Organizational Results Research Report

August 2009
OR10.005

Quick Test for Percent of Deleterious Material

Prepared by Missouri University of Science and Technology and Missouri Department of Transportation

TECHNICAL REPORT DOCUMENTATION PAGE.

1. Report No.:

2. Government Accession No.:

3. Recipient's Catalog No.:

4. Title and Subtitle:

Quick Test for Percent of Deleterious Material

5. Report Date:
August 28, 2009

6. Performing Organization Code:

7. Author(s): David N. Richardson

8. Performing Organization Report No.: RI07-052

9. Performing Organization Name and Address:
Missouri University of Science and Technology
1870 Miner Circle
Rolla, MO 65409

10. Work Unit No.:

11. Contract or Grant No.:

12. Sponsoring Agency Name and Address:

Missouri Department of Transportation
Organizational Results
PO Box 270, Jefferson City, MO 65102

13. Type of Report and Period Covered:
Final Report.

14. Sponsoring Agency Code:

15. Supplementary Notes:
The investigation was conducted in cooperation with the U. S. Department of Transportation, Federal Highway Administration.

16. Abstract: The Missouri Department of Transportation (MoDOT) is considering the replacement of its deleterious materials test method (TM-71) with test methods that are more objective. MoDOT contracted with the Missouri University of Science and Technology (Missouri S&T) to develop a system of test methods. Nine quarry/ledge samples representing seven geologic formations (four limestones and three dolomites) were supplied by MoDOT. The ledge samples represented three aggregates each for use in concrete, asphalt, and granular base. Samples with controlled contamination were also tested, bringing the total to 18. The aggregates were subjected to fifteen test methods. This data, coupled with MoDOT historical specific gravity, absorption, and TM-71 deleterious materials data of the samples formed the basis of the test study dataset. Multiple linear regression was used to produce 15 models of varying accuracy and complexity for TM-71 predictions. The TM-71 deleterious data were used as the response (dependent) variables. The best models entailed test methods not normally performed by MoDOT, such as sieved slake durability, point load strength, vacuum saturated bulk specific gravity/absorption, and aggregate crushing value, along with the more. familiar micro-Deval and plasticity index. Model adjusted-R^2 values ranged from 0.603 to 0.895. Thus, three to four options (models) were open to MoDOT for consideration for each type of deleterious material (Total Deleterious Material, Total Deleterious Material Plus Hard Chert, Deleterious Rock Plus Soft Chert, and Shale). As an alternate to the regression models, a threshold-limits method was presented. The models themselves were not exact enough to predict the various deleterious contents with the level of accuracy required for routine decisions concerning aggregate product acceptance or rejection. As a result, a method of baseline ledge-specific initial calibration of the models was developed to enable MoDOT inspectors to make acceptability decisions on a routine basis without the necessity of performing TM-71.

17. Key Words:
deleterious material; aggregate

18. Distribution Statement:
No restrictions. This document is available to the public through National Technical Information Center, Springfield, Virginia 22161.

19. Security Classification (of this report):
Unclassified.

20. Security Classification (of this page):
Unclassified.

21. No of Pages:
166

22. Price:

Form DOT F 1700.7 (06/98).

FINAL REPORT
RI07-052

QUICK TEST for PERCENT OF DELETERIOUS MATERIAL

Prepared for the

Missouri Department of Transportation
Organizational Results

By

David N. Richardson, PE
Missouri University of Science and Technology

August 28, 2009

ACKNOWLEDGEMENTS

The author wishes to thank the Missouri Department of Transportation (MoDOT) for sponsoring this work and Paul Hilchen, Will Stalcup, and Jennifer Harper for their coordination and support. On the Missouri University of Science and Technology (Missouri S&T) side, thanks go to Gary Davis, Karl Beckemeier, and Michael Keaton for their many hours in the laboratory, and to Michael Lusher for his contributions both in the laboratory and with statistical computation assistance.

EXECUTIVE SUMMARY

The Missouri Department of Transportation (MoDOT) is considering the replacement of its deleterious materials test method (TM-71) with test methods that are more objective. MoDOT contracted with the Missouri University of Science and Technology (Missouri S&T) to develop a method of approximation of various deleterious materials contents based primarily on systems of standard tests which would augment or replace the deleterious test method TM-71. The system would be comprised of one or more objective tests, depending on the outcome of the research project. Nine different quarry/ledge production materials representing seven geologic formations (four limestones and three dolomites) were sampled by MoDOT and delivered to Missouri S&T. The samples represented three aggregates each for use in concrete, asphalt, and granular base. Samples of controlled contamination were also tested, bringing the total to 18. The aggregates were subjected to fifteen different test methods/method modifications. The test results, coupled with MoDOT historical specific gravity, absorption, and deleterious materials data, formed the basis of the study dataset. The test methods were: Los Angeles abrasion, micro-Deval, wet ball mill, wet ball mill-modified, aggregate crushing value, methylene blue value, sodium sulfate soundness, water-alcohol freeze-thaw soundness, point load strength (dry and wet), vacuum saturated bulk specific gravity, vacuum saturated absorption, sand equivalent, plasticity index, and sieved slake durability. Results from historical MoDOT test methods included gradation, bulk specific gravity, absorption, deleterious rock content, shale content, and chert content.

Multiple linear regression was used to produce 15 models of varying accuracy and complexity for TM-71 predictions. Deleterious data for the same aggregate materials (samples) were used as the response (dependent) variable. The best models entailed test methods not normally performed by MoDOT, such as sieved slake durability, point load strength, vacuum saturated bulk specific gravity/absorption, and aggregate crushing value, along with the more familiar micro-Deval and plasticity index. Model adjusted-R^2 values ranged from 0.603 to 0.895. Thus, three to four options (models) were open to MoDOT for consideration for each type of deleterious material (Total Deleterious Material, Total Deleterious Material Plus Hard Chert, Deleterious Rock Plus Soft Chert, and Shale). As an alternate to the regression models, a threshold-limits method was presented.

The models themselves were not exact enough to predict the various deleterious contents with the level of accuracy required for routine decisions concerning aggregate product acceptance or rejection. As a result, a method of baseline ledge-specific initial calibration of the models was developed to enable MoDOT inspectors to make acceptability decisions on a routine basis without the necessity of performing TM-71.

Unfortunately, MoDOT had no historical data with which to verify the models. This is a vital step and must be done in the future before any of the models are implemented.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

GENERAL

The Missouri Department of Transportation (MoDOT) is considering the replacement of its deleterious materials test method (TM-71) with test methods that are more objective. MoDOT contracted with the Missouri University of Science and Technology (Missouri S&T) to develop a method of approximation of various deleterious materials contents based primarily on systems of standard tests which would augment or replace the deleterious test method TM-71. TM-71 is highly subjective in nature. It was envisioned that the system would take one of several forms, including a predictive regression equation(s) or a system of threshold limits. The system could be comprised of several tests, or a single test depending on the outcome of the research program. It was desired that the tests would easily simulate and quantify the specific deleterious actions of aggregates.

The value of such a system of tests would be to progress toward a more objective method. Additionally, the certification of out-of-state testing personnel would become easier if MoDOT was using nationally-accepted standard tests rather than its own test method.

MoDOT specifications (MoDOT, 2004) distinguish between different forms of deleterious materials and assign levels of concern as to the deleterious materials' presence in various aggregate products in two ways: 1) percent maximum allowable limits in materials specifications, and 2) by inclusion or absence in various material specifications in regard to usage. Table 1 shows the various deleterious types and the MoDOT specifications that include maximum limits in order of apparent concern and frequency. The table shows five different uses of aggregate, such as granular base. Some uses are not sensitive to certain deleterious materials, thus not all deleterious materials are limited by all aggregate specifications. Aggregate specifications limit deleterious materials by maximum allowable percent by weight. Table 1 shows nine specific types of deleterious materials as defined by TM-71. An “x” denotes that the specification limits the particular deleterious material. Deleterious material can be either inherent to the parent aggregate material or come from contamination, both natural or artificially generated. Typically, “other foreign material” (OFM) and “mud balls” would be included in the contamination category. All other deleterious materials types are intrinsic to the parent aggregate.

Table 1: Deleterious Material Types and Section 1000 Specifications for Coarse Aggregate

Deleterious Material	**1007: Granular Base**	**1006: Surfacing (Unbound Material)**	**1004: Bituminous Surface (Blade Mix)**	**1002: Superpave 1003: Seal Coats**	**1005: Concrete**
Shale	x	x	x	x	x
Soft rock	x	x	x	x	x
Mud balls	x	x	x	x	x
OFM (coal, lignite, sticks, etc)		x	x	x	x
Shaly rock			x	x	x
Cap + 20%				x	x
Soft chert				x	
Chert in limestone					x
Dispersed clay			x		

Portland cement concrete (PCC), hot mix asphalt (HMA) mixtures, and unbound aggregate base (UAB) materials can suffer from many aggregate-related performance problems, as shown in Table 2.There are 10 aggregate deleterious actions that can cause these material performance problems. In Table 2 are shown various test methods that can be associated with the performance problems and deleterious actions. Throughout this report the following abbreviations will be used: AASHTO (American Association of State Highway and Transportation Officials, ASTM (American Society of Testing and Materials, ACV (aggregate crushing value), PLS (point load strength), PI (plasticity index), SE (sand equivalent), MB (methylene blue), LAA (Los Angeles abrasion), WBM (wet ball mill), I_{sd2} (sieved slake durability), ΔPLS (delta point load strength), $NaSO_4$ (sodium sulfate soundness), WAFT (water-alcohol freeze-thaw soundness), BSG (bulk specific gravity), Abs (absorption), VSBSG (vacuum saturated bulk specific gravity), and VSAbs (vacuum saturated absorption). All abbreviations are listed in the “Glossary” section of the report.

Actions that are deleterious to a given material such as concrete or HMA mixtures or granular base materials can be divided into eight categories: 1) breakdown from handling, e.g. impact or attrition from dropping onto a stockpile or into a bin, or mixing action, 2) breakdown from crushing, such as being driven on, from the dead weight in a stockpile, or from compaction, 3) breakdown or destructive swelling and shrinking from wetting (precipitation) and drying, 4) breakdown or destructive expansion from freezing and thawing, 5) asphalt-aggregate bond interference, 6) water adsorption by fines causing decreasing workability, 7) cement paste-aggregate bond interference, 8) loss of material stability due to lubrication by clay, 9) adverse chemical reactions, such as interference with chemical reactions and iron compound oxidation, and 10) staining.

Table 2: Material Performance Problems, Causes, Relationships to Deleterious Materials, and Test Methods

Material Performance Problems	Primary Cause	Aggregate Deleterious Characteristic	Underlying Cause	Test Method
Lower Strength/Stability of PCC, HMA, UAB	Aggregate crushing/cracking under static or dynamic service loading of PCC	Weak aggregate	Porous, weakly cemented, laminated, cleaved structure, weathered particle surface	ACV, PLS
	Poor bond with asphalt binder (stripping) or portland cement	Poor aggregate surface	Coated with clay, dust	PI, SE, MB, minus #200
			Encrustations, weathered surface	Petrographic analysis
	High water demand in PCC	Excess fines	Impact breakage & abrasion during handling	LAA, MD, WBM, Isd2
			Excess dust in gradation	Minus #200
		Highly plastic fines		PI, SE, MB
		Poor aggregate shape	Flat & Elongated	Flat & Elongated
	PC hydration interference	Organic matter		Organic Impurities
	Poor HMA volumetrics	Poor aggregate shape	Flat & Elongated	Flat & Elongated
	Poor grain-to-grain contact of HMA and UAB from high fines content	Weak, abrasion-prone aggregate	Impact breakage & abrasion during handling	LAA, ACV, MD, WBM
	Poor grain-to-grain contact of HMA and UAB from loss of drainability	Weak, abrasion-prone aggregate	Impact breakage & abrasion during compaction	LAA, ACV, MD, WBM
	Poor grain-to-grain contact of HMA and UAB from clay lubrication	Highly plastic fines		PI, SE, MB
Lower Durability of PCC, HMA, UAB (Unsound aggregate)	Swelling/shrinking from wetting/drying	Water absorptive clay		PI, SE, MB, Isd2, WBM, MD, ΔPLS
	Expansion/contraction from freezing/thawing	Poor pore structure		NaSO4, WAFT, BSG, Abs, ACV, PLS, MD, WBM
	Expansion from gypsum reaction	Presence of gypsum		Petrographic analysis
	Expansion from oxidation	Presence of iron compounds		Petrographic analysis

TABLE CONTINUED ON NEXT PAGE

Material Performance Problems	Primary Cause	Aggregate Deleterious Characteristic	Underlying Cause	Test Method
Lower Durability of PCC, HMA, UAB (Unsound aggregate) [continued]	Excessive thermal expansion	Excessive thermally expansive aggregate		CTE, Petrographic analysis
	Raveling of HMA due to poor bond with binder	Poor aggregate surface	Coated with clay, dust	PI, SE, MB
			Encrustations, weathered surface	Petrographic analysis
Poor Appearance of PCC & HMA	Popouts from expansion from freezing or swelling/ shrinkage and break down from wetting/drying	Poor pore structure		MD, Isd2, WBM, ΔPLS, PI, MB, SE
Poor Appearance of PCC	Staining	Organic matter presence		Petrographic analysis
		Iron compounds presence		Petrographic analysis
Loss of Workability of PCC & HMA	High water and asphalt binder demand from increased fines and gradation change	Weak, abrasion-prone aggregate	Impact breakage & abrasion during handling	LAA, Isd2, MD, WBM
			Excess dust in gradation	Minus #200
	Poor particle shape		Flat & Elongated	Flat & Elongated
Excess Surface Wear of PCC & HMA		Weak, non-abrasion-resistant	Porous, laminated, cleaved structure, weathered particle surface	PLS, LAA, Isd2, MD, WBM

Because the objective of this study is to provide a system of tests to estimate deleterious materials as used by MoDOT TM-71, the test methods that MoDOT already specifies will not be part of the estimation system. These methods are Flat and Elongated, Minus #200 Sieve, and Organic Impurities.

The primary deleterious materials sensitive to water are clay-bearing materials, such as mud balls, shale, shaly rock, “cap+20%”, and to a lesser extent, soft rock and some forms of OFM. Materials sensitive to handling and crushing would be weak materials, which include most of the deleterious materials discussed above. Thus, it may be necessary to include several types of tests for predicting each of the nine types of deleterious materials (Table 1). During the course of the study it became apparent that several of the nine types could be combined, such as is already done for “deleterious rock”. MoDOT may want to consider simplifying the assignment of deleterious types across the five types of aggregate products (MoDOT Standard Specifications sections 1002-1007). Looking at Table 1, and from the results of the testing program, soft shale, soft rock, mud balls, and soft chert seem to offer similar problems to construction materials, and respond in a similar manner to the test methods that emerged in this study. On the other hand, hard chert and shaly rock tend to cause different problems and respond differently to specific test methods than the softer materials in the higher quality end products.

The products of this research project would be one or more simple equations (to be placed in a spreadsheet) into which the results of objective tests would be entered. The resulting factors might be termed the “Shale Factor” (SF) and the “Total Deleterious Materials Factor” (TDMF), as two examples.

The form of the relationships would resemble:

$$TDMF = a_0 + a_1 x_1 + a_2 x_2 + \ldots\ldots\ a_n x_m \qquad (1)$$

Where a_i = regression constants; i = 0, 1,…n
x_i = test results; i = 1, 2,…m

The left-hand side of the equation would be the predicted values of MoDOT’s TM-71 method. The right-hand side of the equation will be the predictors of the left-hand side by a combination of the results of objective tests.

Soft shale/clay characteristics could be defined by several, but certainly not all, of the following test methods, which would be somewhat gentle and most likely water-related: wet ball mill (MoDOT), micro-Deval (AASHTO T 327-06), delta point load index (ASTM D 5731-07) [the delta point load test is a before-and-after water-soaking strength test], and sieved slake durability index. Assistance in identification of the plasticity of the materials could come from: sand equivalent (AASHTO T 176-02), plasticity index (AASHTO T 89-02 and T 90-00), or methylene blue (AASHTO T330).

The hard shale/deleterious rock characteristics would be represented by somewhat harsher tests such as LA abrasion (AASHTO T 96-02) and aggregate crushing value (BS 812-110: 1990). Other tests that may find their way into the regression equations could include sodium sulfate soundness (AASHTO T 104-99), water-alcohol freeze-thaw soundness (MoDOT T-14), specific gravity, and absorption (AASHTO T 85-91). It is possible that the equations may have most of the same test types in them, and/or there may be only one test in each equation. The goal would be to have as few tests involved as possible.

In this manner, the deleterious testing method would retain the strength that it presently has, which is: not only is the type of deleterious material determined, but the amounts (percents of each type of deleterious material) as well.

RESEARCH PROJECT AGGREGATE TESTING

Researchers from Missouri S&T were to perform aggregate testing on a variety of aggregates, chosen by MoDOT to reflect a range in quality and use. The experimental testing plan was limited in scope to include three different MoDOT Section 1000 materials (1002, 1005, 1007), with three different ledges per aggregate-use type, along with two aggregate levels of quality. These two levels of quality would be represented by 1) the as-delivered condition and 2) the as-delivered amount of deleterious augmented by some additional deleterious material seeded into the aggregate to achieve a lower quality level. Each of the nine aggregates were to be subjected to a battery of aggregate tests (as presented above), and the results were to be used to produce the prediction equations.

MoDOT CONTRIBUTION

MoDOT personnel were to sample the production stone stockpiles from each ledge and blend the replicate bags of material prior to delivery. MoDOT personnel were to perform the TM-71 deleterious materials tests and report the results to Missouri S&T researchers. MoDOT was also charged with supplying deleterious material specific to each ledge. Other historical data associated with the materials was to be supplied.

POTENTIAL PROBLEMS

Because the final prediction system may include test methods for which MoDOT does not currently have data, then it is possible that no verification of the prediction model could occur. Verification (and possible model adjustment) would have to come after implementation of the new test methods by MoDOT.

A second problem may be that some of the MoDOT aggregate specifications limit certain deleterious materials, such as shale content, to very small amounts. The threshold levels may be too low for detection by the aggregate test methods to be used in this study.

A third problem could be that some of the parent aggregate may not have certain deleterious materials associated with it.

OBJECTIVE

The objective of this study is to establish a replacement of the existing MoDOT TM-71 deleterious materials method with a more objective system of test methods which would cover the various controlling behavior factors that the TM-71 method represents.

LITERATURE REVIEW

DELETERIOUS MATERIALS

Deleterious materials are defined as materials that are extraneous to the parent material and diminish the optimum use of the aggregate product. Examples are shale, clay balls, soft rock, coal, lignite, wood, organic matter, minus #200 sieve material, soft chert, hard chert, and anything that would fall under the category of lightweight pieces. The literature contains numerous references to the negative action of various deleterious materials (Lang, 1931; Swenson and Chaly, 1956; Bloem, 1966). It has been shown that small amounts of deleterious material can result in poor performance even for aggregates with good field performance (Marks and Dubberke, 1982). “Deleterious material” is a relative term. A certain type of material at a certain content may be deleterious in some applications but not so in others. Due to the limited scope of the present project, deleterious materials not included in the following discussion include those that cause harmful chemical reactions and unsightly staining and efflorescence, such as organic impurities, soluble alkalis, reactive silica, and iron compounds, or have poor particle shape characteristics. These types of deleterious materials are handled by other MoDOT specified tests and policies, so they will not be considered below.

DELINEATION OF DELETERIOUS MATERIALS

There have been a number of attempts to organize deleterious materials into systems (Lang, 1938; Walker and Bloem, 1950; Swenson and Chaly, 1956). Three types have emerged; each of the three is based on one of the following: 1) type of deleterious material, e.g. shale, 2) effect on PCC, HMA, or UAB, such as freeze/thaw damage, and 3) characteristics of aggregates that adversely affect the PCC, HMA, or UAB, such as toughness. MoDOT’s present system (TM-71) delineates the type of deleterious material.

A common way in which deleterious materials are controlled is to prescribe certain test methods, then compare results to published acceptance limits (such as AASHTO M 80) for various classes of deleterious materials. Typical AASHTO test methods include clay lumps and friable particles (T 112), coal and lignite (T 113), low specific gravity chert (T 113), and material finer than #200 sieve (T 11). Other test methods relate to both deleterious materials and to the parent material. Examples of these methods are those that quantify toughness (Los Angeles abrasion T 96), soundness (sulfate soundness T 104), and absorption (T 85). Usually, deleterious materials fare worse in toughness and soundness tests than the parent rock, thus these methods can also be used for delineation of deleterious materials. The method used by MoDOT is MoDOT TM-71, which is a visual examination of particles, a rudimentary form of a petrographic analysis. Had some other method of delineation of deleterious materials been used in this study, the prediction of deleterious materials would probably show different results in the relative importance of different aggregate test methods.

Various deleterious actions and some commonly associated identifying test methods (as presented in the Introduction of this report) are discussed below.

DELETERIOUS ACTIONS

Impact and Abrasion Action

Deleterious action by impact and abrasion of aggregate can occur during handling, stockpiling, bin loading, hauling, mixing, and abrasion across abutting pavement cracks and joints. Particles rubbing against each other or impacting each other or other objects can break down loose or unbound aggregate, changing gradation and increasing fines content, thus decreasing concrete and asphalt mixture workability, decreasing the ability to entrain air in concrete, and causing a loss of stability in aggregate base materials (Gray, 1962; Krebs and Walker, 1971; Folliard and Smith, 2003; Rangaraju and Edlinski, 2008). Abrasion from tire wear of concrete slabs and asphalt pavements can result in loss of surface texture and skid resistance (Senior and Rogers, 1991). The ability to resist impact and abrasion is referred to as toughness. Several test methods have been examined for characterization of toughness, such as Los Angeles abrasion, micro-Deval, wet ball mill, and sieved slake durability (Krebs and Walker, 1971; Richardson, 1985; Senior and Rogers, 1991, Saeed et al., 2001; Cooly and James, 2003; Meininger, 2004; Meininger, 2006; Rangaraju and Edlinski, 2008). Friable particles are subject to impact, resulting in breakdown into smaller particles or even a contribution to fines content. Soft particles are different—they are more prone to just abrasion (Forster, 2006).
The following methods are considered tests of impact and abrasion.

Los Angeles Abrasion

The Los Angeles Abrasion (LAA) test (AASHTO T 96) involves a two-fraction coarse aggregate specimen in a dry state being subjected to impact and abrasion by tumbling steel balls and aggregate particles inside a revolving drum (AASHTO, 2002). Resistance to impact and abrasion is called toughness. Toughness, as measured in the LAA method, is related to asphalt pavement stability (Krebs and Walker, 1971) and concrete aggregate resistance to degradation (Meininger, 2006) although the results of the test do not correlate directly with field performance (Krebs and Walker, 1971; Senior and Rogers, 1991). Some authors consider the LAA as both an impact and abrasion test (Cooly and James, 2003), while others felt it is mainly an impact test (Senior and Rogers, 1991; Rangaraju and Edlinski, 2008). It has been observed that sometimes weaker materials can actually exhibit lower losses due to their ability to absorb impact through elastic accommodation and that deteriorated material in the drum may also absorb some of the impact (Meininger, 2006). Also, the lack of water in the test method may lead to poor field performance correlation because of the lack of interaction of impact/abrasion and water sensitivity (Senior and Rogers, 1991). In a review of aggregate test methods, LAA was evaluated as having merit in prediction of aggregate breakdown, but was limited in prediction of PCC pavement performance (Folliard and Smith, 2003).

Eighty percent of the state DOT’s have LAA recommended limits for HMA of 40-45 percent loss (Kandhal and Parker, 1998). AASHTO M 80 limits LAA to 50 for PCC aggregates (AASHTO, 1999). MoDOT limitations are 50 for HMA aggregates, PCC crushed stone, and seal coat (section 1003) aggregates; 45 for PCC gravels; 55 for bituminous surface blade (section 1004) materials, and 60 for unbound surface (section 1006) aggregate (MoDOT, 2004).

Micro-Deval

The micro-Deval (MD) test (AASHTO T 327) subjects a coarse graded material to revolving in a drum with steel balls (AASHTO, 2006), but the action is mainly abrasion, not impact (Cooly and James, 2003; Rangaraju and Edlinski, 2008). Also, because water is present, the MD test is also a measure of a material’s sensitivity to water and is related to weatherability. So, the test should be applicable to HMA, unbound base, and PCC aggregates. The test is purportedly more applicable to field performance than the LAA method, such as wearing of aggregate from tire wear (Senior and Rogers, 1991). The MD method has been shown to have a greater precision than LAA (Senior and Rogers, 1991). Several studies have shown that a strong correlation between MD and LAA does not exist (Kandhal and Parker, 1998; Cooly and James, 2003; Meininger, 2004; Rangaraju and Edlinski, 2008). It has been postulated that grading of the aggregate specimen is more important to MD than LAA (Rangaraju and Edlinski, 2008). Strong correlations have been found between MD and magnesium sulfate soundness and wet ball mill by some (Kandhal and Parker, 1998; Jayawickrama et al., 2001) while others have disagreed (Meininger, 2004). The MD method was selected as a superior test for evaluation of granular base, asphalt mixture, and portland cement concrete aggregates (Senior and Rogers, 1991; Kandhal and Parker, 1998; Saeed, et al., 2001; Folliard and Smith, 2003; Meininger, 2004; White et al., 2006).

Recommended limits for HMA surface and binder courses of 17 and 20, respectively, have been reported (Kandhal and Parker, 1998). A level of 15 percent loss has also been suggested for HMA (White et al., 2006). For unbound granular base, Saeed et al. (2001) proposed a sliding scale of MD threshold values based on traffic level, moisture availability, and frost action. For an area of high moisture availability and frost potential, the maximum MD value for medium and high traffic levels was 5; for low traffic: 15; for less severe conditions: up to 45.

Wet Ball Mill

The wet ball mill (WBM) test (Tex-116-E) is similar to an LAA test with the addition of water (TexDOT, 2000). Thus, all three destructive factors discussed above are present: impact, abrasion, and water's contribution to both actions. The WBM method was developed as a test method for assessing aggregate for base material. The wet ball mill test method has been in use for aggregate quality testing in various forms for a number of years and for a variety of aggregate end-use purposes, including railroad ballast. Various designations include Mill Abrasion (Clifton et al., 1987; Clifton et al., 1987(2); Selig and Boucher, 1990; UP&BNSFR, 2001) and Texas Wet Ball Mill (Texas DOT, 2000). A good correlation has been found between MD and WBM. However, the method has exhibited greater precision than the MD method (Jayawickrama et al., 2001). One state's recommended upper limit for granular base is 55 percent loss (Texas DOT, 2000).

Sieved Slake Durability

The sieved slake durability (I_{sd2}) test was adapted from ASTM D 4644 to rate shale for applicability as embankment, subgrade, and subbase materials in regard to durability (Richardson, 1984; Richardson, 1985; Richardson and Long, 1987). The test involves the tumbling of particles in a mesh drum in water, with a subsequent evaluation of degradation via a sieve analysis. The action mainly involves sensitivity to water, but there is some abrasive action, thus the method's inclusion in this section. I_{sd2} values of shale have been reported to range from 2 to 90 percent (Richardson, 1984).

Crushing/Cracking During Loading Action

Another destructive action on aggregate that is similar to impact and/or abrasion is a crushing action under static or dynamic load, such as the weight of a stockpile or the compactive effort during construction. Cracking action could occur during service loading of a concrete structure. Breakdown of loose aggregate is somewhat a function of particle shape, where a more elongated angular shape tends to break more easily. Also, a more well-graded aggregate will break down less easily because of the support offered by the smaller particles. Like impact and abrasion, crushing results in a finer gradation and a reduction in desired physical properties (Gray, 1962; Senior and Rogers, 1991; Lade et al., 1996). In concrete, shale and soft sandstones have resulted in significant losses of strength (Lang, 1927; Emmons, 1930; Walker and Bloem, 1950; Dolar-Mantuani, 1978; Richardson and Whitwell, 2009).

Two test methods are thought to represent the action of aggregate under static or dynamic loading: aggregate crushing value and the point load strength.

Aggregate Crushing Value

The aggregate crushing value (ACV) was developed as a standard aggregate quality test (BS 812, 1990) in Britain for a variety of aggregate end-uses. The aggregate crushing value test method (British Standards Institution BS 812: Part 110) consists of subjecting a compacted specimen of aggregate particles to a static load, then measuring the amount of breakdown (BSI, 1990). The aggregate particles bear on each other and are subjected to point contact loads (thus to an indirect tensile load) as well as abrasion action as the particles slide past each other. Being subjected to internal tensile loading would make the test a measure of both tensile strength and elastic response to load. ACV results correlate well with Los Angeles abrasion results (BSI, 1998; Kandhal and Parker, 1998; Saeed et al., 2001; Williamson et al., 2007). Saeed et al. (2001) have found a fair correlation of ACV with MD. Rodgers et al. (2000) have found good correlation of the ACV with field performance of unbound aggregate pavement surfaces. They also noted additional degradation when the test was performed wet as opposed to dry. It has been singled out as a good measure of the strength of aggregate in a graded aggregate setting (Folliard and Smith, 2003). The recommended ACV limit for HMA of 30 percent loss has been reported by Kandhal and Parker (1998).

Point Load Strength

Crushing at a local level within an aggregate particle relates to tensile strength. The measurement of tensile strength of geologic materials has seen several approaches. One is the indirect tensile strength test, also known as the Brazilian test. In this method, a rock core (or concrete cylinder or asphalt puck) is placed on its side with a line load applied diametrically. The Point Load Index test (ASTM D 5731-07) was developed as a quick test method to estimate the indirect tensile strength of rock cores (ASTM, 2007). It is similar to the indirect tension method, but instead of applying a line load, a point load is used. This allows a smaller load and thus a smaller, simpler loading device. Specimens can also be loaded axially; likewise, irregular lumps can be tested (Broch and Franklin, 1972; Bieniawski, 1975). Major advantages of the method include the ability to test irregular lumps, a small load frame requirement, and quickness of testing, resulting in a potential for testing a larger number of specimens. Specimen size affects the outcome, so the results need to be converted to a standard equivalent size (typically 50 mm). Strength decreases as specimen size increases (Hardin, 1985; Richardson, 1989; McDowell and Bolton, 1998; Lade et al., 1996). ASTM D 5731-07 recommends testing specimens no smaller than 30 mm, primarily to assure that the specimen fails in tension rather than compression (ASTM, 2007). One study showed that even for specimens less than 10 mm, results were valid as long as the specimens failed in tension, as opposed to crushing. This concept works for harder aggregates (Lobo-Guerrero and Vallejo, 2006). The point load strength (PLS) has been used to evaluate the durability of shale (Richardson, 1985).

Swelling/Shrinkage and Breakdown from Wetting/Drying

Shale, clay lumps, coal, and lignite are known to be sensitive to wetting and drying cycles. Disintegration in bases, subbases, and subgrades can cause loss of strength and possible swelling, resulting in the loss of stability in pavement structures. Durability rating systems for shale have been developed (Richardson, 1984; Richardson and Wiles, 1990).

Shale, clay lumps, coal, and lignite also disintegrate or swell in concrete slabs or even asphalt pavements, leading to popouts and pitting, or micro-cracking of concrete (Forster, 2006). Unfortunately, it has been found that creation of specifications to control damage from shale has met with limited success due to the wide variation in shale characteristics (Walker and Proudley, 1932).

Because shale and other types of soft rocks fail by different mechanisms, a wide variety of tests have been utilized to assess susceptibility to degradation in the presence of water. Among these are the sieved slake durability index, wet ball mill, micro-Deval, plasticity index (PI), sand equivalent (SE), methylene blue (MB), and delta point load strength. MB values have been linked to degradable aggregate (Bjarnason, et al., 2000).

The ***sieved slake durability, wet ball mill, and micro-Deval*** methods have been discussed earlier. Clay content and activity have been shown to relate to the durability and swelling characteristics of shale (Richardson, 1984), thus, measures of clay characteristics could have some correlation with deleterious action. Typical tests that would represent this sort of activity would include ***PI, sand equivalent, and methylene blue***. These will be discussed in more detail later in the report.

Delta Point Load Strength

The aforementioned point load test can be performed on both dry and wet specimens. The difference between the dry and wet strengths is called the delta point load strength (ΔPLS). The ΔPLS test method was developed to quantify the loss in strength from soaking. As ΔPLS increases, durability has been shown to decrease. Hard shales of intermediate durability have exhibited ΔPLS values as low as 13 percent (Richardson and Wiles, 1990).

Freeze/Thaw Action

Deleterious particles in concrete can lead to several types of distress, including popouts from hard chert, pitting from softer materials, map cracking and D-cracking (Krebs and Walker, 1971). Walker and Bloem (1950) identified deleterious materials in this regard to include porous chert, weathered rock, laminated rock, argillaceous rock, and shale. As little as a five percent content of certain soft stones and shale caused significant losses of freeze-thaw durability.

Walker and Proudley (1932) also included chert as a deleterious material, and rated shale and chert as the most deleterious to concrete. Lang (1931) divided deleterious materials into those that undergo volume change (shale and certain cherts), and those that were soft or weak. Aggregate expansion can caused D-cracking damage to concrete, and popouts in both concrete and asphalt pavements. Freezing/thawing action also broke down aggregate in stockpiles, leading to the above-mentioned problems of increased fines and changed gradation.

Poor performance of inferior aggregate (deleterious) materials has been linked to the particle's pore characteristics, elastic accommodation, and mineralogy (Verbeck and Landgren, 1960). These three factors are discussed in the next section.

Pore Characteristics

Pore characteristics include pore size, distribution, and shape. Pore size and distribution relates to permeability, the ability of water to enter and pass out of aggregate particles. Pore shape affects the ease of which water can escape a pore. A variety of aggregate properties and associated test methods have been used for assessment of aggregate frost susceptibility, including absorption, bulk specific gravity, and soundness tests: water-alcohol freeze-thaw soundness and sulfate soundness.

Tests that relate to pore characteristics are presented below.

Absorption

Absorption, typically measured by AASHTO T 85 (AASHTO, 2000), has been considered a viable indicator of frost susceptibility. It typically is one of the better stand-alone tests for correlation with durability, although the correlation is not high. However, the test is easily and commonly performed (Dolch, 1966; Senior and Rogers, 1991). Aggregates with low absorption (less than 0.3%) frequently show acceptable resistance to frost damage. Upon exposure, there is insufficient water available to cause damage. However, absorption does not accurately measure the ease of water entry and exit as affected by pore shape and distribution. It has been postulated that a more accurate assessment would come from a combination of absorption and permeability (Dolch, 1959). Others have found a good correlation between absorption and AASHTO T 161 Method B "Resistance of Concrete to Rapid Freezing and Thawing" (AASHTO, 2000). Absorption values less than 1.5 percent indicated durability factors (DF) greater than 75, while absorptions greater than two percent were associated with inferior DFs (Koubaa and Snyder, 1996; Richardson, 2009). There are highly porous aggregates that exhibit good durability during freezing and thawing because of large pores that drain easily (Cordon, 1948).

MoDOT absorption percent limits are: 1) for HMA: 4.0 for crushed stone and 5.5 for gravel, 2) for PCC crushed stone (paving): 2.0, 3) for PCC masonry: 3.5 for crushed stone and 4.5 for gravel, 4) section 1003: 6.0, and 5) for section 1004: 7.0 (MoDOT, 2004).

Bulk Specific Gravity

Bulk specific gravity (BSG), also determined in AASHTO T 85, is a function of internal porosity and mineralogy (specific gravity of the solids). Traditionally, it has been thought that absorption is the more direct indicator of freeze-thaw susceptibility compared to specific gravity, and because the two are correlated and in fact are values produced by the same test method, specific gravity has not been considered the primary parameter of the two. However, some studies have shown that for carbonate aggregates, a certain relationship exists between specific gravity and durability. Bulk specific gravities greater than 2.60 or 2.65 exhibited superior durability and had a good correlation with DF (Koubaa and Snyder, 2001; Harman et al., 1970; Richardson, 2009). Low specific gravity chert is limited in AASHTO M 80 to 3.0 percent for paving and bridge deck concrete (AASHTO, 1999). Low specific gravity (less than 2.40) has been associated with poor freeze-thaw resistance (Sweet, 1940). However, some aggregates with very low specific gravities (2.24-2.35) and large absorptions have been shown to be quite durable–a fact explained by a large diameter pore system, which prevented the build-up of pressure (Harman et al., 1970) and possibly a lower elastic modulus, allowing greater elastic accommodation. BSG has been found to be useful in prediction of T 161 DF via regression analysis (Richardson, 2009).

Vacuum Saturated Absorption

Subjecting aggregate to vacuum will increase the amount of absorption of water into pores that are more difficult to enter. Some studies have indicated that vacuum saturated absorption (VSAbs) correlates well with T 161 Method A for aggregates with either high or low DF values (Larson et al., 1965; Larson and Cady, 1969; Richardson, 2009). Others have shown that vacuum saturated absorptions of greater than two percent exhibit excessive dilation or reduction in transverse frequency during T 161 Method A testing (Harman et al., 1970; Williamson et al., 2007).

VSAbs has been found to correlate better with both elastic accommodation tests (LAA, MD, ACV) and soundness tests. Of the three elastic accommodation tests, MD correlated best with VSAbs (Williamson et al., 2007; Richardson, 2009).

VSAbs has been found to be useful in prediction of T 161 DF via regression analysis (Richardson, 2009). VSAbs has also been put forth as a primary screening test for aggregate durability (Williamson et al., 2007; Richardson, 2009).

In general, aggregates with intermediate values of absorption or vacuum saturated absorption (1.5 to 2.5 percent) are problematic in the predictive ability of frost susceptibility.

Vacuum Saturated Specific Gravity

Again, when the absorption of vacuum saturated aggregates is determined, vacuum saturated bulk specific gravity data is also generated. VSBSG has been found to correlate with T 161 results. VSBSG has been found to be useful in prediction of T 161 DF via regression analysis, and has also been suggested as a primary screening test for aggregate durability (Richardson, 2009).

Water-Alcohol Freeze-Thaw and Sulfate Soundness

Both water-alcohol freeze-thaw soundness (AASHTO, 2007) and sulfate soundness (AASHTO, 2003) testing involve water penetration into aggregate pores, thus, these methods involve an element of ease of water entry. The methods are discussed in more detail in a subsequent section.

Elastic Accommodation/Strength

Elastic accommodation is the ability of the particle to expand upon the onset of water freezing without fracture.

Reaction can take the form of either sufficient strength to resist fracture, or elastic accommodation of the pressure. The ideal aggregate would have high tensile strength to resist stress due to expansion, but have a low modulus of elasticity to deflect elastically to accommodate the stress. A high Poisson's ratio would prevent stress from being transmitted laterally in other directions, thus limiting stress (and limiting an increase in pore pressure) in pores in those directions (Verbeck and Landgren, 1960).

Although reports have identified failure as a function of the stress exceeding the tensile strength (Powers, 1955; Verbeck and Landgren, 1960), attempts to quantify aggregate tensile strength in relation to aggregate freeze/thaw durability have not been reported. Unfortunately, high tensile strength and low modulus in brittle materials are usually mutually exclusive. Thus, interpretation of various test method results is difficult; e.g. does a high tensile strength result also indicate low elastic accommodation behavior, or not?

Freeze/thaw-type tests that utilize aggregate in an unconfined state do not consider the effect of confinement by the concrete paste.

The following are tests that reflect some aspect of the manner of the aggregate's reaction to internal pressure.

Aggregate Crushing Value and Point Load Strength

Aggregate crushing value and point load strength test methods have been presented previously. Both methods have been found to be useful in prediction of T 161 DF via regression analysis (Richardson, 2009). Walker and Bloem (1950) reported that soft deleterious aggregate lowered concrete flexural strength and freeze/thaw resistance.

Los Angeles Abrasion

The LAA test method (AASHTO T 96) subjects the aggregate specimen to abrasion and impact loading (AASHTO, 2002). The impact portion could be considered as an indirect measure of tensile strength and elastic accommodation. Unfortunately, harder, stronger aggregates may exhibit lower LAA values because of a lack of accommodation of impact loading, thus, making interpretation of results difficult (Meininger, 1978). LAA results for flat and/or elongated particles are also open to interpretation (Woolf, 1966).

Micro-Deval

Degradation action in the micro-Deval (MD) test (AASHTO T 327) is primarily due to slaking and abrasion, but not impact, as in the LAA test (AASHTO, 2006). Thus, the MD test is limited in its ability to measure tensile strength or elastic accommodation important to freeze/thaw resistance. It does have merit for use as a general quality indicator. Several studies have shown that MD results correlate with service records of durability of asphalt aggregate (Wu et al., 1998a, 1998; Kandahl and Parker, 1998). There have been mixed results reported in the literature in regard to the correlation of MD with other toughness tests, such as LAA and ACV (Kandahl and Parker, 1998; Saeed et al., 2001; Wu et al., 1998b, 1998; Richardson, 2009). MD has been found to be useful in prediction of T 161 DF via regression analysis, and has also been suggested as a primary screening test for aggregate durability (Richardson, 2009).

Water-Alcohol Freeze-Thaw Soundness

It is difficult to decide under what category to place soundness testing, because soundness assesses: 1) the ability for water to enter the aggregate’s pore system, 2) the reaction to wetting, 3) the tensile resistance to expansion and hence to tensile stress (tensile strength and elastic accommodation), and even 4) interactions with the mineralogy of the aggregate.

Various state DOTs and other agencies specify some version of the water-alcohol freeze-thaw soundness method (Forster, 2006). The AASHTO T 103 Water-Alcohol Freeze-Thaw (WAFT) method (AASHTO, 2000) has not been shown to have a strong relationship with frost resistance (Thompson et al., 1980; Mindess et al., 2003; Wu et al., 1998; Wu et al., 1998), and does not correlate particularly well with other soundness tests (Rogers, 1989; Hossain et al., 2007). However, it has been shown to have better precision than other soundness tests (Rogers, 1989). Also, it has been shown to correlate with durability better than sulfate soundness (Brink, 1958). Used in concert with either absorption or MD, WAFT has been successful at identifying marginal aggregates (Senior and Rogers, 1991). It has been noted that the degree of saturation during WAFT testing is important. Non-uniform saturation can explain the lack of agreement between WAFT results and service performance records. It was recommended that 85 percent saturation be achieved via one hour of evacuation followed by 23 hrs. of immersion prior to freeze-thaw testing (Sweet, 1940).

MoDOT's TM-14 (2007) is a hybrid of AASHTO T 103 methods B and C (MoDOT 2007). Method B correlates best with service records. MoDOT percent limits for various applications are PCC: crushed stone (paving): 16.0; and for masonry (crushed stone or gravel): 18.0 (MoDOT, 2004). Former specifications limited TM-14 to 10.0 percent for Gradation F (D-cracking prone) materials.

Magnesium and Sodium Sulfate Soundness

Probably the most commonly specified soundness test is one of the two versions of AASHTO T 104 sulfate soundness, using either magnesium or sodium sulfate (AASHTO, 2003). Like WAFT, the method employs an artificially-induced expansion, with failure measured as a change in gradation of the fabricated gradation. Thus, sulfate soundness could be considered a measure of tensile strength or elastic accommodation.

Sodium sulfate soundness ($NaSO_4$) has been found to be useful in prediction of T 161 DF via regression analysis, and has also been suggested as a primary screening test for aggregate durability (Richardson, 2009). Maximum recommended limits for sodium sulfate soundness as applied to HMA are 11 to 15 percent (about 60 percent of state DOTs) and 25, 30, and 10 percent for Methods A, B, and C, respectively for T 103 (Kandhal and Parker, 1998). Several studies have indicated a preference of magnesium sulfate soundness over sodium sulfate soundness (Kandhal and Parker, 1998; Saeed et al., 2001; White et al., 2006).

Sulfate soundness has not been shown to be an accurate predictor of frost susceptibility in PCC aggregates, either from slow cooling testing or service records. Several reasons for this include the difference in destructive mechanisms and the lack of precision of the methods (Walker and Proudley, 1932; Swenson and Chaly, 1956; Harman et al., 1970; Marks and Dubberke, 1982; Cady, 1984).The method also does not correlate well with WAFT (Brink, 1958). Some studies have reported mixed success in prediction (Paxton, 1982; Chamberlain, 1981), while in others, magnesium sulfate soundness ($MgSO_4$) has been recommended as a preferred method for relating to HMA raveling, potholes, and popouts (Kandhal and Parker, 1998), and for unbound granular base (Saeed et al., 2001). Magnesium and sodium sulfate methods do not necessarily agree. Magnesium sulfate is sometimes preferred to sodium sulfate because the solubility of the magnesium salt is less sensitive to temperature than the sodium salt, and the $MgSO_4$ crystals are more uniform, thus, $MgSO_4$ soundness results tend to be less erratic (Walker and Proudley, 1932). In general, sulfate soundness prediction of freeze-thaw durability has had mixed success, and the method suffers from imprecision. Soundness has been shown to correlate better with MD than LAA does with MD (Cuelho et al., 2007).

For unbound granular base, Saeed et al. (2001) proposed a sliding scale of $MgSO_4$ threshold values based on traffic level, moisture availability, and frost action. For an area of high moisture availability and frost potential, the maximum $MgSO_4$ value for medium and high traffic levels was 13 percent loss; for low traffic: 30; for less severe conditions: up to 45. A level of 20 has also been suggested for HMA (White et al., 2006). For PCC aggregate, AASHTO M 80 limits loss by $NaSO_4$ and $MgSO_4$ to 12 and 18 percent, respectively.

Wet Ball Mill

The wet ball mill (WBM) test method is similar to the LAA test in that aggregate is subjected to impact and abrasion by steel balls picked up on a shelf and dropped in a rotating drum plus the impact and abrasion from other aggregate particles (TexDOT, 2000). The method is similar to the micro-Deval test in that water is also present. The testing action suggests that the results could be used as a measure of tensile strength and elastic accommodation, as well as the resistance to water-induced reduction of aggregate strength. WBM results have been found to be useful in prediction of T 161 DF via regression analysis, and the method has also been suggested as a primary screening test for aggregate durability (Richardson, 2009).

Mineralogy

Trypolitic chert in carbonate aggregate has caused aggregate to disintegrate while undergoing T 161 freeze/thaw testing (Dubberke, 1983). Clay minerals are known to increase water demand in concrete, induce stripping in HMA, and lower stability of unbound granular base material. In a comparison to illites and kaolinites, smectites are the most damaging, having a greater fineness and surface activity.

Asphalt-Aggregate Bond Interference

Deleterious materials, in the form of dust or coatings, can interfere with the bond between asphalt binder and aggregate particle surfaces. Thus, stripping of binder can be the result. Presence of clay can also cause spontaneous emulsification, another cause of stripping (Stuart, 1986; Kandhal, 1992; Kandhal et al, 1998).

The following test methods relate to asphalt-aggregate bond interference.

Plasticity Index

The PI test method involves several test designations: MoDOT TM-79 (MoDOT, 2004), AASHTO T 89 (AASHTO, 2002) and T 90 (AASHTO, 2000). FHWA Technical Advisory T5040.27 (1988) indicated that the presence of clay fines can contribute to stripping. It was recommended that aggregate information in the mix design report should include PI and sand equivalent (SE) values. Suggested limits on SE and amount of deleterious material (clay lumps and friable particles) were given. Both ASTM D 1073 (Standard Specification for Fine Aggregate in Bituminous Paving Mixtures) and D 242 (Standard Specification for Mineral Filler for Bituminous Paving Mixtures) limit the PI of the minus #40 fraction of material used in bituminous mixtures to a maximum of 4. For mineral filler, most state DOTs reference AASHTO M17, which specifies a maximum PI limit of 4. A survey that targeted state DOTs that use limestone in hot mix asphalt conducted by the Missouri Limestone Producers Association (MLPA) revealed that about half the responding DOTs specified a limiting value for PI (MLPA, 2001). Kandhal and Parker (1998) stated in their literature review that a reported correlation between PI and field performance of HMA could not be found in the published literature. However, they recognized the PI is determined for materials that contain minus #40 to plus #200 material and that PI limits should be developed just for material passing the #200. The study indicated that little research has been done relating PI of minus #200 and HMA performance. There is a contention that a material can show plastic properties in the absence of clay content. The report also stated that the liquid limit (LL) and plastic limit (PL) tests are subjective and based on the experience of the tester. In a study of 10 fine aggregates, four of which were seeded with clay, Kandhal et al. (1998) evaluated the PI, sand equivalent, and methylene blue methods by comparing to results of AASHTO T 283 (AASHTO, 2003) and the Hamburg Wheel-Tracking Device (HWTD). Upon testing the minus #40 material, all 10 aggregates were non-plastic; however, when testing the minus #200 material, five were considered to be highly plastic. In almost all cases, those sands with high PI values for the minus #200 material were the worst performers in both the T 283 and HWTD results. MoDOT (2001) has stated that its position on PI is that each aggregate fraction of a common ledge (source) should be tested separately rather than as a blend because the coarser (gritty) size materials will not allow a thread to be rolled, yet there could be deleterious material present which could cause stripping.

For granular base material, Gray (1962) has shown that there is a three percent loss in triaxial shear strength per one percent increase in PI. MoDOT’s percent limits for granular base (section 1007) materials are 6 or 8 depending on the type of unbound base (MoDOT, 2004).

Sand Equivalent

Hveem (1953) developed the sand equivalent (SE) method as a rapid field correlation test to assign a relative amount, fineness, and character of clay-like material in an aggregate sample. Other states were quick to recognize the value in substituting the SE for the more time-consuming traditional PI-and-minus #200 combination as a field test (O' Harra, 1955). The SE method is AASHTO T 176 (AASHTO, 2002) and ASTM D2419 (ASTM, 2002). The SE is a rapid, simple test to perform requiring minimal equipment, training and experience (Kandhal and Parker, 1998). Hveem (1953) and Clough and Martinez (1961) showed that SE decreases with increasing amounts of dust and increasing activity of the dust. However, Gaynor (1968) found little correlation between SE and percent minus #200 material. Hveem also noted a decrease in SE with increasing fineness of dust. FHWA T5040.27 recommended a minimum of 45 percent for the SE (1988). For cleanliness assurance, the 1994 Superpave methodology recommended various levels tied to design traffic which MoDOT has adopted for HMA (MoDOT, 2004). Various studies have indicated that the SE test method is promising in regard to prediction of HMA moisture sensitivity. Clough and Martinez (1961) used specially prepared asphalt mixtures seeded with different types of fines. They found a good correlation between SE and immersion-retained Marshall stability and visual stripping test results. Aschenbrener (1992) also indicated that the SE has a good correlation to HMA resistance to stripping and moisture sensitivity. Kandhal et al. (1998) also found a relationship between SE and T 283 and HWTD results. However, Cross and Voth (2001) did not find a significant correlation between SE, MB, T 283, or Asphalt Pavement Analyzer (APA) rut depths. Heidebrecht (1964) did not find a significant correlation between SE and PI, but asserted that this may have been due to the differing amounts of minus #200 in the test specimens.

Studies have shown a relationship between SE and water demand in concrete mixtures (Dolar-Mantuani, 1966). In regard to concrete, Buth et al. (1967) report that a decrease in SE of 20 percent resulted in a corresponding 16 percent loss of strength and an increase in shrinkage of 15 percent, although there was no change in durability.

MoDOT has reported that because SE does not require a pre-soak, the test method does not adequately identify “shale” content (MoDOT, 2001). Lusher (2004) has pointed out the difficulties in interpreting the results of an angular, coarse graded material.

MoDOT percent limits for Superpave HMA vary from 40 to 50, depending on traffic load (MoDOT, 2004).

Methylene Blue

There are several methods for estimation of the amount and nature of deleterious materials such as clay and organic matter. One of the simplest is AASHTO T 330, the methylene blue test (AASHTO, 2007). Methylene blue is a cationic dye that is adsorbed by clay surfaces due to cationic exchange; the test is really a measure of the cation exchange capacity of the material, and is an indication of surface activity. The MB method measures the amount and nature of potentially detrimental material: greater MB means more clay and/or clay with greater activity. In regard to type of rock, igneous rocks tend to have greater MB values due to the montmorillonite (smectite) content (Kandhal and Parker, 1998). There is evidence that the MB test can be used to assess strength reductions in concrete due to the presence of various clay types (Pike, 1992; Yool et al., 1998).

The International Slurry Seal Association recommends the methylene blue test for quantifying the amount of clays, organic matter, and iron hydroxides in fine aggregate (ISSA, 1989). Kandhal and Parker (1998) correlated both SE and MB results with T 283 and HWTD results and found that the MB method had a greater correlation than SE. The recommendation was to replace PI and SE with MB for control of stripping of HMA. However, White et al. (2006) reported poor stripping predictability by MB because of a poor/fair correlation with T 283 results. Aschenbrener and Zamora (1995) also found that the MB correlated better to T 283 and field performance than the SE. Although Cross and Voth (2001) did not find a significant correlation between MB, SE, T 283, or APA rut depths, they recommended MB as a supplementary test.

Bjarnason et al. (2000) have found MB to be useful in quantification of deleterious fines, which indicated aggregate that is prone to breakdown. Yool et al. (1998) warn that the MB results are not in proportion to the damaging effects on concrete. The damage ratio is less than the MB ratio of the material.

MoDOT has stated (2001) that the MB method gives inconsistent results, and is problematic in that there is no pre-soak requirement. It is recommended that in addition to a dry shaken material, the adherent fines should also be tested (Kandhal and Parker, 1998).

Cement-Aggregate Bond Interference

As with asphalt mixtures, a key factor that affects concrete properties is the bond between the cement paste and the aggregate. Interference by dust (Pike, 1992; Gullerard and Cramer, 2003; Richardson and Whitwell, 2009) and coatings (Goldbeck, 1932; Buth et al., 1964; Shah and Chandra, 1968; Darwin and Slate, 1970; Dolar-Mantuani, 1978; Schmitt, 1990; Popovics, 1998; Richardson and Whitwell, 2009) can lower the bond strength, and in turn, lower the strength of the concrete. Goldbeck reports losses of 1.5 to 2.0 percent per one percent dust.

The action of fines is a function of the amount and nature of them. A small amount of non-plastic fines may actually enhance the properties of the concrete. The strength of the bond to the aggregate can best be determined by strength tests of the concrete, plus a post-test examination (Forster, 2006).

Water Absorption by Highly Plastic Fines

Absorption of water by highly plastic fines will increase the water demand of concrete mixtures, which lowers the workability due to both the activity of their surfaces and their extremely fine nature (Yool et al., 1998). The MB value increases with increasing fines content and hence will cause greater water demand (Stewart et al., 2007). If the water demand is satisfied by the addition of water, strength will decrease: there was an inverse relationship between liquid limit and both compressive strength and modulus of rupture (Buth et al., 1964), and between MB and compressive strength (Stewart et al., 2007). Satisfaction of water demand also lowers durability and increases shrinkage potential. A more plastic material will cause greater problems: e.g. as the montmorillonite (smectite) content increases, there will be more swelling (Swenson and Chaly, 1956). Pike (1992) reported ratios of percent strength loss per increase in MB for kaolinites, illite, and smectite as follows: 1:1, 2:1, and 4:1, respectively.

Clay Lubrication

Presence of clay in aggregate base material and asphalt mixtures can cause a loss of stability, with the type and volume of the clay being the main factors (Hveem, 1953; MoDOT, 2001). As PI increases, triaxial shear strength decreases (Gray, 1962).

SYSTEM ESTIMATION OF AGGREGATE DELETERIOUS MATERIAL CONTENT

The estimation of construction aggregate durability has been successfully accomplished for low quality select material, mainly used for embankment and highway subbase material. The approach was to rate durability in terms of loss of shear strength upon wetting, then approximate the loss rating via a regression equation. The main effects in the regression equation were the results of numerous aggregate quality test methods (Richardson, 1984; Richardson, 1985; Richardson and Long, 1987; Richardson and Wiles, 1990). In a similar manner, T 161 DF of concrete has been predicted with regression of various aggregate test methods (Richardson, 2009).

SUMMARY

There are a variety of deleterious materials that cause problems in PCC, HMA, and UAB. Friable particles, such as weakly cemented sandstones and mud balls, are weak so they break down, creating fines or they stay intact and weaken the PCC, HMA, or UAB. Weak particles, such as some shales, coal and lignite, and clay lumps may also disintegrate and cause surface PCC and HMA pitting. Soft particles abrade, creating fines. “Soft” and “weak” do not necessarily mean the same thing. Unsound particles, such as chert and some shales, may be weak or not, but they expand upon freezing or wetting and cause disruptive forces, and end up being surface popouts in PCC and HMA or causing cracking in PCC.

Friable, weak particles can be detected by impact tests, such as LAA and WBM, and by strength tests such as PLS and ACV. Soft particles can be identified by abrasion tests, such as LAA, WBM, MD, and perhaps ACV and I_{sd2}. Soft and weak particles such as clay balls and shale may contain clay, and so may be identified by PI, MB, and SE. The greater the clay activity, the greater the detrimental effect. Unsound particles can be detected by soundness tests that cause expansive pressure, such as sulfate soundness and WAFT, or by methods that detect pore characteristics, such as Abs, BSG, VSAbs, and VSBSG. There will be some cross-over detection due to the correlation of behavior, such as MD and soundness or LAA and ACV.

Within these subsets of behavior, some tests correlate well with each and some do not. Sometimes the literature reports mixed results. Expectations are:

- LAA correlates well with ACV and VSAbs, but just fair with MD.
- LAA does not correlate well with pavement performance.
- Impact tests can be “fooled” by some soft but resilient materials

- MD may correlate well with WBM and MgSO4 (or it may not), but just fair with ACV (or perhaps good),
- MD has better precision than LAA.
- MD is held up as a superior overall evaluation method.

- ACV has been singled out as a good method for graded aggregate evaluation.

- PLS is a simple way of assessing rock strength.

- WAFT does not correlate well with $NaSO_4$ or freeze-thaw tests
- WAFT in concert with other tests such as MD or Abs correlates well with pavement performance.

- $NaSO_4$ has poor precision.
- $MgSO_4$ is considered a superior test to $NaSO_4$

- $NaSO_4$ does not correlate well with pavement performance

- VSAbs correlates well with Abs, LAA, MD, ACV, VSBSG, and BSG.
- Low BSG (less than 2.4-2.5) is usually associated with poor performance.
- High Abs (greater than 2-3 percent) is usually associated with poor performance.

- PI, SE, and MB do not correlate well with each other, partly because of sample preparation differences.
- The SE procedure is flawed.
- The PI procedure is flawed.

- I_{sd2} is a good test for shale durability.

TECHNICAL APPROACH

GENERAL

Experimental Design

The proposed testing matrix included three levels of material type, three different ledges per material type, and two levels of quality (unseeded and seeded), for a total of 18 sample types. Each of the 18 sample types was to be subjected to a battery of aggregate tests and the results used to produce the TM-71 predictive equation.

Thus, the predictive regression equations would possibly contain one or more terms as determined from a suite of aggregate tests. This full factorial experiment (3x3x2) resulted in 18 different combinations.

Replicate Specimens

Normally, three replicate specimens were tested per test method. The results were analyzed for precision and identification of outliers. The replicate test results were averaged before entry into the correlation and regression analyses.

MATERIALS

MoDOT Construction and Materials (Physical Laboratory Central Laboratory) chose the specific aggregate materials. Sampling was performed by either MoDOT District or Central Laboratory personnel. Central Laboratory personnel delivered the bagged samples to the Missouri S&T Civil, Architectural, and Environmental Engineering (CArE) aggregate laboratory. The actual materials delivered are shown in Table 3.

Table 3: Aggregate Materials

Section (Quality)	Study ID No.	County	Formation
1007 Aggregate base Low	83MA0370	Ralls	Kimmswick Limestone
	85DGG014	Camden	Gasconade Dolomite
	88MA0073	Dallas	Jefferson City-Cotter Dolomite
1002 Asphalt concrete Medium	8MPEH300	Shelby	Burlington/Chouteau Limestone
	85RDP044	Osage	Jefferson City Dolomite
	83MA0234	Knox	Chouteau Limestone
1005 Portland cement concrete High	86L2R034	St. Charles	Plattin Limestone
	85RDP041	Moniteau	Burlington Limestone
	85DGG015	Pettis	Burlington Limestone

Samples came from nine ledges (different quarries). The geologic types were limited to seven formations: four of limestone and three of dolomite.

Typically, material was delivered in two forms: production stone (material completely processed, ready for use) or as material for use in the point-load test. The point load material was supposed to be of a larger size to accommodate the test method (1 to 2 in); however, many times it was no coarser than the nominal maximum size (NMS) of the production stone.

Typically, about 10 bags of production stone were delivered to the CArE aggregate laboratory per aggregate type. This material was then mixed using a Gilson Quartermaster then rebagged. The material was then tested for the as-delivered gradation. Subsequently, the remaining material was mechanically shaken through sieves for 5 to 10 minutes to separate it into various fractions. These stock sizes were then used to build the various test specimens as required by the specific test methods prior to testing.

As-delivered gradations are shown in Tables 4-6.

Table 4: Section 1007 As-Delivered Gradation Percent Passing

Formation	Kimmswick	Gasconade	Jefferson City/Cotter
ID	**83MA0370**	**85DGG014**	**88MA0073**
Sieve	**1007 Type 5**	**1007***	**1007****
1 ¼ in.	100	100	100
1	100	100	100
¾	97	81	97
½	79	41	44
3/8	68	22	13
#4	49	11	2

* ~1005 Gradation B **~1005 Gradation D

Table 5: Section 1002 As-Delivered Gradation Percent Passing

Formation	Burlington/Chouteau	Jefferson City	Chouteau
ID	**8MPEH300**	**85RDP044**	**83MA0234**
Sieve	**1002***	**1002***	**1002****
1 ¼ in.	100	100	100
1	100	100	100
¾	82	83	88
½	44	20	49
3/8	27	6	22
#4	11	3	4

* ~1005 Gradation B ** ~1005 Gradation D

Table 6: Section 1005 As-Delivered Gradation Percent Passing

Formation	Plattin	Burlington	Burlington
ID	**86L2R034**	**85RDP041**	**85DGG015**
Sieve	**1005***	**1005***	**1005***
1 ¼ in.	100	100	100
1	100	100	100
¾	95	92	91
½	63	38	54
3/8	30	14	33
#4	3	1	5

* 1005 Gradation D

MoDOT DATA

Data associated with each of the nine ledges was furnished by MoDOT in the form of Quarry Ledge Information Summaries and from deleterious material testing of the specific samples that were supplied to Missouri S&T. The information was useful for obtaining the overall picture of an aggregate's characteristics. Specific information was used in the correlation and regression analyses reported later in this report. MoDOT aggregate test results for LAA,

$NaSO_4$, WAFT, and AASHTO T 85 BSG and Absorption were also used for correlation with Missouri S&T results for verification that delivered samples were representative of the ledge material.

MoDOT personnel from the Central Laboratory tested representative samples from each of the nine aggregates in this study for deleterious materials content (TM-71). The results are shown in Table 7.

Table 7: Percent of Deleterious Materials in Study Aggregates

Section	**ID**	**Del Rock**	**Shale**	**Soft Chert**	**Hard Chert**	**OFM**	**TDM***
1002	8MPEH300	1.66	0.13	0.04	3.68	0.00	1.83
	85RDP044	0.82	0.50	0.00	1.22	0.00	1.32
	83MA0234	2.64	0.25	0.00	0.01	0.00	2.89
1005	86L2R034	0.61	0.04	0.06	0.23	0.00	0.71
	85RDP041	1.79	0.17	0.00	0.26	0.00	1.96
	85DGG015	0.83	0.00	0.00	0.57	0.00	0.83
1007	83MA0370	14.27	0.00	0.00	0.00	0.00	14.27
	85DGG014	4.34	0.03	0.00	2.28	0.00	4.37
	88MA0073	1.50	0.54	0.00	4.68	0.00	2.04

* TDM here does not include hard chert

TM-71 consists of a visual examination of a 3000 g sample of plus #4 material. The deleterious material particles were identified and classified into the above groups and weighed. Section 1002 “deleterious rock” is defined as the total of soft/porous rock, shaly rock, soft chert, and cap+20 (a non-deleterious particle with at least 20% being a cap of deleterious material). However, for this study, soft chert was quantified separately. Section 1005 “deleterious rock” is defined as the same as 1002 “deleterious rock” without the soft chert. Soft chert plus hard chert is a separate category in section 1005. Again, for this study, soft chert and hard chert were kept separate. Section 1007 deleterious rock is just soft/porous rock. OFM is “Other Foreign Material”, such as sticks.

DELETERIOUS MATERIAL SEEDED SAMPLES

In order to expand the data set to include a wider range of deleterious contents, the samples from the original as-delivered condition (which already were contaminated with some level of deleterious material) were further contaminated by adding varying amounts of additional deleterious materials. This procedure was termed “seeding”. Two decisions had to be made: 1) the type and origin of seed material, and 2) the amount of each seed material.

Type and Origin of Seed Material

Although MoDOT characterizes deleterious materials into nine kinds, the number can be reduced in regard to response to the test methods examined in this study. The actions of the tests involve wetting, impact, abrasion, compression or tension loading, and soundness-type applied stress (internal expansion). Thus, it could be expected that shale and mud balls would respond to wetting tests, while soft rock (including soft chert) would respond to loading-type tests. Aggregate prone to soundness issues (e.g. hard chert) would respond to soundness tests. Shale and deleterious rock were the only deleterious materials available that were common across the 1002, 1005, and 1007 aggregate types. Soft chert was lacking in most of the aggregates, and hard chert is considered deleterious only in 1005 materials. So, the types of deleterious materials used for seeding were shale and deleterious rock. There was sufficient hard chert in the as-delivered material to span the allowable spectrum. There was essentially no “Other Foreign Material (OFM)” in the samples, and very little soft chert.

Shale means many things to many people. In a summary of the various definitions of shale that are in use, Richardson (1984) concluded that shale includes siltstone, mudstone, mudshale, clayshale, arenaceous shale, calcareous shale, siliceous shale, bituminous shale, and gypsiferous shale. On a spectrum of behavior, this definition would include material that is classified anywhere from compaction shales to cemented shales (soft to hard, non-durable to durable). As was stated in 1932 in a report of shale in concrete (Walker and Proudley), “Shales also range into sandstones and limestones...it is how a substance acts in concrete that we are most interested in, not what its local name may be.” However, MoDOT calls very hard shales “Shaly Stone” and includes it in the “Deleterious Rock” category, not in the “Shale” category, for certain classes of stone, such as sections 1002 and 1005. For the purposes of seed material, “shale” as used here would include MoDOT’s classifications of only shale, while shaly stone would be placed in the deleterious rock (DR) seed.

In general, deleterious seed material was the material that was associated with the production material, whenever possible. On several occasions, there were two kinds of shale or deleterious rock available for a given production stone. In those cases, decisions were made to use one or the other, or a combination weighted in accordance with the amounts present. In two other cases, shale seed material was not available, and other surrogate shale materials were used. Decisions as to which deleterious seed materials to use were based on the desired balance of soft, medium, and hard shale and deleterious rock that were present in all 18 samples. In other words, it was desired to have a reasonable representation of soft, medium, and hard seed materials in the data set. In the end, based on the soaked PLS results (shown in Figs. 1-2), for shale seed there were four soft, two medium, and three hard shales. For deleterious rock, there were two soft, two medium, three mixtures of soft and hard, and two hard materials. Table 8 shows the allocation of the character of the seed materials.

Because of a labeling problem, one material that was used as a shale seed material had actually been classified by MoDOT personnel as a shaly stone and thus was classified as DR. So, in effect, that particular aggregate ultimately had no shale seed and actually had extra DR seed. The correct values were used in the regression studies.

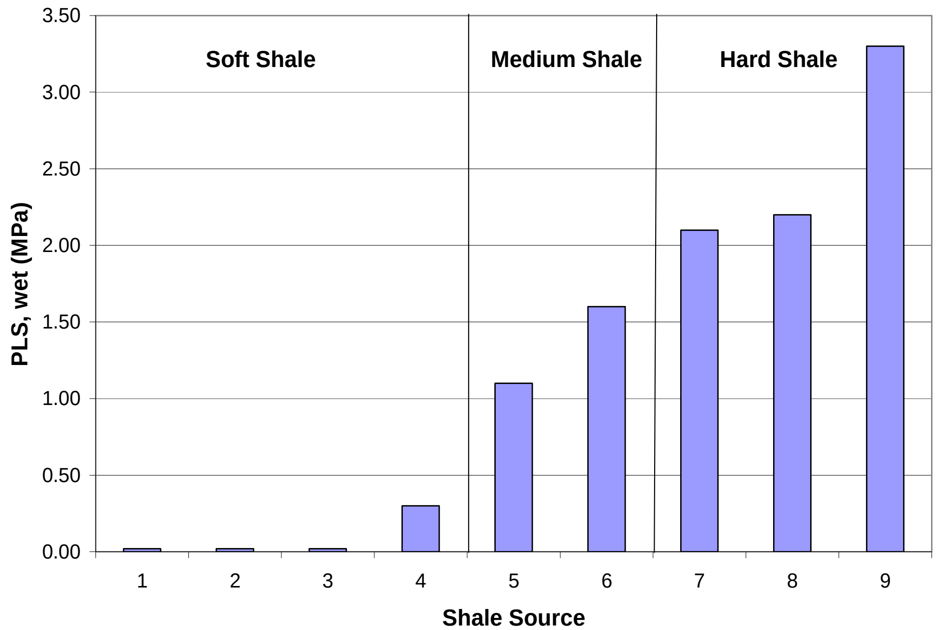

Figure 1: Shale Seed Material Hardness

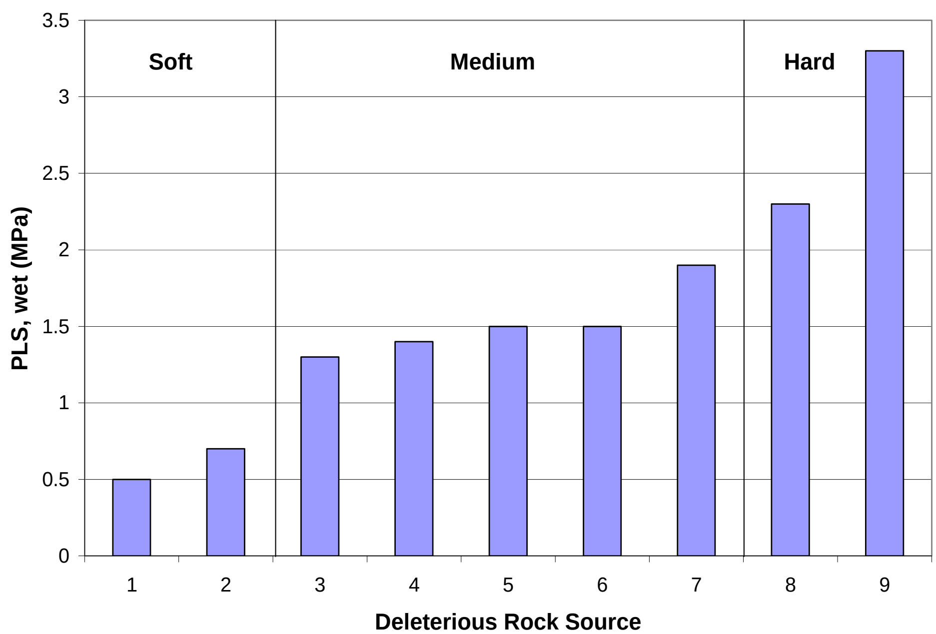

Figure 2: Deleterious Rock Seed Material Hardness

Table 8: Deleterious Material Used As Seed Material

Section	ID	Material	Description	Use
1002	8MPEH300	del rock	mostly hard some soft	used both in proportion
		shale	shaly stone*	use
	85RDP044	del rock	hard	use
	"	shale	hard	surrogate
	83MA0234	del rock	soft (50%) hard (50%)	used both in proportion
		shale	hard	use
1005	86L2R034	del rock	soft (most) hard (some)	used both in proportion
		shale	soft	use
	85RDP041	del rock	soft	use
		shale	soft	use
	85DGG015	del rock	medium	use
		shale	soft	use
1007	83MA0370	del rock	medium	use
		shale	soft	use
	85DGG014	del rock	soft	use
		shale	medium	surrogate
	88MA0073	del rock	hard (most)	use
		shale	medium	use

* Actually was "deleterious rock"

Amount of Seed Material

Each of the nine aggregates in this study was supplemented with additional amounts of seed deleterious material. The amount of seed material was tied to the allowable amount of each kind of deleterious material in MoDOT's specifications for each of the three end-use materials in this study (sections 1002, 1005, 1007). Specified allowable limits for shale, deleterious rock, and total deleterious material for section 1002 materials are 1.0, 8.0, and 8.0 %, respectively. For 1005 material, the limits are 1.0, 6.0, and 6.0%, respectively, with the additional stipulation that total chert cannot exceed 4.0%. Section 1007 material is allowed simply 15% total deleterious. After some preliminary testing and calculations, it was decided to add seed material to the as-delivered material in the following amounts: 1) 1005 material: 2.0 % shale and 4.0 % deleterious rock, 2) 1002 material: 2.0% shale and 6.0% deleterious rock, and 3) 1007 material: 5.0% shale and 10.0% deleterious rock. Coupled with the as-delivered amounts, the quantity spectrum on each material was a well-distributed range, as shown in Figs. 3-7. "Total-deleterious-material-including-hard-chert" (TDMHC) was calculated as the sum of deleterious rock, shale, soft chert (SC), and hard chert. Total Deleterious Material (TDM) was calculated as the sum of deleterious rock, shale, and soft chert for 1002, 1005, and 1007 materials. Dashed lines show the allowable limits for the 1002, 1005, and 1007 specifications.

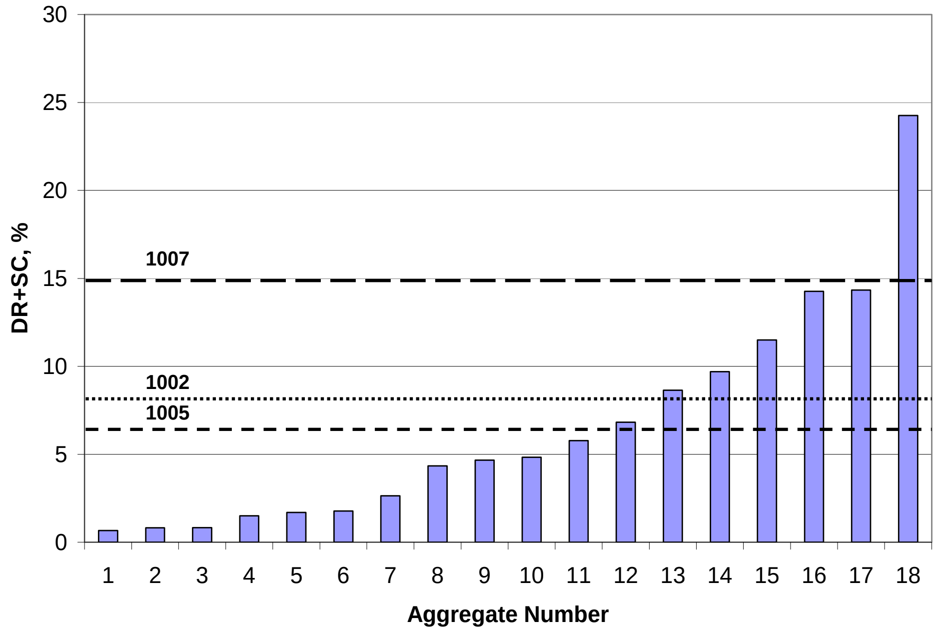

Figure 3: Distribution of Deleterious Rock Soft Chert in Samples

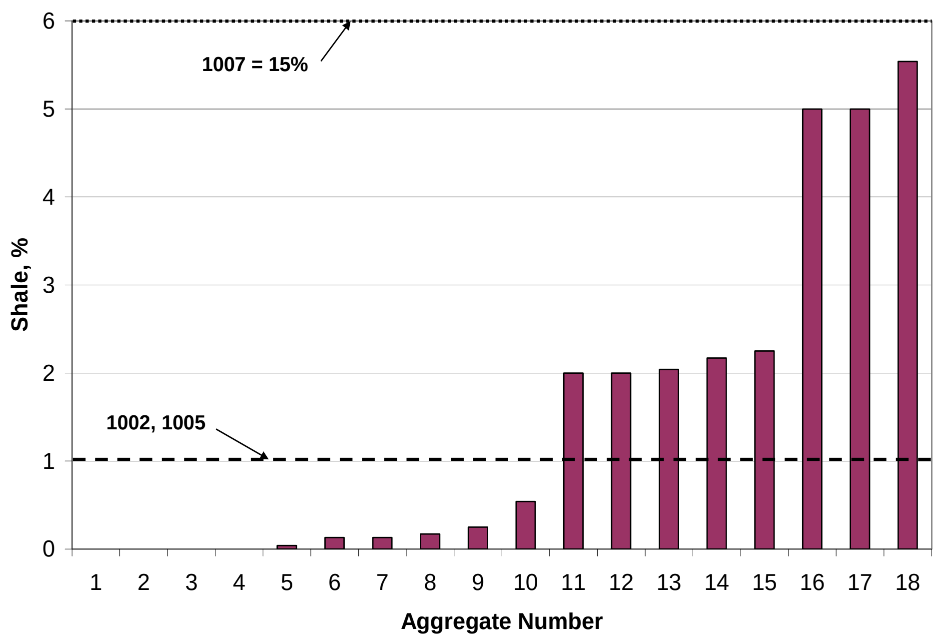

Figure 4: Distribution of Shale in Samples

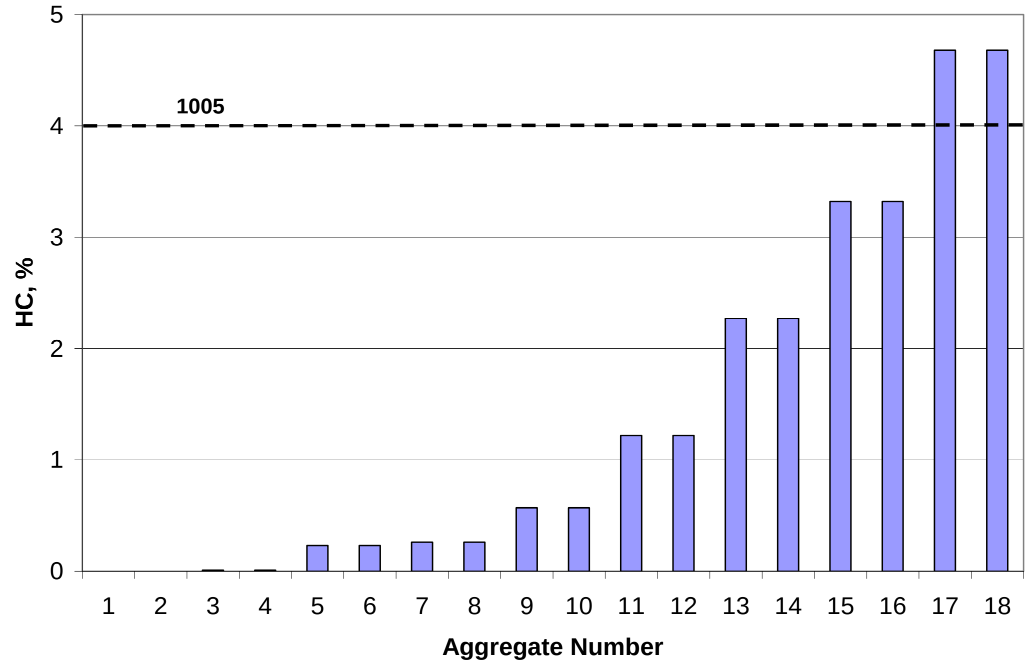

Figure 5: Distribution of Hard Chert in Samples

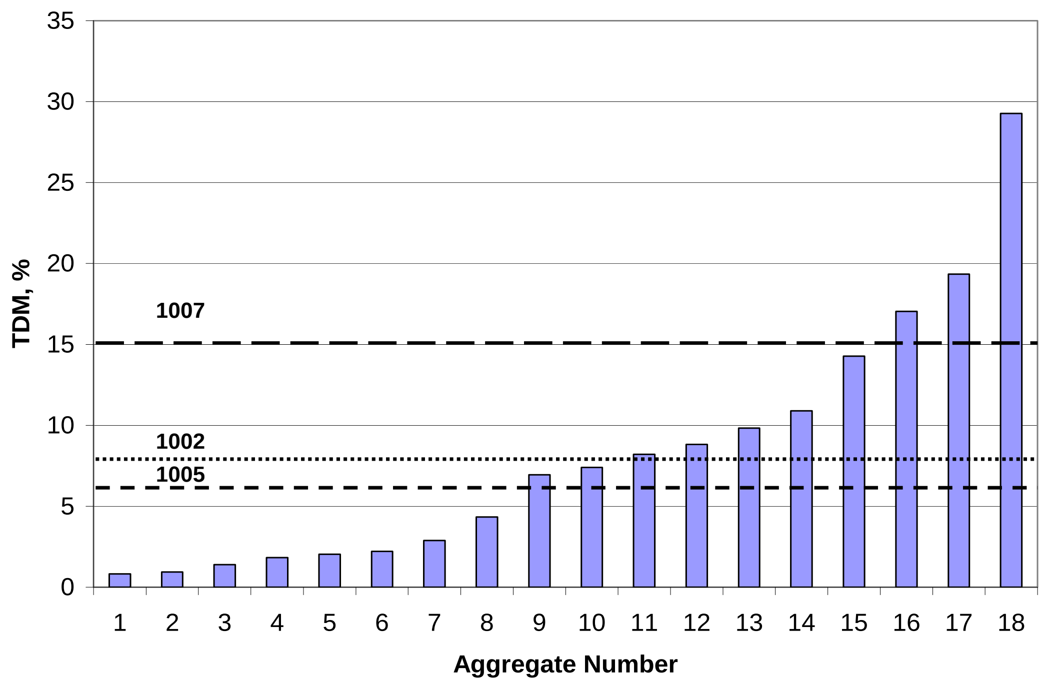

Figure 6: Distribution of Total Deleterious Material in Samples

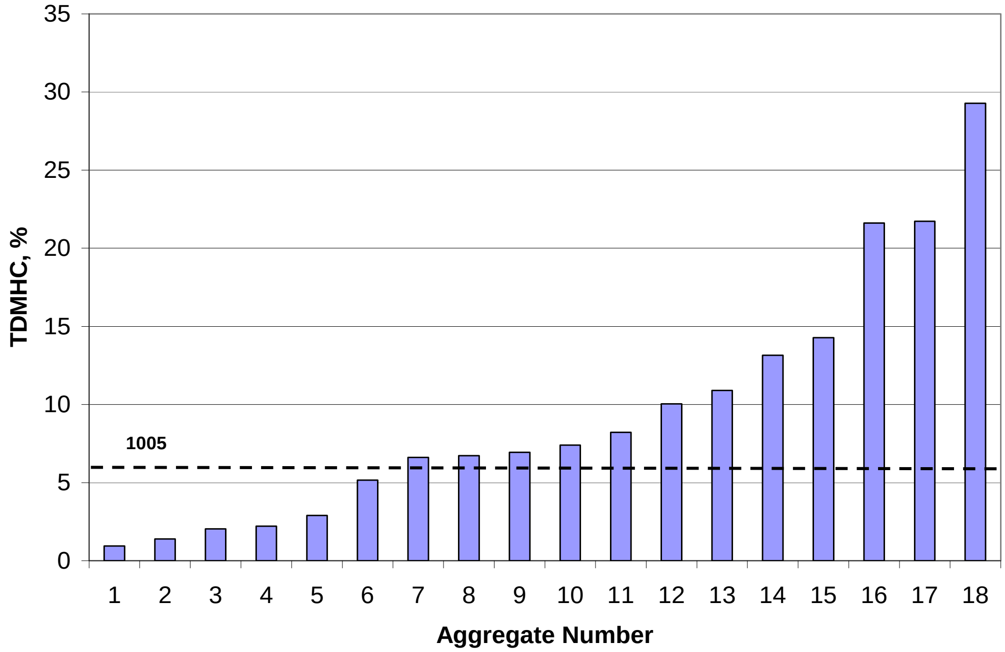

Figure 7: Distribution of Total Deleterious Material Hard Chert in Samples

TEST PROCEDURES and EQUIPMENT

The test procedures and equipment used were a mix of traditionally specified test methods and some non-traditional methods, which are discussed in the following sections.

Seeding

Both shale and deleterious rock seed material were handled in the same manner. Deleterious rock was mainly soft material, but did not include soft chert, OFM, or hard chert. The seeded material was reduced in size by use of a hammer and a steel plate. The particles were then sieved for one minute using a mechanical shaker. Because the amount of deleterious seed material that was available was limited, care was taken to not over-degrade the particles. Once the particles of various sizes were produced, test specimens were fabricated by adding the appropriate mass of seed material to the production stone on a sieve-by-sieve size basis. The seed material was a certain percent of the total specimen mass per sieve (production stone plus seed material). The seed amounts were 2% shale and 4% deleterious rock, 2% shale and 6% deleterious rock, and 5% shale and 10% deleterious rock for 1005, 1002, and 1007 materials, respectively. Details of the seeding process are included in the sections below. In general, the target seed masses were easily met for the larger specimens, but for test methods such as I_{sd2} which entail small specimen sizes but large particles, one shale particle may have satisfied the required seed amount. Judgment had to be used to try to keep the quality of seed particles the same from replicate to replicate and from test method to test method. Some variability was thus inherent to the seeding procedure. Also, when necessary, it was important to thoroughly distribute the seed material throughout the production stone, yet not degrade the soft material during the homogenization process.

Impact Breakage and Abrasion

Los Angeles Abrasion

The LAA method is considered to impart both impact and abrasion action. AASHTO T 96-02 was followed, with one exception. The specimen was not initially washed nor was it wet-sieved at the conclusion of the test because the effect of wetting would interfere with the determination of the deleterious material quantity in the specimen. The initial specimen grading followed the recommendations of the method (LAA grading is a function of the as-received gradation of the material). Thus, LAA Grading B was used for all aggregates. After the prescribed number of rotations, the material was dry-sieved over a #12 sieve and the loss recorded.

Micro-Deval

The MD method is considered to impart mostly abrasion action, as modified by the presence of water. AASHTO T 327-06 was followed for this part of the study. However, the specimen was not initially washed because the effect of wetting would interfere with the determination of the deleterious material quantity in the specimen. A Geneq, Inc. three-tiered model micro-Deval device was used. The initial specimen grading followed the recommendations of the method (MD grading is a function of the as-received gradation of the material). Thus, MD grading 8.2 was used for all aggregates. The test method calls for an initial oven dry period of 24 hrs followed by a one hour soaking period prior to rotation. After the required rotation time was achieved, the material was wet-sieved over a #16 sieve, oven dried for 24 ± 6 hrs, and the loss calculated.

Wet Ball Mill

The WBM method is considered to impart both impact and abrasion action, as modified by the presence of water. A method developed by the MoDOT Central Laboratory was utilized in this study. It is an adaptation of Texas DOT test method Tex-116-E (TexDOT, 2000). The details of this method entail the use of six steel balls and 600 revolutions of the drum, with a 2500 g specimen (plus #4 material) in water. The device used is manufactured by the Rainhart Co. and is shown in Fig. 8. The specimen was not initially washed because the effect of wetting would interfere with the determination of the deleterious material quantity in the specimen (they were soaked for 24 hrs). The specimens were wet-sieved over a #10 sieve, oven dried, and then mechanically shaken over a nest of sieves for five minutes. Details of the method can be found in Appendix A.

Figure 8: Wet Ball Mill Device

Several adjustments to the method were instituted in order to increase the precision of the method. First, specimen size was kept constant at 2500 g, rather than just achieving a minimum of 2500 g. Second, rather than assuming that the gradation of a specimen was the same as the as-delivered gradation, the specimens were actually built sieve-by-sieve to duplicate the as-delivered gradation (plus #4 sieve material). Both of these steps helped increase the precision of the replicate specimen test results.

A second reason for actually building an initial gradation was to make possible a true modification of the test method: to determine the final gradation after the standard testing was complete. The change in gradation brought about by the action of the balls, aggregate, and water was quantified by the method developed in previous research (Richardson, 1984; Richardson, 2009). The new method is termed herein as the “Wet Ball Mill-Modified” (WBMM). Details are included in Appendix A. WBMM can be calculated either on a #4 sieve basis or a #10 sieve basis. In this study, the #4 sieve basis is reported. Future studies should include the #10 basis method.

Sieved Slake Durability

The sieved slake durability test is a modified version of ASTM D 4644-04. The method consists of placing 500 g of the largest particles available (oven-dry) into a #10 mesh drum that is partially immersed in a trough of water. The drum is rotated for 10 minutes at 20 revolutions/minute. The material is oven dried at 110 ± 5°C (230 °F) for 24 ± 6 hrs, then the process is repeated. The gradation of the specimen is determined and quantified with a gradation index known as the aggregate gradation modulus, which weights the calculated sieved slake durability index I_{sd2} more heavily for a greater degree of break down. The greater the index (on a scale of zero to 100), the more durable the aggregate. The testing device is shown in Fig. 9. The full procedure is discussed in Appendix B.

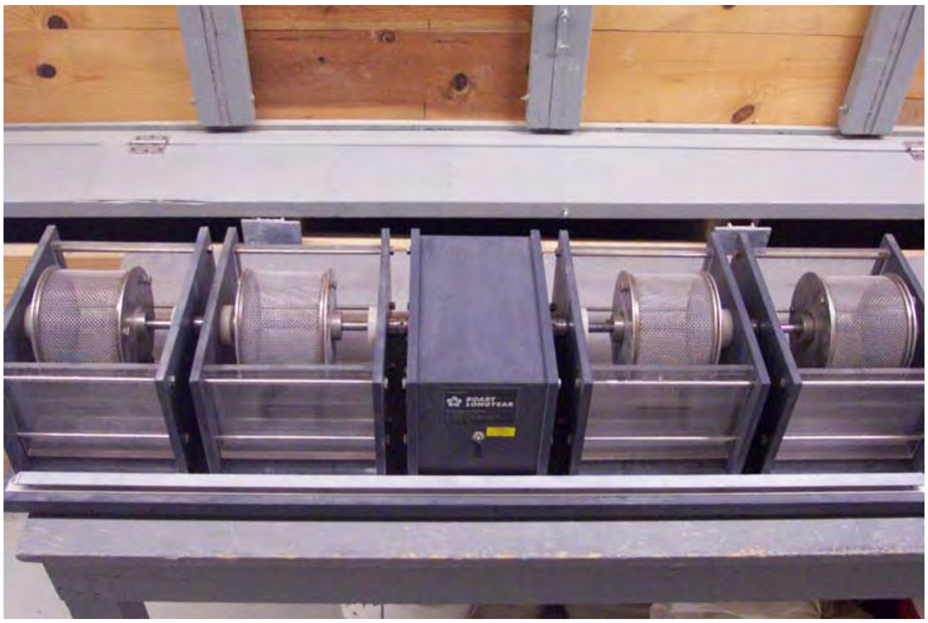

Figure 9: Sieved Slake Durability Device

Crushing Under Loading

Aggregate Crushing Value

The ACV is a direct-compression type of test which entails lightly compacting an unwashed oven dry (24 ± 6 hrs) graded sample (usually passing a 0.52 in. (13.2 mm) sieve and retained on a ⅜ in. (9.5 mm) sieve) into a heavy steel mold with a rod and subjecting the material to a hydraulically–applied compression load via a plunger. The material is then sieved over a #8 sieve and the percent loss is calculated. The method used in this study followed BS 812:110. The mold and plunger were fabricated to meet the required specifications; all other equipment was commercially available. The load was applied with a 200,000 lb. compression machine, which typically is used for breaking concrete cylinder specimens. The tamping rod essentially meets specifications for a concrete slump tamping rod. Fig. 10 depicts the Missouri S&T compaction mold, plunger, and rod.

Figure 10: Missouri S&T ACV Mold, Rod, and Plunger

The material is gently compacted into the mold by dropping the tamping rod 25 times from a height of one in. per each of three layers. The compression load is then applied over a period of 10 minutes, increasing constantly until an ultimate value of 89,924 lbs. is reached. The dry material is then mechanically shaken over a #8 (2.36 mm) sieve and the loss is calculated as the ACV. The full procedure is reported in Appendix C.

Point Load Strength

The PLS method is basically a tensile-failure type test. ASTM D 5731-07 was followed with several deviations. The method calls for testing 20 pieces of oven-dried (24 ± 6 hrs) aggregate at least 30 mm in size. Each piece is placed between the testing machine's platens (points) and loaded to failure. The final load and the distance between the points at failure are recorded. The point load strength is mathematically corrected to a standard 50 mm size. Because the purpose of this test in the context of this study is to identify small percentages of soft and water-sensitive materials, the standard procedure of discarding the two greatest and two smallest values was omitted. Any of the 20 pieces that disintegrated prior to testing in the load frame were assigned a strength of zero and were included in the data set. The point load device is shown in Fig. 11.

Figure 11: Point Load Device

The device is a manually operated MATEST digital point load tester. For very low loads (shale), a different device was used that had a lower load capacity readout, which was the Geotest S5840 Multi-Loader using a 1000 kg (2200 lb) load cell.

Special large-size PLS samples were requested from MoDOT. Obtaining 1½ to 2 in. material that matched the production stone characteristics proved to be difficult; in many cases the average delivered specimens were smaller than the required 30 mm size. Other than the standard correction to 50 mm, no further attempt was made to analyze possible effects this may have had on the PLS results. The full procedure is reported in Appendix D.

It was decided to test the production size material rather than the larger 1½ to 2 in. material for two reasons: first, it was difficult to obtain large specimens from every aggregate type, and second, the deleterious materials test is a quality control type of test, thus samples of production stone would actually be tested in practice. Thus the tested particle size ranged from 0.4-0.5 in. (10-13 mm).

The handling of the seeding procedure for PLS was different from all the other test methods. Because the PLS specimen was comprised of 20 particles, attaining small percentages of deleterious materials was impossible. Thus, a different approach was required. For this method, the production stone, shale, and deleterious rock were all tested separately from each other. The results were combined mathematically via weighted averages.

Breakdown from Wetting/Drying (Swelling/Shrinking)

Sieved Slake Durability

This method has been discussed previously. The aggregate specimen is subjected to two cycles of wetting and drying in addition to a tumbling action.

Wet Ball Mill

This method has been discussed previously. The aggregate specimen is subjected to one cycle of wetting and drying in addition to tumbling and impact actions.

Micro-Deval

This method has been discussed previously. The aggregate specimen is subjected to one cycle of wetting and drying in addition to a tumbling action.

Delta Point Load Strength

Point load strength in a dry condition (PLS_{dry}) has been discussed previously. The loss in strength due to soaking is determined by testing a second set of particles after soaking in water 16 ± 2 hrs to obtain PLS_{wet}. Pieces that disintegrated during any phase of soaking or testing were considered to have zero strength and were included in the calculation of average strength. The procedure of eliminating the two highest and lowest values was also omitted. The difference between the dry and wet PLS as a fraction of dry PLS was considered the percent change-in (Delta) PLS.

Plasticity Index

The Plasticity Index (PI) is the difference between the liquid limit (LL) and the plastic limit (PL) of the minus #40 sieve material. Specimens were prepared in accordance with MoDOT TM-79, and LL and PL tests were performed in accordance with AASHTO T 90-00 and T 89-02. Three points were produced for each liquid limit replicate. Three LL replicates were produced along with three PL replicates.

The seeding procedure consisted of dry-shaking the shale and deleterious rock over a #40 sieve and then combining the above-prepared minus #40 production stone material with minus #40 shale and deleterious rock in the proper proportions.

Methylene Blue

The Methylene Blue Value is a measure of the presence of certain clay minerals. The test method followed AASHTO T 330-07. Fine production stone material (minus #40 sieve) from the preparation of the PI test material was dry sieved over a #200 sieve, as was the shale and deleterious rock seed material. The three materials were then blended in the proper proportions. A slurry was made with the material, then titrated with methylene blue solution. The full procedure is reported in Appendix E.

Sand Equivalent

Sand Equivalent testing was performed in accordance with AASHTO T 176-02 (Method 1 Air Dry), utilizing the SE mechanical shaker device. The specimen was prepared by separating the as-delivered material over a #4 sieve. The plus #4 material was cleaned by rubbing the material between the hands, as per ASTM D 2419-02 (ASTM, 2002); the minus #4 material produced in that manner was then added to the material that had already passed the #4 sieve. Then, the combined minus #4 material was reduced by riffle splitting down to a specimen size that would fill a moisture-type tin.

Special care was exercised when adding the seed material to the production stone so that the seed material did not segregate prior to and during the addition process: 150 g specimens were built according to the seed percentages, then homogenized prior to placing in the specimen tin.

Expansion/Contraction from Freezing/Thawing

Damage from freezing/thawing has been linked to four contributors: 1) aggregate pore characteristics, 2) aggregate pore length, 3) mineralogy, and 4) elastic accommodation/strength.

Aggregate Pore Characteristics

The following are tests that reflect some aspect of the manner of the aggregate's ability to take in water and to expel water, disregarding pore length as a variable. Pore size, distribution, and shape are included.

Absorption and Bulk Specific Gravity

AASHTO T 85 BSG is a function of mineralogy (specific gravity of the solids) and porosity. In the past, MoDOT has used a threshold minimum allowable BSG for certain concrete applications. Absorption is a commonly specified property for aggregate quality and has been used by MoDOT as an acceptability criterion. MoDOT personnel performed the tests in accordance with AASHTO T 85. The material tested would be all plus #4 sieve size. The data was obtained from the Quarry Ledge Information Summaries, thus was not specific to the samples tested in this study.

Vacuum Saturated Absorption and Bulk Specific Gravity

The test method in its final form was derived from methods reported in the literature from the Wisconsin DOT (Williamson et al., 2007), the Iowa DOT IM 380 (IDOT, 2004), MCHRP 86-1 (MoDOT, 1993), the maximum theoretical specific gravity of asphalt mixtures (Rice) method AASHTO T 209 (AASHTO, 2005), and AASHTO T 85-02 (AASHTO, 2002). The level of vacuum is essentially the same as in T 209 and Iowa's method, and slightly greater than the Wisconsin method. The 30 minute vacuum period is the same as Iowa's and is greater than the other three methods. The specimen is not initially washed because the effect of wetting would interfere with the determination of the deleterious material quantity in the specimen. In essence, ungraded oven-dried material (plus #4 sieve) is subjected to a vacuum of 27.5 ± 2.5 mm mercury absolute pressure for five minutes. Water is introduced under vacuum and eventually covers the aggregate. The specimen is then subjected to agitation for a total of 30 minutes under vacuum (including the initial five minutes). The material is allowed to stand submerged at atmospheric pressure for 24 hrs. At that point, the balance of the procedure follows the T 85 procedure. The full procedure is reported in Appendix F. Fig. 12 depicts the Missouri S&T vacuum saturation station.

Figure 12: Vacuum Saturation Workstation

Care was taken to minimize loss of material once the saturated, surface dry (SSD) weight was obtained. However, some loss of material could have occurred prior to weighing during the saturation and soaking steps. Thus the specimen that finally went through the weighing steps may not have contained the full amount of deleterious material.

Water-Alcohol Freeze Thaw

MoDOT's TM-14 (modified from AASHTO T 103-07, Method B) was followed. The initial specimen gradation was built to a standard gradation, consisting of three fractions: #4 to ⅜ in., ⅜ to ½ in., and ½ to ¾ in. The specimen was not initially washed because the effect of wetting would interfere with the determination of the deleterious material quantity in the specimen. After 16 cycles of freezing and thawing, the specimens were wet-sieved over a #8 sieve; the plus #8 material was oven dried, cooled, and mechanically sieved for five minutes over a #8 sieve.

Freezing and thawing cycle durations were initially determined by use of thermocouples placed in specimens undergoing freezing and thawing cycles, with the freezer and thawing tank loaded with the expected number of specimens.

It was especially important to get a good distribution of seed material in the test specimens because it was observed that the material in the bottom of the pans experienced a greater amount of degradation due to the water that was left in the pan bottoms during the freeze-thaw cycles.

Sodium Sulfate Soundness

The test methodology followed AASHTO T 104-03. However, the specimen was not initially washed because the effect of wetting would interfere with the determination of the deleterious material quantity in the specimen. All aggregate specimens were built to the standard gradation except Ash Grove, which lacked sufficient material for the ¾ - 1 in. size. The soaking cycle lasted 16 hrs. The drying time interval for all samples was established as per the test protocol to be six hours. After the five cycles were concluded, the specimens were flushed, dried, and mechanically shaken for four minutes over the appropriate sieve.

Pore Length

Length of pores was not addressed in this study because all samples had the same nominal maximum size, thus holding pore length essentially constant.

Mineralogy

Methylene Blue, Plasticity Index, and Sand Equivalent

These methods have been discussed previously.

Water-Alcohol Freeze-Thaw

This method was previously discussed. Response to freezing and/or ordering of water molecules at cold temperatures has been shown to be related to mineralogy of aggregates, hence the inclusion of the method in the Mineralogy section.

Elastic Accomodation/Strength

The following are tests that reflect some aspect of the manner of the aggregate's reaction to internal pressure. Reaction can take the forms of being sufficiently strong to resist fracture or elastic enough to accommodate the pressure.

Aggregate Crushing Value, Los Angeles Abrasion, Micro-Deval, Point Load Strength, Wet Ball Mill

These methods were previously discussed.

Water-Alcohol Freeze-Thaw

This method was previously discussed. Elastic and plastic response to the expansion and contraction during freezing and thawing ties this test into the Elastic Accomodation/Strength section of this study.

Asphalt Binder Bond Interference

Methylene Blue, Plasticity Index, and Sand Equivalent

These methods were previously discussed.

Water Absorption

Methylene Blue, Plasticity Index, and Sand Equivalent

These methods were previously discussed.

Concrete Paste Bond Interference

Methylene Blue, Plasticity Index, and Sand Equivalent

These methods were previously discussed.

Clay Lubrication

Methylene Blue, Plasticity Index, and Sand Equivalent

These methods were previously discussed.

RESULTS AND DISCUSSION

PRECISION AND OUTLIER ANALYSIS

Three replicate specimens were tested for every test sample/method. Standard deviations, coefficients of variation (CV), and ranges of CV were computed. The allowable d2s range (as published by AASHTO or ASTM) for each test method's results was determined, and a comparison was made between the results of the precision calculations and the allowable range. Also, each set of three replicate specimens' results were examined for outliers in accordance with ASTM E 178 (ASTM, 2008). Out of 810 results examined, only three sets were outside the recommended d2s ranges, and only two sets exhibited outliers. However, due to the low test values involved, it was decided that the possibility of an actual problem existing was remote and could be considered a statistical anomaly. Altogether, the replicate testing was quite precise. Table 9 shows the coefficient of variation of the data in this study for each test method, averaged across all materials.

Table 9: Precision of Unseeded Test Methods

Test Method	CV (%)*
VSBSG	0.2
I_{sd2}	0.4
MB	0.9
ACV	1.5
LAA	1.7
MD	1.8
VSAbs	2.1
LL	3.2
WBM	4.6
PL	4.7
SE	4.8
WBMM	6.0
$NaSO_4$	9.3
WAFT	9.9
PI	19.4
PLS	NA

*Single operator

TEST RESULTS

Deleterious Materials Testing

Results of MoDOT Central Laboratory testing of TM-71 Deleterious Materials content have been shown in Table 7.

Aggregate Testing

Nine different ledge materials were subjected to 15 types of aggregate tests by Missouri S&T and three by MoDOT. Results from one of the test methods were expressed in several different ways to bring the total number of test method/major effects studied to 19.

Ranges of test values in the final results data set varied from test to test. A large range is preferable in developing a regression equation in order to be able to predict a wide range of behavior of Missouri aggregates. Based on typical data from the literature, those test methods that could be characterized as having a wide range of test results included MD and $NaSO_4$. Those with a moderate range included PI, WAFT, bulk specific gravity, absorption, SE, LAA, WBM, ACV, and PLS. Those with a more narrow range were I_{sd2} and MB.

In a subjective sense, test methods could be rated in terms of ease of testing. This comes in to play when choosing methods for a predictive or threshold acceptance system, which will be discussed later. Test methods considered as fairly easy to perform include specific gravity, absorption, VSBSG, VSAbs, LAA, MD, MB, PLS, SE, and I_{sd2}. More arduous methods are $NaSO_4$, PI, WAFT, and WBM (if initial and final gradations are built).

Tables 10-13 depict the averages of all aggregate test results. Fifteen test methods were performed at Missouri S&T, while results of two more (T 85 BSG and Absorption) were extracted primarily from MoDOT's Quarry Ledge Information Summaries. Except for MoDOT data, in almost every case, each result is the average of three replicates. Results of MoDOT-determined deleterious material testing for deleterious rock (DR), shale (Shale), soft chert (SC), and hard chert (HC) are also shown.