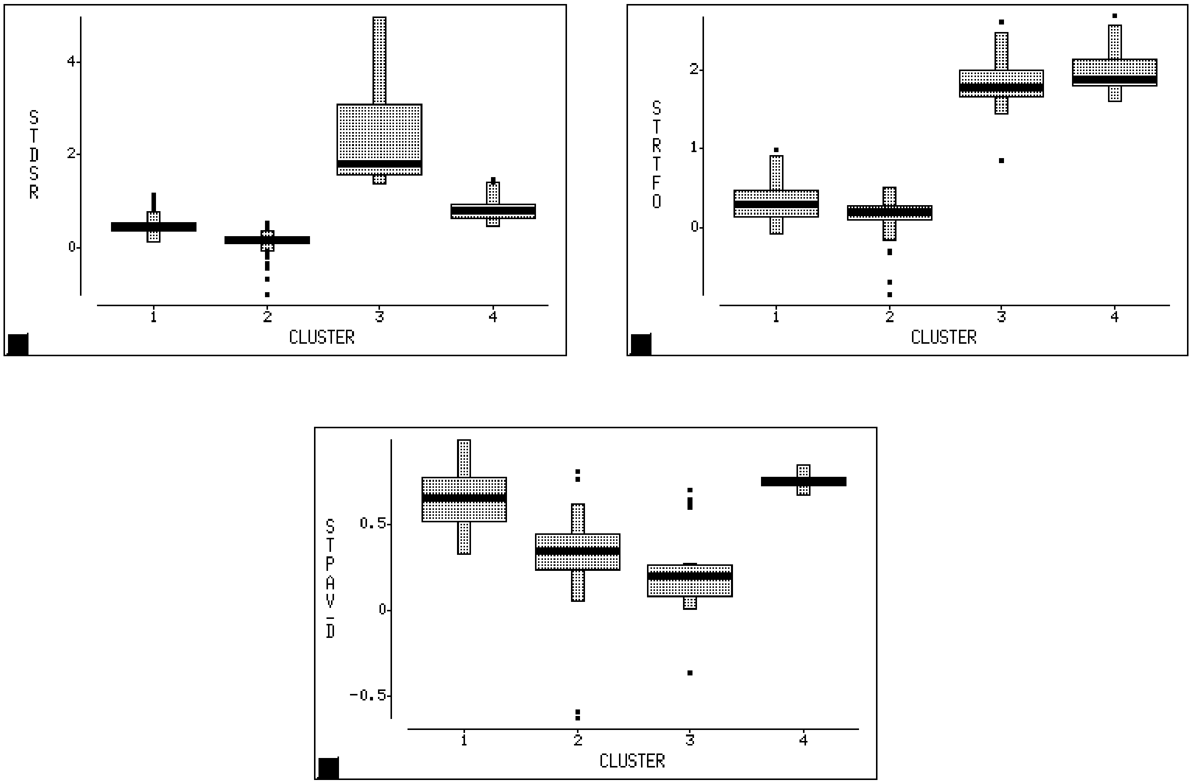


Figure 4. Box-plots for Each Variable by Cluster.

By comparing descriptive statistics for each cluster to those for the entire data set, researchers offer the following observations:

- Cluster 1 contains below-average STDSR values, below-average STRTFO values, and above-average STPAV values.
- Cluster 2 contains STDSR and STRFO values far below average and below-average STPAV values.
- Cluster 3 contains STDSR and STRFO values far above average and STPAV values far below average.
- Cluster 4 contains above-average STDSR values and STRTFO and STPAV values far above average.

In summary, the best cluster is Cluster 4 and the worst cluster is Cluster 2 based on the number of failures or results not passing the specification.

Table 32 shows the number of failures for the three variables by cluster. Cluster analysis could not locate all failures into one cluster. The first three clusters have observations with failures. Cluster 2 has all 10 STDSR failures; one-third (seven out of 21) of all STRTFO failures are in Cluster 1, and the remaining two-thirds (14 out of 21) are in Cluster 2. Two-thirds of all STPAV failures (two out of three) are in Cluster 2 with one-third (one out of three) in Cluster 3.

Table 32. Number of Failures by Cluster.

Variable	Cluster				
	1	2	3	4	Total
STDSR	0	10	0	0	10
STRTFO	7	14	0	0	21
STPAV	0	2	1	0	3

Table 33 shows suppliers sorted by cluster. Based on Table 33, approximately 35.5 percent of all tests were grouped into Cluster 1, 40 percent into Cluster 2, 9.5 percent into Cluster 3, and 15 percent into Cluster 4. Most of the suppliers have observations in each cluster. Almost all suppliers, except G, K, and M, have the majority of observations in the first two clusters, ones that reflect bad (compared to other clusters) performance for STDSR and STRTFO. For Suppliers A and C, more than 50 percent of the observations are in Cluster 2, the worst cluster. Some suppliers, like Supplier F, have a significant percentage in every cluster, which might indicate unstable performance (test results vary significantly). This can be explained by the fact that for some suppliers, performance changes by binder. Therefore, researchers also conducted cluster analysis for each binder separately and then compared the results.

Table 33. Suppliers Sorted by Cluster

Supplier	Cluster				
Frequency	1	2	3	4	Total
A	18	52	13	8	91
B	9	0	0	1	10
C	27	60	1	1	89
D	17	13	1	9	40
E	6	14	3	2	25
F	29	43	19	13	104
G	20	9	0	29	58
H	23	11	0	5	39
J	48	32	0	2	82
K	7	0	15	12	34
L	1	0	0	0	1
M	0	0	3	0	3
Total	205	234	55	82	576

Table 34 shows how many samples were included from each supplier by binder. The major suppliers are Suppliers A, C, and F, providing almost one-half of the samples. Most samples (approximately 84.5 percent) are from Binders 1, 2, 5, and 6. Notice that some suppliers like Supplier K specialize only in one binder, and some produce several binders. Binders 1, 2, 5, 6, and 8 had enough data to perform cluster analysis. Only Supplier J provided samples of Binder 3. Also, it appears that Binder 4 is not widely used. There were only a total of 2 samples, one from Supplier E and one from Supplier J. For Binder 7, there were 17 samples: 8 samples (47 percent) from Supplier D, 6 samples (35 percent) from Supplier E, and one sample (6 percent) from each of Suppliers A, C, and H.

Table 34. Suppliers Sorted by Binder.

Supplier	Binder								
Frequency	1	2	3	4	5	6	7	8	Total
A	16	42	0	0	15	4	1	13	91
B	1	1	0	0	8	0	0	0	10
C	37	8	0	0	42	0	1	1	89
D	2	15	0	0	0	2	8	13	40
E	2	8	0	1	3	3	6	2	25
F	17	28	0	0	12	16	0	31	104
G	20	29	0	0	8	0	0	1	58
H	0	38	0	0	0	0	1	0	39
J	0	21	10	1	7	43	0	0	82
K	0	0	0	0	0	34	0	0	34
L	0	0	0	0	0	1	0	0	1
M	0	0	0	0	0	3	0	0	3
Total	95	190	10	2	95	106	17	61	576

Tables 35 through 37 present descriptive statistics for Binders 3, 4, and 7. There were no failures for Binders 3, 4, and 7.

Table 35. Descriptive Statistics for Binder 3.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	10	0.500	0.078	0.380	0.600
STRTFO	5	0.460	0.105	0.318	0.568
STPAV	5	0.637	0.009	0.628	0.652

Table 36. Descriptive Statistics for Binder 4.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	2	1.320	0.948	0.650	1.990
STRTFO	2	1.707	1.218	0.845	2.568
STPAV	2	0.725	0.041	0.696	0.754

Table 37. Descriptive Statistics for Binder 7.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	17	0.712	0.297	0.100	1.410
STRTFO	6	0.250	0.146	0.005	0.418
STPAV	6	0.805	0.046	0.762	0.866

Table 38 shows descriptive statistics for each variable for Binder 1.

Table 38. Descriptive Statistics for Binder 1.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	95	0.849	0.622	0.422	1.870
STRTFO	39	0.959	0.706	0.159	2.082
STPAV	36	0.406	0.231	0.008	0.842

For Binder 1, researchers chose the number of the clusters to be three based on several statistics that measure cluster homogeneity. Table 39 shows descriptive statistics for each variable by cluster.

Table 39. Descriptive Statistics for Binder 1 by Cluster.

Cluster #	Variable	N	Mean	Std. Dev.	Minimum	Maximum
1	STDSR	44	0.318	0.222	-0.422	1.160
	STRTFO	24	0.430	0.208	0.159	0.995
	STPAV	21	0.505	0.082	0.296	0.690
2	STDSR	36	1.591	0.180	1.280	1.870
	STRTFO	12	1.770	0.197	1.427	2.082
	STPAV	12	0.129	0.090	0.008	0.252
3	STDSR	15	0.622	0.069	0.480	0.780
	STRTFO	3	1.936	0.169	1.741	2.036
	STPAV	3	0.819	0.020	0.803	0.842

Figure 5 shows box-plots for each variable by cluster.

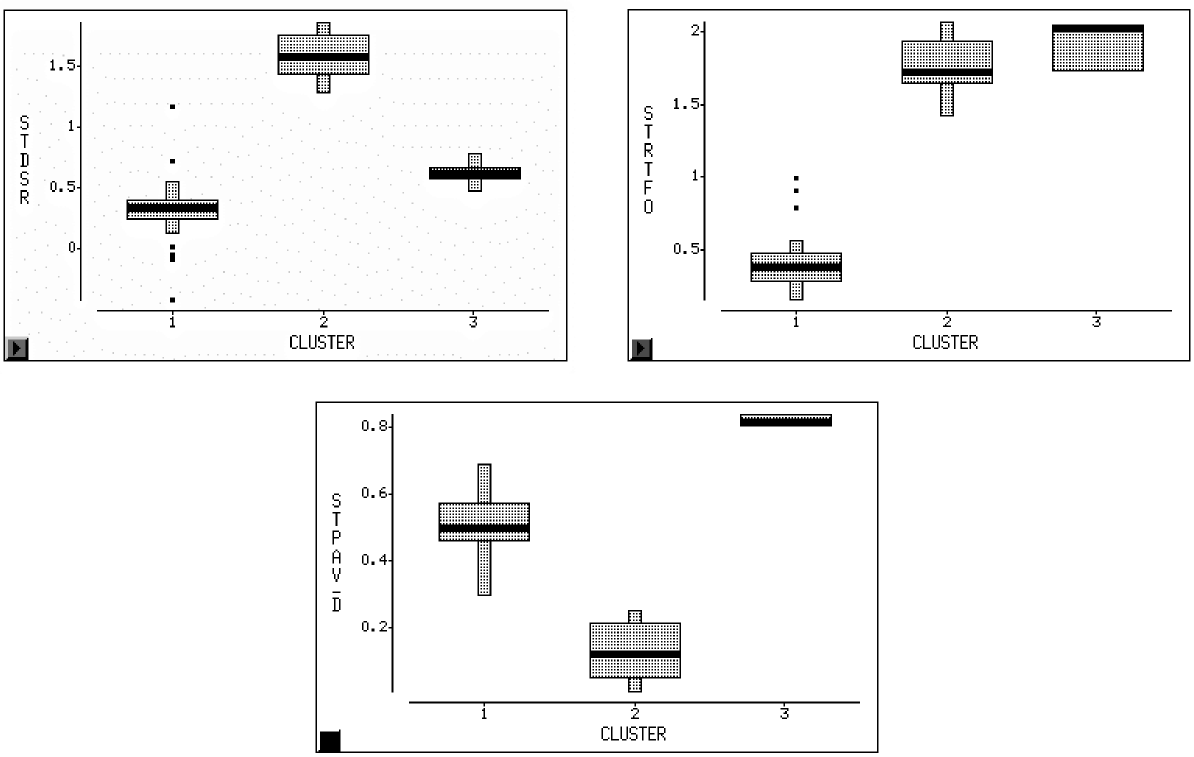


Figure 5. Box-plots for Each Variable by Cluster for Binder 1.

Based on Tables 38 and 39 and Figure 5, researchers offer the following observations:

- Cluster 1 contains STDSR and STRTFO values far below average and a little above-average STPAV values.
- Cluster 2 contains STDSR and STRTFO values far above average and below-average STPAV values.
- Cluster 3 contains STDSR values below average, STRTFO and STPAV values far above average.

Table 40 presents the number of failures for the three variables by cluster. For this binder, there are failures only for the first variable, STDSR, and all of them were in Cluster 1.

Table 40. Number of Failures by Cluster for Binder 1.

Variable	Cluster			
Frequency	1	2	3	Total
STDSR	3	0	0	3
STRTFO	0	0	0	0
STPAV	0	0	0	0

Seven out of twelve suppliers manufacture Binder 1. Table 41 shows the results separated by supplier. Based on Table 41, approximately 46 percent of all tests belong to Cluster 1, 38 percent to Cluster 2, and 16 percent to Cluster 3. The majority of suppliers, except for Supplier G, mainly belong to one cluster. Supplier C has 97.3 percent of its tests in Cluster 1; Supplier F has 94 percent in Cluster 2; Suppliers A, D, and E have 100 percent of their observations in Cluster 2. For Supplier G, dates when samples had been received were investigated. Almost all samples from Cluster 1 were received earlier than those from Cluster 3, possibly indicating some improvement in performance for that supplier for Binder 1. As for the analysis of the entire data set, Supplier C has a lot of observations in clusters corresponding to low performance for STDSR and STRTFO. The same tendency is observed for Binder 1. On the other hand, Supplier A was moved to the cluster that characterizes suppliers with high STDSR and STRTFO performance. Supplier F for this analysis shows stable performance for Binder 1, unlike the results from analyzing the entire data set.

Table 41. Suppliers Sorted by Cluster for Binder 1.

Supplier	Cluster			
Frequency	1	2	3	Total
A	0	16	0	16
B	0	0	1	1
C	36	0	1	37
D	0	2	0	2
E	0	2	0	2
F	1	16	0	17
G	7	0	13	20
Total	44	36	15	95

Table 42 shows descriptive statistics for each variable for Binder 2. Researchers also noted that the relationship among variables is different for Binders 1 and 2. Correlation between STDSR and STPAV did not change significantly from 0.833 for Binder 1 to 0.722 for Binder 2, but the correlation between STDSR and STRTFO and the correlation between STPAV and STRTFO changed significantly from -0.778 to 0.163 and from -0.468 to 0.500, respectively. This also supports the conclusion that a separate analysis for each binder is needed.

Table 42. Descriptive Statistics for Binder 2.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	185	0.342	0.239	0.040	1.660
STRTFO	68	0.558	0.638	0.095	2.691
STPAV	71	0.362	0.231	0.366	0.999

To be consistent, researchers set the number of clusters for Binder 2 at three. Table 43 contains descriptive statistics for each variable by cluster.

Table 43. Descriptive Statistics for Binder 2 by Cluster.

Cluster #	Variable	N	Mean	Std. Dev.	Minimum	Maximum
1	STDSR	114	0.196	0.078	0.040	0.480
	STRTFO	42	0.251	0.116	-0.095	0.495
	STPAV	45	0.253	0.092	0.048	0.484
2	STDSR	43	0.463	0.128	0.150	0.950
	STRTFO	15	0.408	0.207	-0.086	0.777
	STPAV	15	0.491	0.224	0.292	0.999
3	STDSR	28	0.753	0.224	0.530	1.660
	STRTFO	11	1.931	0.335	1.586	2.691
	STPAV	11	0.630	0.332	-0.366	0.776

Figure 6 shows box-plots for each variable by cluster.

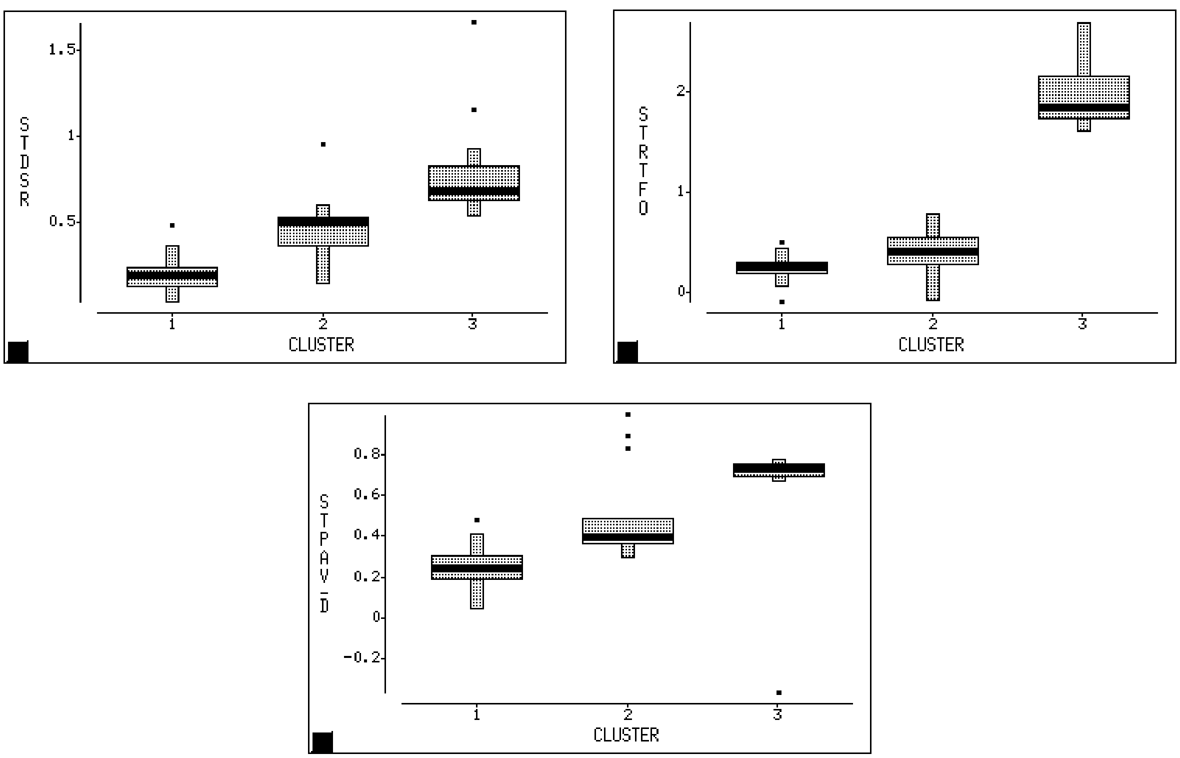


Figure 6. Box-plots for Each Variable by Cluster for Binder 2.

Researchers describe the clusters as compared to the average for each test for Binder 2 as follows:

- Cluster 1 contains STDSR values far below average and STRTFO and STPAV values below average.
- Cluster 2 contains above-average STDSR values, below-average STPAV values, and above-average STRTFO values.
- Cluster 3 contains STDSR and STPAV values far above average and above-average STRTFO values.

Table 44 presents the number of failures for the three variables by cluster. Each cluster has one failure. Cluster 1 and Cluster 2 have a failure in STRTFO; Cluster 3 has a failure in STPAV.

Table 44. Number of Failures by Cluster for Binder 2.

Variable	Cluster			
Frequency	1	2	3	Total
STDSR	0	0	0	0
STRTFO	1	1	0	2
STPAV	0	0	1	1

Nine suppliers produce Binder 2. Table 45 shows the result of separation of suppliers by cluster. Approximately 62 percent of the data are in Cluster 1, 23 percent in Cluster 2, and 15 percent in Cluster 3. Most of the suppliers, except for G and H, have counts mostly in one cluster. Supplier A has approximately 83 percent, Supplier D has 80 percent, Supplier F has 89 percent, Supplier C, Supplier E, and Supplier J have 100 percent of their observations in the first cluster, and Supplier B has 100 percent (1 of 1) of its observations in the second cluster.

Compared to the analysis for Binder 1, Supplier C is still in the cluster that is low in STDSR and STRTFO performance along with Supplier A, unlike for Binder 1. Supplier F is stable in performance, but for Binder 2 it is in the worst cluster in terms of performance.

Table 45. Suppliers Sorted by Cluster for Binder 2.

Supplier	Cluster			
Frequency	1	2	3	Total
A	35	5	2	42
B	0	1	0	1
C	8	0	0	8
D	12	2	1	15
E	8	0	0	8
F	24	2	1	27
G	0	9	20	29
H	10	24	4	38
J	21	0	0	21
Total	118	43	28	189

Table 46 shows descriptive statistics for each variable for Binder 5.

Table 46. Descriptive Statistics for Binder 5.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	91	0.213	0.371	0.323	3.420
STRTFO	53	0.165	0.206	0.327	0.804
STPAV	46	0.448	0.103	0.092	0.624

Table 47 shows descriptive statistics for each variable by cluster.

Table 47. Descriptive Statistics for Binder 5 by Cluster.

Cluster #	Variable	N	Mean	Std. Dev.	Minimum	Maximum
	STDSR	40	0.098	0.101	-0.323	0.250
1	STRTFO	31	0.067	0.188	-0.327	0.259
	STPAV	25	0.463	0.069	0.318	0.574
	STDSR	13	0.668	0.844	0.280	3.420
2	STRTFO	8	0.354	0.198	0.200	0.805
	STPAV	8	0.545	0.039	0.504	0.624
	STDSR	38	0.178	0.074	-0.162	0.350
3	STRTFO	14	0.276	0.098	0.127	0.518
	STPAV	13	0.361	0.121	0.092	0.522

Figure 7 shows box-plots for each variable by cluster.

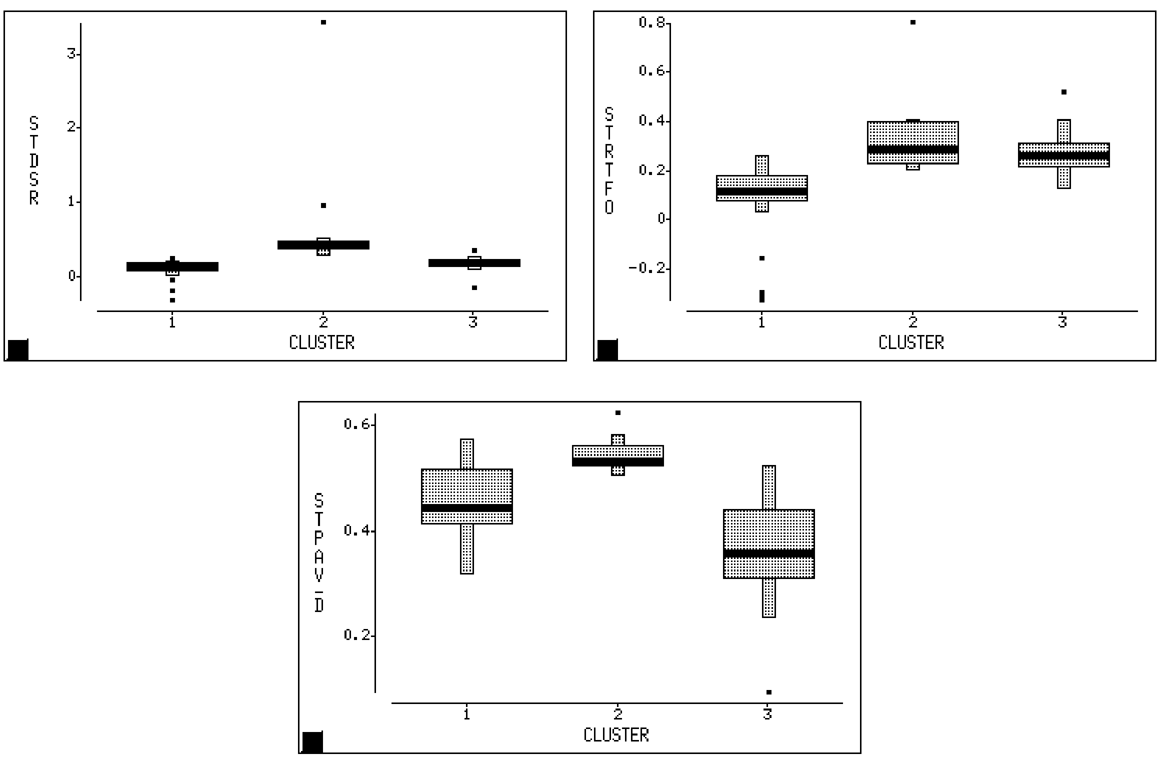


Figure 7. Box-plots for Each Variable by Cluster for Binder 5.

Researchers describe the clusters as compared to the average for that binder as follows:

- Cluster 1 contains STDSR and STRTFO values below average and a little above-average STPAV values.
- Cluster 2 contains STDSR and STRTFO values far above the average and above-average STPAV values.
- Cluster 3 contains a little below-average STDSR values, above-average STRTFO values, and below-average STPAV values.

Table 48 presents the number of failures for the three variables by cluster. Approximately 75 percent of failures (3 out of 4) for STDSR are in Cluster 1 and 25 percent (1 out of 4) are in Cluster 3. All six failures for STRTFO are in Cluster 1; there are no failures for STRTFO.

Table 48. Number of Failures by Cluster for Binder 5.

Variable	Cluster			
Frequency	1	2	3	Total
STDSR	3	0	1	4
STRTFO	6	0	0	6
STPAV	0	0	0	0

Seven suppliers produce Binder 5. Table 49 shows the result of separation of suppliers by cluster. Approximately 46.5 percent of the data are in the first cluster, 13.5 percent in the second, and 40 percent in the third. Based on Table 20, there is no good separation of suppliers, i.e., each supplier (except for Supplier B and Supplier E) has results in more than one cluster.

Table 49. Suppliers Sorted by Cluster for Binder 5.

Supplier	Cluster			
Frequency	1	2	3	Total
A	9	0	6	15
B	0	8	0	8
C	24	3	15	42
E	0	0	3	3
F	6	0	6	12
G	2	2	4	8
J	3	0	4	7
Total	44	13	38	95

Table 50 shows descriptive statistics for each variable for Binder 6. There are no failures for this binder.

Table 50. Descriptive Statistics for Binder 6.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	106	1.082	1.181	0.150	4.610
STRTFO	32	0.483	0.668	0	2.614
STPAV	32	0.64	0.060	0.514	0.776

Table 51 shows descriptive statistics for each variable by cluster.

Table 51. Descriptive Statistics for Binder 6 by Cluster.

Cluster #	Variable	N	Mean	Std. Dev.	Minimum	Maximum
1	STDSR	61	0.438	0.094	0.150	0.620
	STRTFO	21	0.173	0.116	0	0.395
	STPAV	21	0.686	0.053	0.580	0.776
2	STDSR	28	0.999	0.286	0.470	1.870
	STRTFO	8	0.572	0.127	0.405	0.786
	STPAV	8	0.662	0.076	0.514	0.736
3	STDSR	17	3.531	1.010	1.800	4.610
	STRTFO	3	2.420	0.231	2.164	2.614
	STPAV	3	0.620	0.025	0.600	0.648

Figure 8 shows box-plots for each variable by cluster.

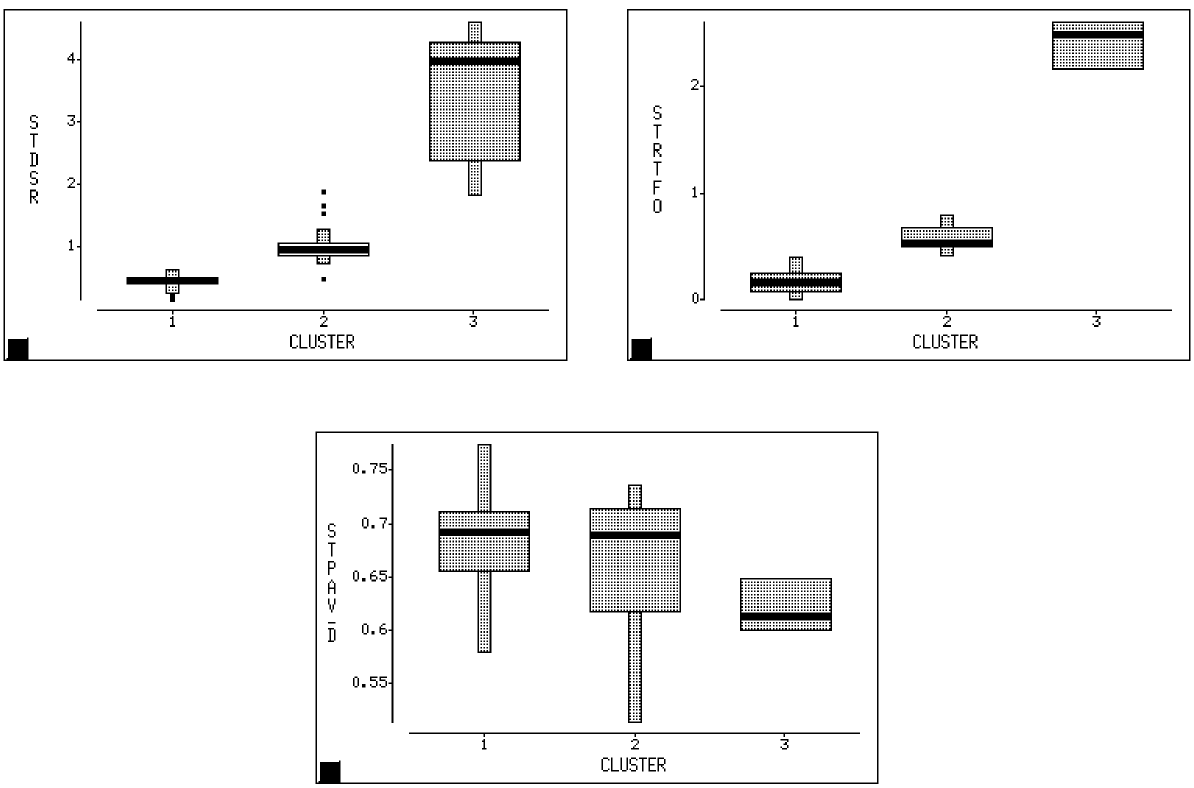


Figure 8. Box-plots for Each Variable by Cluster for Binder 6.

Researchers describe the clusters as compared to the average for the binder as follows:

- Cluster 1 contains STDSR, STRTFO, and STPAV values above average.
- Cluster 2 contains below-average STDSR values, above-average STPAV values, and STRTFO values far below average.
- Cluster 3 contains STDSR and STPAV values far below average and STRTFO values far above average.

Table 52 presents the number of failures for the three variables by cluster and shows no failures for this binder.

Table 52. Number of Failures by Cluster for Binder 6.

Variable	Cluster			
Frequency	1	2	3	Total
STDSR	0	0	0	0
STRTFO	0	0	0	0
STPAV	0	0	0	0

Eight suppliers produce Binder 6. Table 53 shows the result of separation of suppliers by cluster. Cluster 1 contains approximately 57.5 percent, Cluster 2 contains 26.5 percent, and the Cluster 3 contains 16 percent of all observations. Three out of eight suppliers (Suppliers F, J, and K) provided approximately 88 percent of all samples. For this binder, more than two-thirds of all observations have results for only one QA test (DSR).

Table 53. Suppliers Sorted by Cluster for Binder 6.

Supplier	Cluster			
Frequency	1	2	3	Total
A	1	3	0	4
D	1	1	0	2
E	3	0	0	3
F	10	6	0	16
J	43	0	0	43
K	2	18	14	34
L	1	0	0	1
M	0	0	3	3
Total	61	28	17	106

Tables 54 and 55 show descriptive statistics for each variable for Binder 8 and for each cluster.

Table 54. Descriptive Statistics for Binder 8.

Variable	N	Mean	Std. Dev.	Minimum	Maximum
STDSR	49	0.480	0.343	0.998	1.030
STRTFO	34	0.007	0.231	0.854	0.514
STPAV	24	0.709	0.415	0.630	0.887

Table 55. Descriptive Statistics for Binder 8 by Cluster.

Cluster #	Variable	N	Mean	Std. Dev.	Minimum	Maximum
1	STDSR	34	0.601	0.157	0.260	1.030
	STRTFO	30	0.047	0.119	-0.127	0.514
	STPAV	20	0.840	0.075	0.570	0.887
2	STDSR	12	0.423	0.113	0.140	0.550
	STRTFO	2	0.180	0.029	0.159	0.200
	STPAV	2	-0.613	0.024	-0.630	-0.596
3	STDSR	3	-0.662	0.341	-0.998	-0.317
	STRTFO	2	-0.777	0.107	-0.854	-0.702
	STPAV	2	0.713	0.135	0.618	0.808

Figure 9 shows box-plots for each variable by cluster.

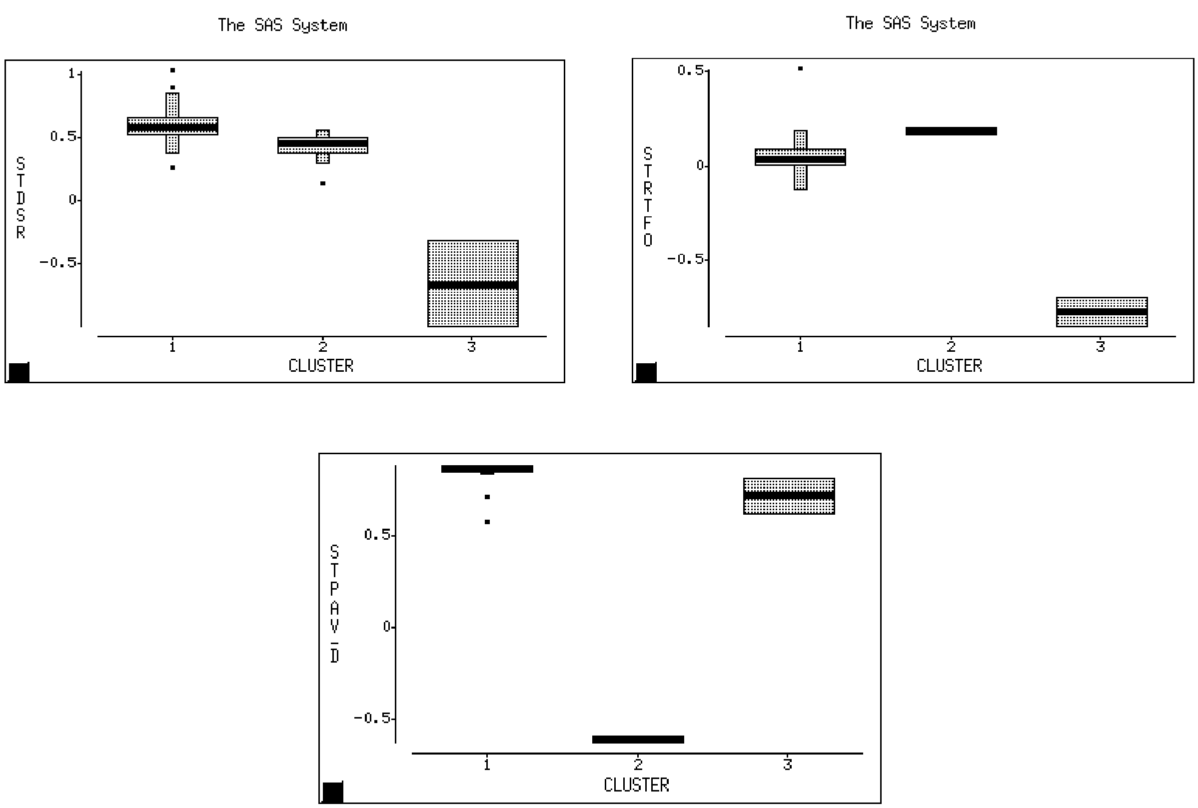


Figure 9. Box-plots for Each Variable by Cluster for Binder 8.

Based on descriptive statistics in Tables 54 and 55 and box-plots in Figure 9, researchers describe the clusters for Binder 8 as follows:

- Cluster 1 contains STDSR, STRTFO, and STPAV values above average.
- Cluster 2 contains below-average STDSR values, above-average STPAV values, and STRTFO values far below average.
- Cluster 3 contains STDSR and STPAV values far below average and STRTFO values far above average.

Table 56 presents the number of failures for the three variables by cluster. Cluster 1 has 78 percent (7 out of 9) of all failures for the STRTFO, Cluster 2 has 100 percent (2 out of 2) of all failures for STPAV, and Cluster 3 has 100 percent of all failures for STRDSR and 22 percent (2 out of 9) for STRTFO.

Table 56. Number of Failures by Cluster for Binder 8.

Variable	Cluster			
Frequency	1	2	3	Total
STDSR	0	0	3	3
STRTFO	7	0	2	9
STPAV	0	2	0	2

Six suppliers produce Binder 8. Table 57 shows the result of separation of suppliers by cluster. Cluster 1 contains approximately 75 percent, Cluster 2 contains 20 percent, and Cluster 3 contains 5 percent of all observations.

Table 57. Suppliers Sorted by Cluster for Binder 8.

Supplier	Cluster			
Frequency	1	2	3	Total
A	10	3	0	13
C	1	0	0	1
D	11	1	1	13
E	2	0	0	2
F	21	8	2	31
G	1	0	0	1
Total	46	12	3	61

Cluster analysis results in a good separation of suppliers (i.e., observations from one supplier belong mainly to one cluster) if there is high correlation among variables in the data set (e.g., see, Table 41 for Binder 1). When the correlation among variables is low, cluster analysis does not seem to be very useful in that there is not a good separation of suppliers (i.e., observations from one supplier evenly split among two or more clusters) as can be seen in Table 49 for Binder 5 or Table 53 for Binder 6. In fact, for Binder 5, approximately one-half and for Binder 6 almost two-thirds have only one variable, DSR. Also, the correlations between DSR and PAVDSR are low for both binders, 0.267 for Binder 5 and -0.347 for Binder 6, as opposed to -0.778 for Binder 1 for which a good separation of suppliers was obtained.

In summary, researchers separated suppliers into three well-defined groups using statistical clustering methods for each binder. In each group, measured DSR values for all three aging states (original, after RTFO, and after RTFO and PAV) were similar. Thus, researchers found groups of suppliers more likely than others to be out of specification for a particular binder. With this result, the Colorado archived data provided useful information about the Colorado PG binders and suppliers.

Researchers recommend clustering by binder because for some suppliers performance in terms of specification compliance changes by binder. In addition, this type of analysis may contribute to the definition of a formal classification scheme, indicating rules for assigning new binders to clusters for identification and diagnostic purposes.

OREGON

Researchers also evaluated Oregon emulsion data to determine if cluster analysis could be used to identify materials and corresponding suppliers that consistently fail specific property requirements. Unfortunately, for the data set evaluated, all emulsion test results met specifications so cluster analysis was not pursued.

TEXAS

In contrast, Texas data cannot be easily used in a binder QA program. After extensive effort to archive data in a usable form, researchers summarized the statistical information that could be extracted and analyzed the data using CART. They analyzed PG64-22 and PG70-22 binder data including critical selected properties measured in the DSR (DSR on unaged binder, RTFO-DSR, and PAV-DSR) and the BBR (BBR stiffness *S* and *m*-value). For CRS2 and CRS2P emulsions, researchers selected Saybolt viscosity measured at two temperatures, demulsibility, penetration of the residue, and ductility of the residue as critical properties.

There were 322 data records from 20 suppliers for the PG64-22 data with some missing values for each variable and all but three records labeled Pass, Fail, or For Information Only. CART analysis produced a classification tree with six classes. Class 6, with a PAV-DSR value greater than 3.5 MPa, contained all three of the Fail values from two of the suppliers, one of 27 For Information Only values, and five of 289 Pass values.

There were 543 data records from 21 suppliers for the PG70-22 data with some missing values for each variable and all but 17 records labeled Pass, Fail, or For Information Only. CART analysis produced a classification tree with three classes. Class 1, with a STRTFO value of less than 0.002 (or a STDSR value less than 0.009 for missing STRTFO values), contained the bulk of the Failures and For Information Only values (8 of 11 Fail and 54 of 79 For Information Only) and only one of the 436 Pass values. Class 2 contained two more of the 11 Fail values, seven additional For Information Only values, and no Pass values. Class 2 required STRTFO values greater than 0.002 (or STDSR values greater than 0.009 for missing STRTFO values) and standardized *m*-values (STM) values greater than -0.002. Class 3 contained the remaining values, including all but one of the Pass values, one Fail value, and 18 For Information Only values. Conclusions from this analysis point to Fail classification based on low RTFO and DSR values. Most For Information Only values grouped with the Fail values, and some suppliers produced an unusually large percentage of Fail and For Information Only values.

There were 273 data records from 15 suppliers for the PG76-22 data with a typical record labeled Pass (216 values), Fail (1 value), or For Information Only (55 values). The PG76-22 data were not analyzed using CART due to the small number of failures.

There were 134 data records from 9 suppliers for the CRS2 data with a typical record labeled Pass (108 values), Fail (19 values), or For Information Only (5 values). Two records labeled Meets Specifications Only were not analyzed. CART analysis produced a classification tree with three classes. Classes 1 and 3 combined contained 13 of 19 Fail values and two of five For Information Only values. Class 1 required Saybolt viscosity values at 122 °F (50 °C) less than 144.5 s if data were available. Class 3 required Saybolt viscosity values at 122 °F (50 °C) greater than 493 s Conclusions from this analysis point to classification of a Failure based on low or high Saybolt viscosity values. Researchers identified a single supplier with both the largest number (13 of 19) and largest percentage (68 percent) of Fail values. The other Fail values were distributed over five other suppliers (one of eight samples, two of 22 samples, two of 40 samples, and one of three samples).

There were 297 data records from 13 suppliers for the CRS2P data, but the records were labeled Pass (248 values), Fail (25 values), For Information Only (9 values), Meets Specifications Only (14 values), and Variation from Specifications is Immaterial (1 value). Analysis of the CRS2P data did not produce meaningful classification rules, possibly due to a significant number of data records that were categorized with labels other than Pass, Fail, or For Information Only.

LABORATORY TESTING PROGRAM

Based on discussions with TxDOT personnel and their field experience, researchers selected specific factors highlighted in Table 1 for inclusion in a laboratory testing program. They designed two types of experiments to identify factors that have the most impact on RTFO-DSR that may change between production and use during construction: an extensive laboratory experiment and a limited field experiment. They selected this binder property based on its direct relationship with performance in terms of resistance to rutting in the early life of an asphalt concrete pavement, frequent use as a QA parameter by other state DOTs, and equipment availability in the TxDOT districts. The laboratory experiment utilized supplier samples and simulation of storage conditions and contamination. The field experiment involved obtaining field samples and their corresponding supplier test results. This section describes the experimental designs and the results and analysis to date for both types of experiments.

Laboratory Experiment

The factors for asphalt cements in the laboratory experiment were Modifier (with 2 levels: modified PG76-22 (L1) and unmodified PG64-22 (L2)), Contamination (with 3 levels: no contamination (L1), contamination of transport truck (L2), and contamination of contractor tank (L3)), Storage Time (with 3 levels: 1 week (L1), 1 month (L2), and 2 months (L3)), and Storage Temperature (with 2 levels: 335 °F (168 °C) (-1) and 375 °F (191 °C) (1)). In addition to these factors, Supplier (with 2 levels: Supplier 1 and Supplier 2) was introduced as a block to remove excess variation due to differences in manufacturing process among suppliers. Each factor-level combination corresponds to a different treatment, and researchers plan to test two replicate samples (with two replicate measurements on each sample) for each combination. Prior to treatment (storage at elevated temperature), researchers fabricated each asphalt cement sample by pouring a small amount of the material into an ointment tin, flushing the tin with nitrogen to simulate storage in a closed tank by precluding aging at the surface, and sealing the lid with a stiff asphalt cement. After treatment, the response variable was measured as either (1) the difference in RTFO-DSR before and after each treatment was applied or (2) the relative difference based on the initial value. Test runs corresponding to treatments were randomized to average out the effects of nontreatment factors on the responses. This resulted in a D-optimal design shown in Table 58. When testing is completed next year, this design will allow for estimation of all main effects and two-way interactions.

The factors for emulsions in the laboratory experiment were Modifer (with 2 levels: modified CRS-2P (L1) and unmodified CRS-2 (L2)), Contamination (with 2 levels: no contamination (L1), contamination of transport truck (L2)), Storage time (with 3 levels: 2 days

(L1), 1 week (L2), and 1 month (L3)), and Storage temperature (with 2 levels: 150 °F (66 °C) (-1) and 180 °F (82 °C) (1)). Supplier (with 2 levels: Supplier 1 and Supplier 2) was used as a block in the design shown in Table 59. Again, researchers plan to test two replicate samples (with two replicate measurements on each sample) for each factor level combination, and they will estimate all main effects and two-way interactions when testing is completed next year. For the emulsion samples, water will be removed to produce a residue by the stirred-can method developed during TxDOT Research Project 0-1710 (*25*). These samples will not be sealed because the water vapor released during storage at elevated temperature will preclude aging and simulate storage in a closed tank.

Table 58. Laboratory Experimental Design for Asphalt Cements.

Rows	Modifier	Contamination	Time	Temperature	Supplier	Response Variable
1	L2	L2	L1	-1	1	.
2	L2	L3	L3	1	1	.
3	L2	L2	L2	1	1	.
4	L1	L2	L1	-1	1	.
5	L2	L1	L2	1	1	.
6	L1	L2	L2	1	1	.
7	L1	L3	L3	-1	1	.
8	L2	L3	L2	-1	1	.
9	L2	L1	L1	-1	1	.
10	L1	L3	L1	1	1	.
11	L1	L1	L2	-1	1	.
12	L2	L1	L3	1	1	.
13	L1	L2	L3	1	2	.
14	L1	L2	L2	-1	2	.
15	L1	L3	L2	1	2	.
16	L1	L3	L1	-1	2	.
17	L1	L1	L3	-1	2	.
18	L2	L1	L2	-1	2	.
19	L2	L2	L1	1	2	.
20	L2	L2	L3	-1	2	.
21	L2	L3	L1	1	2	.
22	L2	L3	L3	-1	2	.
23	L1	L1	L3	1	2	.
24	L1	L1	L1	1	2	.

Table 59. Laboratory Experimental Design for Emulsions.

Rows	Modifier	Contamination	Time	Temperature	Supplier	Response Variable
1	L1	L1	L2	-1	1	.
2	L1	L2	L1	-1	1	.
3	L1	L2	L3	1	1	.
4	L2	L1	L3	-1	1	.
5	L1	L1	L3	-1	1	.
6	L2	L2	L2	-1	1	.
7	L1	L1	L2	1	1	.
8	L2	L2	L1	-1	1	.
9	L2	L1	L1	1	1	.
10	L2	L2	L3	1	1	.
11	L1	L2	L2	1	1	.
12	L1	L1	L1	1	1	.
13	L2	L1	L3	1	2	.
14	L1	L2	L3	-1	2	.
15	L1	L2	L2	-1	2	.
16	L2	L1	L2	-1	2	.
17	L2	L2	L2	1	2	.
18	L2	L1	L1	-1	2	.
19	L2	L2	L1	1	2	.
20	L2	L1	L2	1	2	.
21	L1	L1	L3	1	2	.
22	L2	L2	L3	-1	2	.
23	L1	L1	L1	-1	2	.
24	L1	L2	L1	1	2	.

Laboratory Experiment Results

Researchers analyzed the laboratory experimental RTFO-DSR data collected to date for asphalt cements under various factor-level combinations to identify the important factors that affect a change in this property. The factors investigated with the partial data set include Modifier with 2 levels (L1: PG 76-22, L2: PG 64-22), Contamination with 3 levels (L1: no contamination, L2: 100 of 6,000 gallons, L3: 500 of 20,000 gallons), Time with 2 levels (L1: 1 week, L2: 1 month), Temperature with 2 levels (-1: 335 °F (168 °C), 1: 375 °F (191 °C)), and Supplier with 2 levels (1: Supplier 1, 2: Supplier 2). It should be noted that during testing, the actual measured low temperature level was 340 °F (171 °C). The intermediate analysis described represents two-thirds of the laboratory experiment shown in Table 58. Researchers will complete this experiment along with the experiment shown in Table 59 in the second year of the project. Analysis and conclusion results will change as more data is gathered.

Table 1 contains the row numbers corresponding to factor-level combinations in Table 58 and RTFO-DSR test results reported to the nearest 0.1 kPa. Researchers obtained test results by measuring $G^*/\sin\delta$ at each factor-level combination for two or three samples (with two replicate measurements on each sample). The average test results over those two or three samples are reported as the RTFO-DSR test value (Y1) for each factor-level combination. Researchers compared these average test results against the results from the control samples (which were not stored at elevated temperatures) for the corresponding (Modifier, Supplier) combination. Control samples were assumed contamination-free at the supplier location to correspond with contamination locations simulated in the experiment in the transport truck and the contractor tank. The average (over two samples) RTFO-DSR test results for the control samples are given as Y0 values in Table 60. Note that only Modifier and Supplier are relevant factors for control samples, which resulted in four different Y0 values. Thus researchers estimated the change in RTFO-DSR based on the differences between average test results for control samples (Y0) and average test results for the stored samples obtained by simulation of storage conditions and contamination (Y1). The difference Y1 - Y0 and the relative difference (Y1 - Y0)/Y0 were both considered as response variables simulating a change in the RTFO-DSR property between supplier and field samples. The relative difference represents the percent change (after multiplying by 100) for the stored sample compared to that of the control sample.

Researchers encountered one specific testing anomaly when measuring the results presented in Table 60. A majority of both unmodified and modified samples formed a thick, extremely stiff crust after storage at elevated temperatures. This effect was noted after both 1-week and 1-month storage times. When the material was prepared for the RTFO, homogeneity was difficult to achieve, and small stiff flakes remained in the material. Care was taken to avoid these flakes when preparing the DSR sample for testing, but some results may not be representative due to the presence of this stiff material. The cause of the formation of the crust is unknown at this time, and investigation into the cause and possible remedy is currently ongoing.

Table 60 Partial Laboratory Experiment Results for Asphalt Cements (to nearest 0.1).

Rows	Y0 (kPa)	Y1 (kPa)	Y1 - Y0 (kPa)	(Y1 - Y0)/Y0
1	2.4	41.7	39.3	16.2
3	2.4	78.9	76.4	31.6
4	3.4	8.6	5.2	1.5
5	2.4	31.6	29.2	12.1
6	3.4	8.4	5.0	1.5
8	2.4	155.6	153.2	63.3
9	2.4	21.9	19.4	8.0
10	3.4	37.4	34.0	10.0
11	3.4	29.4	26.0	7.7
14	4.2	7.3	3.0	0.7
15	4.2	146.7	142.4	33.8
16	4.2	13.2	9.0	2.1
18	1.6	50.2	48.6	31.2
19	1.6	31.4	29.9	19.2
21	1.6	87.7	86.1	52.3
24	4.2	7.8	3.6	0.9

With these limitations in mind, Figure 10 presents the overlay chart for the difference Y1 - Y0 and the relative difference (Y1 - Y0)/Y0. Note that there is some discrepancy between the pattern of Y1 - Y0 and that of (Y1 - Y0)/Y0 for row numbers 15 and 21. The value for row number 15 is greater than that of row number 21 for Y1 - Y0 but vice versa for (Y1 - Y0)/Y0. In addition, for the factor-level combination corresponding to row number 8, there was high sample-to-sample variability for the RTFO-DSR test results (ranging from 102.7 to 206.2 kPa). Researchers thus obtained the average RTFO-DSR test value (Y1) for row number 8 based on five samples rather than two or three samples for the other factor-level combinations.

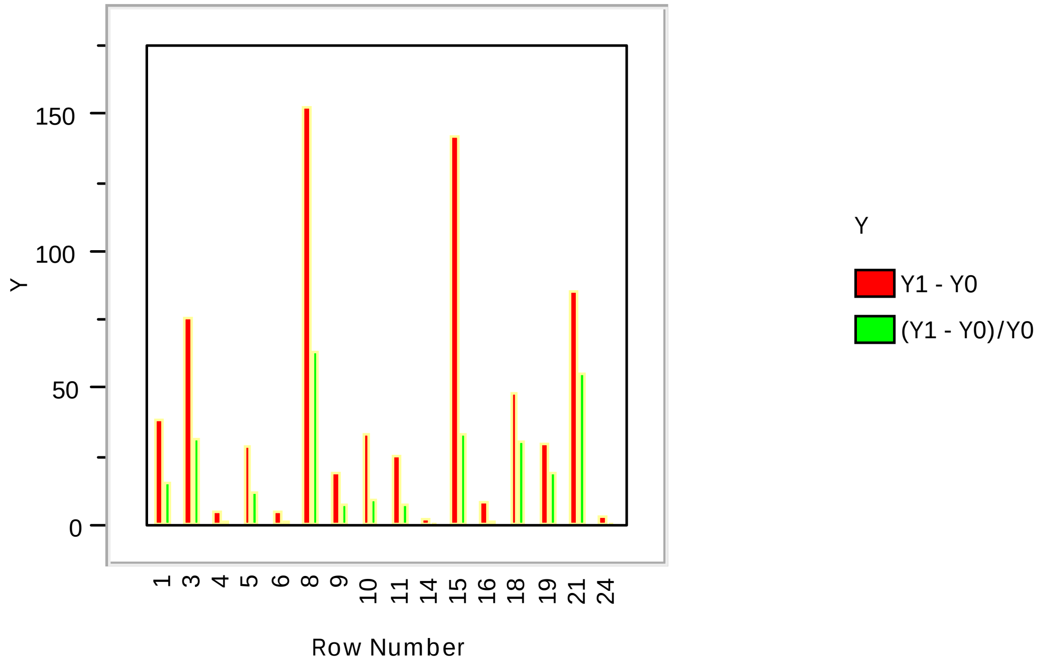


Figure 10. Overlay Chart for Differences and Relative Differences.

Tables 61 and 62 contain the analysis of variance (ANOVA) results for each type of response variable, Y1 - Y0 and (Y1 - Y0)/Y0. Researchers considered a model having Modifier, Contamination, Time, Temperature, and Supplier as main effects, and Modifier * Time, Modifier * Temperature, Contamination * Time, Time * Temperature, and Temperature * Supplier as interaction effects.

Table 61. Analysis of Variance for Differences Y1 - Y0.

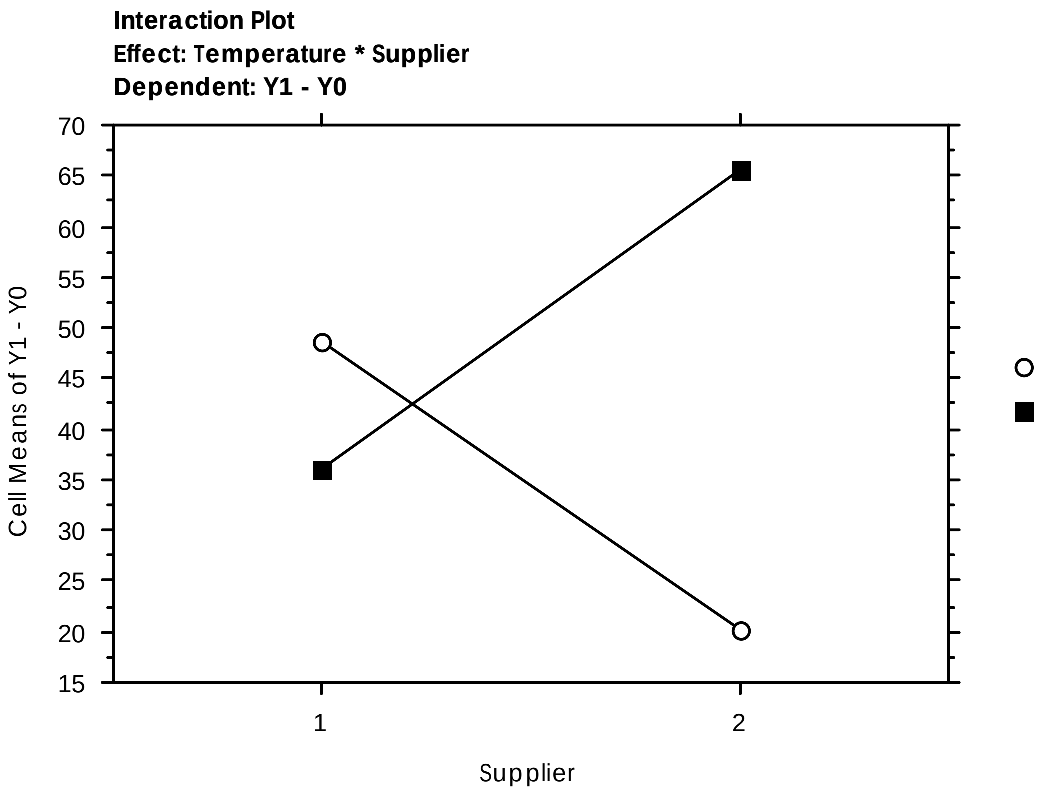
Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Value	P-Value
Modifier	1	4692.32	4692.32	13.89	0.0336
Contamination	2	13109.17	6554.59	19.41	0.0192
Time	1	7468.62	7468.62	22.11	0.0182
Temperature	1	252.44	252.44	0.75	0.4509
Supplier	1	150.93	150.93	0.45	0.5517
Modifier * Time	1	587.00	587.00	1.74	0.2790
Modifier * Temperature	1	352.85	352.85	1.04	0.3820
Contamination * Time	2	15.21	7.61	0.02	0.9779
Time * Temperature	1	105.45	105.45	0.31	0.6153
Temperature * Supplier	1	1436.90	1436.90	4.25	0.1312
Residual	3	1013.22	337.74		

Table 62. Analysis of Variance for Differences Y1 - Y0/Y0.

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Value	P-Value
Modifier	1	2334.89	2334.89	137.83	0.0013
Contamination	2	1769.50	884.75	52.23	0.0047
Time	1	1166.53	1166.53	68.86	0.0037
Temperature	1	3.51	3.51	0.21	0.6797
Supplier	1	114.45	114.45	6.76	0.0804
Modifier * Time	1	298.91	298.91	17.65	0.0246
Modifier * Temperature	1	135.13	135.13	7.98	0.0665
Contamination * Time	2	232.25	116.13	6.86	0.0761
Time * Temperature	1	3.60	3.60	0.21	0.6762
Temperature * Supplier	1	572.92	572.92	33.82	0.0101
Residual	3	50.82	16.94		

Table 61 (based on Y1-Y0) shows that the change in RTFO-DSR is significantly affected by the main effects of Modifier, Contamination, and Time at a 5 percent significance level (α = 0.05). In terms of the relative differences, (Y1-Y0)/Y0, researchers reached a somewhat different conclusion. Table 62 shows that the interaction effects Modifier * Time and Temperature * Supplier were significant at a 5 percent significance level (α = 0.05), suggesting that the individual factor effects of Modifier, Time, Temperature, or Supplier can only be assessed conditional on each level of the other factor. For example, the effect of Temperature on the response variable needs to be determined for each level of Supplier separately, since the effect of Temperature varies with Supplier. Figure 11 shows the interaction plots between Temperature and Supplier. For one of the suppliers (1), the change was larger at the lower temperature (○), but for the second supplier (2), the change was larger at the higher temperature (■). The main effect of Contamination was also statistically significant at a 5 percent significance level (α = 0.05).

As more data (observations corresponding to 2-month storage time) are obtained, researchers expect to clarify ambiguous factor effects that are at the margin (corresponding to P-values between 0.05 and 0.1).



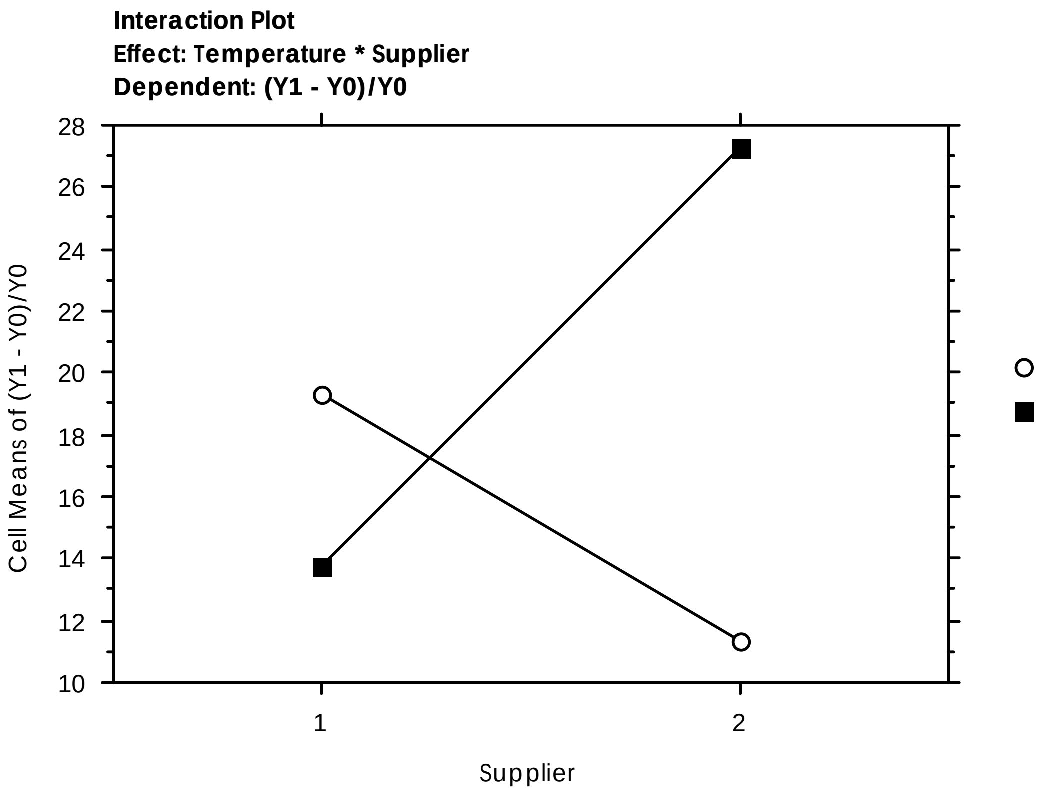


Figure 11. Interaction Plots between Supplier and Temperature

(o: 335 °F (168 °C), ■: 375 °F (191 °C)).

Field Experiment

For the field experiment, researchers could not include all of the factors in the laboratory experiment in the design because some of the factors such as Contamination and Storage temperature were uncontrollable in the field. Also, they needed to restrict the number of test runs to a small number due to the difficulty of obtaining the samples from the field with their corresponding test results from the supplier tank. Thus, researchers proposed a screening design shown in Tables 63 and 64. The factors in the field experiments were Modifier (with 2 levels: modified PG76-22 (L1) and unmodified PG64-22 (L2) for the asphalt cements or modified CRS-2P (L1) and unmodified CRS-2 (L2) for the emulsions) and Storage time (with 2 levels: more than 1 week (1) and less than 1 week (-1)). Storage time was taken as the sum of the storage times at both the supplier and contractor locations. Storage temperature was used as a covariate (an uncontrollable variable that influences the response but is itself unaffected by any other experimental factors) in contrast to the laboratory experimental design where Storage temperature was also one of controllable factors. Supplier (with 2 levels (1) and (2) to be determined as field samples are identified) was again used as a block to increase precision in the estimation of factor effects. As for the laboratory experiment, the response variable will be measured as either (1) the difference in the RTFO-DSR before and after each treatment was applied or (2) the relative difference based on the initial value.

Table 63. Field Experimental Design for Asphalt Cements.

Rows	Modifier	Storage Time	Supplier	Storage Temperature	Response Variable
1	L1	1	1		
2	L2	-1	1	.	.
3	L1	1	1	.	.
4	L1	-1	1	.	.
5	L1	-1	1	.	.
6	L2	1	1	.	.
7	L2	1	2	.	.
8	L2	1	2	.	.
9	L1	-1	2	.	.
10	L2	-1	2	.	.
11	L1	1	2	.	.
12	L2	-1	2	.	.

Table 64. Field Experimental Design for Emulsions.

Rows	Modifier	Storage Time	Supplier	Storage Temperature	Response Variable
1	L1	1	1	.	.
2	L2	-1	1	.	.
3	L1	1	1	.	.
4	L1	-1	1	.	.
5	L1	-1	1	.	.
6	L2	1	1	.	.
7	L2	1	2	.	.
8	L2	1	2	.	.
9	L1	-1	2	.	.
10	L2	-1	2	.	.
11	L1	1	2	.	.
12	L2	-1	2	.	.

Field Experiment Results

To date, an extensive effort by researchers to locate field samples where all data and storage information is available has produced only approximately half of the required asphalt cement samples and only one emulsion sample. If possible, researchers may utilize field samples identified for use in TxDOT Project 0-1710 to increase the number of emulsion samples for this project. Obtaining field samples that meet the requirements of the experimental designs was also hampered by the peak of construction season where materials were utilized in less than a week after production. For these reasons, laboratory testing and analysis of field sample results will continue in the next year.

As a preliminary example of the potential problem, two PG76-22 field samples showed increases in RTFO-DSR of 18.7 percent and 33.0 percent (17.3 percent and 28.4 percent of the mean value, respectively). Two PG64-22 field samples showed smaller increases in the same parameter, of approximately 0 percent and 13.6 percent (12.7 percent of the mean value). Based on the repeatability cited in AASHTO TP5 (11 percent of the mean value), three of the four samples exhibited a substantial change in RTFO-DSR. Researchers will explore the implications of these differences in the second year of the project as the field and laboratory experiments are completed. In addition, they will carefully reexamine survey results from TxDOT districts where field samples have shown significantly different properties than corresponding samples taken from supplier tanks.

CHAPTER 4. RECOMMENDED CHANGES

TxDOT samples and approves asphalt materials at the source, and these materials are utilized in highway projects without consideration of possible changes in properties that may occur subsequent to approval. Toward the primary goal of evaluating the current TxDOT QA program for binders and recommending revisions as necessary, this interim report documents an initial assessment based on (1) an extensive information search and review that included two detailed surveys of TxDOT districts and nine other state DOTs and (2) partial results from a comprehensive laboratory testing program. This assessment produced the following preliminary recommendations toward improving the TxDOT binder QA program:

- There exists a definite need to conduct a comprehensive evaluation of the current TxDOT binder QA program. This type of evaluation would require funding and time resources beyond the scope of this project. TxDOT may utilize a future research project or statistical support contract to accomplish this substantial task.
- Following a comprehensive evaluation, researchers expect implementation of revisions. Researchers recommend the appointment of a binder QA program manager. In addition, they recommend education of all employees on all aspects of the revised binder QA program to ensure maximum benefit at the least cost.
- Researchers recommend that the binder QA program established by TxDOT be only one tool in a system aimed at improving quality of the materials utilized during pavement construction and thus prolonging pavement life. Other recommended tools include required QC plans for both binder suppliers and asphalt paving contractors. They also suggest training programs for all binder technicians and personnel responsible for taking samples. Researchers recommend a round-robin program to establish the testing variability for selected binder QA parameters across multiple laboratories as another tool in the system.
- Data collected in the binder QA program should be stored in a user-friendly database that can be accessed by TxDOT district personnel. In addition, the number of labels for data records should be reduced to three, if possible, to facilitate the production of meaningful statistical results.
- Researchers also strongly recommend that data be organized and analyzed frequently to

detect problems or show historical specification compliance for different binders and suppliers. TxDOT may use historical data to set field sampling rates by binder and supplier on an annual basis. Implementation of this recommendation will require time to educate suppliers, contractors, and TxDOT personnel.

- When field samples are taken, contractors or TxDOT personnel must label them with the corresponding acceptance laboratory number based on the supplier sample. With this information and a readily accessible database, statistical analysis can be used to gather further evidence of the potential problem of binder properties changing subsequent to acceptance.
- Based on the partial results from the laboratory experiment, preliminary analysis indicated that modifier, time, and contamination produce a significant change in the selected binder property (RTFO-DSR). With this result, researchers recommend the inclusion of special handling requirements in QC plans for both suppliers and contractors. Contractors may need to check for both specification compliance of supplier and/or field samples and total storage time at elevated temperatures.
- Data for a particular binder shipment should include storage times and storage temperatures for both the supplier and contractor locations. Researchers recommend that this information, along with specification compliance of supplier and/or field samples, be stored in the same database as pavement performance data throughout the life of the pavement. This may help in forensic investigations and allow future research projects to examine the effect of binder noncompliance on pavement performance. Development of these types of models is urgently needed.

In the next year of this project, researchers will complete the laboratory and field experiments and statistically analyze the results as described in this report. Researchers will also assess recommended changes including the possibility of field sampling and testing in district laboratories based on different field sampling rates for each binder/supplier combination.

REFERENCES

(1) Mitra, A., *Fundamentals of Quality Control and Improvement*, 1998, New Jersey: Prentice Hall.

(2) Summers, D. C. S., *Quality*, 1997, New Jersey: Prentice Hall.

(3) Aguayo, R., *Dr. Deming: The American Who Taught the Japanese About Quality*, 1990, New York: Simon & Schuster.

(4) Gabor, A., *The Man Who Discovered Quality: How W. Edwards Deming Brought the Quality Revolution to America: The Stories of Ford, Xerox, and GM*, 1990, New York: Times Books/Random House.

(5) Deming, W. E., *Out of the Crisis*, 1986, Cambridge, Massachusetts, Massachusetts Institute of Technology.

(6) American Association of State Highway and Transportation Officials (AASHTO), *AASHTO Provisional Standards*, April Edition, 2000, Washington, D.C.: American Association of State Highway and Transportation Officials.

(7) Northeast Center of Excellence for Pavement Technology (NECEPT), *Northeastern Superpave Center News*, Vol. 2, No. 2, Summer 2000.

(8) Anderson, D., "QA Specification for Binders: An Update," presentation at NEAU/PG Meeting, Washington, D.C., February 14, 2001, http://www.superpave.psu.edu.

(9) Northeast Center of Excellence for Pavement Technology (NECEPT), *NEAU/PG Split Sampling Program* – Version as of June 13, 2001, http://www.superpave.psu.edu.

(10) Northeast Center of Excellence for Pavement Technology (NECEPT), *PGAB Specification* – Version as of March 22, 2001, http://www.superpave.psu.edu.

(11) The Asphalt Institute, *Performance Graded Asphalt Binder Specification and Testing*, Superpave Series No. 1 (SP-1), 1994, Lexington, Kentucky: The Asphalt Institute.

(12) Bahia, H. U., D. I. Hanson, M. Zeng, H. Zhai, M. A. Khatri, and R. M. Anderson, *Characterization of Modified Asphalt Binders in Superpave Mix Design*, NCHRP Report 459, 2001, Washington, D.C.: National Cooperative Highway Research Program, Transportation Research Board, National Research Council.

(13) University of California at Berkeley, Oregon State University, Austin Research Engineers, Inc., and SWK Pavement Engineering, *Stage 1 Validation of the Relationship Between Asphalt Properties and Asphalt-Aggregate Mix Performance*, SHRP-A-398, 1994, Washington, D.C.: Strategic Highway Research Program, National Research Council.

(14) American Society for Testing and Materials (ASTM), *1999 Annual Book of ASTM Standards, Vol. 04.03: Road and Paving Materials; Pavement Management Technologies*, 1999, Philadelphia: American Society for Testing and Materials.

(15) Roberts, F., P. Kandhal, E. R. Brown, D. Lee, and T. Kennedy, *Hot Mix Asphalt Materials, Mixture Design, and Construction*, 1991, Lanham, Maryland: National Asphalt Pavement Association Education Foundation.

(16) Charmot, S., *Pay Adjustment Factors for Superpave Performance Graded Asphalt Binders*, M.S. Thesis, University of Nevada, Reno, 1997.

(17) Monismith, C., J. Epps, F. Finn, R. G. Hicks, and R. Leahy, Design, Construction, and Performance of Asphalt Paving Mixes, Course Notes, University of California, Institute

of Transportation Studies and University Extension, Richmond, California, December 9-11, 1993.

(18) Texas Department of Transportation, Construction Division, Materials Section, Procedure for Asphalt Inspection at the Source, October 1998.

(19) Texas Department of Transportation, Construction Division, Materials Section, *Manual of Testing Procedures*, 2001.

(20) Behrens, M., Quality Assurance Program for Asphalt Binders, Memo to All District Engineers, May 12, 1999.

(21) Texas Department of Transportation, Construction Division, Materials Section, Draft Version 1 of Special Specification for Quality Control / Quality Assurance of Asphalt Binders for Bituminous Mixtures, February 1996.

(22) Texas Department of Transportation, Construction Division, Materials Section, Draft Version 2 of Special Specification for Field Acceptance of Performance Graded Asphalt Binders for Bituminous Mixtures, October 1997.

(23) Texas Department of Transportation, Construction Division, Materials Section, Draft Version 3 of Special Specification for Field Acceptance of Performance Graded Asphalt Binders for Bituminous Mixtures, June 1998.

(24) Sharma, S., *Applied Multivariate Techniques*, 1996, New York: Wiley & Sons.

(25) Epps, A. L., C. J. Glover, and R. Barcena, *A Performance-Graded Binder Specification for Surface Treatments*, Research Report 1710-1, 2001, College Station, Texas: Texas Transportation Institute.

APPENDIX A: SURVEYS

TXDOT DISTRICT SURVEY

BINDER QA PROGRAM EVALUATION SURVEY

This survey is being conducted as part of TxDOT Research Project No. 0-4047, Analysis and Development of Asphalt Quality Assurance Procedures, under the supervision of Darren Hazlett. The purpose of this survey is to help evaluate and then recommend improvements to the current Quality Assurance (QA) program for asphalt binders (asphalt cements, emulsions, and cutbacks) in Texas. This project does not address asphalt concrete mixture QA. Currently the binder QA program involves sampling, testing, and approval of binders from the supplier tanks. Testing and approval is not required for binders as they enter hot mix asphalt plants. To evaluate and possibly create a program that is both useful and informative, we are sending out this survey to all TxDOT districts and the suppliers of asphalt materials in the state of Texas. With the results of this survey we will attain an evaluation of the current QA program. We would appreciate your participation in helping us to attain the goal of long-life asphalt concrete roadways by making sure that all asphalt binders used to produce asphalt concrete pass the specifications used by TxDOT. If there are any questions concerning this survey or this project you may contact Dr. Amy Epps of the Texas Transportation Institute (TTI) at (979)-862-1750. Once again we appreciate your help and assistance.

Please provide the following information.

District Name: ___________________
Contact Name: ___________________
Phone (____)-_____-_________
Fax (____)-_____-_________

QA Satisfaction

1. Are you satisfied with the current TxDOT QA program?
   _____Yes ____No

2. Do you feel that the TxDOT QA program is fair?(Why or why not)_________________________________
   ____
   ____

3. Do you feel that the current binder QA program achieves its primary goal of producing asphalt concrete that lasts its intended design life? Please explain the answer.
   ____
   ____
   ____

4. Who is held responsible when an asphalt concrete mixture fails prematurely?

   _______Binder supplier, ______ hot mix asphalt plant, _______ other.

Suppliers and Contractors

5. Who is your major supplier of asphalt? ___ Koch Materials; ___Texas Fuel; ___ Coastal refining; ___Wright Asphalt; ___Fina; ___Texaco; ___Gulf State; ___ Chevron; ___Exxon; ____Other.

6. How did you choose that supplier? ___ You did not; ____

____
____

7. Who are your major contractors for asphalt concrete projects? ___________________________________

8. How many mix plants do they own? _____ # of mix plants.

9. How many mix plants do they have in your district? ____ # of mix plants

Sampling

10. Do you take field samples of binders used in asphalt concrete jobs? (If no, please skip to the end).
    ___Yes ___No
    Do you sample all the suppliers ___ OR specific suppliers ___?
    Please list suppliers ____.

11. Who takes the samples? ___contract employee, ___district personnel, ______other.

12. Are they trained to know what they are doing? If yes please explain.
    ____
    ____

13. Where in the chain do you sample the material? ____Delivery truck ____contractors tank, ___ in line for Hot mix asphalt plant ____all the above.

14. How often do you sample material for testing from each location? ______

15. How many different individuals handle a sample before testing? _______.

Equipment

16. List the asphalt testing equipment used in your district laboratory ________________________________
    ____
    ____

17. How often do you test the equipment to make sure that it falls within the specific requirements?
    ____
    ____

    Who does the checks to make sure that the equipment is within specifications?
    ____ Technician; ____Calibration officer.

18. Are all individuals who test asphalt binder certified to operate the asphalt binder equipment? ____ Yes ____No

Testing

19. Do you perform any asphalt binder testing? ____Yes ____No. (If no please skip to the end).

20. What tests are performed in your district on asphalt binders sampled in the field (please list all)?
    ____
    ____
    ____

21. How often are asphalt binders tested from each sampling location? _____

23. How many replicates are performed for each test?

24. What materials do you test most often?
____PG. graded asphalt cements. _____Emulsions _____Cut backs

25. What percentage of the binders fail their respective specification based on supplier sample testing?___0-3% ___3-6% ____6-10% ____>10%

26. What percentage of the binders fail their respective specification based on field sample testing?
___0-3% ___3-6% ____6-10% ____>10%

27. What specific test or tests based on supplier sample testing fail most often?
____

28. What specific test or tests based on field sample testing fail most often?
____

29. If one replicate or sample fails the specification, is additional testing performed? ___Yes,____ No
(If yes, please explain your procedure)____
____
____

30. How many different individuals test the samples?
____ # For PG. sample, ____ # for emulsion sample ____ # for cut back asphalt

31. How many individuals run a specific test for each type of material/specification? _______________
____
____

32. How many employees are qualified to run asphalt binder tests? ________

33. What is the average experience of these individuals (in years)? ________

34. What is the number of man-hours per week spent in testing for each individual?
___ 10-20hrs, ___ 20-30hrs, ___ 30-40 hrs, ___more than 40 hrs.

35. How many weeks of the year do you test? ____

36. How many samples <u>can</u> you test in a day?

- ac ___
- PG. ___
- Emulsions ___
- Cutbacks____
- Total ____

37. How many samples <u>do</u> you test in a day on average?

	Summer	Winter
ac	___	___
PG.	___	___
Emulsions	___	___
Cutbacks	___	___
Total	___	___

38. When you test how much in agreement are you with TxDOT Austin test results. ___0-10%; ____10-20%; ___20-30%; ___ 30-40%; ___40-50%; ___ 50-60%; ___60-70%; ___70-80%; ___ 80-90%; ___90-100%.

Thank you for your time and effort in completing this survey. The results will aid us in formulating recommendations toward improving the binder QA program in Texas, which will result in improved asphalt concrete and longer pavement life.

SUPPLIER SURVEY

BINDER QA PROGRAM EVALUATION SURVEY

This survey is being conducted as part of TxDOT Research Project No. 0-4047, Analysis and Development of Asphalt Quality Assurance Procedures, under the supervision of Darren Hazlett. The purpose of this survey is to help evaluate and then recommend improvements to the current Quality Assurance (QA) program for asphalt binders (asphalt cements, emulsions, and cutbacks) in Texas. This project does not address asphalt concrete mixture QA. Currently the binder QA program involves sampling, testing, and approval of binders from the supplier tanks. Testing and approval is not required for binders as they enter hot mix asphalt plants. To evaluate and possibly create a program that is both useful and informative, we are sending out this survey to all TxDOT districts and the suppliers of asphalt materials in the state of Texas. With the results of this survey we will attain an evaluation of the current QA program. We would appreciate your participation in helping us to attain the goal of long-life asphalt concrete roadways by making sure that all asphalt binders used to produce asphalt concrete pass the specifications used by TxDOT. If there are any questions concerning this survey or this project you may contact Dr. Amy Epps of the Texas Transportation Institute (TTI) at (979)-862-1750. Once again we appreciate your help and assistance.

Please provide the following information.

Supplier Name: ___________________
Contact Name: ___________________
Phone (____)-_____-_________
FAX (____)-_____-_________

QA Satisfaction

1. Are you satisfied with the current TxDOT QA program?
   _____Yes ____No

2. Do you feel that the TxDOT QA program is fair?(Why or why not)
   ____
   ____
   ____

3. How many districts do you supply? ____ # of districts.

4. Which districts do you supply? ___Abilene; ___Amarillo; ___Atlanta; ___Austin; ____Beaumont; ___Brownwood; ___Bryan; ___Childress; ___Corpus Christi; ___Dallas; ___El Paso; ___Fort Worth; ___Houston; ___Laredo; ___Lubbock; ___Lufkin; ___Odessa; ___Paris; ___Pharr; ___San Angelo; ___San Antonio; ___Tyler; ___Waco; ___Wichita Falls; ___Yoakum.

5. What products do you supply to each of the districts (asphalt cements, emulsions, cutbacks, etc.)? ____

____
____
____

6. Do you know where your lab numbers for each sample go? ____ Yes ____No

7. How many times a year do you seek advanced approval? ___Estimate of the number of times per year.

8. Do you find advanced approval difficult to get? ____ Yes ____No; If yes please explain ____
____

9. Are pay factors or payment schedule based on QA test results used in other states where you supply asphalt binder? _____ Yes ______ No.

10. Please comment on pay factors in general.
____
____
____

Testing

11. Do you conduct any QC testing on asphalt binders? _____Yes ____No
(If yes, please list name of lab: __________) (If no, Skip to question 22).

12. How often do you sample material for testing? ___Once per tank, ___ once per tank load ______ other.

13. How often is the material tested? ______.

14. Which tests do you run?(please list them all) ____
____
____

15. How many replicates are performed for each test? ________

16. What does it cost to run the different tests? $ per test _____$

17. Which specific tests that you run fail most often? ___________________________________

____

18. How many different people run a test? ____ # of people per test.

19. Are they certified to run the test? ____ Yes ____No

20. Is it required to have someone who is certified to run the test? ____ Yes ____No

21. If a binder fails specification, what happens to the product? ___You discard it; ___You test it again; ____ you sell it for some other purpose_______ other.

22. When the material passes specification and is approved by TxDOT, how long does it stay on site before it is shipped to a hot mix asphalt plant? ___ days, ___months, ____years.

Transportation and Accountability

23. How do you transport the material?___ Company truck; ___Private shipping
___ hot mix asphalt plant truck.

24. For transportation to hot mix asphalt plants how many different companies do you use? ____ the same company all the time; ___ different companies per batch.

25. Do you check the trucks to make sure there are no other materials in the truck?
___Yes ___No

26. When do you feel that you are no longer responsible for the material and its properties and performance? ___when the material is in the truck,___ at the Hot mix asphalt plants, ___never, ___always.

Thank you for your time and effort in completing this survey. The results will aid us in formulating recommendations toward improving the binder QA program in Texas, which will result in improved asphalt concrete and longer pavement life.

STATE SURVEY

BINDER QA PROGRAM EVALUATION SURVEY

This survey is being conducted as part of the Texas Department of Transportation (TxDOT) Research Project No. 0-4047, Analysis and Development of Asphalt Quality Assurance Procedures, under the supervision of Darren Hazlett (512-232-1902). The purpose of this project is to evaluate and then recommend improvements to the current Quality Assurance (QA) program for asphalt binders (asphalt cements, emulsions, and cutbacks) in Texas. This project does not address asphalt concrete mixture QA.

Currently the binder QA program in Texas involves sampling, testing, and approval of binders from the supplier tanks. Testing and approval is not required for binders as they enter hot mix asphalt plants or are spray applied. To evaluate and possibly create a program in Texas that is both useful and informative, we are sending out this survey to a number of state DOTs with binder QA programs. Using the results of this survey, we will evaluate each QA program and determine which specific aspects might be applied successfully in Texas.

We would appreciate your participation in helping us to understand your binder QA program. If there are any questions concerning this survey or this project, please contact Amy Epps of the Texas Transportation Institute (TTI) at 979-862-1750. Once again we appreciate your help and assistance.

State: __________________
Contact Name: __________________
Phone (____)-_____-_________
Fax (____)-_____-_________

QA Program

Briefly describe your current binder QA program in terms of sampling, testing, specification compliance, and any pay factors. Specific questions related to each subtopic are contained in subsequent sections of this survey. Please skip any question that has already been addressed in the description of your program.
** Please indicate any references we might obtain to help better understand your binder QA program.

QA Satisfaction

Are you satisfied with your current binder QA program? _____Yes ____No If no, please explain.

Do you feel that your binder QA program is fair? (Why or why not?)

What is the primary goal of your current binder QA program? Do you feel that your current binder QA program achieves its primary goal? Please explain the answer.

Who is held responsible when an asphalt concrete mixture or spray application fails prematurely due to the inadequate performance of the binder?

_______ binder supplier ______ contractor _______ other

Suppliers and Contractors

Who are your major suppliers of asphalt binders? How were these suppliers selected?

Who are your major contractors for asphalt concrete or spray applications? How many mix plants does each major contractor own?

Comment on the impact of your current binder QA program (including delivery schedule & any pay factors) on suppliers.

Comment on the impact of your current binder QA program (including construction schedule & any pay factors) on contractors.

Sampling

Do you take binder samples from supplier tanks or transports if blended? ___Yes   ___No   How often?
What is the size of the sample? Lot/sublot?

Do you sample all the suppliers ___ OR specific suppliers ___? Please list suppliers.

Do you take field samples of binders used in asphalt concrete or spray applications? ___Yes   ___No
What is the size of the sample? Lot/sublot?

Do you sample all the suppliers ___ OR specific suppliers ___? Please list suppliers.

Who takes the samples? ___contract employee ___DOT personnel ______other
Are they trained? If yes, please explain.

Where in the binder's journey do you take field samples? ____delivery truck ____contractor tank ___ in line at hot mix asphalt plant ____distributor truck
How often do you sample material for testing from each location?

How many different individuals handle a sample before testing?

Equipment

How many laboratories conduct binder testing as part of your current binder QA program?

List the asphalt testing equipment used in your central and/or field laboratories.

At each laboratory, how often do you calibrate the equipment and/or test to make sure it meets specifications?
Who checks to make sure that the equipment is within specifications?

Testing

Are suppliers required to submit binder test data to you? What binder tests are required? At what sampling/testing frequency?

What binder tests are performed on supplier samples? How many replicates are performed for each test?

What binder tests are performed on field samples? How many replicates are performed for each test?

How often are asphalt binders tested from each sampling location?

What materials do you test most often? ____PG-graded asphalt cements _____Emulsions _____Cutbacks _____Other

How many different individuals test the samples?

How many individuals run a specific test for each type of material/specification?

How many employees are qualified to run asphalt binder tests? What is the average experience of these individuals (in years)? Describe any certification program.

Specification Compliance & Pay Factors

Define specification compliance for each type of material (asphalt cements, emulsions). Are tolerances allowed? If so, how are they determined?

What percentage of the binders fail their respective specification based on supplier sample testing?
___0-3% ___3-6% ____6-10% ____>10%

What specific test or tests based on supplier sample testing fail most often?

What percentage of the binders fail their respective specification based on field sample testing?
___0-3% ___3-6% ____6-10% ____>10%

What specific test or tests based on field sample testing fail most often?

If one replicate or sample fails the specification, is additional testing performed? ___Yes____ No
(If yes, please explain procedure.)

Are pay factors included in your current binder QA program? Are penalties used? Are bonuses an option? How are pay factors determined?

Resources & Costs

What is the number of person-hours per week spent testing binders?

How many weeks of the year do you test?

How many samples of each material type (asphalt cements, emulsions) <u>can</u> you test in a day?

How many samples of each material type (asphalt cements, emulsions) <u>do</u> you test in a day on average?

Estimate testing costs, equipments costs, and staff costs for your current binder QA program.

Benefits

What are the benefits of your current binder QA program?

How are the benefits assessed? OR How is the binder data analyzed and used?

What decisions are based on results of the analysis of the binder data?

For each binder material (asphalt cements, emulsions), do you obtain any estimates of reliability or confidence that the corresponding specification is met?

Do you track the binder's possible contribution to performance?

Has any Benefit/Cost ratio study been conducted to justify the costs of your current binder QA program?

Any additional comments.

Thank you for your time and effort in completing this survey. The results will aid us in formulating recommendations toward improving the binder QA program in Texas. Again, please indicate any references we might obtain to help better understand your binder QA program.

APPENDIX B: TxDOT SURVEY RESPONSE SUMMARY

District	Contact	Satisfaction/Goals/Shortcomings	Responsibility	"Size"
Bryan	Pat Williams 979-778-9753	Not satisfied/ Not fair because infrequent testing frequency/ Does not achieve primary goal because not enough sample taken/	TxDOT Hot-mix plant	# technicians 3 # lab 1 Supplier- Fina, Eagle, Wright Contractor-Young Materials, Paver Supply, C.D.S Enterprises Colorado Materials, A.L. Helmcamp
Lufkin	Cheryl Flood 936-633-4364	Satisfied/ Fair because inline testing/ Achieve goal	Binder supplier Contractor	# technicians 3 # lab 1 Suppliers- Texaco, Fina, Exxon Major contractors- Moore Bros.
Childress	Ronald F. Hatcher 940-937-7161	Not satisfied/ Not fair because asphalt received fails specs 85%/ Do not achieve goal because not receiving the asphalt that are been paid for	TxDOT	# technicians 1 # lab 1 Supplier- Koch, Wright, Fina, Diamond Shamrock Contractor- Jordan paving, J. Lee Milligan, Gilbert Texas
Lubbock	John E. Rantz 806-748-4463	Not satisfied/ No opinion about fairness/ Do not achieve goal because materials are not consistent and fail	No one is responsible	# technicians 0 # lab 0 Supplier- Fina Contractor- Granite Cons. Co. Amarillo Road Company
Wichita Falls	Joe Anderson 940-720-7716	Satisfied/ Fair because tests assure quality of received asphalt/ Do not achieve goal because larger amount not meeting specs nationwide	TxDOT	# technicians (not available) # lab 1 (not available) Supplier- Koch Contractor- Zack Burkett, Duinick Bros.

District	Contact	Satisfaction/Goals/Shortcomings	Responsibility	"Size"
Amarillo	Lonnie Parr	Satisfied/ Fair because of tests from plant and refinery, also tests from each agency/ Achieve goal because specs need to be satisfied for approval.	Supplier, contractor, Aggregate producer	# technicians 5 # lab 1 Supplier- Koch, Fina Contractor- Gilvin Terril, J. Lee Milligan, Duinick, Gilbert Texas, La Fuller & Sons, Holmes, E.D. Baker
Corpus Christi	Mario Garza 361-808-2223	Satisfied/ Fair because assure good product and even conditions for all suppliers/ Achieve goal/	Contractor design	# technicians 2 # lab 1 Supplier- Koch, Eagle Contractor- Bay Ltd., Haas-Anderson Cons.
Waco	C. A. Stan 254-867-2782	Not sure/ Not enough information to tell about fairness/ Not enough information to tell about achieving goal/	Hot-mix plant	# technicians 1 # lab 1 Supplier- Texas fuel, wright, Fina, Gulf State Contractor-Young Contractors
El Paso	Tomas A. Saenze 915-790-4350	Not satisfied/ No direct opinion about fairness but depend too much on supplier test results/ Do not achieve goal because no guideline to address the failed materials, failed materials can only be detected when it is in hot-mix plant	TxDOT	# technicians 1 # lab 1 Supplier-Chevron Contractor-Jobe Concrete Materials
Paris	Bobby Jones 903-732-9321	Satisfied	Hot-mix asphalt	# technicians (not available) # lab (not available) Suppliers-Fina and others Contractor-Boster Paving, Rushing Paving

District	Contact	Satisfaction/Goals/Shortcomings	Responsibility	"Size"
Atlanta	Miles R. Garrison 360-709-5421	Not satisfied/ Should represent what goes into mix/ Not adequate	Contractor	# technicians 1 # lab 1 Suppliers-Lion Oil Contractor-APAC, Madden, Martin Marietta, Duinick Bros.
Dallas	Al Argmoon 214-320-6191	Satisfied/ Does not provide assurance that end product contains correct PG/ Current binder program is better than before	Contractor	# technicians 2 # lab 1 Suppliers-Koch, total Contractor-APAC, Austin Asphalt
Houston	Stanley F. Yin 713-802-5211	Not Satisfied/ Not fear to the state as many things can happen between source and plant	Contractor	# technicians 3 # lab 1 Suppliers-Koch, Coastal Refining, Wright, Fina, Gulf State Contractor-Baytown, Martin Marietta, American Materials, Williams Brothers, Paver Supply
Beaumont	Ronnie van Pelt 409-898-5762	Not Satisfied/ Should be sampled from contractor tank/	Contractor	# technicians (not available) # lab 1 (not available) Suppliers-Koch, Coastal Refining, Fina, Exxon Contractor-Baytown, Martin Marietta, APAC (90%), Matheus Const.

District	Sampling, Testing, & Handling	DOT Sampling & Testing of Field Samples	Equipment	Specification Compliance
Bryan	Developing a program to test liquid asphalt from hot-mix plant	No samples taken.		
Lufkin	Field test is done for all the suppliers Contractor and state employee witnesses the sampling. Employees are trained. PG 6 sample/ 1 (summer), 1 (winter)	Contractor tank, in-line to hot-mix plant Sampling is done every two weeks after start. DSR,RTF, pin/ 10 working days sampling is done/ if fail double check	DSR, RTF, PIN/ Once a year/ Certified calibration officer	0-3% supplier samples fail. 0-3% field samples fail. 90-100% field tests agree with TxDOT Austin test result.
Childress	Field test is done for all the suppliers District personnel take sample. Employees are trained. PG 2 sample/ 2 (summer), 2 (winter)	Delivery tank. Sampling is done every load. RTFO, DSR Daily sampling/	RTFO, DSR, Rotational viscometer	NA/ >10% field samples fail. 90-100% field tests agree with TxDOT Austin test result
Lubbock	Field test is done for all the suppliers District personnel take sample. Employees are trained.	Delivery tank. Sampling is done every load	NA	
Wichita Falls	Field test is done for all the suppliers District personnel take sample. Employees are trained. PG 1 sample/	Contractor tank. Sampling is done daily. RTFO, DSR Daily random selection at 1:5 ratio/	RTFO, DSR, Rotational viscometer	
Amarillo	Field test is done. For the supplier, by request from area engineer, pending policies. Contractor and state employee witnesses the sampling. District personnel take sample. PG 4 PG, 4 ac/ pending.	Delivery tank, Contractor tank, In-line or hot-mix plant Sampling is done as requested. RTFO, DSR, ABSON 211-F, and pen 502C/ Sampling is done by request	RTFO, DSR, ABSON 211-F, and pen 502C	NA

District	Sampling, Testing, & Handling	DOT Sampling & Testing of Field Samples	Equipment	Specification Compliance
Corpus Christi	Field test is done. For the supplier, by request from area engineer, pending policies. District personnel take sample. Employees are trained. PG 5 PG, 5 ac/ 2 PG and 2 ac (summer), 1 PG and 1 ac (winter)	In-line or hot-mix plant. Sampling is done every production day. Unaged and aged property is tested/ Sampling is 1:9/	DSR, RTFO	0-3% supplier samples fail. 0-3% field samples fail. 90-100% field tests agree with TxDOT Austin test result.
Waco	No field sample. PG 6 sample/ 2 PG in summer	DSR, RTFO Once per sampling location/ More replicate if fails.	DSR, RTFO, Technician test the machine prior to test	0-3% supplier samples fail. 0-3% field samples fais. 90-100% field tests agree with TxDOT Austin test result.
El Paso	Field test is done for all the suppliers District personnel take sample. Trained employee. DSR, RTFO on aged and unaged materials. Once a week during production. PG 2 sample/ 2 sample during summer	In-line or hot-mix asphalt plant. Sampling is done once a week during production . No asphalt binder test is performed	DSR, RTFO, Brookfield viscometer Technician test the equipment once a year	0-3% supplier samples fail. 0-3% field samples fail. 90-100% field tests agree with TxDOT Austin test result.
Paris	Field test is done for all the suppliers District personnel take sample.	In-line or hot-mix asphalt plant Sampling is done 1 per project	DSR Calibration Officer annually test the machine	

District	Sampling, Testing, & Handling	DOT Sampling & Testing of Field Samples	Equipment	Specification Compliance
Atlanta	Field test is done for all the suppliers Contractor employee takes sample TxDOT tries to witness. Employees are not trained enough. PG 4 sample/ 1 PG in Summer and Winter	In-line or hot-mix asphalt plant. Sampling is done 1 per week per course, if fails daily until satisfaction. DSR, RTFO/ Sampling information not available/ Repeat test is performed when sample fails.	DSR, RTFO, PEN, ABSON Recovery Calibration Officer annually test the machine	NA/ >10% for Fina and 0-3% for others. 70-80% field tests agree with TxDOT Austin test result.
Dallas	Field test is done for all the suppliers. Sample is not collected by contractor or district employee. PG 6 sample/ 1 PG in Summer and Winter	In-line or hot-mix asphalt plant. Sampling is done randomly, once or twice depending on the job. DSR, RTFO Random sampling/ Repeat test is performed when sample fails.	DSR, RTFO Technician (Austin) annually test the machine	0-3% field samples fail. 90-100% field tests agree with TxDOT Austin test result.
Houston	Field test is done, but not for all the suppliers. District personnel take the sample. Employees are trained by material section. PG 4 sample/ 2 PG in Summer and 1 in Winter	In-line or hot-mix asphalt plant. Sampling is done once a month DSR, before and after RTFO rotational viscosity, penetration/ Once a month asphalt are tested/ 3 replicates	DSR, RTFO, Rotational viscometer, pen, Superpave gyratory compactor, core-lock device, Hveem stabilometer, Abson recovery Maintenance: DSR = before test RTFO = every 6 month Others annually	0-3% supplier samples fail. 0-3% field samples fail. 90-100% field tests agree with TxDOT Austin test result.
Beaumont	No field test is done.	NA	NA	NA

APPENDIX C: STATE SURVEY RESPONSE SUMMARY

Highlighted references can be obtained from corresponding DOT literature.

State	Contact	Satisfaction/Goals/Shortcomings	Responsibility	"Size"	Impact
AZ	Dan Anderson 602-712-8214	Y – matl spec on road / fair to contractors – pass penalty to suppliers, ?fair to suppliers – lack of QC @ contractor location / contamination prob	DOT – penalty system in place	4 PG suppliers 4 emulsion (emul) suppliers 6 major AC contractors 6 major spray contractors 1 central + 3 regional labs 3 full-time + 2 summer techs @ central 1-2 techs/regional lab	Some supplier disputes w/contractor None on contractors
CA	Mike Cook 916-227-7300	Y – matl spec on road / fair – does not cause delays	DOT – after acceptance, before can shut down construction or revoke supplier cert	11 suppliers 40-50 major contractors 1 central lab 3 full-time techs	No delays unless trend of failing results
CO	Bill Schiebel 303-757-9235	Y – matl spec on road / price reduction perceived as fair, contractor passes penalty to supplier	DOT – penalty system in place	5 suppliers 84 total contractors 1 central lab 2.5 full-time + 0.5 summer	Supplier management plan – requires QC plan w/ AASHTO tests & results, sample/cert @ beg of season, proficiency samples, inspect if necessary, inspect transports, procedure for failures, procedure for de-certif None on contractors unless shutdown job Reqd project specific QC plan for contr soon w/ QC plan including transp/storage/handling, tank inspection, any samp & test
MD	Gloria Burke 800-477-7453 ext 103	Y – matl spec on road / fair in sampling freq & penalties statewide, pass penalties from contractors to suppliers	Contractor - first yr DOT – after first yr	Part of NE user-producer group (~20 suppliers, ~35 labs) 6 suppliers 6 major contractors 1 central + 2 regional labs (western, eastern) 3 full-time @ central 2 full-time @ western 1 full-time @ eastern	Supplier QC plan including testing reqd – AASHTO PP26 Contractors can be reqd to stop production, mill & replace, or be penalized

State	Contact	Satisfaction/Goals/Shortcomings	Responsibility	"Size"	Impact
MN	Jim McGraw 651-779-5548	Y, definitely – matl spec on road, save time & effort through shared cert & inspection / fair through good communication, more responsibility to suppliers – cert in 6 states	DOT – penalty system in place	Part of Combined States 10 PG supplier locations in MN 7 emul suppler locations in MN 52 major AC contractors 1 central lab 4 full time (in summer)	Reqd QC plan from suppliers No impact on contractors – must supply samples
NV	Wayne Brinkmeyer 775-888-7879	Y – matl spec on road / fair to contractors – pass penalty to suppliers, ?fair to suppliers – lack of QC @ contractor location – working towards reqd QC for both supplier & contractor	DOT – penalty system in place	5 suppliers 3 major AC contractors (10 minor) 1 N lab, 1 S lab 5 full-time techs + 1 summer / lab	Penalty may be passed to suppliers unless show contractor prob No impact on contractors - shut down job if continuous failing
OR	Bruce Patterson 503-986-3052	Y – matl spec on road / ?fair to supplier – lack of QC @ contractor location – contamination, mishandling – lack of responsibility – requires too many resources for DOT to check	DOT – penalty system in place	5 suppliers 3-5 major AC contractors 3-5 spray contractors (some same) 1 central lab 3 full-time techs + 1 summer	Only sample suppliers during mix design No impact on contractors – rare to shut down if series of failures
TX	Jerry Peterson 512-232-1913	N – matl spec on road & consistent / fair to all, min effort reqd / ineffective, easy to circumvent, needs improvement – some suppliers w/same goal	DOT – approved based on supplier samples, assumed o.k. in field	18 total suppliers 40-50 major AC contractors 90 total 1 central lab (only certified results) 25 district labs 4 full-time techs + 2 @ 25% in central	Some delay to suppliers w/out advance acceptance (qual mngmnt plan) Supplier advance acceptance includes QC plan w/daily abbrev samp & test, monthly complete samp & test, report prob to DOT no impact on contractor – binder previously accepted
UT	Cameron Petersen 801-965-4296	Y – Asphalt Binder Qual Mngmnt System (ABQMS) – matl spec on road / ?fair to supplier – more effort than contractor but passed penalty / need to increase contractor QC responsibilities to include testing	DOT – penalty system in place	5-6 suppliers 5-6 major contractors 1 central lab 2 full-time techs + 1 summer intern	Supplier certif program w/reqd QC plan including ≥2 PG tests, truck inspection, DOT reports if requested Contractor QC plan reqd w/handling transp inspection, storage, & samp (& test but not reqd)
WA	Joe Devol 360-709-5421	Y – matl spec on road / fair – risk distributed / contractor passes penalty to supplier / need QC plan for contractors – more responsibility	DOT – penalty system in place Can leave open up to 2 yrs to pay	9 PG suppliers 9 spray suppliers (some same) ~ 15 major contractors 1 central lab 2 + 1 prep full-time techs	Reqd QC plan for PG suppliers + annual inspection, access to data No impact on contractor now – may shut down or remove & replace reqd QC plan for contractors in 03

State	Sampling, Testing, & Handling	Supplier Testing	DOT Sampling & Testing of Supplier Samples	DOT Sampling & Testing of Field Samples	Equipment
AZ	Contractor w/DOT witness AASHTO T40 for emul No tech certification – future AASHTO 11 mos/yr 10-12 PG/day 1-2 emul/wk 85% PG / 15% emul testing 1 sample tested by ~ 3 techs Test after use	Y – cert of comp for all tests for stnd products/ shipment Emul cert of comp good for 30 days Cert of analysis from approved lab w/QC plan & all test results/ shipment less common	No samples taken	All suppliers Table 3 – Sampling Guide 2 gallons ac (1-lab, 1-job) X 2/day in line PG (ac): complete test #2, #40, every 40^{th}, limited test every 4^{th} after #2 (test most likely out of spec – f(grade, mod, climate; ex. DSR-RTFO + ORIG, add BBR w/mod or DSR-PAV in winter) 1 replicate unless test requires more occasionally 2X 2L emul / dist truck	AASHTO – 2 QA people Regional: DSR, RTFO to test 76-10, 70-10, 64-16, 64-10, 58-16 Central: all other PG, mod, excessive regional results
CA	DOT Caltrans Indpt Assurance Program – training / qual for sampling & testing AASHTO April – Oct 20 ac/day (complete + non-complete) 10-15 emul/day 75% ac / 25% emul testing Test after use 2,875 samples in 2000	Y – cert of comp for all tests /shipment Results avail for review Sect 6-37, 6-39, p. 8-02-4 Const Manual Sect 6-1.07, 92, 93, 94 Stnd Spec Suppliers certified by demonstrating production of matls that meet spec + QC program Stnd products – automatic certification New products – samples to DOT @ beg of season	All suppliers of new products – only at beg of season	All suppliers p. 8-01-12 Const Manual 1 L ac X 2/day in line (sample #1, #2) test 1/2 Complete (140F & 275F visc, pen-ORIG & RTFO, duct, G & sol (if requested)) for #1, every 5^{th} Noncomplete (140F visc) other samples 2 L emul X / shipment dist truck or cont tank Test all	AASHTO

State	Sampling, Testing, & Handling	Supplier Testing	DOT Sampling & Testing of Supplier Samples	DOT Sampling & Testing of Field Samples	Equipment
CO	Contractor w/DOT witness In-house inspection of techs AASHTO Not all techs trained for all tests April-Oct/Nov 12-16 PG/wk 85% PG / 15% emul testing	Y – certif @ beg of season (split samp w/DOT) + & complete spec comp 1/month + QC testing in reqd plan investigate if split samp results not within precision or 2 stnd dev of RR results decertif if 3 consecutive noncomp, no RR participation, inadequate records or QC plan, unsatisfactory inspection	All suppliers Only @ beg of season 1 L	All suppliers 1 L ac X 1/1000 tons AC in line or contractor tank PG (ac): complete test lot #1 (random sample 1/5 in lot), lot #4, every 3rd lot, limited test (DSR-ORIG) all other lots (random 1/5) 1 L emul X 1/truckload or 6000 gallons – dist truck or contractor tank	AASHTO – 1 QA person
MD	Contractor w/DOT witness 2+1 @ central & 2 @ western through tech cert program @ UConn, Penn St AASHTO @ central, western April – early Nov ~20 fingerprint samples/day ~4 complete samples/2 days Test PG @central, western Test emul @ eastern	Y – spec comp cert / lot (batch or tank)	No samples taken	All suppliers 1 Qt ac X 1/6000 tons AC in line or contractor tank for complete testing 1 Qt ac X 1/day @ random time in line or contractor tank for fingerprint testing (Brookfield @ eastern or Brookfield + DSR-ORIG @ central, western)	AASHTO @ central, western central, western: complete eastern: fingerprint (only Brookfield)

State	**Sampling, Testing, & Handling**	**Supplier Testing**	**DOT Sampling & Testing of Supplier Samples**	**DOT Sampling & Testing of Field Samples**	**Equipment**
MN	Contractor w/DOT inspector Minimal sampling cert classes AASHTO T40 AASHTO Superpave binder course May – Oct 3-5 dates/PG ~2 dates/emul ~85% PG (67% complete) / 15% emul testing	Y – combined states share supplier cert for PG including QC program, open records & facilities, transport inspection, facility inspection by DOT each spring (review samp & test, QC, facility changes, tank inspect, storage), participation in RR, problem solutions cert each shipment lose if 3 consecutive noncomp field samples, no RR, no reqd records, no QC plan Annual prior to season ½ wks Daily – pen, visc, or DSR Bi-weekly – complete Emul: trial program this yr (Saybolt daily) + 3-4 complete/season on supplier sample	Option in special situations	All suppliers PG: 1 L X 1st shipment + 1/1000 AC tons from contractor delivery truck Complete on #1, DSR-ORIG on others Emul: 1 gal X 1/50,000 gal for chip seals (other apps specific samp by job) @ random from supplier Complete on one, abbreviated (Saybolt) on others 1 replicate unless test requires more	AASHTO – 3 techs

State	Sampling, Testing, & Handling	Supplier Testing	DOT Sampling & Testing of Supplier Samples	DOT Sampling & Testing of Field Samples	Equipment
NV	ac: contractor – DOT not required Emul: dist truck driver – DOT not reqd No reqd samp training, some through supplier AASHTO in-house training May – Oct 87% ac / 11% emul testing 1-50 ac/day 1-12 emul/day Test PG/ac after use Test emul – Saybolt before use	Y – cert w/all tests /tank if noncomp – can revoke certif, require pretesting	No samples taken	All suppliers PG: 1 Qt X 1/100 mtons binder in line complete test 1/10 & DSR-ORIG on others ac: 1 Qt X 1/23 mtons binder in line complete 1/5 & visc-ORIG on others Emul: 1 Qt X / dist truck Complete test all emul 1 replicate unless test requires more adjust samp freq if consistently close to spec, history of non-comp, reduce testing time	AASHTO – 1/lab for QA, 3 certif/calib for matls div

State	Sampling, Testing, & Handling	Supplier Testing	DOT Sampling & Testing of Supplier Samples	DOT Sampling & Testing of Field Samples	Equipment
OR	Contractor – DOT not reqd, 10% witnessed when indpt samples taken AASHTO T40 samp training class for contractors & DOT AASHTO April – Oct 70% ac / 30% emul testing FOR ONE TECH: 5-6 limited ac/day OR 3-4 complete ac/day x 2 techs + 1 summer 4-5 abbrev emul/day OR 3 complete emul/day x 1 tech	N No precertif	No samples taken – only during mix design, proposed for acceptance recently – resources & storage prob	All suppliers ac: 1 Qt X 1/1000 Mg AC contractor tank complete testing on indpt witnessed samples (1/10,000 Mg AC) Limited (ORIG-abs visc, RTFO-abs visc, RTFO-4C pen) on 1/5 of other samples (1/1000 Mg AC) Emul in AC: 1/1000 Mg AC Emul in chip seals: 1/50 Mg before dilution dist or transp truck Complete testing on indpt witnessed samples (1/10,000 Mg AC) Limited (Saybolt, % res, % dist, 25C res pen) on 1/5 1 replicate unless test requires more	AASHTO – 1 QA person
TX	DOT contract employee Tex –500-C No samp training AASHTO tech certif only in central lab March – Nov >20 PG/day 30 emul/day	Y – cert comp Daily abbrev testing if daily off target, show complete comp advance acceptance (qual mngmnt plan) w/3 consecutive spec comp includes daily abbrev samp & test, monthly complete samp & test ac: approved for 60 days Emul: approved for 30 days	w/advance acceptance (qual mngmnt plan) PG: 1 Qt X 1/month for approval + 3 random/month Complete on 1/month DSR-ORIG or no testing on others Emul: 1/tank + 1/truck Complete 1/tank Abbrev/check (Saybolt) 1/truck 1 replicate unless test requires more	Suggested, some districts on regular basis or if prob suspected 1 gallon into 3 qts In-line, contractor tank, or dist truck Complete testing or as specified (forensic) Resources not available	AASHTO – 1 QA person Central: complete District: DSR, RTFO

State	**Sampling, Testing, & Handling**	**Supplier Testing**	**DOT Sampling & Testing of Supplier Samples**	**DOT Sampling & Testing of Field Samples**	**Equipment**
UT	Contractor w/DOT witness AASHTO T40, ASTM D3665 AASHTO May – Nov most PG testing 2 complete PG/wk 1 emul/wk	Y – complete for 3 consecutive lots (job by job - ~ tank), reduce to 1/2 lots if all tests comp, may reduce further, min 1/month	May sample all suppliers PG: split 1 L X 1/month tank Emul: split X 1/month tank	All suppliers PG: 3 L X 1/day @ random (2–DOT, 1–contr) in line Sublot = day; lot = week or 3-8 sublots complete if concerned or abbrev (DSR) X 1 sublot/lot Emul: no samples taken unless prob	AASHTO – 2 techs for QA
WA	Contractor w/DOT witness AASHTO T40+ state samp training program AASHTO May – Oct (4/15-10/15 in spec) emul: July-Sept 75% PG / 25% emul testing 6 complete PG/2 dys 15 abbrev PG/day	Y – complete spec comp @ beg of season Supplier certif program/ qual products list Option for comparison testing by DOT	Option to take samples – previous or w/DOT witness	All suppliers PG: 2 X 1Qt. X second ac sample + 1/2 ac samples (800 tons AC) in line or contractor tank Test 1^{st}, 3^{rd}, 5^{th}, every 5^{th} samples Complete test 1^{st}/contract, abbrev (DSR-ORIG if consistent, good record) @ discretion, dept on resources Emul: 1/2 shipments dist truck Test all samples Complete 1^{st} /day + Saybolt for others	AASHTO – 1 QA person

State	Specification Compliance	Pay Factors / Penalties	Costs	Analysis / Benefits
AZ	PG: MP1 Emul: must pre-cert no tolerances ~15% failures on field samples PG: DSR-RTFO, m-value Emul: Saybolt, sieve Failure: test other sample from same day	Y – only on worst prop If 2nd sample fails, addl testing (ADOT or 1 of 2 indpt labs – less common) to determine quantity out of spec Tables 1005-1, 1005-3	vs private – few yrs ago lab cost $325 ADOT vs. $1000 private	Acceptance Database – review quarterly / supplier & communicate, some forensic only qual confidence No B/C
CA	No tolerances Caltrans 0-3% failure on field samples ac: 140F visc Emul: residue by evap, sieve Failure: repeat test & avg, then test other sample from same day	N	Not available	No elec database – near future for research, track suppliers w/paper database – track new product perf, Forensic No B/C
CO	No tolerances ~5% PG failures on field samples 22% emul failures on field samples PG+: DSR-RTFO & ORIG, toughness & tenacity, ductility Emul: penetration Failure: rerun test, run complete on other 4/5 from same lot	Y – price reduction / lot – price reduction factors & equations (Section 105.03 + amend + examples) – cumulative (P total) for Ptotal < 3 no penalty, reduced price for Ptotal 3-25, Ptotal > 25 may remove & replace – PG – only DSR-RTFO + m-value + PG+ prop	No justification reqd $350 for staffing only on avg job [50,000 tons AC = 50 samples = 10 lots = 4 complete + 6 abbrev (DSR-ORIG)]	Acceptance Database – ID repeated prob w/specific supplier, forensic as needed Only qual confidence – manage resources w/ good representation of matl qual No B/C

State	Specification Compliance	Pay Factors / Penalties	Costs	Analysis / Benefits
MD	Tolerances for PG based on recent AMRL Proficiency results from >80 labs (memo) w/understanding that suppliers must exceed spec reqmnt by ½ COV of results and DOT will accept out of spec matl if within ½ COV of results NE user-producer est interlab variability through split sampling 1/2 months – these are not used currently Emul – ASTM precision & bias for tolerances If fingerprint & supplier outside supplier range – further eval of other region's data on same lot, possible complete testing 3-6% failures on complete field samples DSR-ORIG & RTFO, m-value failure: fingerprint test again & possibly retest surrounding days, complete resample and retest & check QC's	No formal system – job by job negotiated w/contractor as f (tonnage, binder type, facility), may pass onto supplier, up to 20% Use AASHTO Superpave QC/QA Software to calculate pay factors	Complete \$89.52/sample Fingerprint \$14.25/sample 2 techs (6 hrs testing total) = \$90/day + equip costs	Acceptance Database accessible by all MD DOT – check inconsistencies Track particular suppliers all over state – ID if need more attention/training Review trends including changes in binder use track costs to MD Constant attention to product reinforces expectation of quality Increased communication w/ suppliers only qual confidence – based on historical data, failures – sometimes reduce amount of fingerprint testing Track performance mainly on special projects/matls No B/C NE user-producer – common analysis cert
MN	Tolerances for field samples based on RR No tolerances for supplier results Must exceed tolerance before price reduction ~0% failures on field samples DSR-RTFO Emul residue, 25C pen Failure: retest, investigate & determine quantity, increase samp freq, report, possible penalty & corrective actions	Y – based on max reduction from single test if multiple tests fail, price reduction schedule for specific property ranges and additional property ranges that require 25% price reduction or determined by engineer, remove & replace also possible not uniform in combined states	Testing including salaries + benefits + OH: PG \$163 (complete + abbreviated), Emul \$79 billed to cities/counties	Acceptance Database – forensic, compare products in same grade, track results by grade & supplier, biweekly inventory of lab samples, lab perf measures, research on PG use Only qual confidence – determine samp freq based on resources & emp evidence/sense of confidence No B/C – no justification reqd, generates revenue

State	Specification Compliance	Pay Factors / Penalties	Costs	Analysis / Benefits
NV	Tolerances based on AASHTO repeatability & rejection limit Exceed tolerance before penalty prorated between tolerance & rej limit PG+ & NDOT 6% PG & ac failures on field samples 9% emul failures on field samples ductility, toughness & tenacity Failure: repeat test, do not avg, check w/ AASHTO repeatability	Y – cumulative incremental demerits based on all tests, max = 21 (remove & replace), job by job (location, climate, pol power) demerit schedules Penalty if 2 failures – assess demerits on result closest to spec	Not available – no justification reqd	Acceptance Database – track suppliers and products, forensic, summary / job, set limits for new spec & eval impact, track trends / failures by test No confidence est No B/C – no justification reqd
OR	No tolerances RR test results currently only used for overall eval of testing & eqpmnt PBA – PG next yr <2% ac failures on field samples Abs visc-ORIG, 4C pen ~11% emul failures on field samples Saybolt, residue pen Failure: retest, test surrounding to determine quantity	Y – based on one property, min = 25%, 50% for specific matls/prop/climates	Not available – no binder testing labs available	Acceptance Database – forensic, RR to eval testing & equipment Only qual confidence – good perf, PBA exceeds PG climate reqmnts No B/C
TX	No tolerances ~3% failures on all supplier samples >10% failures on field samples (23% PG in dist w/prob, 18% other matls tested) DSR-ORIG & RTFO Saybolt Failure: retest if close, sometimes stop w/remaining tests in spec, look @ any supplier split sample results or resample & retest	N – but supplier must pay test costs if fails spec	FY93 for staff only $0.5 mill DOT vs $8.7 mill private	Acceptance Old database w/out PG results (new, online soon)– forensic, track % failures, prod totals simple, concentrated responsibility Only qual confidence – in PG, continuous prod can't track based on supplier samples district may track perf from field samples

State	Specification Compliance	Pay Factors / Penalties	Costs	Analysis / Benefits
UT	Tolerances = compliance & rejection limits (509.5.1) If DOT & supplier QC results don't match – investigation No prob w/emul supplier samples distillation PG+ Only 1 sample failed last yr based on field samples DSR-ORIG, toughness & tenacity Failure: appeal w/supplier QC results + complete, if accepted – test random or surrounding 2 sublots for prob tests & avg, last resort – referee testing of all other sublots for prob tests by indpt lab – contr pays if non-comp	Y – 0-25% from linear interpolation between compliance & rejection limits, based on all test prop, max 25% or remove & replace Will consider any binder QC results from contractor in penalty assessment	Not available Lab pricing $500/PG $1200/PG for private	Acceptance Database – online soon, track suppliers & products for ASAP communication, basic summary stats, forensic Only qual confidence – in PG No B/C Assure qual w/ min testing
WA	~10% tolerances 4% PG failures on field samples DSR-PAV <5% emul failure on field samples Saybolt Failure of supplier sample: review QC & certif status Failure of field sample: test surrounding samples to determine quant	Y – reject if more than 10% out OR less than 10% w/2 surrounding samples non-comp OR leave w/ min 5% penalty penalty job by job based on which prop & est effect on perf Max penalty 75% (remove & replace)	Not available Indiv test costs	Acceptance Older database – new soon: forensic, chart/track suppliers & products by abbrev PG tests, research – validate binder "bumping" only qual confidence – balance limited resources w/history of perf & database to track data No B/C