CENTER FOR TRANSPORTATION INFRASTRUCTURE AND SAFETY

Quick Test for Concrete Durability Factor Estimation

by

David N. Richardson

NUTC
R219

A National University Transportation Center
at Missouri University of Science and Technology

Disclaimer

Technical Report Documentation Page

1. Report No. NUTC R219	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Quick Test for Concrete Durability Factor Estimation		5. Report Date March 2009
		6. Performing Organization Code
7. Author/s David N. Richardson		8. Performing Organization Report No. 00020297
9. Performing Organization Name and Address Center for Transportation Infrastructure and Safety/NUTC program Missouri University of Science and Technology 220 Engineering Research Lab Rolla, MO 65409		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DTRT06-G-0014
12. Sponsoring Organization Name and Address U.S. Department of Transportation Research and Innovative Technology Administration 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period Covered Final
		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 use of the AASHTO T 161 Durability Factor (DF) as an end-result performance specification criterion for evaluation of paving concrete. However, the test method duration can exceed 75 days before results are known. MoDOT contracted with the Missouri University of Science and Technology (Missouri S&T) to develop a method of approximation of DF based primarily on aggregate testing that would be of a shorter duration. Nineteen different ledge samples representing 18 ledges, 10 geologic formations (nine limestone and one dolomite) were sampled by MoDOT and delivered to Missouri S&T. The ledge samples represented DFs of 28 to 95 and nominal maximum aggregate sizes of ⅜ to 1 in. The aggregates were subjected to twelve different test methods. This information, coupled with MoDOT historical gradation, specific gravity, absorption, and deleterious materials data, formed the basis of the test study dataset. Multiple linear regression was used to produce seven models of varying accuracy and complexity for DF prediction. Historical T 161 DF data for the same aggregate materials (different samples) was used as the dependent variable. Model R2 values ranged from 0.804 to 0.974. Thus, seven options were open to MoDOT for consideration. As an alternate to the regression models, a threshold-limits method was presented.

17. Key Words Freeze-thaw durability, D-cracking, concrete mixtures	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.		
19. Security Classification (of this report) unclassified	20. Security Classification (of this page) unclassified	21. No. Of Pages 124	22. Price

Form DOT F 1700.7 (8-72)

FINAL REPORT
RI07-042

QUICK TEST for DURABILITY FACTOR ESTIMATION

Prepared for the

Missouri Department of Transportation
Organizational Results

By

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

March 12, 2009

The opinions, findings, and conclusions expressed in this report are those of the principal investigator and the Missouri Department of Transportation. This report does not constitute a standard, specification, or regulation.

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 Justin Carr, 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 use of the AASHTO T 161 Durability Factor (DF) as an end-result performance specification criterion for evaluation of paving concrete. However, the test method duration can exceed 75 days before results are known. MoDOT contracted with the Missouri University of Science and Technology (Missouri S&T) to develop a method of approximation of DF based primarily on aggregate testing that would be of a shorter duration. Nineteen different ledge samples representing 18 ledges, 10 geologic formations (nine limestone and one dolomite) were sampled by MoDOT and delivered to Missouri S&T. The ledge samples represented DFs of 28 to 95 and nominal maximum aggregate sizes of ⅜ to 1 in. The aggregates were subjected to twelve different test methods. This information, coupled with MoDOT historical gradation, specific gravity, absorption, and deleterious materials data, formed the basis of the test study dataset. The test methods were: Los Angeles Abrasion, Micro-Deval, wet ball mill, wet ball mill modified, Aggregate Crushing Value, Iowa Pore Index, Methylene Blue Value, sodium sulfate soundness, water-alcohol freeze-thaw soundness, point load strength, vacuum saturated bulk specific gravity, and vacuum saturated absorption. Results from historical MoDOT test methods included gradation, bulk specific gravity, absorption, deleterious rock content, and chert content.

Multiple linear regression was used to produce seven models of varying accuracy and complexity for DF prediction. Historical T 161 DF data for the same aggregate materials (different samples) were used as the dependent variable. Some models entailed test methods not normally performed by MoDOT, such as vacuum saturated bulk specific gravity, wet ball mill, Aggregate Crushing Value, and point load strength. Other models included 28 day concrete compressive strength. Less accurate models contained more familiar test methods, such as gradation, bulk specific gravity, sodium sulfate soundness, Micro-Deval, and Iowa Pore Index. Model R^2 values ranged from 0.804 to 0.974. Thus, seven options were open to MoDOT for consideration. As an alternate to the regression models, a threshold-limits method was presented.

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

Recent or proposed changes in the Missouri Department of Transportation (MoDOT) concrete pavement acceptance specifications are end-result in nature. It has been proposed that concrete will be accepted, in part, on the results of the Durability Factor (AASHTO T 161) determination from paving concrete sampled on-site. Unfortunately, the T 161 test duration is quite lengthy (75 or more days), and results may lag construction progress so much that mid-course corrections would be impossible to achieve in a timely manner. Miles of out-of-specification concrete could be placed and go undetected for several months. Thus, it would be useful to have a quicker answer. This would be especially useful if the Durability Factor (DF) ever became a Quality Level Analysis (QLA) pay factor. The solution could be achieved by being able to establish an approximation of DF that could be determined within a short time after sampling. The approximation of DF would alert the construction inspector that concrete might have durability problems.

Researchers at the Missouri University of Science and Technology (Missouri S&T) Department of Civil, Architectural, and Environmental Engineering (CArE) proposed to MoDOT to be allowed to develop an evaluation system that would estimate DF in a more timely manner. It was envisioned that the system could take one of several forms, including a predictive regression equation(s) or a system of threshold limits.

In regard to the regression equation(s), it was initially proposed that the terms of the predictive equation would include both aggregate and concrete mix characteristics. The aggregate characteristics would have reflected all the major variables that affect freeze-thaw durability of aggregate: water escape path length, pore size and distribution, strength and elastic accommodation, potentially available water for freezing and ordering of water molecules, and availability of clay minerals. The reason that any single test has traditionally failed to correlate well with concrete durability is that all major mechanisms of freeze-thaw attack must be represented, and single test methods do not do that. These mechanisms could be represented by several, but certainly not all, of the following test methods: maximum aggregate size [smallest sieve through which 100% passes] (MAS) or nominal maximum aggregate size [largest sieve that retains an appreciable amount of material](NMS), Iowa pore index (IPI), vacuum saturated bulk specific gravity and absorption (VSBSG/VSAbs) or AASHTO T85 bulk specific gravity/absorption (BSG/Abs), sodium sulfate soundness ($NaSO_4$), water-alcohol freeze-thaw soundness (WAFT), methylene blue (MB), Micro-Deval (MD), wet ball mill (WBM), Los Angeles Abrasion (LAA), point load strength (PLS), and aggregate crushing value (ACV). The concrete mix characteristics would have included air system quality and paste strength.

Alternatively, it was also proposed that an equation could be developed to predict T 161 results based strictly on coarse aggregate qualities, ignoring paste contributions to durability. The predicted durability would be termed the "Aggregate Durability Factor" (ADF). A possible form of the equation could resemble the following:

$$ADF = a_1 \cdot (MAS \text{ or } NMS) + a_2 \cdot (IPI \text{ or } VSBSG/VSAbs \text{ or } BSG/Abs \text{ or } WAFT \text{ or } NaSO_4) + a_3 \cdot (ACV \text{ or } PLS \text{ or } LAA \text{ or } WBM \text{ or } MD \text{ or } NaSO_4) + a_4 \cdot (MB \text{ or } IPI) \quad (1)$$

Where a_i = regression coefficients

Ultimately, the general form of the equation might look like the following, which would reflect both the aggregate and the concrete paste contribution to durability:

$$DF = a_1 \cdot (ADF) + a_2 \cdot (\%\ air) + a_3 \cdot (air\ spacing\ factor) + a_4 \cdot (w/c \text{ or } strength) \quad (2)$$

However, after discussions with MoDOT personnel, it was decided to narrow the scope of the work to include only aggregate characteristics, thus only Eq. 1 was to be developed.

Fortunately, toward the end of the project, MoDOT was able to supply some related strength data that was successfully brought into the results, so the final report became scaled up to a hybrid somewhere between a strictly aggregate-based approximation system and a total mix-related one.

FIELD USE

The DF equation will be established in this study. Subsequently, during construction, the aggregate characteristics could be pre-determined from project records or from Quality Control (QC) results of aggregate at the concrete plant. Thus, within a short period of time (depending on the test methods that are included) of sampling, the DF would be calculated (and thus the T 161 DF could be approximated), and if trouble was indicated, corrective action could be undertaken instead of waiting for an extended period of time.

RESEARCH PROJECT CONCRETE TESTING

The Durability Factors, necessary for the left hand side of either equation (Eqs.1 and 2), were to come from MoDOT T 161 concrete data. Several options for producing T161 results were initially proposed by Missouri S&T, termed ***Option A*** or ***Option B***, or some combination of the two.

Option A would have involved MoDOT laboratory personnel casting specimens and performing T 161 tests on a variety of mixes that would represent the major factors controlling freeze-thaw durability of concrete: aggregate type (quality), MAS, paste quality (strength, *w/c*), and air void system quality (e.g. Air Void Analyzer(AVA)). Option A was not chosen by MoDOT for a variety of reasons.

The major disadvantage of Option A was that MoDOT would have to undertake a special testing program which would involve making and testing numerous laboratory mixes. Secondly, MoDOT was experiencing problems with its AVA and thus may have had problems controlling/measuring the air void system of fresh concrete.

The advantages of Option A would have been that the resulting DF predictive regression equation would reflect not only the contribution of aggregate durability, but also the other variables that affect concrete durability. Thus, in the field, the equation could spot a situation where the aggregate may be acceptable, but for a given day's pavement placement, the *w/c* may be a bit high, the air content a little low, and the air void system a little marginal due to interactions with the other mix constituents, thus the concrete durability may actually be unacceptable, or result in a low pay factor.

Option B was chosen by MoDOT. ***Option B*** entailed no additional concrete testing by MoDOT, with just a reliance on existing DF data. Thus the predictive DF equation will be similar to Eq. 1. The advantage is, of course, that no further concrete testing by MoDOT was required. The disadvantage is that the existing DF data will reflect essentially no ranges of air content, air quality, and *w/c*, because these variables have been purposely held constant in past MoDOT T 161 testing. Thus the predictive equation will be much more limited in accuracy and applicability.

RESEARCH PROJECT AGGREGATE TESTING

Missouri S&T was to perform aggregate testing in the as-delivered condition (e.g. gradation) for a variety of aggregate ledges. The specific ledges (subsets of specific geologic formations within a specific quarry) were chosen by MoDOT to reflect a range in D-cracking (freeze/thaw) susceptibility and NMS. The D-cracking susceptibility would be already known from MoDOT's previous T 161 data (and any service experience available). Ideally, for each ledge sample, there would be data available for three gradations based on varying MAS.

The proposed matrix of testing included the following: three levels of MAS (NMS), three levels of quality (DF), and at least two different ledges per DF level, for a total of 18 ledges. Each of the 18 aggregates was to be subjected to a battery of aggregate tests, and the results used to produce the ADF equation (Eq. 1) for use in the field.

MoDOT CONTRIBUTION

MoDOT was to choose the aggregate ledges, supply the corresponding test results (e.g. T 161 and other mix-related data), and supply samples of aggregate. Aggregate test data was also to be supplied for verification that the aggregate tested at Missouri S&T in this study matched the aggregate that was used in the T 161 test specimens. It was recommended that aggregate samples be tied as closely as possible to the actual aggregate used in the T 161 testing that produced the data that were to be used in this study. The aggregate was to be production stone. Additionally, for gradations D and F, larger pieces were required in addition to the regular samples because the Point Load test method requires individual aggregate specimens that are at least 1 in. in diameter. MoDOT personnel were to sample and blend the replicate bags of material.

POTENTIAL PROBLEMS

It was recognized that there might be several potential problems. First, there might not be available data for all the possible combinations of DF and MAS (NMS) that were required. Second, there might not be aggregate samples available that could be tied directly to the aggregate that was used in past DF testing, thus rendering the estimation equation(s) less accurate. In other words, if the aggregates used in this study did not come from the same samples as the aggregates that went into the T 161 beams, then the amount of error in the regression equations will most likely increase. Third, because the final prediction system may include test methods for which MoDOT does not currently have data, then no verification of the prediction model can occur. Verification (and possible model adjustment) would have to come after implementation of the new test methods.

OBJECTIVES

The objective of this study is to establish an evaluation system of concrete freeze-thaw durability based on concrete mix properties (specifically, the coarse aggregate) which would correlate with the results of AASHTO T 161. T 161 test results are expressed in terms of the Durability Factor (DF). The evaluation system could take one of two forms: 1) a statistical regression-type relationship between the concrete Durability Factor and various quickly-determined aggregate characteristics, or 2) a threshold-limits system of pass-fail criteria related to specific test results.

The regression product of this study would be a simple equation (say, that was entered into a spreadsheet) into which the results of quickly-determined tests would be entered. The resulting estimated DF could be used to make adjustments in the construction process in a much-reduced time frame. The form of the relationship would resemble:

$$DF = a_1x_1 + a_2\,x_2 + \ldots\ldots. a_nx_n$$

Where a_n= regression constants determined in this study
x_n= test results

A second product of the study would be a simple listing of threshold limits for various test results which, taken together, would predict if a given aggregate could be expected, in most cases, to result in a concrete mixture's T 161 DF to exceed 75, providing that the paste portion was frost-resistant.

LITERATURE REVIEW

FREEZE-THAW DAMAGE THEORIES

A variety of theories have been put forth to explain freeze-thaw damage of concrete: Powers' hydraulic pressure theory (Powers, 1945, 1955; Verbeck & Landgren, 1960); Powers and Helmuth's osmotic pressure theory (Powers & Helmuth, 1953; Powers, 1975); Litvan's relative humidity water movement theory (Litvan, 1972, 1973, 1975, 1978); Dunn and Hudec's adsorbed water theory (Dunn & Hudec, 1965, 1966, 1972); and Larson and Cady's dual mechanism theory (freezing and adsorption of water) (Larson & Cady, 1969). Generally speaking, pressure is created within the cement paste pores and aggregate pores, causing rupture.

ROLE OF AGGREGATE IN FREEZE-THAW DAMAGE

Freeze-thaw damage of concrete falls into three categories: paste failure, aggregate failure, and paste-aggregate interface failure. Aggregate characteristics affect the latter two categories, and are the subject of this study.

AASHTO T 161

AASHTO T 161 ”Resistance of Concrete to Rapid Freezing and Thawing” (AASHTO, 2000) has been shown to be a good predictor of frost susceptibility of concrete, and in general, correlates well with service records (Thompson et al., 1980; Klieger et al., 1974; Chamberlain, 1981). The Durability Factor is one of the parameters calculated from T 161 that is used to quantify or predict frost susceptibility of concrete. If the paste is designed to be frost-resistant via a good air void system and sufficient strength, then T 161 becomes a tool for assessing frost-susceptibility of aggregate. It must be remembered that the conditions of the test method dictate the level of frost damage. Whether the specimen is frozen in air or water affects the results. Another major factor is the fact that T 161 is considered to subject the specimen to a relatively rapid rate of freezing and thawing. When attempting to predict the frost susceptibility of aggregate, the rate will affect the relationship between aggregate test methods results and DF. Rapid freezing is thought to cause the water to freeze from the outside, progressing toward the interior, and trapping water, thus not allowing water to escape. This trapped water then is available for causing problems, such as transmission of hydraulic pressure. The rapid rate may not be realistic compared to actual field conditions, thus, T 161 may over predict frost damage (Kaneuji, 1978). Another reason for over prediction is that the stresses induced by rapid freeze-thaw testing may be unrealistically high, and not actually seen in the field (Powers, 1949). T 161 may be a better test for acceptance rather than rejection (Chamberlain, 1981).

The method used by MoDOT is T 161 Method B, which calls for both freezing and thawing in water. Had some other method been used, the prediction of DF would probably show a different result in the relative importance of different aggregate test methods.

SYSTEM ESTIMATION OF AGGREGATE DURABILITY

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; Long, 1987; Richardson & Long, 1987; Wiles, 1988; Richardson & Wiles, 1990.

In a limited way, concrete durability has also been approximated by regressing aggregate pore characteristics with several measures of freeze-thaw durability, such as the T 161 Method B Durability Factor or dilation, and the Virginia Polytechnic Institute (VPI) slow–freeze test. Only four aggregate pore-type test methods were used, omitting other damage mechanism test representatives (Koubaa & Snyder, 2001).

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.

There seems to be an interaction between rate of freeze-thaw and pore size distribution. Aggregates with large pores exhibit lower durability in rapid freeze-thaw tests (Kaneuji, 1978). The water in the larger pores freezes first. In a rapid test, the unfrozen water is trapped, causing problems. Test methods that measure large pore volumes should correlate with DF from T 161.

There seems to be a relationship between pore size and frost susceptibility. In general, for damage to occur, pores need to be large enough to permit entry of water into a large fraction of the pore space, but small enough to limit rapid exit Lewis et al., 1953). Aggregates with a high frequency of small pores are prone to damage (Hiltrop & Lemish, 1960; Domaschuk & Garychuk, 1988; Winslow et al., 1982). Pore size alone does not correlate well, but becomes more important when used in conjunction with mineralogy (Hiltrop & Lemish, 1960). Aggregates with a wide range in pore sizes tend to be resistant (Winslow et al., 1982). Pore sizes of 10 µm to 0.1 µm seem especially prone to problems, while aggregates with very small pore diameters of less than 45 angstroms are resistant, possibly because the pores are too small to allow water to freeze (Kaneuji, 1978). Apparently, water in pores of the 10 µm to 0.1 µm range has difficulty in escaping. Conversely, for large pores (greater than 1 μm), the larger the pore, the less sensitive to damage. Above 10 μm, aggregates are frost resistant. The assumption is that the large size allows water to escape, thus pressure cannot become excessive. In another study, average pore diameters in the 0.02 to 0.04 μm diameter range were shown to usually be non-durable (Marks & Dubberke, 1982). However, the occurrence of exceptions point out that the use of a calculated average may not adequately describe the actual size distribution, that is, an aggregate may have a large amount of both small and large sizes which would lead to a calculated average that may not in reality exist. Small and large pore sizes have been shown to be frost resistant, and thus could explain why the middle-size calculated average distribution would inaccurately predict frost susceptibility.

Pore shape also is important to frost susceptibility. Pores that have small openings leading to an enlarged void would allow a relatively large volume of freezable water to enter, but would provide a restricted avenue of egress, thus allowing pressure to build. The small pore openings allow easy saturation via capillary action. These pore shapes have been termed “ink-bottle” pores. Pores with ink bottle shapes are harmful, but also give misleading results when attempting to quantify pore diameter—the frequency of small sizes would be inflated. This type of pore is considered to be the most harmful.

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). Aggregates with low absorption (less than 0.3%) frequently show acceptable resistance to frost damage. 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 T 161 Method B. Absorption values less than 1.5% indicated DFs greater than 80, while absorptions greater than 2% were associated with DFs less than 60 (Koubaa & Snyder, 1996).

Bulk Specific Gravity

Bulk specific gravity, also determined in AASHTO T 85, is a function of mineralogy (specific gravity of the solids) and porosity. 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 of greater than 2.60 or 2.65 exhibited superior durability and had a good correlation with DF (Koubaa & Snyder, 2001; Harman et al., 1970). However, some aggregates with very low specific gravities (2.24-2.35) and large absorptions were 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.

Vacuum Saturated Absorption

Subjecting aggregate to vacuum will increase the amount of absorption 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 & Cady, 1969). Others have shown that vacuum saturated absorptions of greater than 2% 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 elastic accommodation tests (LAA, MD, ACV) than with soundness tests. Of the three, MD correlated best with VSAbs (Williamson et al., 2007).

VSAbs has also been put forth as a primary screening test for aggregate durability (Williamson et al., 2007).

In general, aggregates with intermediate values of absorption or vacuum saturated absorption (1.5 to 2.5%) 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.

Iowa Pore Index

The Iowa Pore Index (IPI) test was developed to provide a simple test method that would identify aggregates that are prone to D-cracking because of their susceptibility to critical saturation. In essence, the aggregate is subjected to water pressure and the amount of water taken in (called the “secondary load”) after a short initial period is considered a measure of the absorptiveness of the ink-bottle or smaller size pores. Large pores are ignored. The results have been shown to correlate well with both T 161 Method B and service records for IPI values less than 20 and greater than 35 (Marks & Dubberke, 1980). Better results are obtained when the aggregate sample is geologically homogeneous and pore sizes are in the 0.04 to 0.1 μm range (Marks & Dubberke, 1982). IPI for carbonate aggregates correlates well with durability factors (Koubaa & Snyder, 1996, 2001). Others have shown that the IPI method suffers when attempting to test gravels (Traylor, 1979) and does not identify shale as being non-durable (Koubaa & Snyder, 2001). IPI has been shown to correlate well with mercury porosimetry results, but without the problems associated with that method (Shakoor & Scholer, 1985). One study indicated the possibility that for a timed test like the IPI, a different size aggregate may yield different results, with slower absorption rates associated with larger aggregates. However, the actual IPI test method was not performed and pressure was not applied, so the results are somewhat conjectural (Winslow, 1987). Past Iowa DOT test method 219-D and specifications are based on an IPI threshold value of 27: aggregates in excess of this are considered non-durable. Current specifications contain several thresholds of 20, 25, and 30 depending on the class of concrete (Iowa DOT, 2000, 2007).

Water-Alcohol Freeze-Thaw and Sulfate Soundness

Both water-alcohol 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.

Pore Length

As water freezes, there is a volume expansion. If the unfrozen water cannot escape easily, pressure is transmitted from the ice to the unfrozen water. If the aggregate size is large enough, the magnitude of the pressure will exceed the aggregate’s tensile strength, and rupture could occur. The maximum size that the particle can be without damage is termed the “critical size” (Powers, 1955). Other factors important to frost resistance besides tensile strength include freezing rate and aggregate permeability (Verbeck & Landgren, 1960). Thus NMS of the aggregate would be important, and crushing the aggregate to a size smaller than its NMS should increase its resistance to freeze-thaw damage. This idea has been supported numerous times in the literature in regard to service records (Stark, 1976; PCA/Stark, 1976; Stark & Klieger, 1973; Stark, 1976; Kliege et al., 1978).

NMS As-Tested T 161

The simplest measure of maximum pore length is the NMS of the aggregate in the mixture. There are a variety of definitions of NMS: perhaps the simplest is the largest sieve that retains at least 5% material.

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 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 an increase in pore pressure in pores in those directions (Verbeck & Landgren, 1960).

Although reports have identified failure as a function of the stress exceeding the tensile strength (Powers, 1955; Verbeck & Landgren, 1960), attempts to quantify aggregate tensile strength in relation to aggregate durability has not been done. 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 stress result also indicate low elastic accommodation behavior, or not?

Tests that utilize aggregate in an unconfined state do not consider the effect of confinement by the concrete paste.

Aggregate Crushing Value

The Aggregate Crushing Value (ACV) test method (British Standards Institution BS 812: Part 110) consists of subjecting a compacted specimen of aggregate particles to a static load, and then measuring the amount of breakdown (BSI, 1990). The aggregate particles bear on each other and are subjected to point contact loads and thus to an indirect tensile load, as well as some 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 (LAA) results (BSI,1998; Saeed et al., 2000; Williamson et al., 2007; Kandahl & Parker, 1998).

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 test (AASHTO T 327) (MD) 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. 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., 1998, 1998; Kandahl & Parker, 1998). Several studies have noted a lack of correlation of MD with other toughness tests, such as LAA and ACV (Kandahl & Parker, 1998; Saeed et al., 2000; Wu et al., 1998, 1998).

Point Load Strength

The Point Load Index test (ASTM D 5731-07) is a measure of indirect tensile strength. It is similar to such indirect tensile strength tests as the line-load loading scenarios for AASHTO T 283 (AASHTO, 2003) and for the so-called Brazilian test method for rock cores and concrete cylinders. Instead of a line load, the load is applied as a point (Bieniawski, 1975). Major advantages of the method include the ability to test irregular lumps, a small load frame requirement, quickness of testing, and thus the potential for testing a larger number of specimens. Specimen size affects the outcome, so the results need to be converted via a standard equivalent size (50 mm). Strength decreases as specimen size increases (Hardin, 1985; McDowell & Bolton, 1998; Lade et al., 1996; Richardson, 1989). 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 & Vallejo, 2006).

Water-Alcohol Freeze Thaw

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.

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., 1998,1998), 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 test (Rogers, 1989). Also, it has been shown to correlate with durability better than sulfate soundness (Brink, 1958).

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.

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, again, sulfate soundness could be considered a measure of tensile strength or elastic accommodation.

Sulfate soundness has not been shown to be an accurate predictor of frost susceptibility, either from slow cooling testing or service records. Several reasons for this include the difference in destructive mechanism and the lack of precision of the methods (Swenson & Chaly, 1956; Marks & Dubberke, 1982; Harman et al., 1970; Cady, 1984).The method also does not correlate well with WAFT (Brink, 1958). Other studies have reported mixed success in prediction (Paxton, 1982; Chamberlain, 1981). Also, magnesium and sodium sulfate methods do not necessarily agree. In general, sulfate soundness prediction of freeze-thaw durability is mixed, and the method suffers from imprecision.

Soundness has been shown to correlate better with MD than LAA does with MD (Cuelho et al., 2007).

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 (Texas DOT, 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. 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 and unbound highway base material. Various designations include Mill Abrasion (Clifton et al., 1987; Clifton et al., 1987; Selig & Boucher, 1990; Union, 2001) and Texas Wet Ball Mill (Texas DOT, 2000). A good correlation has been found between MD and WBM (Jayawickrama et al., 2001).

MINERALOGY and PORE WATER CHEMISTRY

The importance of the mineralogy of the aggregate has been shown to be influenced sometimes by the pore water chemistry (Dubberke, 1983; Bisque &

Lemish, 1958). For carbonate-type aggregates, mineralogy includes the type and amount of cations (calcium and magnesium), the form and amount of silica, type and amount of clay minerals, and associated elements such as strontium, sulfur, and manganese. Pore water chemistry involves, among other things, the presence and type of de-icing salts.

Magnesium Content

Somewhat conflicting information about the role of magnesium content is in the literature. Some studies have shown that calcium-to-magnesium ratios of less than 9.0 showed poor service records, indicating that dolomites should be susceptible to damage (Hiltrop & Lemish, 1960).

Role of Silica

Aggregates with a low insoluble residue content (less than 1.6%) have evidenced good service records (Hiltrop & Lemish, 1960). Greater levels of silica seemed to cloud the ability of the IPI to accurately predict DF, with the threshold being somewhere between 2 to 3%. Additionally, medium and small silica grain size is associated with poor service records (Dubberke, 1983). It is speculated that smaller grain size affords an increase in surface area with which the salt can react. Trypolitic chert in carbonate aggregate has caused aggregate to disintegrate while undergoing T 161 testing (Dubberke, 1983).

Clay Minerals

Certainly, deleterious clay has been shown to be detrimental to concrete (Buth et al., 1964, 1967). However, clay content herein is defined as clay that exists in the aggregate particles themselves, as opposed to free clay present as a deleterious material. Again, results are conflicting. Illite has been shown to be detrimental (Hiltrop & Lemish, 1960). However, other studies have shown the opposite: more illite, better durability. However, it was noted that as the illitic clay content increased, the amount of microscopic silica decreased (which has been shown to be detrimental to durability). Thus, the reason for the decline in freeze-thaw durability was unclear. It was also noted that the illite interfered with IPI results, causing an inverse relationship between IPI and durability, whereas illite content had a positive effect on T 161 results (Dubberke, 1983). It has been postulated that water adsorption by illite within aggregate has caused failure in certain aggregates, even though the water itself did not freeze (Dunn & Hudec, 1966, 1972).

In a related matter, free clay in the form of dust or aggregate surface coating, lowers durability. The greater the liquid limit, the more pronounced is the effect.

Deicing Salts

The detrimental effects of deicing salt (chloride-based) are well documented in regard to cement paste. However, in some cases, certain deicing salts have also been shown to interact negatively with certain aggregates, such as carbonates containing cryptocrystalline chert. Sodium chloride substantially lowered frost resistance of some aggregates, but not others (Dubberke, 1983), both in DF testing and in the field. Furthermore, increased levels of sulfur and manganese in dolomites led to increased susceptibility to deicer attack of the aggregate, possibly due to unstable or impure crystal formation within the aggregate leading to breakdown (Dubberke, 1983).

In limestones, increased levels of strontium have led to increased breakdown in the presence of salt (Iowa DOT, 2000).

Methylene Blue

There are several methods for estimation of clay content and clay type. One of the simplest is AASHTO T 330, the methylene blue test (AASHTO, 2007). The test is really a measure of the cation exchange capacity of the material, and is an indication of surface activity. Thus, MB can be used to estimate the amount of harmful clays present. There is evidence that the MB test can be used to assess strength reductions in concrete due to the presence of various clay types (Yool et al., 1998).

Iowa Pore Index

In response to the effects of mineralogy and pore water chemistry in conjunction with salt susceptibility on the results of the IPI test, guidelines have been developed to augment the IPI method (Iowa DOT, 2000). In essence, a Pore Index Quality Number can be calculated from IPI measurements which will indicate on a scale of zero (good) to ten (poor) how susceptible an aggregate is to chemical reactions in the pore fluid.

Water Alcohol Freeze-Thaw

The Water-Alcohol Freeze-Thaw method has been presented earlier.

AMOUNT OF D-CRACKING-PRONE AGGREGATE IN THE MIX

The extent of the content of large particles has been shown to be important as well as the NMS (Klieger et al., 1978). Blended amounts as low as 15% of poorer quality material were sufficient to significantly lower DF (Marks & Dubberke, 1982). Numerous reports have recognized the importance of the particle size and/or amount of large particles to frost susceptibility (Mindness et al., 2003; Domaschuk & Garychuk, 1988; Pigeon & Pleau, 1995; PCA, 2008).

NMS Percent Volume of Plus ¾ in. Material

The percent by volume of the concrete aggregate that is of a size large enough to be a problem (greater than the critical size) relates to both the concepts of critical size and to the amount of material prone to damage. Traditionally, a common NMS for concrete aggregate is ¾ in., and mixtures that exhibit freeze-thaw damage at this NMS are considered problematic. Thus, the amount of coarse aggregate above ¾ in. would be an indicator of potential durability problems.

Hudson's Ā

The aggregate gradation modulus Hudson's Ā has been used by researchers to characterize the behavior of materials in a variety of settings (Hudson & Waller, 1969; Richardson, 1984; Richardson & Long, 1987; Richardson et al., 1996; Lusher, 2004). This is a value calculated from gradation results, similar to the fineness modulus, except it is the sum of the percent passing of a series of sieves (1 ½, ¾, ⅜ in., #4, #8, #16, #30, #50, #100, and #200) divided by 100.

Index of Crushing

The "Index of Crushing" (IC) is a method to quantify the break down of a granular material as a result of slaking (Aughenbaugh, 1962). The method entails: 1) determination of the mean sieve size for each fraction of interest, 2) calculating the product of the mean sieve size and the individual percent retained for each fraction, and 3) summing the products. The IC is the difference between the before-slaking sum and the post-slaking sum, divided by the initial, expressed as a percent.

Ratio2

Simple gradation indices suffer from the fact that the calculated value could represent a number of different gradations. Ratio2 is a gradation index that was developed to assist in tying a given gradation to a more unique number (Lusher, 2004). The calculations are somewhat rigorous, and are presented in Appendix E.

DELETERIOUS MATERIALS

Deleterious materials are defined as materials that are extraneous to the parent material. Examples are shale, clay balls, soft rock, chert, and anything that would fall under the category of lightweight pieces. The literature has numerous references to the negative action of various deleterious materials (Bloem, 1966). It has been shown that small amounts of deleterious material can result in low DF values even for aggregates with good field performance (Marks & Dubberke, 1982). Presence of chert can cloud conclusions based on the effect of other variables that are being studied (Cramer & Carpenter, 1999). MoDOT has a deleterious materials method as a standard specification, TM-71 (MoDOT, 2006).

TECHNICAL APPROACH

GENERAL

Experimental Design

In Table 1 is shown the proposed testing matrix: three levels of MAS, three levels of quality (DF), and at least two different ledges per DF level, for a total of 18 ledges. Each of the 18 aggregates was to be subjected to a battery of aggregate tests, and the results used to produce the DF equation (Eq. 1) for use in the field.

MAS is usually defined as the smallest sieve though which 100% passes. Nominal maximum size (NMS) is typically understood to be the largest sieve where at least 5 % is retained.

Table 1: Experimental testing plan

MAS	DF	Ledge
1 ½”	<70	1
		2
	70-80	3
		4
	>80	5
		6
1”	<70	7
		8
	70-80	9
		10
	>80	11
		12
½”	<70	13
		14
	70-80	15
		16
	>80	17
		18

Replicate Specimens

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

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 Lab personnel delivered the bagged samples to the Missouri S&T CArE aggregate laboratory. The actual materials delivered are shown in Table 2. Materials were chosen to represent MoDOT's former gradation types B, D, and F. Unfortunately, none of the materials meeting the B gradation actually had a significant amount of plus 1 in. material. It should be understood that the MAS is defined as the MAS used in the T 161 concrete mixture batches. It is not the as-delivered state nor is it as-utilized in the field. For instance, 85RDP040 was supposed to fulfill an as-utilized 1½ in. MAS material, when in actuality it was tested in T 161 as a 1 in. MAS material, so it is shown in Table 2 as a 1 in. (¾ in. NMS) material. Because the regression equations will reflect actual freeze-thaw behavior as-tested, then the NMS identified with any given aggregate should reflect the actual T 161 as-tested size. The table shows a lack of larger MAS materials, even though there were several materials that essentially met gradation B requirements. As mentioned previously, NMS is defined herein as the largest sieve to retain at least 5%.

Table 2: Aggregate materials

MAS	DF	Gradation	County	Formation	Study ID
1½ in. (1 in. NMS)	<70				
	<70				
	70-80				
	70-80				
	>80				
	>80				
1 in. (¾ in. NMS)	54	B	Osage	Jefferson City Dolomite	85RDP040
	64	B	St. Charles	Plattin	86L2R020
	66	D	Pettis	Burlington	85DGG007
	69	D	St. Louis	St. Louis-Salem	86R3M028
	72	B, D	Knox	Chouteau	83MA0234
	73	B	Greene	Burlington	88MA0024
	75	D	Moniteau	Burlington-Choteau	85DLR012
	78	D	Cape Girardeau	Plattin	80MA0051
	87	B, D	Howell	Jefferson City Dolomite	89TCR067
	89	D	St. Charles	Plattin	86R3M029
	95	B	Alton, IL	St. Louis	86R3M025
¾ in. (½ in. NMS)	28	D	Cass	Bethany Falls	84SRE203
	72	D	Andrew	Kereford	81MA0379
	76	E	Cass	Bethany Falls	84SRE039
	81	D	Jasper	Warsaw	87ASM006
	94	E	Jefferson	Plattin	86R3M031
½ in. (⅜ in. NMS)	<70				
	<70				
	70-80				
	70-80				
	89	F	St. Louis	St. Louis	86L2R021
	94	F	Andrew	Amazonia	81MA0292
	95	E	St. Charles	Plattin	86L2R034

The geologic types were limited to 10 formations: nine of limestone and one of dolomite. The number of MoDOT-defined ledges was 18. One of the ledges was tested at two different gradations (86L2R020 and 86R3M029). Thus the number of total data sets was 19.

Typically, material was delivered in two forms: production stone (material completely processed, ready for use for incorporation into concrete mixes) 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 NMS of the production stone.

Typically, about 10 bags of production stone were delivered to the CArE aggregate laboratory. This material was then mixed and rebagged, following a modified version of MoDOT TM-67 windrowing method (MoDOT, 2000). The material was then tested for the as-delivered gradation. Subsequently, the material was mechanically shaken through sieves for 5 to 10 minutes 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 and the gradations used in the T 161 prisms are compared in Tables 3 to 9. As can be seen, the NMS did not agree for four of the aggregates. Two were coarser and two were finer. The gradation designations B, D, and F are no longer in use in the current MoDOT specification for paving concrete. Also, some of the gradations labeled as "F" actually would conform better to MoDOT's "E" gradation.

Table 3: B Gradation Percent Passing

Formation	**Jeff. City Dolomite**		**Plattin**		**Burlington**	
ID	**85RDP040**		**86L2R020**		**88MA0024**	
Sieve	**T 161 Tested**	**As-Delivered**	**T 161 Tested**	**As-Delivered**	**T 161 Tested**	**As-Delivered**
1 ¼ in.	100	96	100	100	100	100
1	100	94	98	96	100	100
¾	93	89	83	76	71	88
½	33	33	40	48	22	46
3/8	13	9	19	29	11	24
#4	3	2	2	7	1	3

Table 4: B Gradation Percent Passing

Formation	St. Louis		Chouteau		Jeff. City Dolomite	
ID	86R3M025		83MA0234		89TCR067	
Sieve	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered
1 ¼ in.	100	100	100	100	100	100
1	100	100	100	100	100	100
¾	85	94	88	88	92	90
½	38	59	50	49	42	33
3/8	18	31	26	22	19	13
#4	3	5	10	4	2	3

Table 5: D Gradation Percent Passing

Formation	Plattin		St. Louis-Salem		Burlington	
ID	80MA0051		86R3M028		85DGG007	
Sieve	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered
1 ¼ in.	100	100	100	100	100	100
1	100	100	100	100	100	100
¾	93	93	90	94	93	90
½	53	35	61	52		54
3/8	13	13	48	27	21	34
#4	2	2	22	3	2	7

Table 6: D Gradation Percent Passing

Formation	Burlington-Chouteau		Plattin		Warsaw	
ID	85DLR012		86R3M029		87ASM006	
Sieve	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered
1 ¼ in.	100	100	100	100	100	100
1	100	100	100	100	100	100
¾	92	89	95	95	100	100
½	33	45	64	53	54	70
3/8	17	23	45	25	20	32
#4	3	3	11	5	1	4

Table 7: D Gradation Percent Passing

Formation	Kereford		Bethany Falls			
ID	81MA0379		84SRE203			
Sieve	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered
1 ¼ in.	100	100	100	100		
1	100	100	100	100		
¾	100	100	100	100		
½	61	88	47	100		
3/8	36	61	16	40		
#4	6	12	1.7	5.7		

Table 8: E Gradation Percent Passing

Formation	Bethany Falls		Plattin		Plattin	
ID	84SRE039		86L2R034		86R3M031	
Sieve	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered
1 ¼ in.	100	100	100	100	100	100
1	100	100	100	100	100	100
¾	100	100	100	100	100	100
½	81	64		97	92	100
3/8	46	32	55	53	47	100
#4	2	3	8	7	3	34

Table 9: F Gradation Percent Passing

Formation	St. Louis		Amazonia			
ID	86L2R021		81MA0292			
Sieve	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered	T 161 Tested	As-Delivered
1 ¼ in.	100	100	100	100		
1	100	100	100	100		
¾	100	100	100	100		
½	96	95	100	92		
3/8	81	70	91	45		
#4	33	20	7	2		

MoDOT DATA

Data associated with each of the 18 ledges was furnished by MoDOT. Five different types of reports were shared: 1) Quarry Ledge Information Summaries, 2) Work Cards, 3) T 161 Evaluation Reports, 4) Concrete Batching Program spreadsheets, and 5) Freeze & Thaw Ledgers. Information from each was useful for obtaining the overall picture of an aggregate's characteristics. Specific information was used in the correlation and regression studies reported later in this report. MoDOT aggregate test results for LAA, MD, $NaSO_4$, WAFT, T 85 BSG and Absorption, and T 161 as-tested NMS as well as DF, pre-test and post-test flexural strength-modulus of rupture (MR) and pre-test compressive strength, sample location, ledge number, and formation were obtained from the Evaluation Reports. Mix volumetric information, w/c, air content, NMS, and T 85 BSG and Absorption were obtained from the batch reports. The Freeze-Thaw reports were useful for beam size for flexural strength calculation as well as verification of DF. Aggregate gradation information was derived from a combination of Work Cards, Evaluation Reports, and Batch spreadsheets to obtain the most accurate estimate of the gradation that was used in the T 161 beams.

In addition to standard reports, two other types of data sets were shared: 1) IPI data and 2) a T 161 Study Tabulation. The T 161 Study Tabulation was useful for fleshing out missing information, and both were used for correlation with Missouri S&T results for verification that delivered samples were representative of the T 161 material.

TEST PROCEDURES and EQUIPMENT

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

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. MoDOT has in the past 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.

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. In essence, ungraded washed and oven-dried material (plus #4 sieve) is subjected to a vacuum of 27.5 ± 2.5 mm mercury absolute pressure for 5 minutes. Water is introduced under vacuum and eventually submerges the aggregate. The specimen is then subjected to agitation for a total of 30 minutes under vacuum (including the initial 5 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 A. Fig. 1 depicts the Missouri S&T vacuum saturation station.

Figure 1: Vacuum Saturation Workstation

Iowa Pore Index

The Iowa Pore index procedure used herein followed the Iowa 219-D procedure. The test apparatus modeled the Iowa DOT (IDOT) version as closely as possible. The Missouri S&T device was a copy of the same piece of equipment used by MoDOT, which was on-loan from the Iowa DOT. About a third of the tests were performed with the Iowa control box, but with Missouri S&T valve/cylinder panel and specimen pressure pot. In the other two-thirds of the tests, all three items were of Missouri S&T fabrication. Fig. 2 depicts the Missouri S&T device.

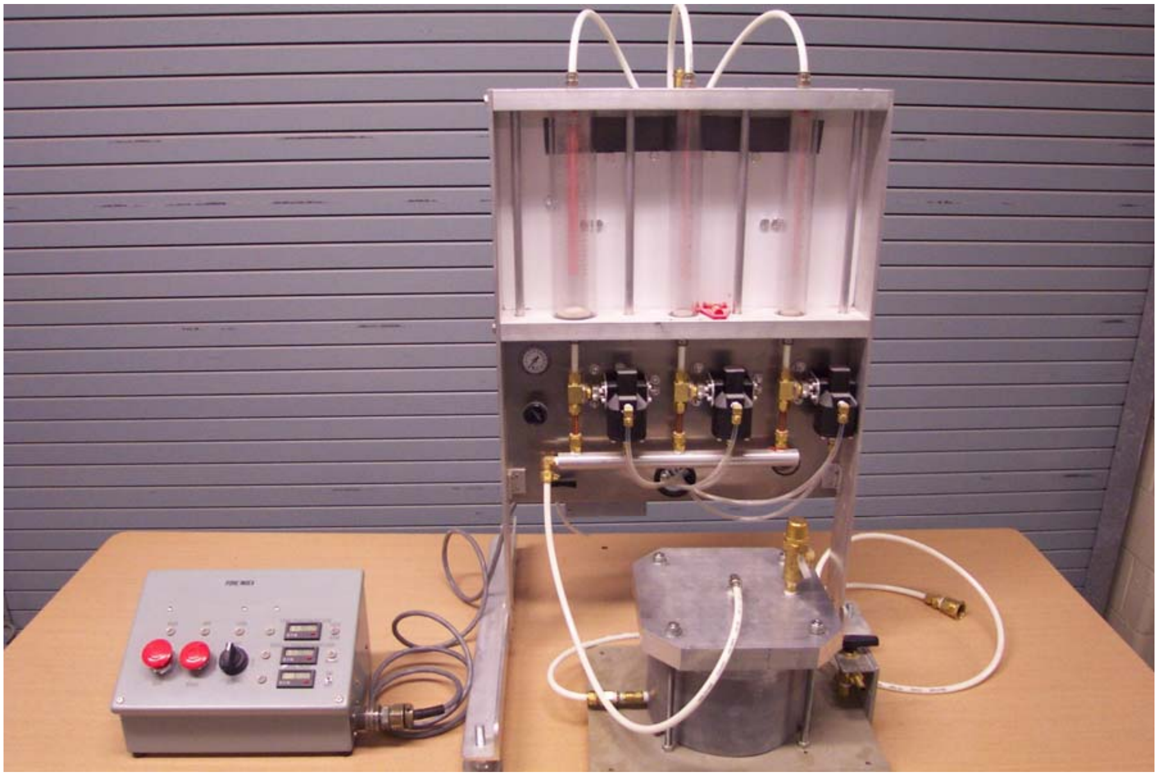


Figure 2: Missouri S&T Iowa Pore Index Device

In essence, 4500 g of oven-dried material is subjected to water under pressure. The volume of water taken in during the first minute is called the primary load, while the water taken in during the subsequent 14 minutes is termed the secondary load. The secondary load, adjusted for specimen size, is the “Iowa Pore Index”.

One modification that was used in this study involved finer gradations. The Iowa method calls for aggregate to be a ½ in. retained to ¾ in. passing size. For MoDOT’s Gradation F material, the maximum aggregate size (100% passing) is ½ in. and the next smaller size (85-100% passing) is ⅜ in. Thus, Gradation F materials would not have any material of the correct size, and could possibly have no material of the ⅜ in to ½ in. size. Thus, the decision was to use a #4 to ⅜ in. size for the Gradation F material. Unfortunately, MoDOT has phased out

Gradation F and a coarser material was delivered to Missouri S&T after testing commenced. Thus, all of the finer gradations that actually had some available material in the standard fraction were tested with that fraction. However, three of the delivered materials did not have sufficient material for the standard fraction, and were tested with the smaller fraction. After discussions with MoDOT and IDOT personnel, and a review of the literature, plus some side-by-side testing of different fractions, the effect of using a finer specimen fraction is still not clear.

Water-Alcohol Freeze Thaw

MoDOT’s TM-14 (modified from AASHTO T 103-07, Method B) was followed, except for specimen gradation in some cases. According to TM-14, the initial specimen gradation was supposed to be built to a standard gradation, consisting of three fractions: #4 to ⅜ in., ⅜ to ½ in., and ½ to ¾ in. However, some as-delivered samples lacked the larger size(s), thus some specimens were built with one or even two of the coarser fractions missing. However, in these cases, the required total specimen weight of 2500 g was still utilized. After 16 cycles of freezing and thawing, the specimens were mechanically sieved for 5 minutes over a #8 sieve.

Freezing and thawing cycle durations were determined by use of thermocouples placed in specimens undergoing freezing and thawing cycles, under the expected specimen loading conditions in the freezer and the thawing tank.

Sodium Sulfate Soundness

The test methodology followed AASHTO T 104-03. 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 cycling was concluded, the specimens were flushed, dried, then hand-shaken over the appropriate sieve.

Pore Length

NMS As-Tested T 161

This determination was made by examining MoDOT’s Evaluation Record, concrete batch spreadsheet, and work card for each material. The as-tested NMS in the T 161 tests was termed NMSDF in this study.

NMS Percent Volume of Plus ¾ in Material

Using the above-determined as-tested (T 161) gradation, the absolute volume of the plus ¾ in. portion of the coarse aggregate portion, as shown on the concrete batch sheet, was calculated. This was termed the “Vol % Plus ¾ in.”

Hudson's Ā

This is a value calculated from gradation results, similar to the fineness modulus, except it is the sum of the percent passing of a series of sieves (1 ½, ¾, ⅜ in., #4, #8, #16, #30, #50, #100, and #200) divided by 100.

MIC

The Index of Crushing (a gradation index) has been discussed previously. The original index has been modified in this study by omitting the step of taking the difference between the "before slaking" and "after slaking" gradations. The new index is termed the "Modified Index of Crushing" (MIC). It is the sum of the products of mean screen size and individual percent retained for each fraction in a sieve analysis for a given gradation.

Ratio2

Ratio2 is another gradation index, previously discussed in another section. Details are in Appendix E.

Mineralogy

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 material (minus #12 sieve) from the completion of the LAA test was dry sieved over a #200 sieve. A slurry was made with the material, then titrated with methylene blue solution. The full procedure is reported in Appendix C.

Iowa Pore Index

This method was previously discussed above. Although a seemingly physical type of test, the IPI has been shown to be linked to certain mineralogical-related phenomena, hence its inclusion under the Mineralogy section of this study.

Water Alcohol Freeze-Thaw

This method was previously discussed above. 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 Accommodation/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 the sufficiency of strength to resist fracture, or to elastically accommodate the pressure.

Aggregate Crushing Value

The ACV is a direct-compression type of test which entails lightly compacting a graded sample into a heavy steel mold with a rod and subjecting the material to a hydraulically–applied compression load via a plunger. 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. 3 depicts the Missouri S&T compaction mold, plunger, and rod.

Figure 3: Missouri S&T ACV mold, rod, and plunger

The specimen is comprised of oven-dried material that passes a 0.52 in. (13.2 mm) sieve and is retained on a ⅜ in. (9.5 mm) sieve. 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 material is then dry-sieved over a #8 (2.36 mm) sieve and the loss is calculated as the ACV. The full procedure is reported in Appendix D.

Los Angeles Abrasion

AASHTO T 96-02 was followed. The initial specimen grading followed the recommendations of the method: both grading B and C were used, depending on the as-received gradation of the material. After the prescribed number of rotations, the material was sieved over a #12 sieve and the loss recorded.

Micro-Deval

AASHTO T 327-06 was followed for this part of the study. A Geneq, Inc. three-tiered model Micro-Deval device was used. The initial specimen grading followed the recommendations of the method: both gradings 8.2 and 8.4 were used, depending on the as-received gradation of the material. After the required rotations were achieved, the material was wet-sieved over a #16 sieve, dried, and the loss calculated.

Point Load Strength

ASTM D 5731-07 was followed with a few deviations. The method calls for testing 20 pieces of aggregate at least 30 mm in size. The specimens are in an oven-dried state. 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 corrected to a standard 50 mm size. The two greatest and two smallest values are discarded and the average PLS is calculated. In Fig. 4, the point load device is shown.


Figure 4: Point Load Device

Special large-size PLS samples were requested from MoDOT. Obtaining 1½ to 2 in. material that matched the production stone characteristics was 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 E.

Water-Alcohol Freeze Thaw

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

Wet Ball Mill

MoDOT Central laboratory supplied the MoDOT method for performing the Wet Ball Mill test utilized in this study, which is an adaptation of Tex-116-E (TexDOT, 2000). The particulars 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. 5.

Figure 5: 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 is the same as the as-delivered gradation, the specimens were actually built sieve-by-sieve to duplicate the as-delivered gradation (plus #4 sieve). 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). The new method is termed herein as the “Wet Ball Mill Modified” (WBMM).

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 684 results examined, only one set was outside the d2s range, and only one set exhibited an outlier. 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.

TEST RESULTS

Concrete Durability Factor

T 161 DF results were supplied by MoDOT. The DF reported herein is, in almost all cases, the average of three replicate beams. From the Batch spreadsheets, it appears that the mix proportions and the paste characteristics were held constant: paste volume, fine aggregate volume, coarse aggregate volume, air content, w/c, and sand source (Missouri River sand, Capital #1). However, air content did vary somewhat (3.9 to 7.5% via pressuremeter). The air system characteristics were not reported. All of the T 161 testing was performed at MoDOT's Central Laboratory.

Aggregate and Concrete Testing

Nineteen different ledge materials representing 18 ledges were subjected to 13 types of aggregate tests by Missouri S&T and two by MoDOT. In regard to the concrete mixtures that were used to make the T 161 beam specimens, two types of test results were used in the correlation and regression analyses (besides DF). Results from several test method types 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 DF and IPI. Those with a moderate range included $NaSO_4$, WAFT, bulk specific gravity, absorption, and NMS. Those with a more narrow range were LAA, MD, WBM, ACV, MB, PLS, compressive strength (Comp), and flexural strength (Flex).

Compressive and flexural strength refer to the specimens that were cast at the time of T 161 beam specimen casting. The compressive strengths were determined at an age of 28 days and one set of beams was tested at 35 days with no freeze/thaw cycling; these are sometimes referred to as “pre-test” strengths. T 161 beams were also broken (“post-test”) and the retained strength ratio is expressed as percent retained flexural strength (RetFlex). RetFlex is not useful for prediction of DF because of the necessity of performing the T 161 procedure.

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, discussed later. Test methods considered as fairly easy to perform include specific gravity, absorption, VSBSG, VSAbs, LAA, MD, MB, IPI, PLS, NMS, compressive strength, and flexural strength. More arduous methods are $NaSO_4$, WBM, and WAFT.

Table 10 depicts the averages of all aggregate test results. Twelve test methods were performed at Missouri S&T, while results of two more (T 85 BSG and Absorption) were extracted from primarily the T 161 Evaluation reports and supplemented by the Batching spreadsheets. Except for MoDOT data, each result is the average of three replicates. Results of MoDOT-determined deleterious material testing for deleterious rock (DR) and chert (Chert) are also shown.

Table 10: Aggregate Test Result Averages

ID	Formation	DF	LAA	MD	WBM	WBMM	ACV	IPI	MB	NaSO4
86R3M025	St. Louis	95	23.9	13.5	19.8	26.6	23.1	21.3	3.8	1.6
86L2R034	Plattin	95	29.8	14.0	25.4	30.2	29.0	24.0	3.7	10.8
81MA0292	Amazonia	94	25.5	18.4	22.7	27.2	27.3	36.0	7.3	10.8
86R3M031	Plattin	94	26.3	13.6	37.4	37.5	31.6	8.7	2.8	5.9
86R3M029	Plattin	89	27.3	14.2	20.8	26.6	25.6	20.3	1.5	6.3
86L2R021	St. Louis	89	24.1	13.2	21.1	23.6	23.3	28.7	4.0	5.5
89TCR067	JC Dolomite	87	41.1	18.3	24.9	32.3	29.2	13.7	3.8	7.3
87ASM006	Warsaw	81	29.4	21.1	23.2	31.0	25.9	24.0	1.2	1.9
80MA0051	Plattin	78	21.2	8.1	11.6	18.0	23.5	14.0	1.5	3.5
84SRE039	Bethany Falls	76	26.3	14.8	19.4	26.5	27.1	46.0	2.3	5.9
85DLR012	Burlington/Chou.	75	28.9	22.2	21.0	27.3	23.4	44.0	3.3	1.6
88MA0024	Burlington	73	36.7	18.4	23.2	25.5	29.0	12.0	0.7	1.1
83MA0234	Chouteau	72	37.5	30.3	32.4	39.0	30.4	33.3	3.3	18.7
81MA0379	Kereford	72	24.7	16.9	24.2	24.4	23.2	43.3	5.8	8.4
86R3M028	St. LouisSalem	69	26.3	22.8	20.5	26.4	24.3	55.3	5.8	12.3
85DGG007	Burlington	66	37.7	18.7	30.3	30.6	29.0	11.3	1.0	3.7
86L2R020	Plattin	64	27.5	13.4	20.3	26.1	25.1	22.0	1.5	11.8
85RDP040	JC Dolomite	54	31.3	18.7	21.1	27.1	23.1	49.0	3.1	12.8
84SRE203	Bethany Falls	28	27.3	20.5	21.5	29.4	26.6	15.0	2.3	15.0

ID	Formation	WAFT	PLS	VSBSG	VSSA	BSG	ABS	DR	Chert
86R3M025	St. Louis	9.9	4.51	2.644	1.19	2.653	0.8	0.3	0.1
86L2R034	Plattin	8.3	3.44	2.629	1.40	2.637	1.2	0.5	0.6
81MA0292	Amazonia	5.3	2.87	2.602	2.16	2.610	1.9	0.4	0.01
86R3M031	Plattin	5.2	2.82	2.606	1.41	2.642	1.0	0.3	0.2
86R3M029	Plattin	2.6	3.83	2.627	1.69	2.596	1.9	0.2	0.01
86L2R021	St. Louis	2.5	4.16	2.571	2.12	2.628	2.0	1.1	0.01
89TCR067	JC Dolomite	10.4	3.05	2.629	2.33	2.656	1.7	1.5	2.8
87ASM006	Warsaw	3.8	3.51	2.586	1.40	2.598	0.8	0.4	0.8
80MA0051	Plattin	0.9	4.23	2.682	0.92	2.675	1.1	0.01	0.01
84SRE039	Bethany Falls	6.5	3.72	2.587	1.89	2.602	1.5	0.1	0.01
85DLR012	Burlington/Chou.	10.5	3.17	2.589	2.49	2.605	1.7	0.6	0.01
88MA0024	Burlington	2.0	3.10	2.640	0.85	2.640	0.8	0.1	0.7
83MA0234	Chouteau	12.9	3.12	2.514	2.97	2.516	2.6	1.7	0.3
81MA0379	Kereford	20.2	4.16	2.579	2.09	2.580	2.0	1.8	0.01
86R3M028	St. LouisSalem	10.9	3.78	2.537	3.08	2.558	2.2	2	0.4
85DGG007	Burlington	1.2	3.52	2.593	1.75	2.616	1.2	3	0.5
86L2R020	Plattin	3.7	3.83	2.624	1.79	2.608	1.7	0.7	0.7
85RDP040	JC Dolomite	18.5	4.86	2.57	3.43	2.563	3.0	1.7	0.8
84SRE203	Bethany Falls	9.1	3.38	2.526	2.90	2.521	2.7	0.7	0.01

In Table 11 is shown mixture-related data that impact or are related to DF. The strength information came from MoDOT's T 161 Evaluation Reports. Likewise, the NMS of the material used in the T 161 tests was extracted from the Evaluation reports. The NMS values were used in the correlation and regression studies reported later, as opposed to the as-delivered material received by Missouri S&T. The percent of mixture volume represented by the plus ¾ in aggregate was computed from gradation information reported on the Evaluation Reports and absolute volumes reported in the Batching spreadsheets. In two cases, some information on certain sieves was missing and had to be interpolated from gradation plots.

Table 11: Mixture related data

ID	Formation	DF	Vol (+3/4")	NMS,DF	HudA	MIC	Ratio2	Flex	Comp	%RetFlex
86R3M025	St. Louis	95	6.3	0.75	2.451	1408	26.150	1014	6030	83.3
86L2R034	Plattin	95	0.0	0.50	3.024	847	2.762			
81MA0292	Amazonia	94	0	0.375	2.552	697	12.056	844	5260	89.7
86R3M031	Plattin	94	0	0.50	3.477	947	9.089	955	5990	85.2
86R3M029	Plattin	89	2.1	0.75	2.369	1056	7.537	940	5410	78.1
86L2R021	St. Louis	89	0	0.50	2.980	574	19.987			
89TCR067	JC Dolomite	87	3.6	0.75	2.262	1351	4.204	877	5870	70.9
87ASM006	Warsaw	81	0	0.50	2.444	1238	16.720	969	6550	71.4
80MA0051	Plattin	78	2.8	0.75	2.851	1156	6.642	916	5330	56.4
84SRE039	Bethany Falls	76	0	0.50	2.430	1000	15.837	946	5640	62.1
85DLR012	Burlington/Chou.	75	3.5	0.75	2.229	1387	13.709	786	5250	67.3
88MA0024	Burlington	73	12.6	0.75	2.228	1616	15.053	845	5540	61
83MA0234	Chouteau	72	4.9	0.75	2.254	1244	6.118	747	5290	57.5
81MA0379	Kereford	72	0	0.50	2.876	1111	6.019	866	5430	55.4
86R3M028	St. LouisSalem	69	4.1	0.75	2.318	1012	17.330	860	4900	50.6
85DGG007	Burlington	66	2.9	0.75	2.470	1294	4.918	919	6030	50.2
86L2R020	Plattin	64	7.1	1.00	2.305	1435	7.348	813	4740	42.5
85RDP040	JC Dolomite	54	2.8	0.75	2.112	1398	6.924	796	4740	46.3
84SRE203	Bethany Falls	28	0	0.50	2.713	1284	4.477	848	4260	33.2

CORRELATION

Interrelated Test Correlations

In the next sections are presented the one-to-one test method correlations. Correlation was done to: 1) check to see if correlations that are expected to exist do indeed exist, 2) look for outliers, 3) look for potential candidates for entry into regression predictive equations, and 4) flag possible future problems of multi-colinearity in regression work (in other words, it is not advisable to put two test methods in a predictive equation that correlate well with each other). The strength of a given correlation is represented by Pearson's correlation coefficient "R". The greater the R, the better the correlation, with "1.000" being perfect.

Correlations of tests were performed for methods within a specific aggregate property set, such as "Aggregate Pore Characteristics". At the end of this section, Table 12 is included which depicts the correlation coefficients ranked in descending order. Only correlations above 0.500 are shown as figures.

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, omitting pore length as a variable. Thus, pore size, distribution, and shape are included. Test methods included are T 85 Abs and BSG, and their vacuum saturated counterparts, IPI, WAFT, and $NaSO_4$.

Absorption (T 85). MoDOT supplied the test results, as reported on the Evaluation Reports and Batching spreadsheets. There was only one replicate tested. No precision information is available.

Bulk Specific Gravity (T 85). See above comments.

BSG is a function of mineralogy (specific gravity of the solids) as well as pore characteristics.

Fig. 6 shows the relationship between T 85 bulk specific gravity (dry) and Absorption. The correlation coefficient R is fairly good (-0.780) and is negative, both of which would be expected.

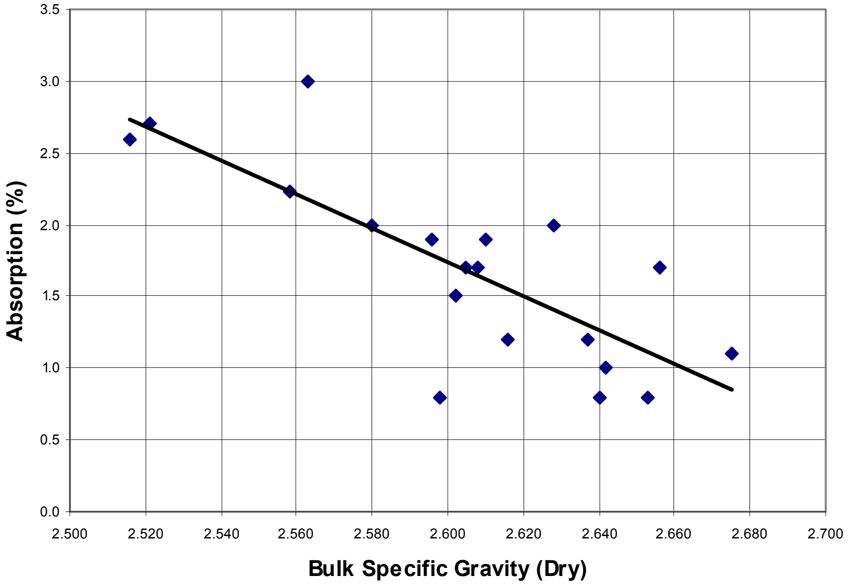


Figure 6: Absorption vs. Bulk Specific Gravity (Dry)

Vacuum Saturated Absorption. Vacuum saturation should result in more water being pulled into the aggregate, compared to the standard T 85 24 hr. soak, thus increasing the absorption value. In most cases, this held true. The change in absorption ranged from -0.2 to + 0.9%, with an average increase of 0.3%. The comparison is shown below in Fig. 7. The correlation factor R was 0.918. A paired t-test showed that the two parameters were statistically different.

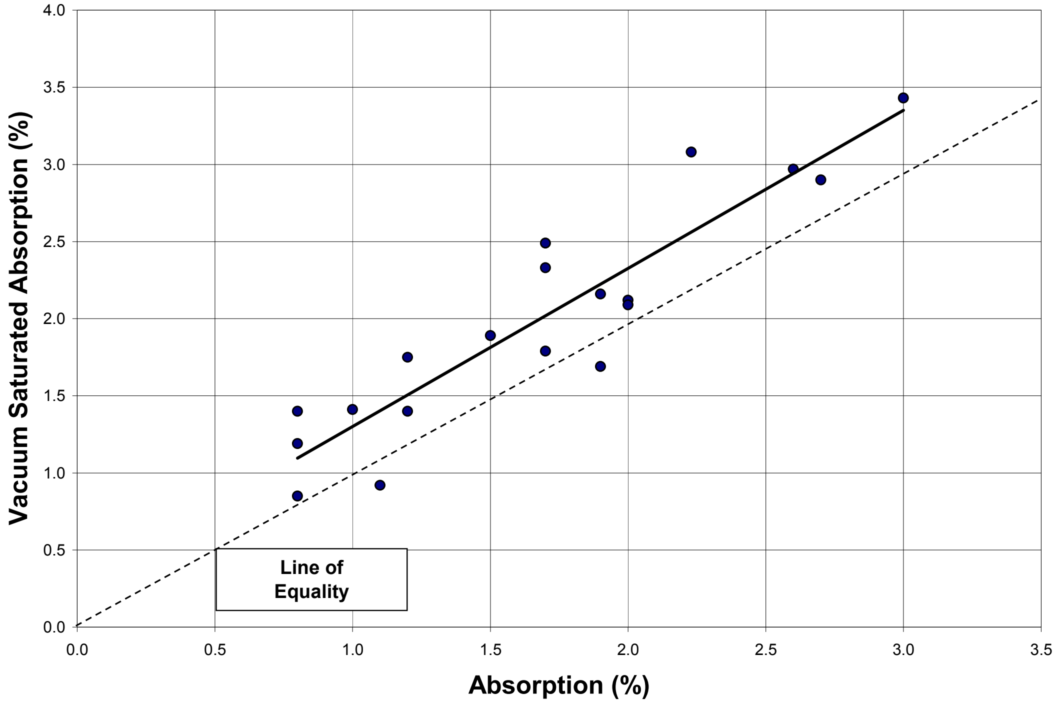


Figure 7: Vacuum Saturated Absorption vs. Absorption

Vacuum Saturated Bulk Specific Gravity. Fig. 8 shows the relationship between VSBSG and VSAbs. The correlation coefficient R is good (-0.806) and is negative, both of which would be expected.

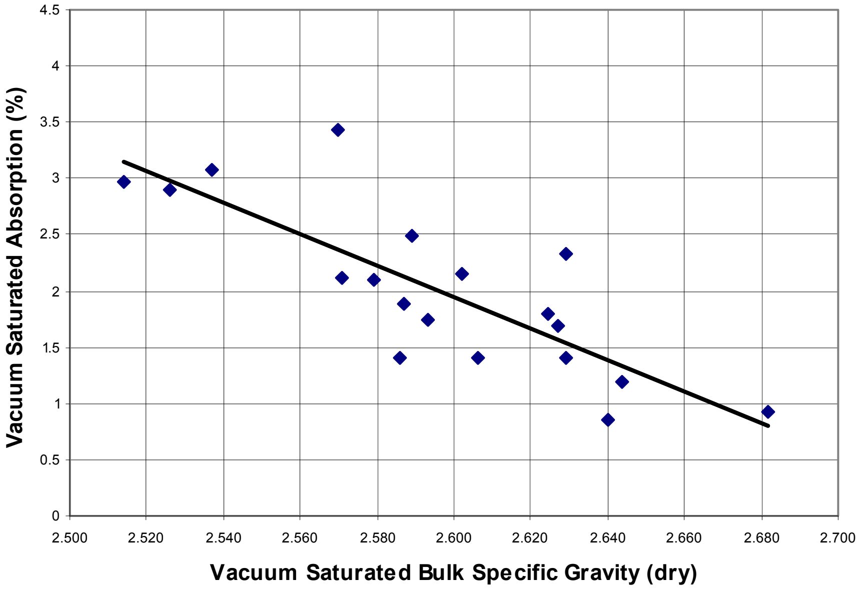


Figure 8: Vacuum Saturated Absorption vs. Vacuum Saturated Bulk Specific Gravity (dry)

Fig. 9 shows the relationship between T 85 BSG and VSBSG. The correlation coefficient R is good (0.896) and is positive, both of which would be expected. A paired t-test showed that there was no statistical difference between the two.

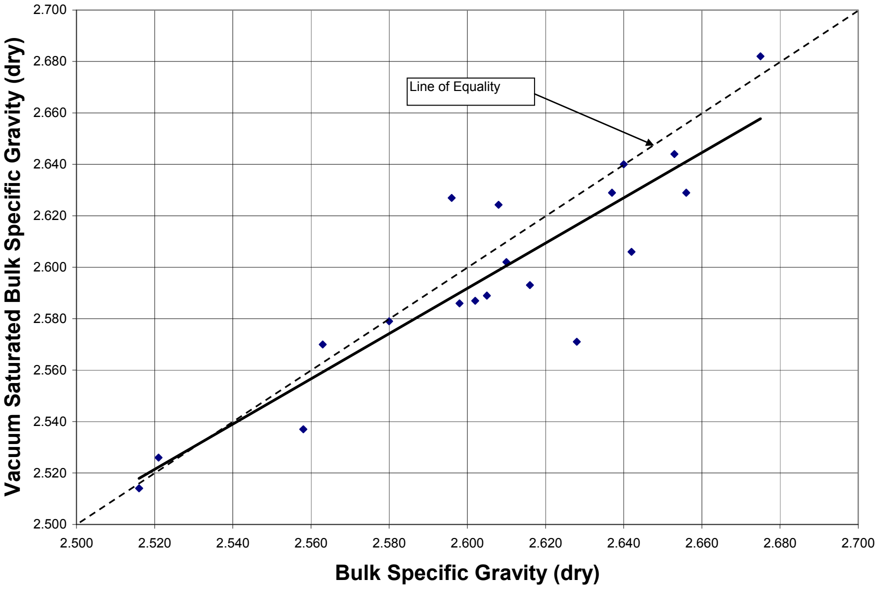


Figure 9: Vacuum Saturated Bulk Specific Gravity (dry) vs. Bulk Specific Gravity (dry)

Fig. 10 shows the relationship between T 85 BSG and VSAbs. The correlation coefficient R is fairly good (-0.776) and is negative, both of which would be expected.

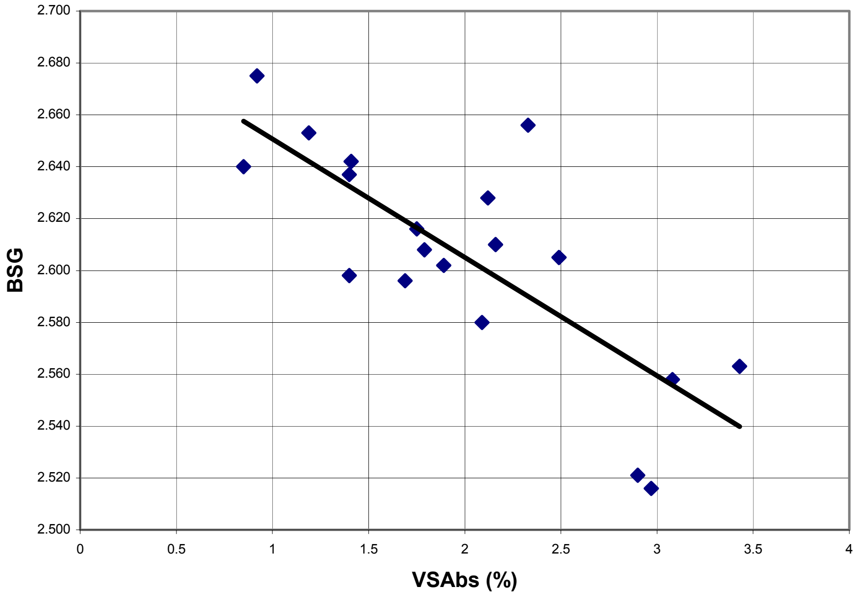


Figure 10: Bulk Specific Gravity (Dry) vs. Vacuum Saturated Absorption

Fig. 11 shows the relationship between VSBSG and Abs. The correlation coefficient R is fair (-0.710) and is negative, both of which would be expected.

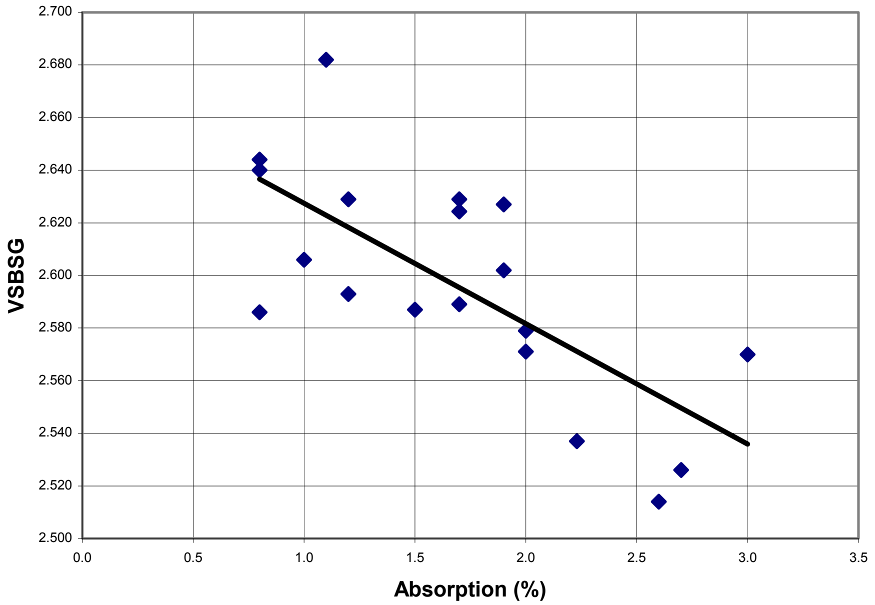


Figure 11: Vacuum Saturated Bulk Specific Gravity (Dry) vs. Absorption

Iowa Pore Index. Fig. 12 shows the relationship between Iowa Pore Index and VSBSG. The correlation coefficient R is fair (-0.527) and is negative, both of which would be expected.

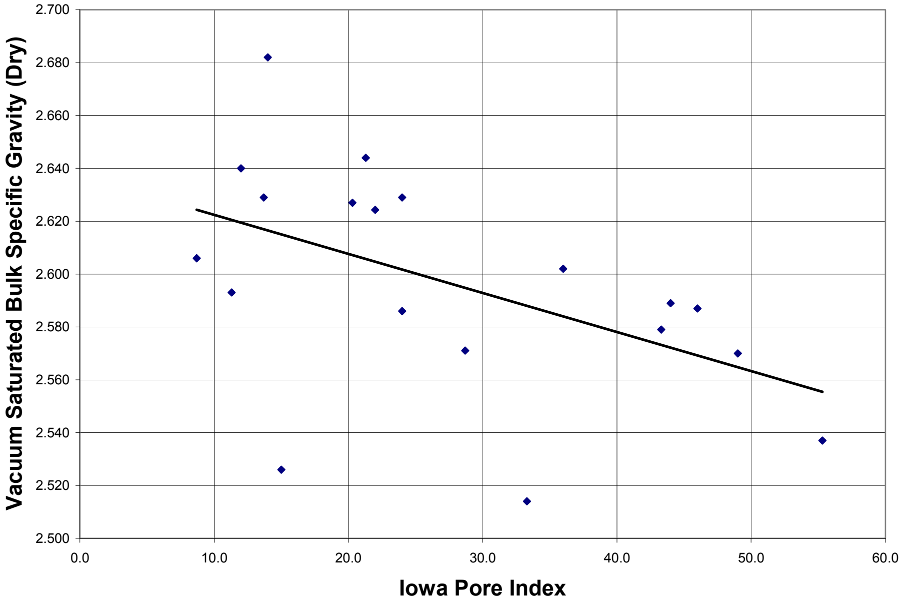


Figure 12: Vacuum Saturated Bulk Specific Gravity (Dry) vs. Iowa Pore Index

Fig. 13 shows the relationship between Iowa Pore Index and Vacuum Saturated Absorption. The correlation coefficient R is fair (0.657) and is positive, both of which would be expected.

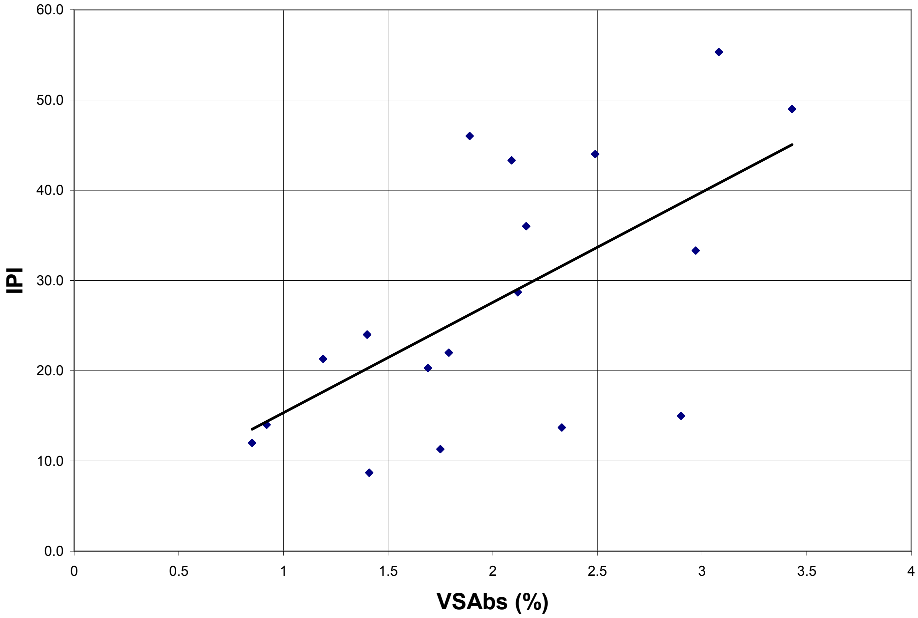


Figure 13: Iowa Pore Index vs. Vacuum Saturated Absorption

Fig. 14 shows the relationship between Iowa Pore Index and Absorption. The correlation coefficient R is fair (0.526) and is positive, both of which would be expected.

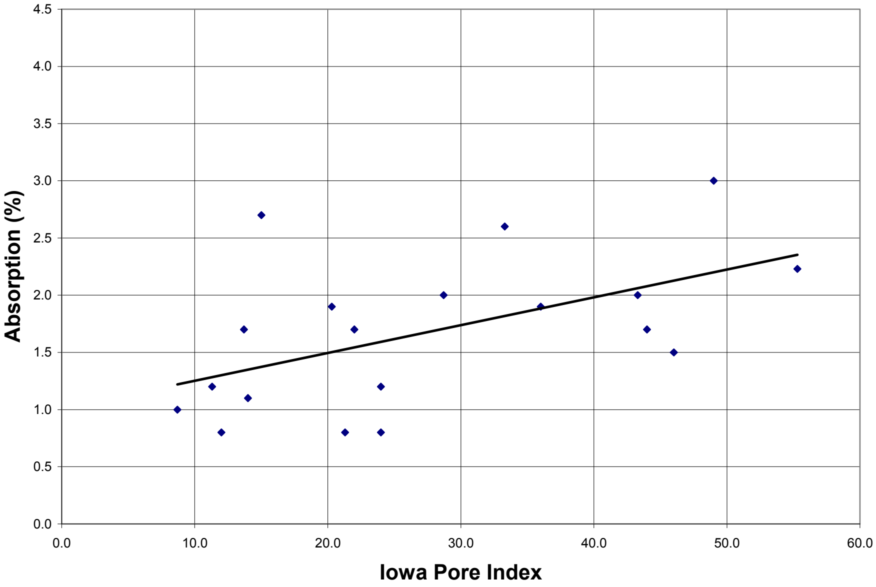


Figure 14: Absorption vs. Iowa Pore Index

Water-Alcohol Freeze Thaw. Fig. 15 shows the relationship between WAFT and Absorption. The correlation coefficient R is fair (0.580) and is positive, both of which would be expected.

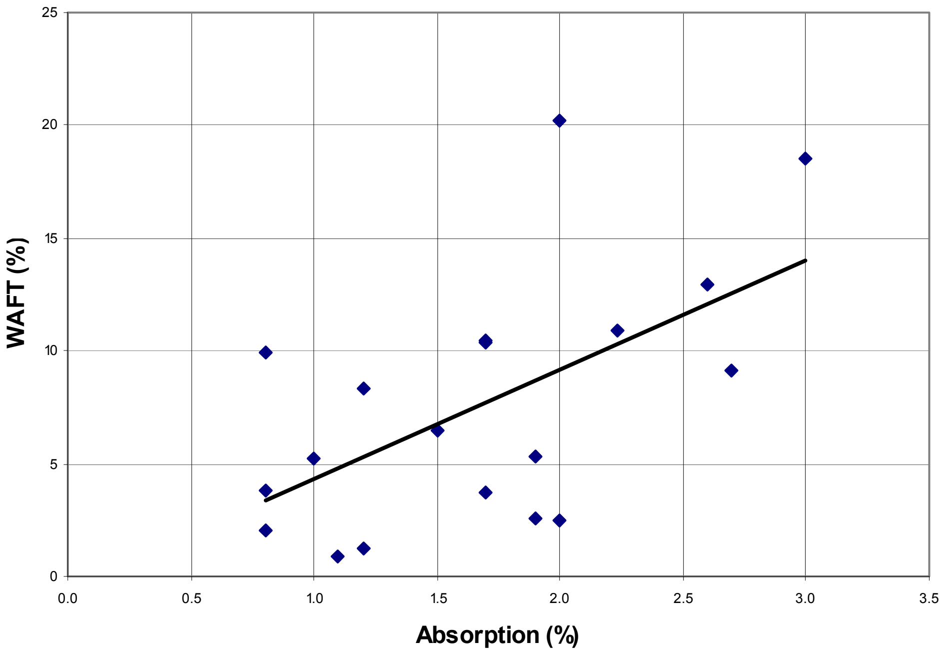


Figure 15: Water Alcohol Freeze-Thaw Soundness vs. Absorption

Fig. 16 shows the relationship between WAFT and VSAbs. The correlation coefficient R is fair (0.644) and is positive, both of which would be expected.

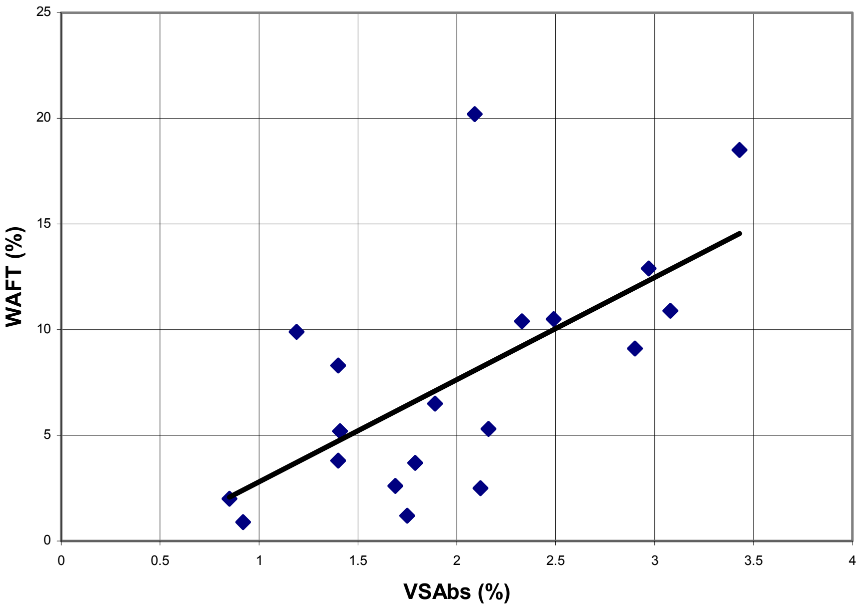


Figure 16: Water Alcohol Freeze-Thaw Soundness vs. Vacuum Saturated Absorption

Sodium Sulfate Soundness. Fig 17 shows the relationship between NaSO4 soundness and VSBSG. The correlation coefficient R is fair (-0.629) and is negative, both of which would be expected.

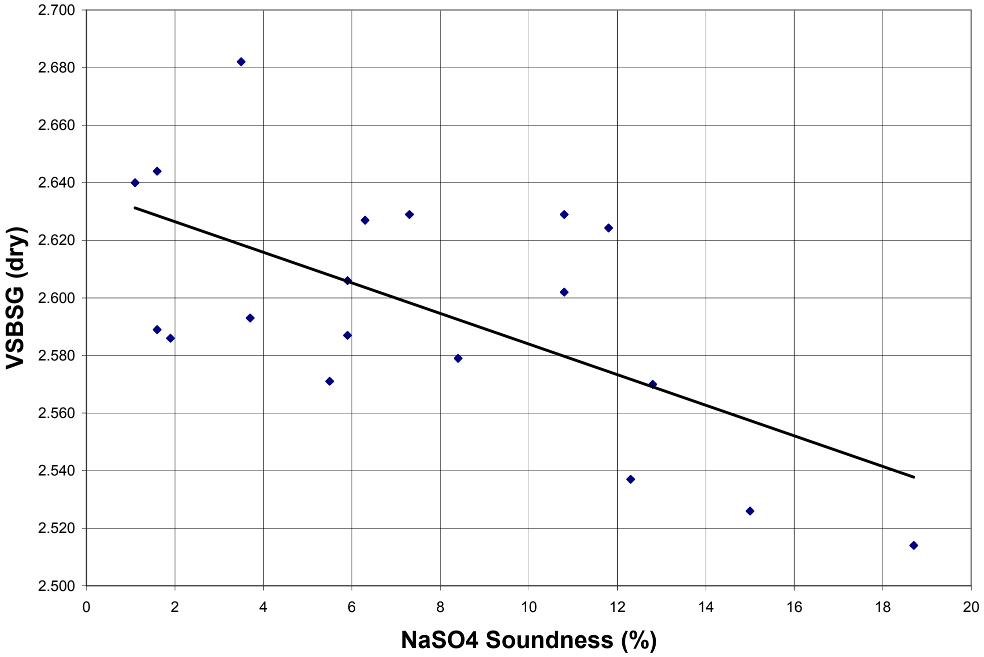


Figure 17: Vacuum Saturated Bulk Specific Gravity (dry) vs. Sodium Sulfate Soundness

Fig 18 shows the relationship between $NaSO_4$ soundness and VSSA. The correlation coefficient R is fairly good (0.706) and is positive, both of which would be expected.

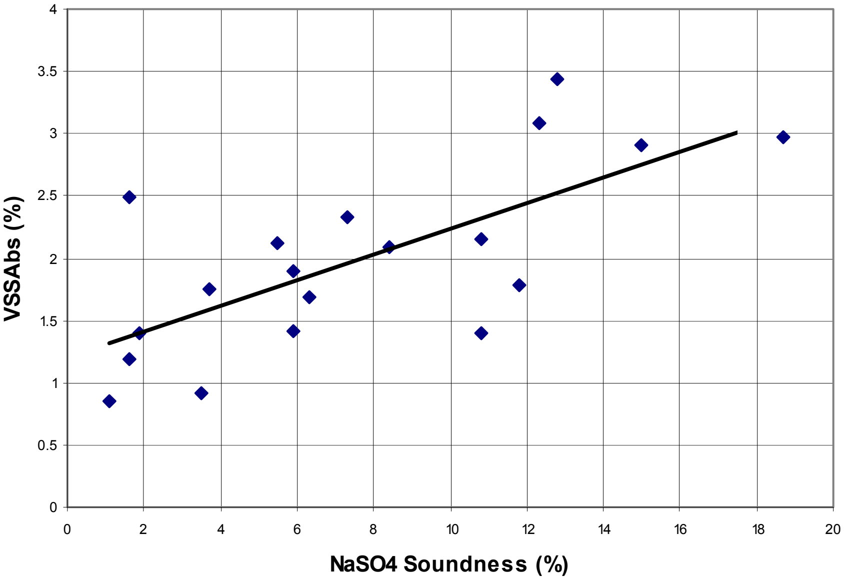


Figure 18: Vacuum Saturated Absorption vs. Sodium Sulfate Soundness

Fig. 19 shows the relationship between $NaSO_4$ soundness and BSG. The correlation coefficient R is fairly good (-0.726) and is negative, both of which would be expected.

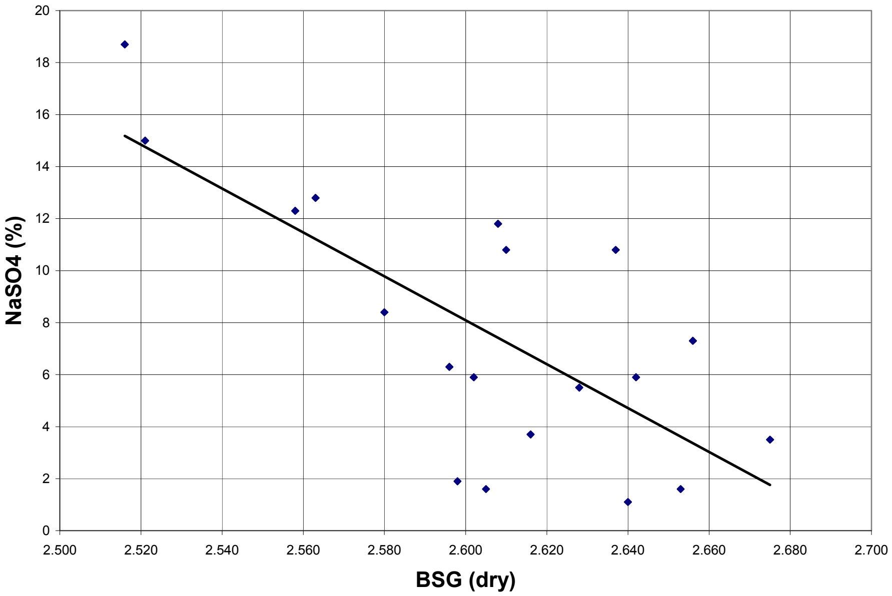


Figure 19: Sodium Sulfate Soundness vs. Bulk Specific Gravity (dry)

Fig 20 shows the relationship between $NaSO_4$ soundness and Abs. The correlation coefficient R is fairly good (0.791) and is positive, both of which would be expected.

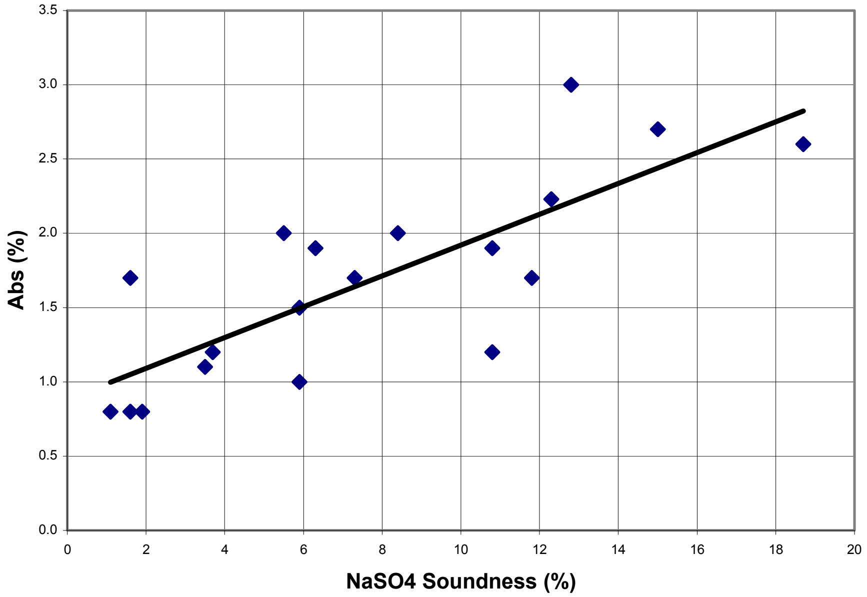


Figure 20: Absorption vs. Sodium Sulfate Soundness

Pore Length

Several measures of pore length were developed: nominal maximum size of the material as-tested in the T 161 tests (NMSDF), volume of plus ¾ in material expressed as a percent of concrete volume as-tested in the T 161 tests (VolPlus3/4), HudA, MIC, and Ratio2. None of the above five gradation parameters correlated well with any individual parameter. However, they did correlate well amongst each other.

Fig. 21 shows the relationship between NMSDF and VolPlus3/4. The correlation coefficient R is fairly good (0.735) and is positive, both of which would be expected.

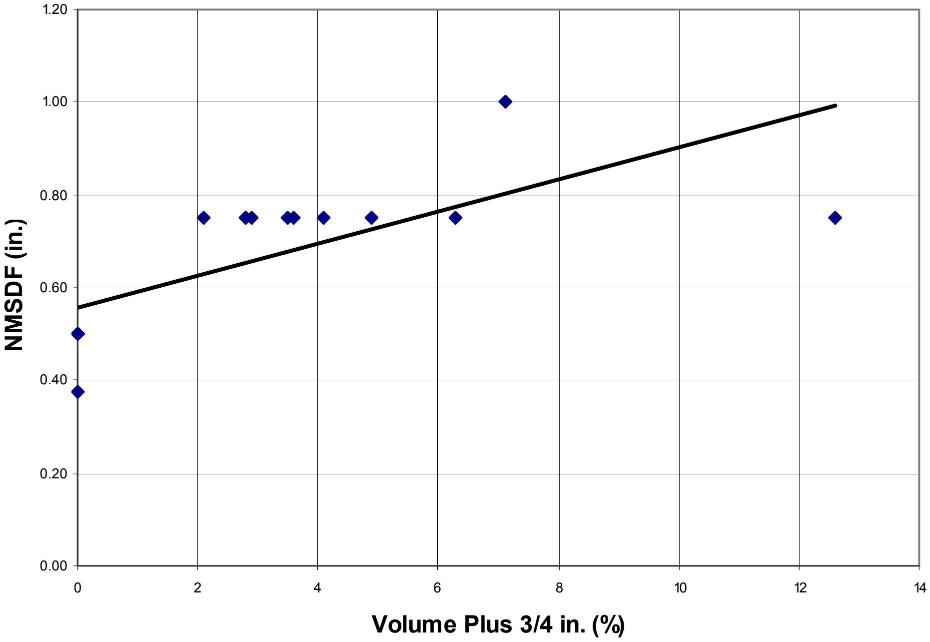


Figure 21: Nominal Maximum Size, DF vs. Volume Plus ¾ in.

Fig. 22 shows the relationship between HudA and MIC. The correlation coefficient R is fairly good (0.897) and is negative, both of which would be expected.

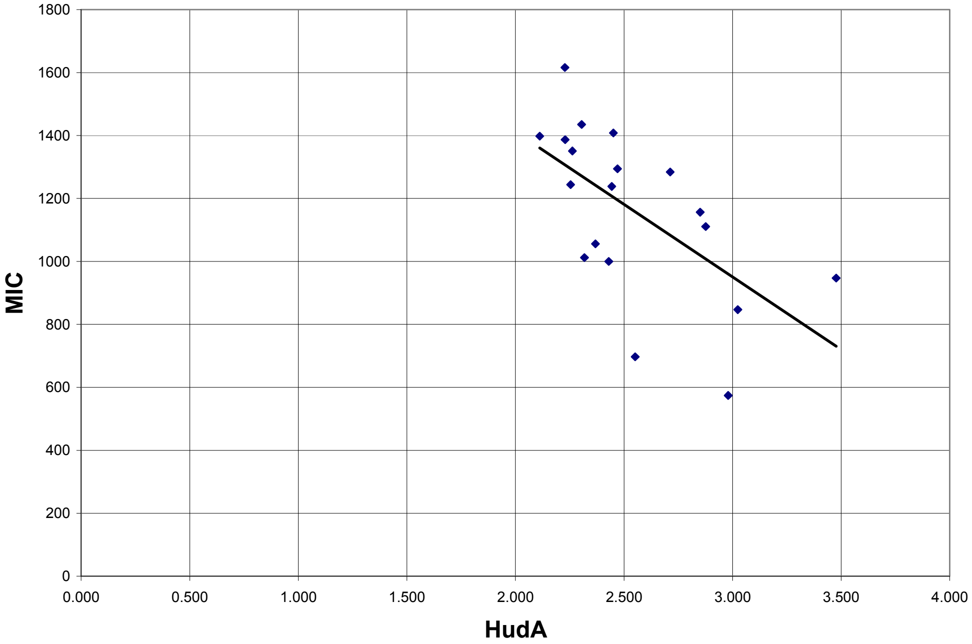


Figure 22: MIC vs. HudA

Fig. 23 shows the relationship between Vol Plus ¾ in. and MIC. The correlation coefficient R is fair (0.753) and is positive, both of which would be expected.

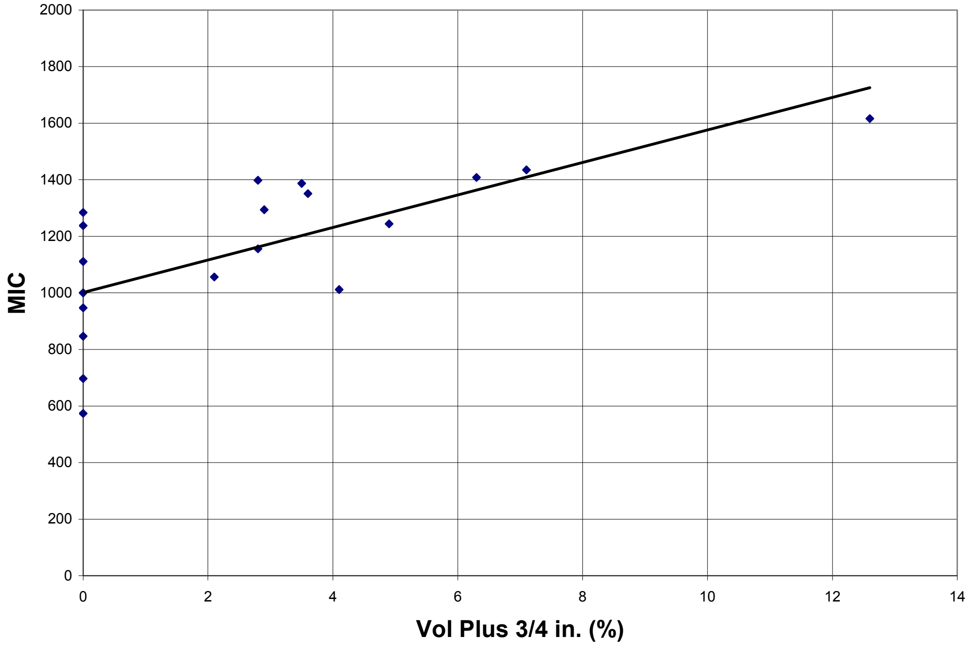


Figure 23: MIC vs. Volume Plus ¾ in.

Mineralogy

Some studies have shown that clay mineralogy within the matrix of a given stone has an effect on the results of T 161 testing. In some cases, the effect is even positive. Methylene blue is an indicator of the presence of certain clay minerals. It has also been shown that the results of pore characteristic-related tests such as IPI are also influenced by clay mineralogy. Likewise, it may be that other pore characteristic-related tests such as WAFT are affected by the presence of clay minerals.

Fig. 24 shows the relationship between MB and WAFT. The correlation coefficient R is fair (0.528) and is positive, both of which would be expected if the type of clay minerals present are harmful to durability.

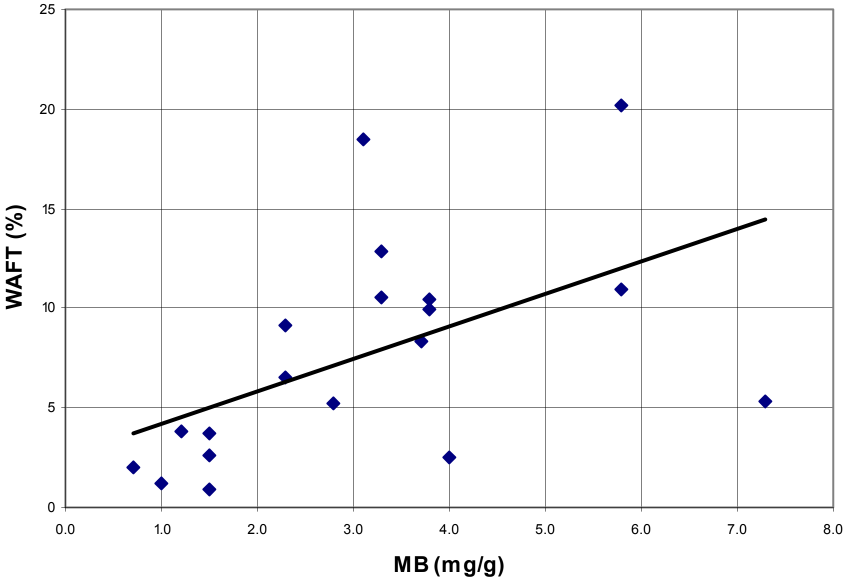


Figure 24: Water Alcohol Freeze-Thaw Soundness vs. Methylene Blue

Elastic Accommodation/Strength

These are tests that reflect some aspect of the manner of the aggregate's reaction to internal pressure. Reaction can take the forms of the sufficiency of strength to resist fracture, or to elastically accommodate the pressure.

Aggregate Crushing Value. Fig. 25 shows the relationship between ACV and PLS. The correlation coefficient R is fairly good (-0.780) and is negative, both of which would be expected.

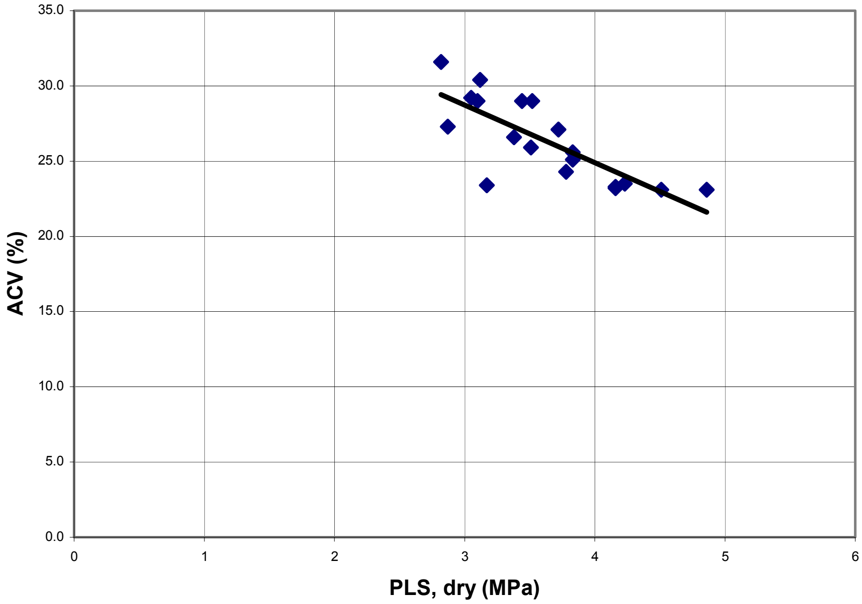


Figure 25: Aggregate Crushing Value vs. Point Load Strength

Los Angeles Abrasion. Fig. 26 shows the relationship between LAA and MD. The correlation coefficient R is fair (0.549) and is positive, both of which would be expected.

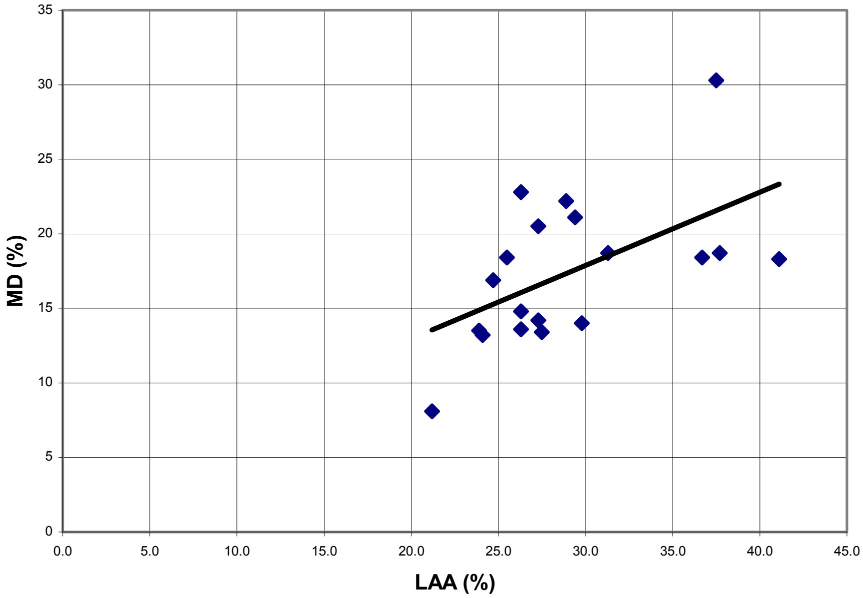


Figure 26: Micro-Deval vs. Los Angeles Abrasion

Fig. 27 shows the relationship between LAA and ACV. The correlation coefficient R is fair (0.602) and is positive, both of which would be expected.

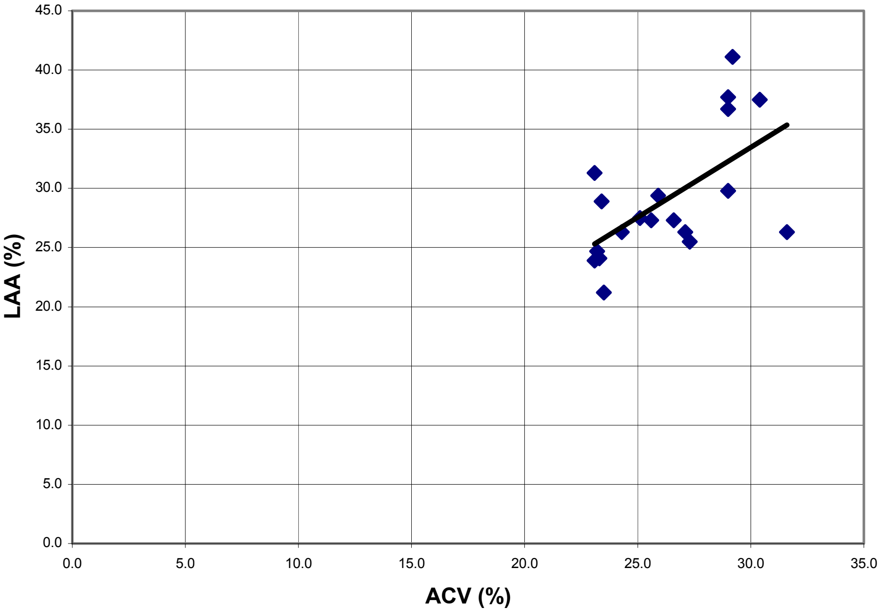


Figure 27: Los Angeles Abrasion vs. Aggregate Crushing Value

Micro Deval. Successful correlations are depicted elsewhere.

Point Load Strength. Successful correlations are depicted elsewhere.

Water-Alcohol Freeze Thaw. Successful correlations have already been depicted above.

Wet Ball Mill. Fig. 28 shows the relationship between WBM and ACV. The correlation coefficient R is fairly good (0.762) and is positive, both of which would be expected.

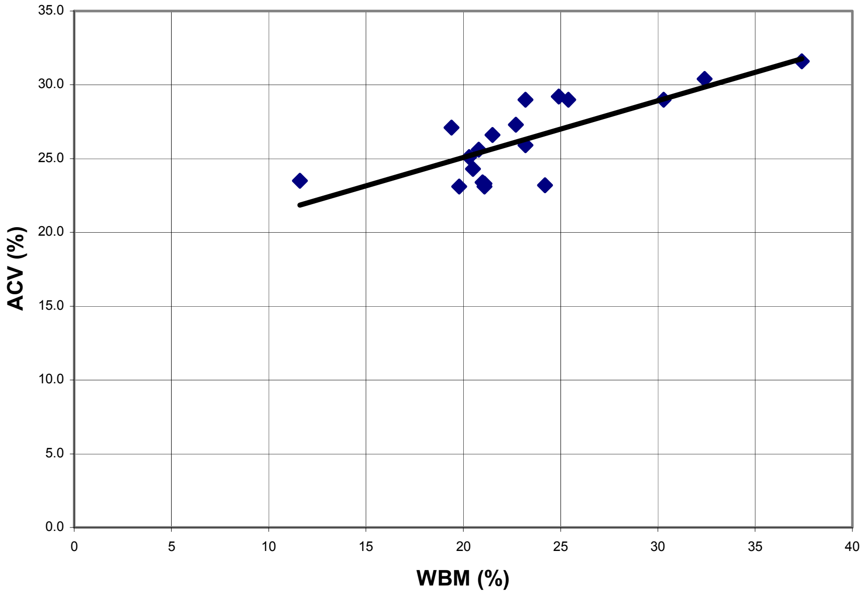


Figure 28: Aggregate Crushing Value vs. Wet Ball Mill

Fig. 29 shows the relationship between WBM and PLS. The correlation coefficient R is fair (-0.562) and is negative, both of which would be expected.

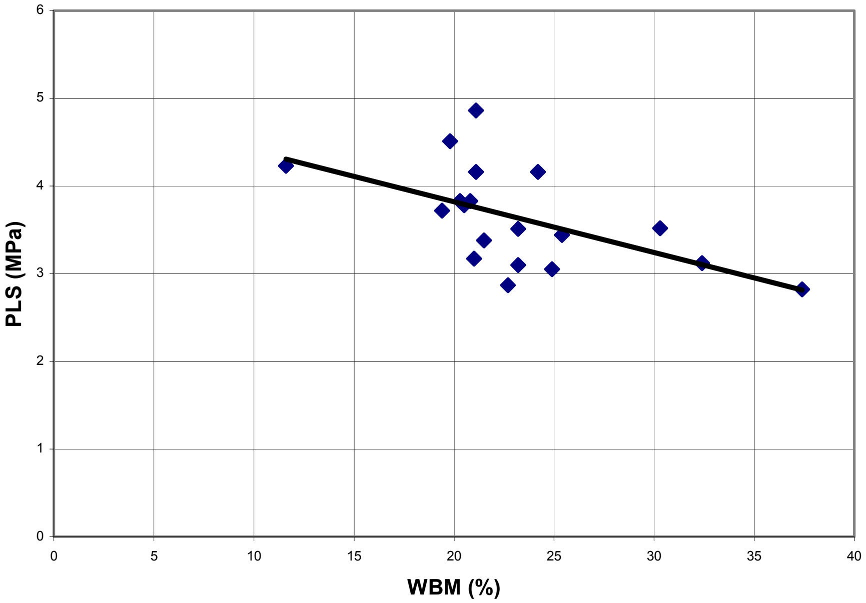


Figure 29: Point Load Strength vs. Wet Ball Mill

Wet Ball Mill-Modified (WBMM). This is a modified version of the standard wet ball mill test. The plus #4 sieve residue of the WBM test is subjected to a gradation analysis, with the breakdown of the material quantified, giving more weight to the finer sizes.

Fig. 30 shows the relationship between WBMM and WBM. The correlation coefficient R is good (0.881) and is positive, both of which would be expected.

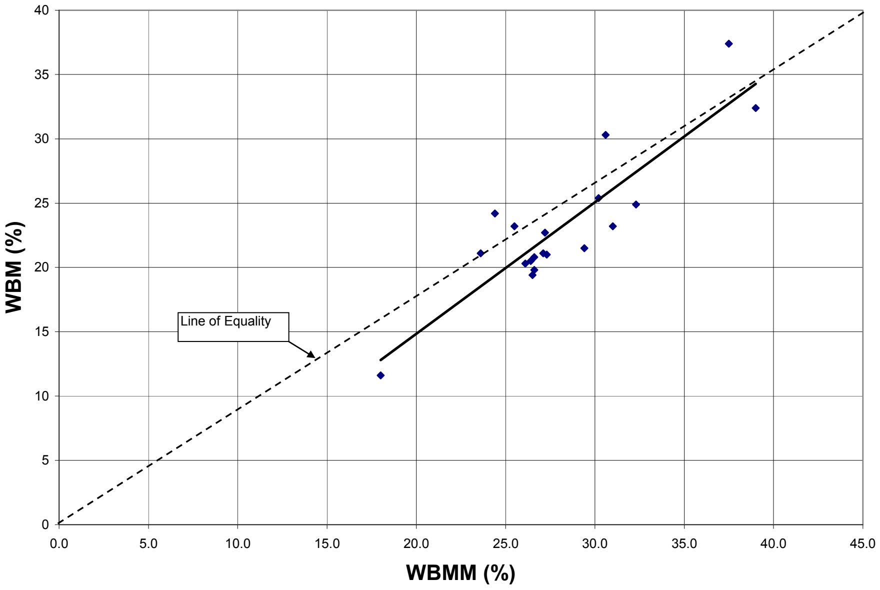


Figure 30: Wet Ball Mill vs. Wet Ball Mill-Modified

Fig. 31 shows the relationship between WBMM and ACV. The correlation coefficient R is fairly good (0.759) and is positive, both of which would be expected.

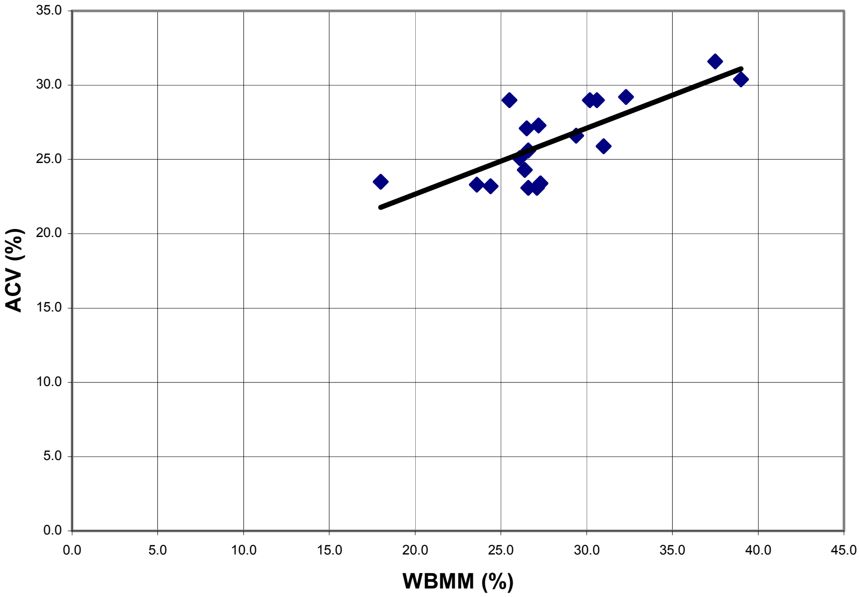


Figure 31: Aggregate Crushing Value vs. Wet Ball Mill-Modified

Fig. 32 shows the relationship between WBMM and LAA. The correlation coefficient R is fair (0.561) and is positive, both of which would be expected.

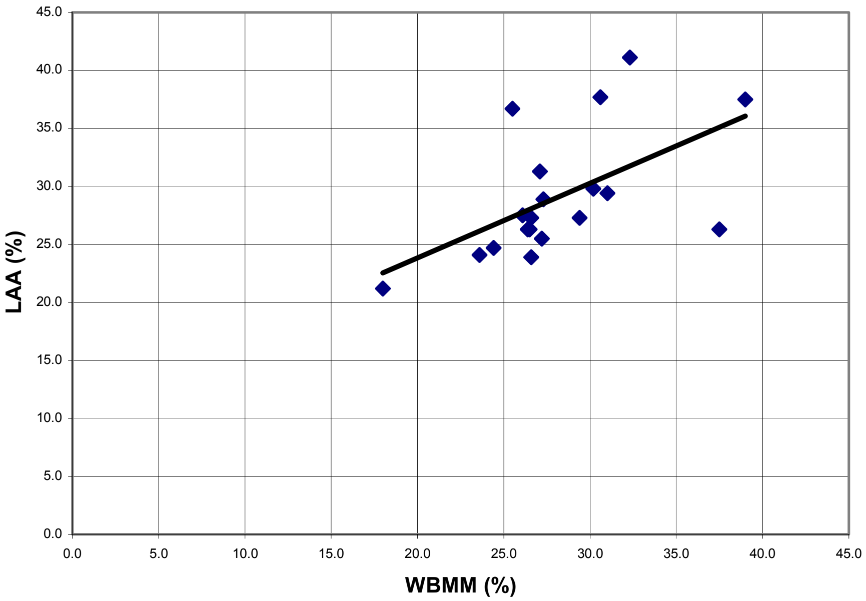


Figure 32: Los Angeles Abrasion vs. Wet Ball Mill-Modified

Fig. 33 shows the relationship between WBMM and MD. The correlation coefficient R is fair (0.576) and is positive, both of which would be expected.

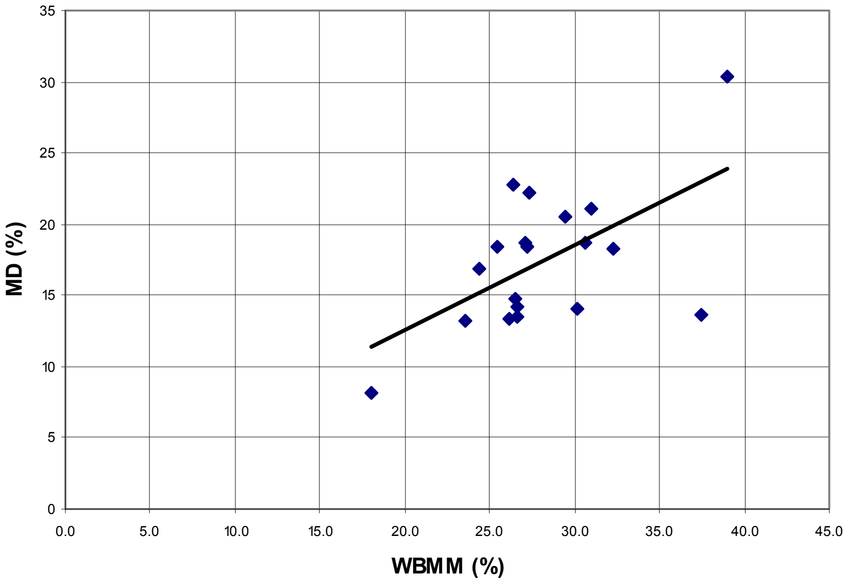


Figure 33: Micro-Deval vs. Wet Ball Mill-Modified

Fig. 34 shows the relationship between WBMM and PLS. The correlation coefficient R is fair (-0.592) and is negative, both of which would be expected.

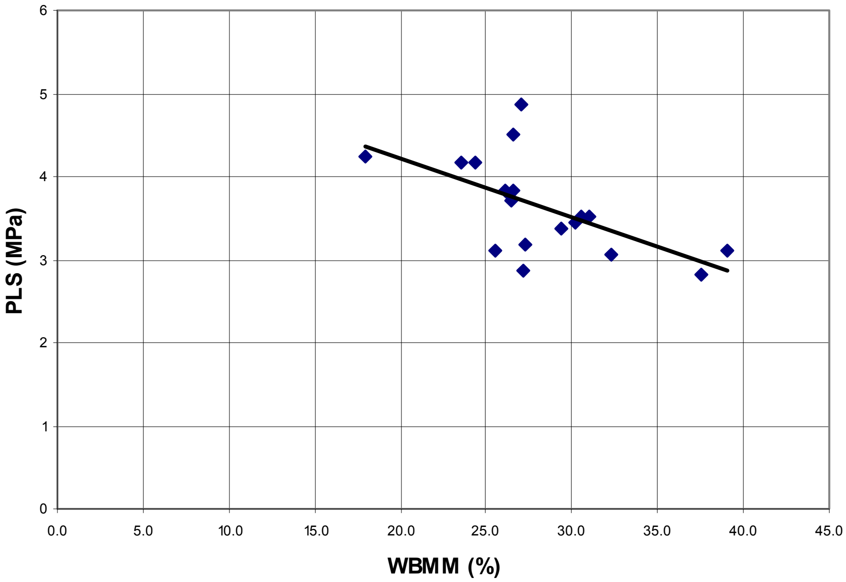


Figure 34: Point Load Strength vs. Wet Ball Mill-Modified

Ranked Interrelated Correlation Coefficients

Below is Table 12, which depicts the correlation coefficients greater than 0.600 ranked in numerical order. Appendix F contains the full correlation matrix.

Table 12: Interrelated correlation coefficients

Test Methods	R
VSAbs vs Abs	0.918
HudA vs MIC	-0.897
VSBSG vs BSG	0.896
WBM vs WBMM	0.881
VSAbs vs VSBSG	-0.806
NASO4 vs Abs	0.791
ACV vs PLS	-0.780
Abs vs BSG	-0.780
VSAbs vs BSG	-0.776
Vol Plus 0.75 vs WP0.5	0.765
ACV vs WBM	0.762
MD vs VSBSG	-0.760
ACV vs WBMM	0.759
NMSDF vs WP0.5	0.753
NMSDF vs Vol Plus 0.75	0.735
MD vs BSG	-0.734
NaSO4 vs BSG	-0.726
Abs vs VSBSG	-0.710
NaSO4 vs VSAbs	0.706
LAA vs Chert	0.699
MIC vs NMSDF	0.680
IPI vs VSAbs	0.657
WAFT vs VSAbs	0.644
MD vs VSAbs	0.637
NaSO4 vs BSG	0.629
IPI vs WAFT	0.614
ACV vs LAA	0.602

.

Correlation with MoDOT Results

To see if Missouri S&T results lined up with historical test data from MoDOT, correlations were performed for tests that were common to both datasets. This involved MD, LAA, NaSO4, WAFT, and to a limited extent, IPI. MoDOT has not performed WBM on any of the study aggregates. Figures 35 through 39 are shown below. In general, considering that these tests were not performed on split samples, rather, the samples were taken months or even years apart, the test results seemed to correlate fairly well. Correlation coefficients ranged from 0.797 to 0.933, and are shown on each plot. Paired t-tests showed that, for each test method , there was no statistical difference between MoDOT's and Missouri S&T's results, with the exception of LAA. However, in regard to the LAA results, there is one data point that appears to be an outlier. Looking at other LAA test values on the Quarry Reports, the results were much more in line with the Missouri S&T value. Thus, if the MoDOT value was not representative, then the correlation with Missouri S&T results would be much stronger. The conclusions are that the materials used in the present study were probably fairly close in nature to the materials used in the T 161 testing.

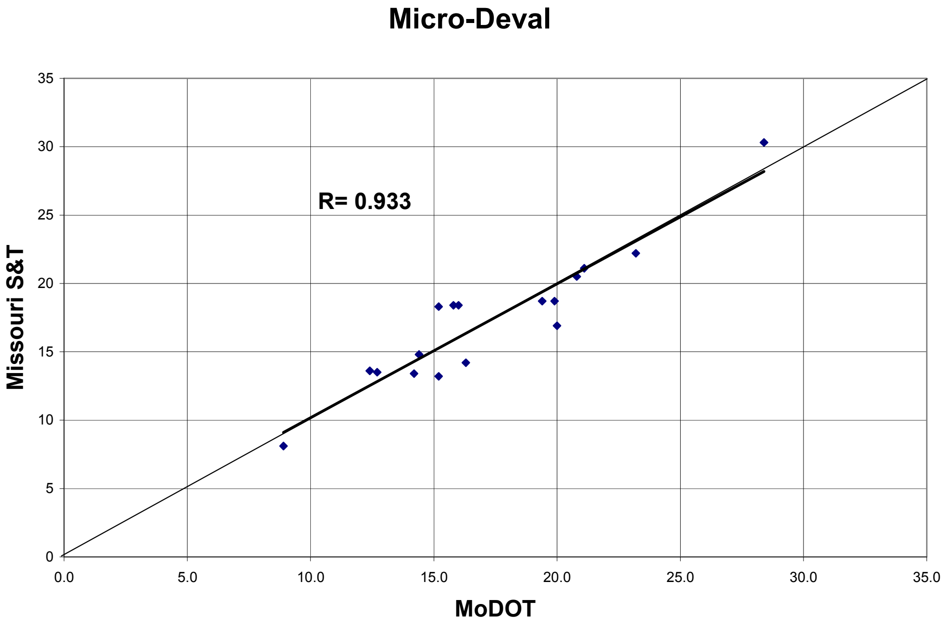


Figure 35: Comparison of MoDOT vs. Missouri S&T Micro-Deval results

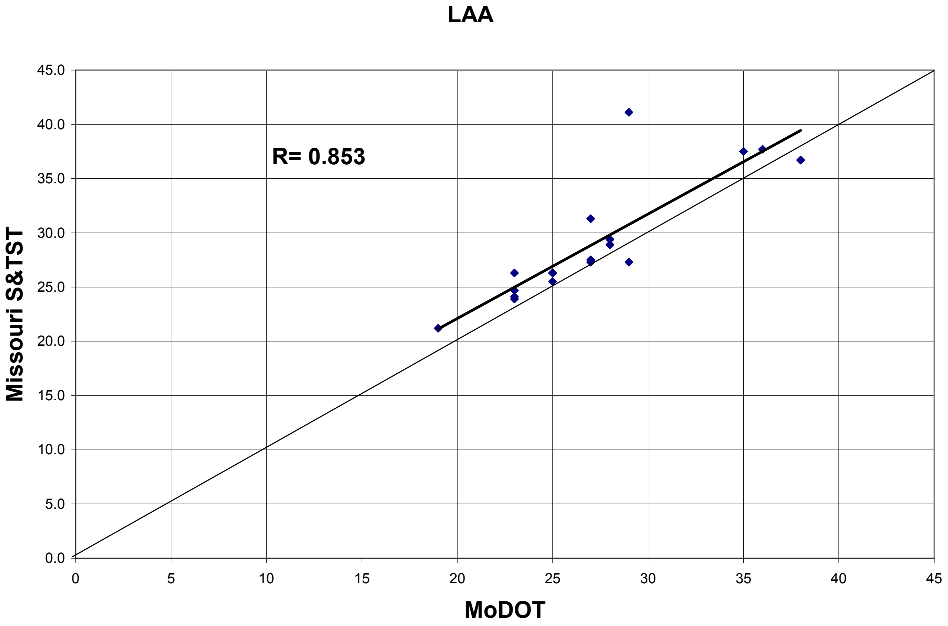


Figure 36: Comparison of MoDOT vs. Missouri S&T LAA results

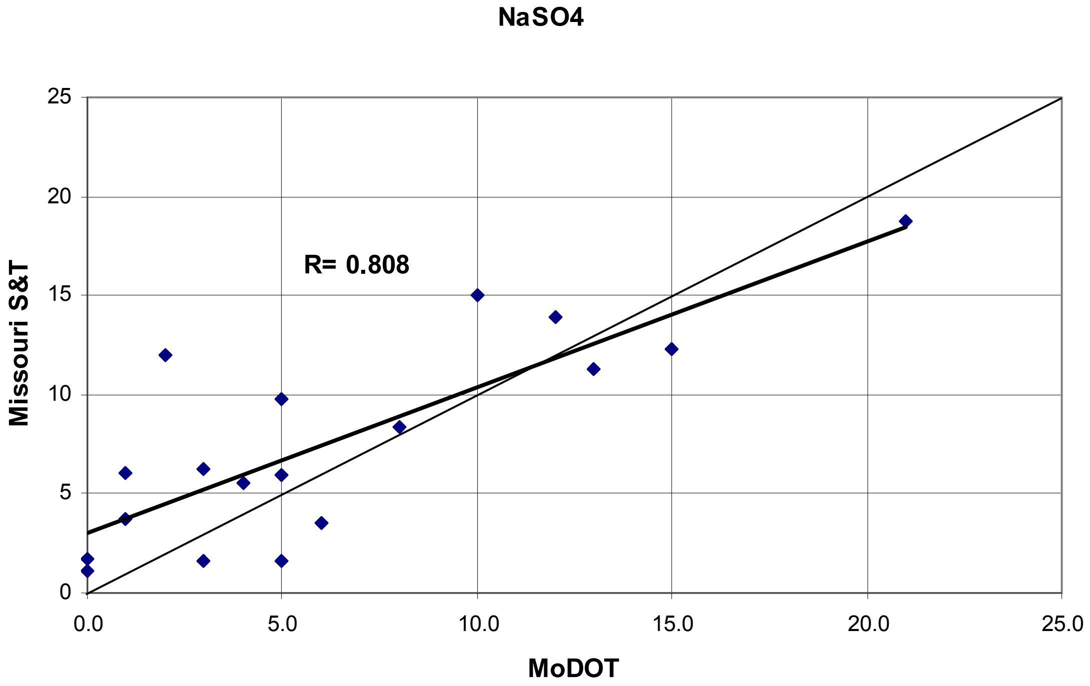


Figure 37: Comparison of MoDOT vs. Missouri S&T $NaSO_4$ results

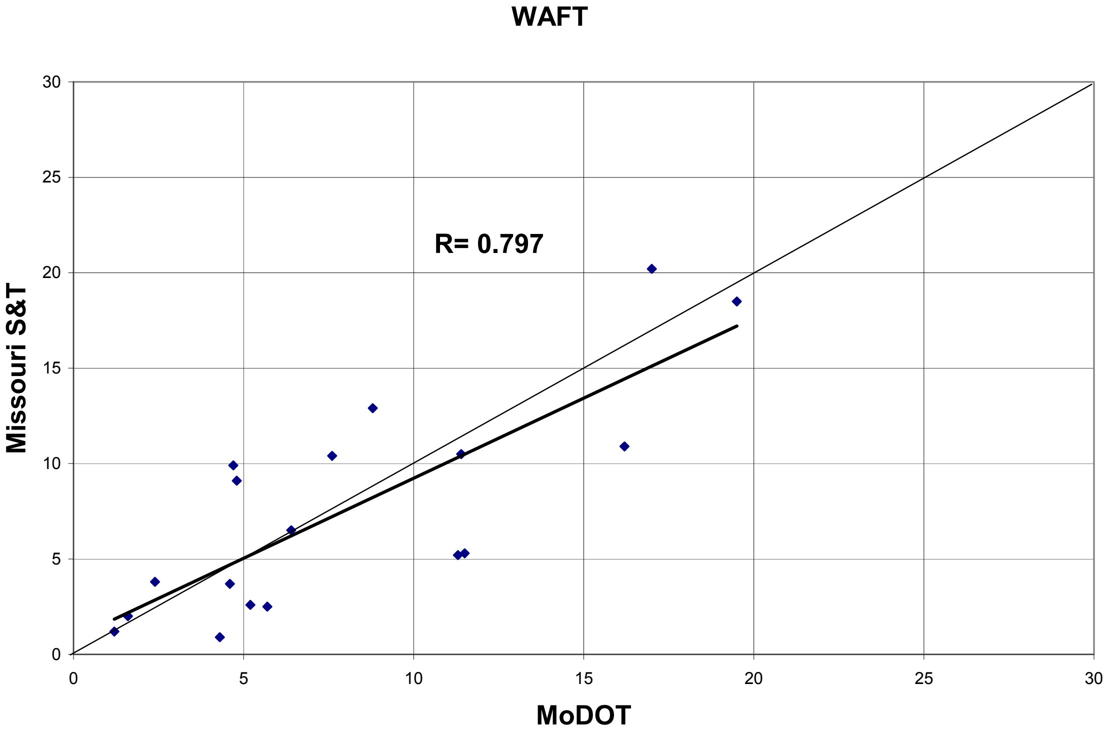


Figure 38: Comparison of MoDOT vs. Missouri S&T WAFT results

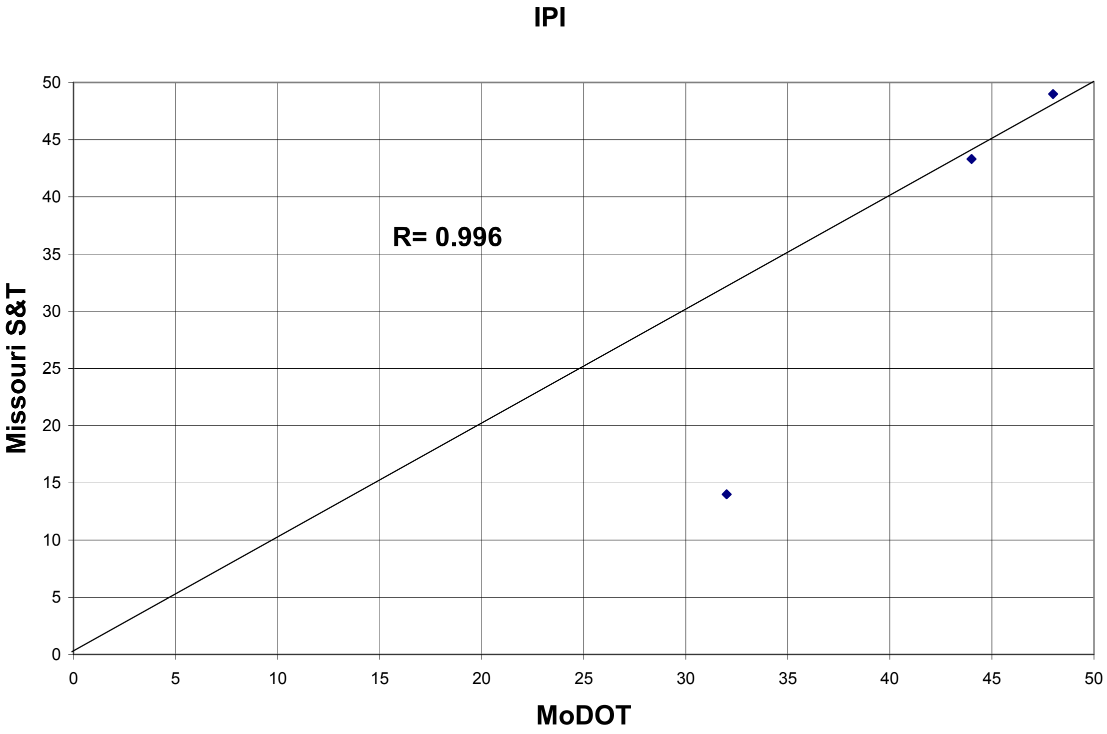


Figure 39: Comparison of MoDOT vs. Missouri S&T IPI results

Correlation of DF to All Individual Test Results

In Tables 13 and 14 are shown the results of correlation of Durability Factor with individual test methods and mix characteristics, respectively. It should be kept in mind that the signs (slope of the curve) may be meaningless for very low correlations.

Table 13: DF correlation with test methods

Test	R
BSG	0.664
Abs	-0.567
VSAbs	-0.530
VSBSG	0.525
NaSO4	-0.429
MD	-0.352
DR	-0.323
WAFT	-0.267
MB	0.255
PLS	-0.159
ACV	0.155
WBM	0.141
LAA	-0.135
IPI	-0.125
Chert	0.059
WBMM	0.054

Table 14: DF correlation with mixture characteristics

Characteristic	Correlation Factor
RetFlex	0.919
Comp	0.698
Flex	0.477
MIC	-0.453
HudA	0.376
NMSDF	-0.206
Vol Plus 3/4"	-0.093
Ratio2	0.054

Figs. 40 through 51 show the strongest relationships between DF and various parameters as listed in Tables 13 and 14. Lines of demarcation show the DF threshold of 75, and other past, present, or possible MoDOT aggregate threshold limits.

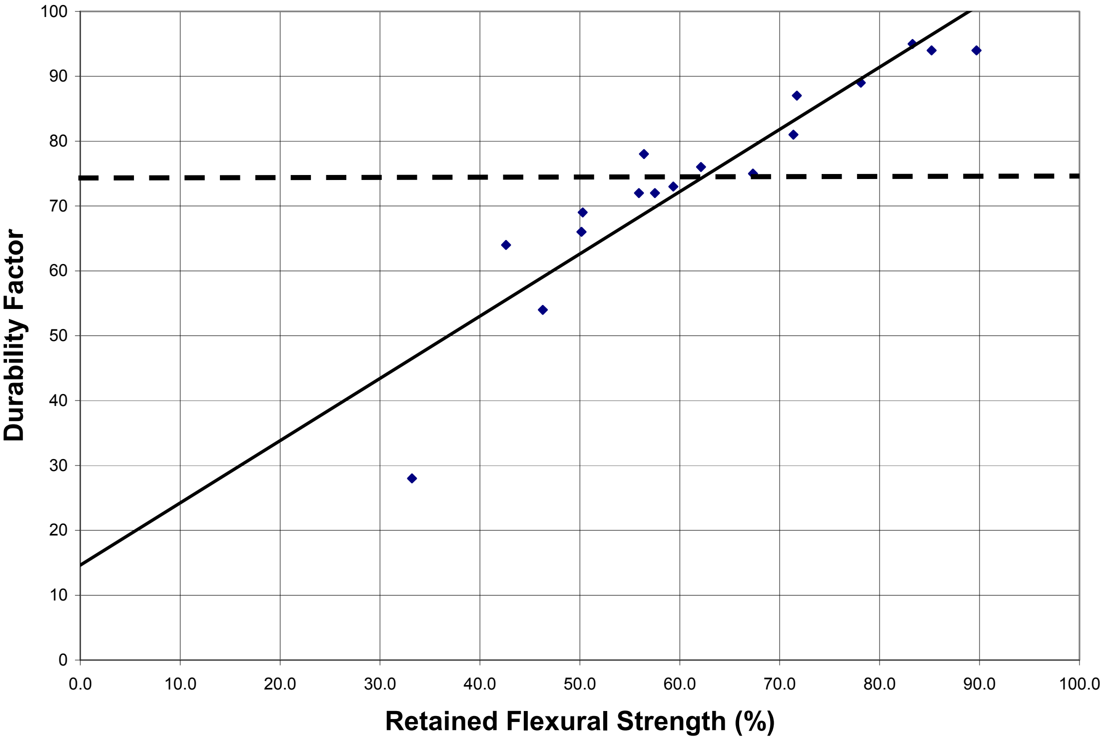


Figure 40: Durability Factor vs. Retained Flexural Strength

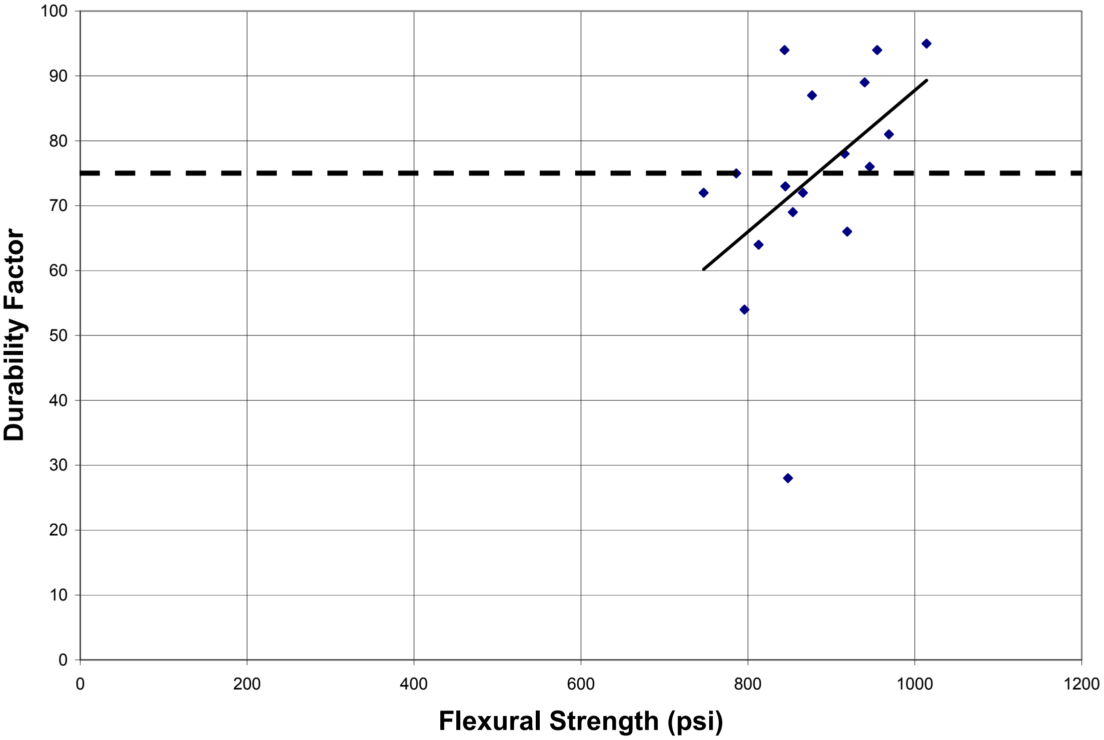


Figure 41: Durability Factor vs. Flexural Strength

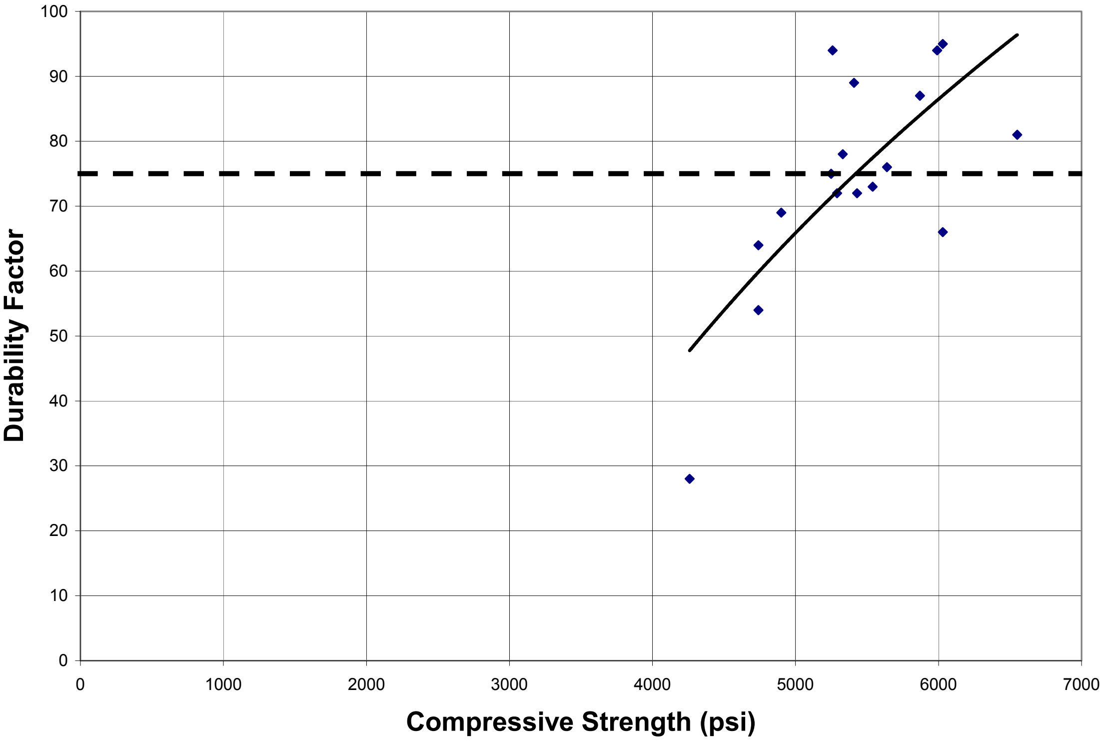


Figure 42: Durability Factor vs. Compressive Strength

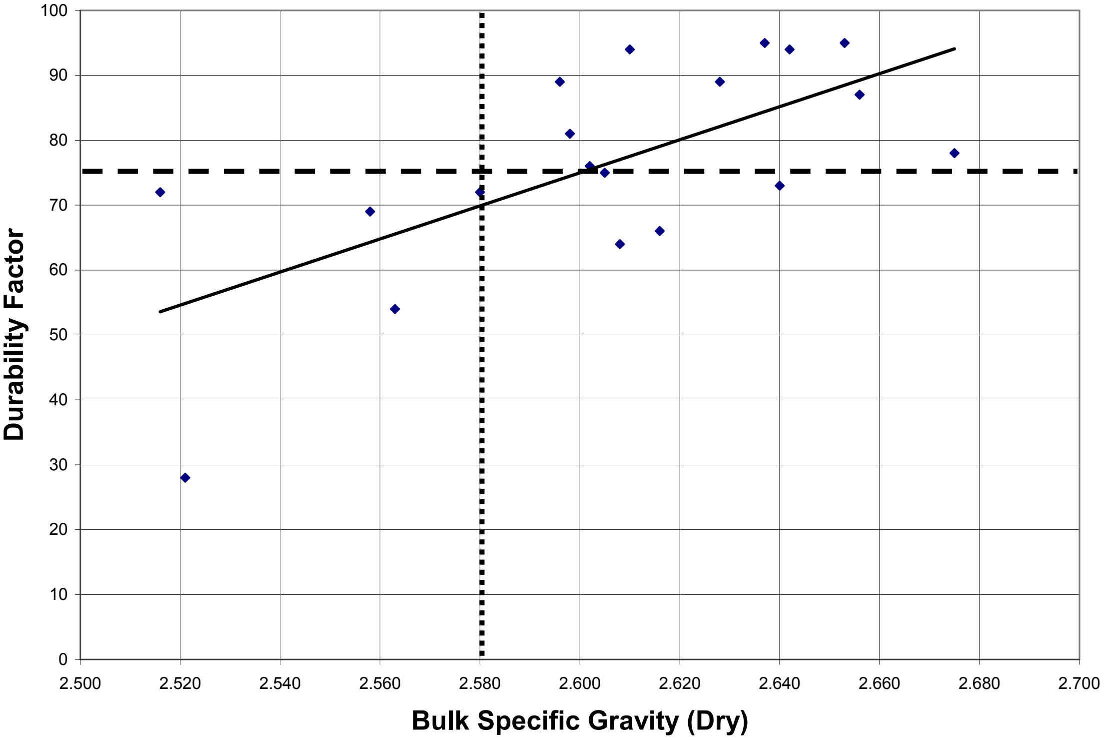


Figure 43: Durability Factor vs. Bulk Specific Gravity (Dry)

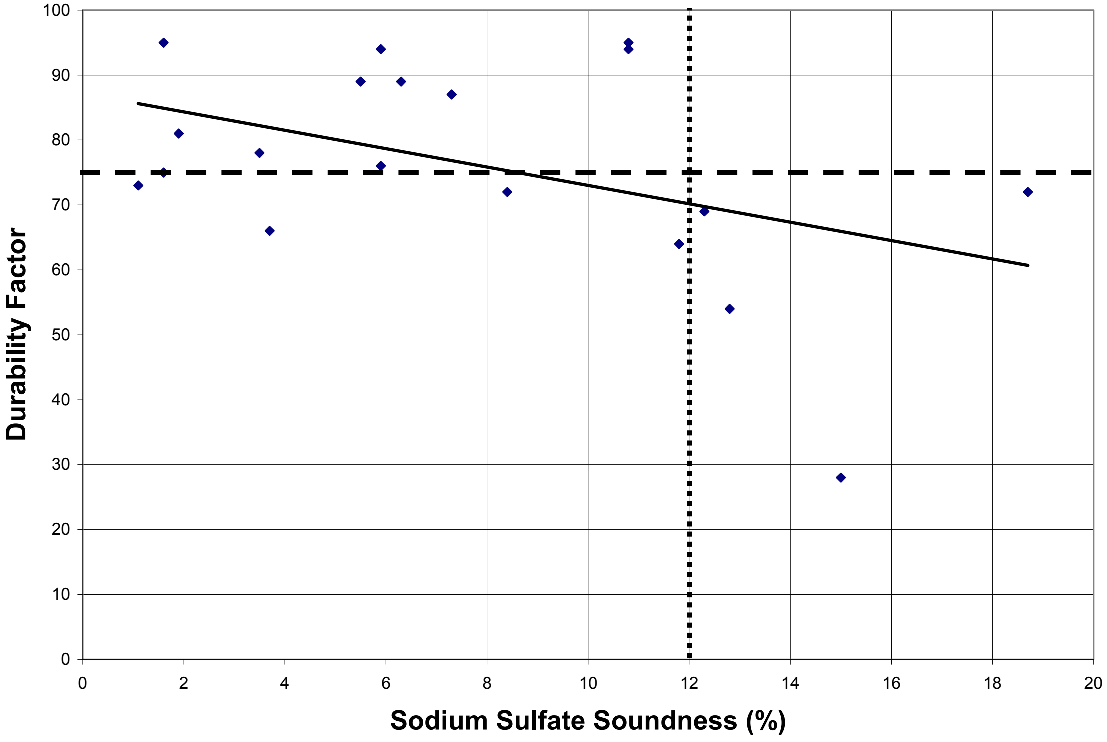


Figure 44: Durability Factor vs. Sodium Sulfate Soundness

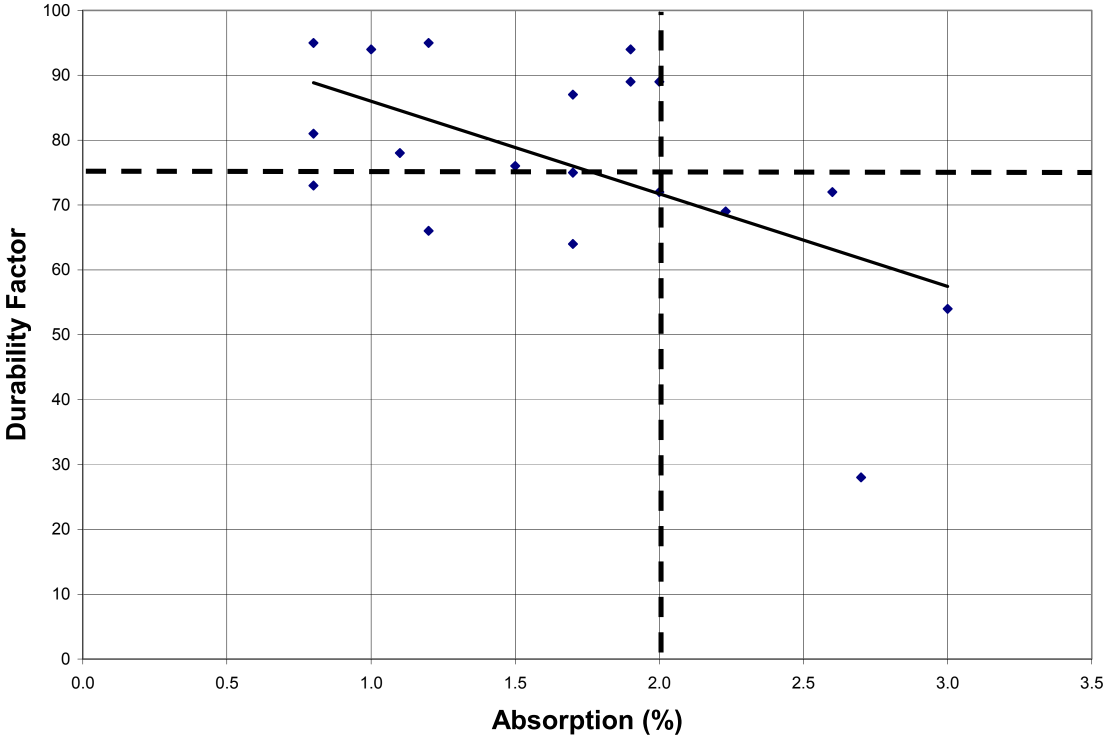


Figure 45: Durability Factor vs. Absorption

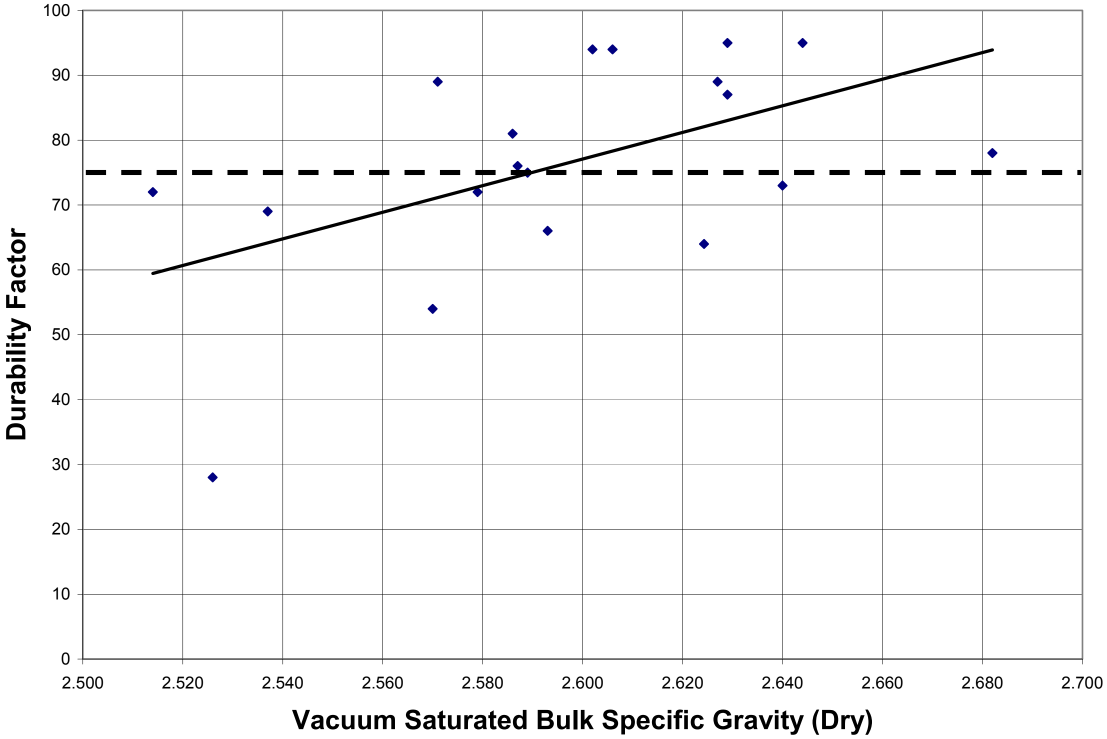


Figure 46: Durability Factor vs. Vacuum Saturated Bulk Specific Gravity (Dry)

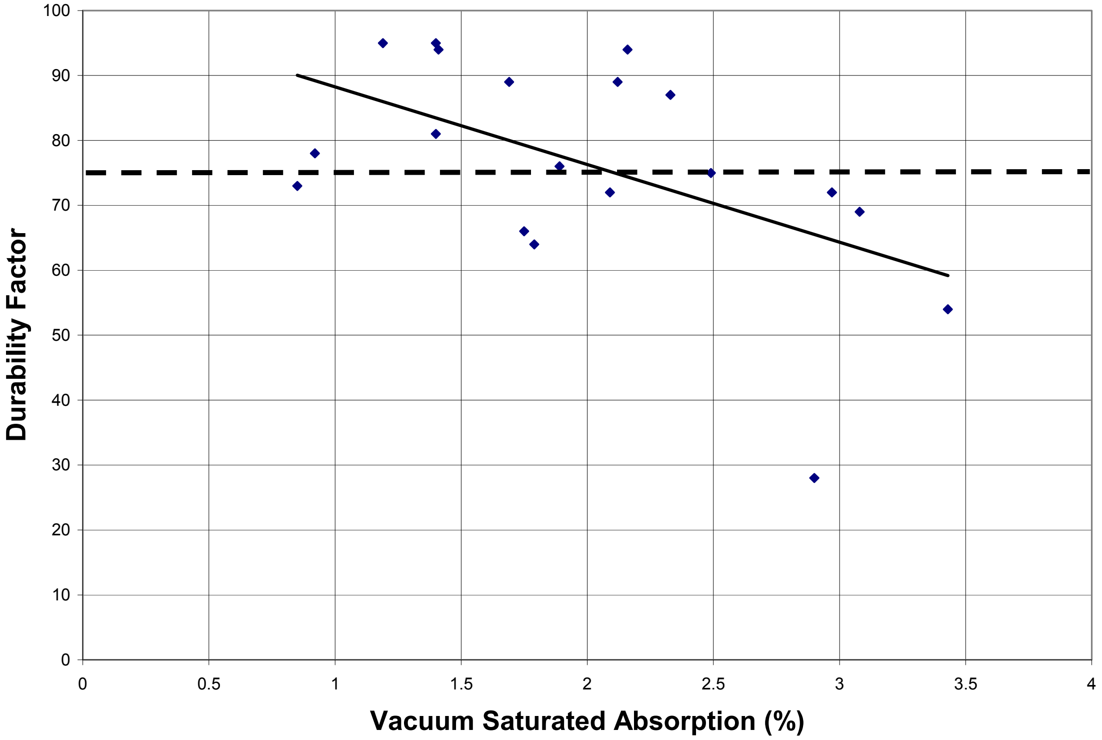


Figure 47: Durability Factor vs. Vacuum Saturated Absorption

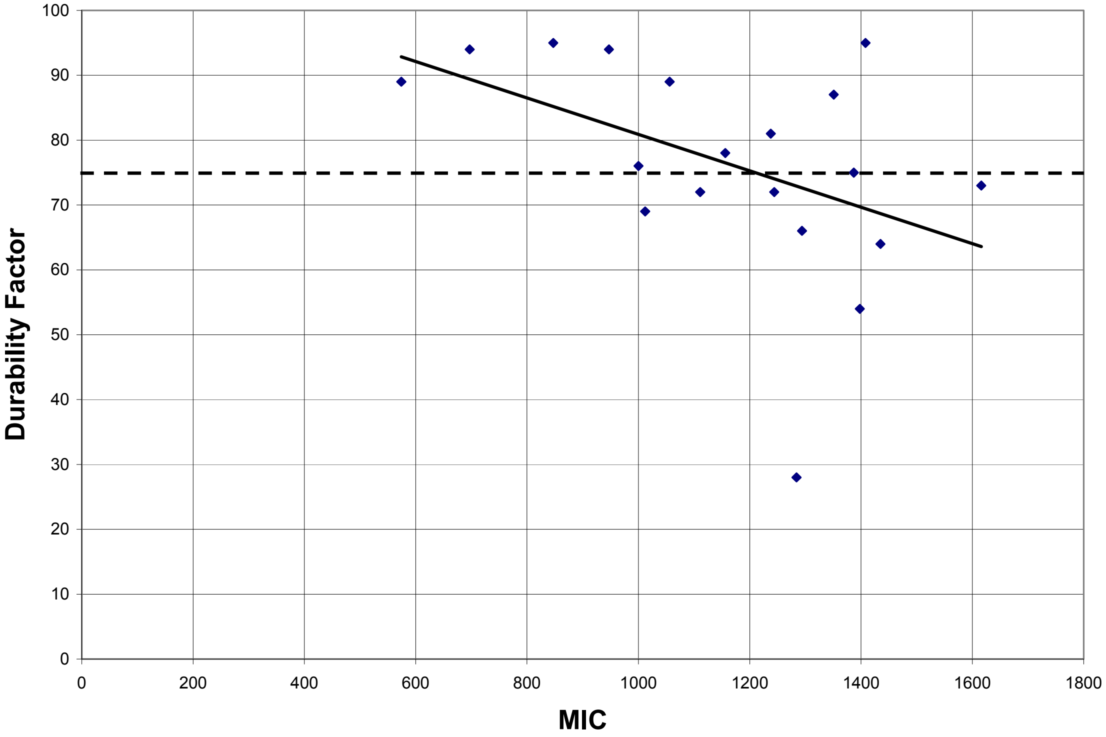


Figure 48: Durability Factor vs. MIC

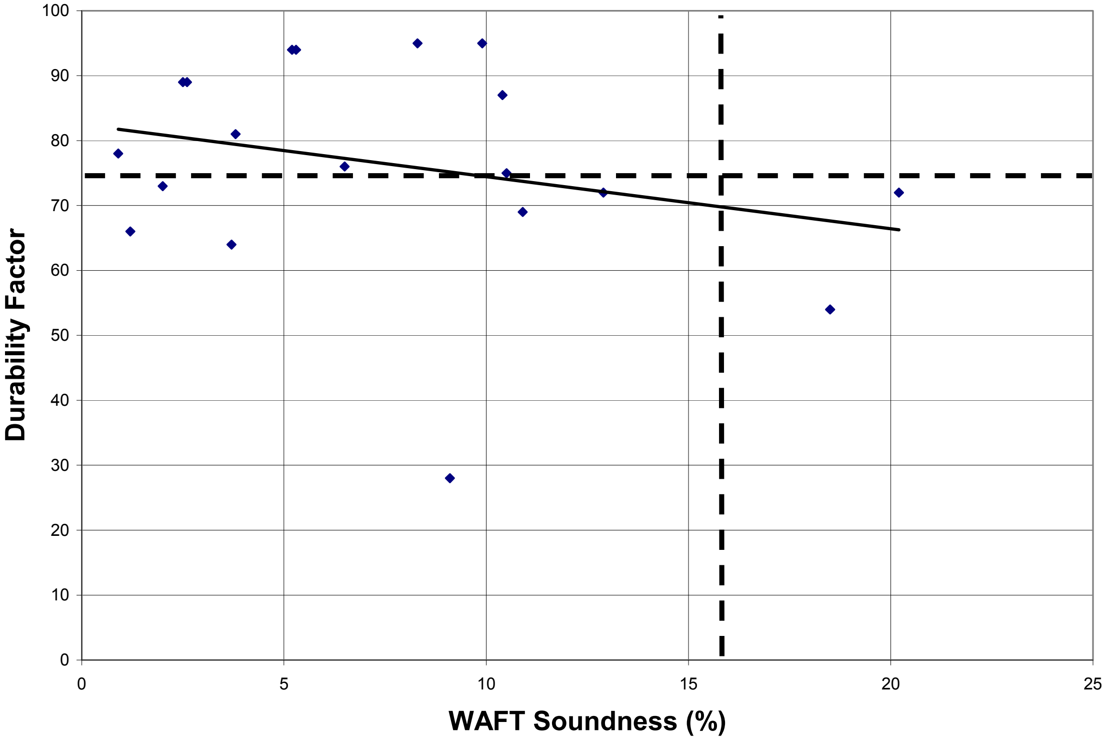


Figure 49: Durability Factor vs. WAFT Soundness

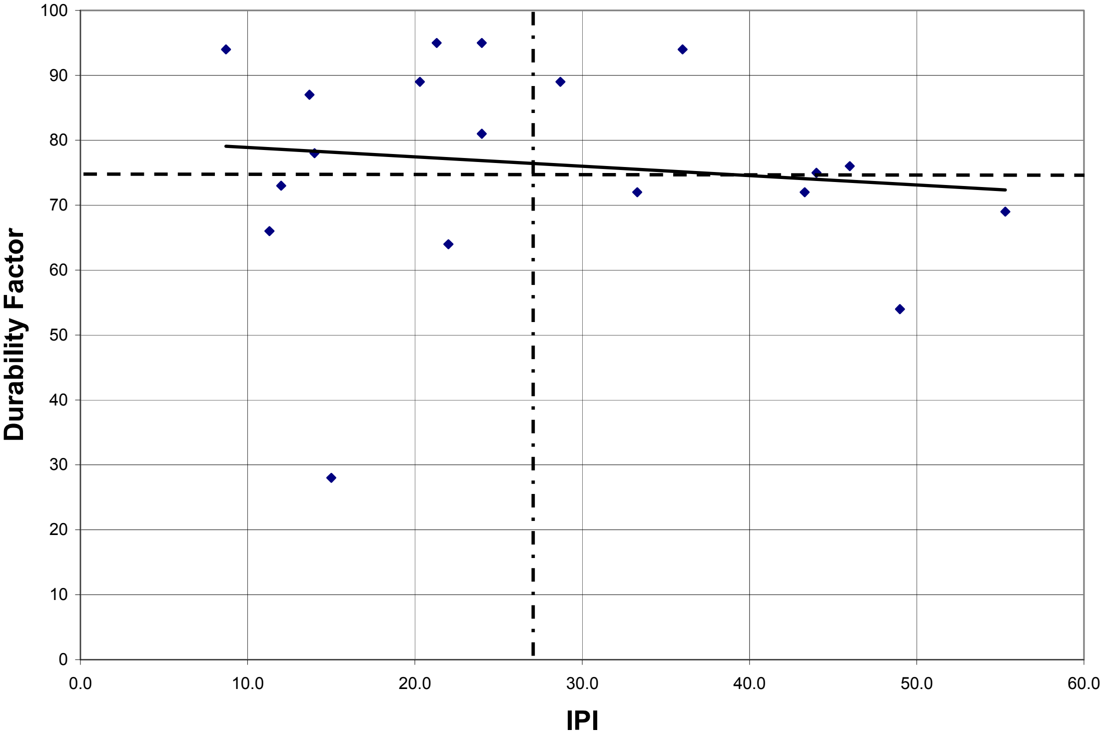


Figure 50: Durability Factor vs. IPI

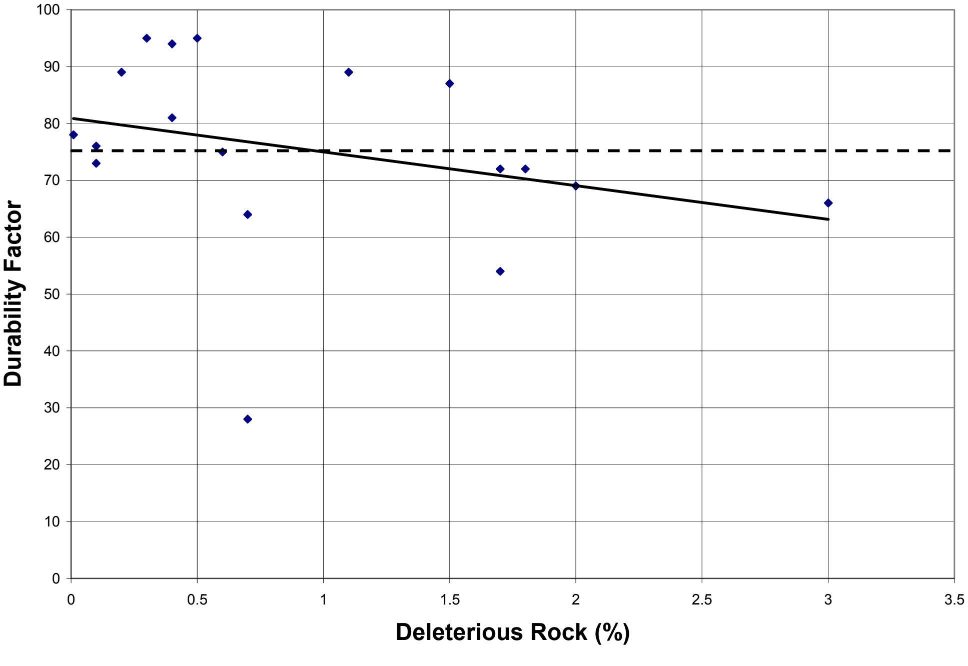


Figure 51: Durability Factor vs. Deleterious Rock

REGRESSION ANALYSIS

Methodology

In this study, regression models were sought that would accurately predict DF by one or more aggregate and/or concrete mixture characteristics. Thus, DF was the dependent variable and the aggregate/ concrete mixture characteristics were the independent variables. The dependent variable is also known as the “response variable”, and the independent variables are also known as “predictors” or “regressors”. If not included in an interaction, independent variables are also known as “main effects”. Several different types of regression models were desired, based on the sort of data that was to be included in each model. For instance, one type of model consisted of aggregate-only independent variables. Usually, model accuracy was sacrificed by using fewer or less definitive (but easier) test methods. The models presented herein are the most accurate within the constraints of each model type, and meeting several statistical acceptance criteria. Several statistics computer packages were used: JMP7®, MiniTab®, SigmaPlot®, and SAS®.

Step-wise regression in JMP7® was used for identification of possible models for further analysis. MiniTab® was also used in initial screening for choosing best models for a variety of number of main effects. The models were then checked in JMP7®, SigmaPlot®, and SAS®. Checking consisted of running certain statistical tests, and comparing the results to appropriate threshold acceptance criteria. The choice of threshold level of acceptance was arbitrary, but conformed to typical practice.

Model Acceptance Criteria

Seven statistical criteria were used for model acceptance: one criteria for ranking models and the other six for checking for possible problems.

R^2

The “R^2“ (coefficient of determination) of a regression model is a measure of the fit with the sample data. It is the proportion of Y variability that can be predicted from X in the sample (Schulman, 1992). As the R^2 increases, the fit of the model improves.

Adjusted R^2

The “adjusted R^2“ of a regression model is a measure of the fit with the population data. Adjusted R^2 is a superior statistic to R^2 during model selection because it takes into account the varying numbers of independent variables so as to not falsely inflate R^2. For each type of model, the one ultimately chosen in this study was the one with the highest adjusted R^2 that met all the criteria listed below. Reviewers who are more familiar with working with R^2 should note that adjusted R^2 values are always lower than R^2 values (predictions in the sample are always better than in the population), so one must adjust one's frame of reference.

Equation Significance

Each equation must show that the developed model fits the data, and thus is significant at a 0.01 level. The analysis of variance F-statistic will indicate this condition.

Term Significance

Each term in a regression equation must be significant at an $\alpha = 0.05$ level.

Multi-Collinearity

Multi-collinearity must be minimized in order to assure stability of the equation. For example, if two or more main effects are highly collinear, then unstable predictions may be made by the equation. Thus, only one of the collinear predictor variables should be allowed to remain in the equation. Multi-collinearity was assessed by two test statistics: Variance Inflation Factor (VIF) and Condition Number (CN). VIFs are measured for each variable in the equation. A threshold level generally preferred is 4 or less, with 5 being an upper limit. CNs are global; one CN is assigned to the entire equation. A desirable CN is 30 or less.

Undue Influence of Single Data Points

Single observations should not be allowed to influence the regression unduly. An observation in regression analysis is defined as all the data that predicts a single response value. In other words, an observation would be a row of data points (i.e. test results) in Table 10 that is associated with the single DF. Thus there were 19 observations in this study (17 observations when regressing with compressive and flexural strength, due to missing data). Any data point within a given observation could cause the excessive influence. Influence is measured by DFFITS (Difference in Fits), which is the change in a given predicted value when the observation being tested is removed from the data set and the model is re-fit to the remaining data. A desirable value of DFFITS used in this study was 2 standard deviations or less. High DFFITS values should be explored to determine if any action is deemed necessary in regard to rejecting an observation. Usually, a conservative approach is to retain the observation unless there is a compelling reason to reject an outlier.

Normality of Test Residuals

A residual is the error (difference) between an actual (observed or measured) single response variable value (e.g. DF) and the associated predicted value for a given regression model. The residuals should be normally distributed. Meeting normality criteria checks the assumption of regression modeling that residuals are indeed normal. Normality was checked with the Kolmogorov-Smirnov test at a significance level of 95%.

Constant Variance of Residuals

Residuals should also be checked to make sure that the magnitudes of the residuals are relatively uniform throughout the entire range of data. Again, this is just a check of another assumption that is part of regression analysis. Constant variance was checked with the Spearman rank correlation test at a significance level of 95%.

Regression Models

The following are the regression models that were considered to have the most application for MoDOT's use. MoDOT can choose the model(s) that will work best under various conditions. Options that are presented involve choosing test methods based on familiarity, willingness to start something new, equipment cost, sensitivity to test duration, and the level of accuracy that is considered acceptable, plus the overall predictive system ease of use.

A common set of test types usually surfaced as the best model for each category of model. Surprisingly, absorption or VSAbs almost never showed up because of being trumped by BSG or VSBSG. Also, efforts to force BSG to trump VSBSG always failed.

It is sometimes surprising which main effects (test methods) show up and which ones do not. A good one-on-one correlation with DF does not guarantee successful inclusion. And, if several main effects are highly correlated, only one will be allowed to remain, otherwise predictive instability may occur. Also, as statisticians caution, both sign and size of regression coefficients for linear equations (multipliers of the independent variables) may be counter-intuitive because of: 1) the scale-dependency of the coefficients, 2) correlations among predictors, and 3) influence of single observations (Schulman, 1992).

Although crossed terms representing interactions between main effects may increase the adjusted R^2, in the final analysis, crossed terms were not left in the final models for fear of creation of instability. With a larger dataset, inclusion of these types of terms (with resulting better-looking adjusted R^2 values) may be a more appropriate time to do so. In the future, if more data becomes available, say, during a verification exercise, use of interactive terms could be explored, with a resultant increase in accuracy.

Highest Adjusted R^2 Any Aggregate Test Method

Considering aggregate test methods only, the model with the greatest adjusted R^2 (0.830) and meeting all model test criteria was the following. MoDOT currently performs most of these tests, either routinely or for research purposes, except for VSBSG. However, the VSBSG initial vacuum saturation step is not considered onerous, and is very similar to the procedure in the Rice specific gravity test. The coefficient ”A” listed below is the intercept of the model , while the rest of the regression coefficients (B to F) are multipliers of the independent variables. For instance, 111.44810 would be multiplied times the log of WBM.

$$DF= A+B(LogWBM)+C(LogIPI)+D(logNaSO4)+E(LogVSBSG)+F(LogHudA) \tag{3}$$

where “Log” refers to the Log_{10} of the variable

Table 15: Statistical Summary: Model 1

Coefficient	Coefficient	p-value	VIF
A (intercept)	-10062.62287	<0.0001	-
B	111.44810	<0.0001	1.57
C	26.66355	0.010	1.77
D	-14.82523	0.020	1.60
E	2173.61536	<0.0001	2.23
F	156.42377	<0.0001	1.32
R^2	0.877	**F-Statistic**	<0.0001
Adj R^2	0.830	**CN(w/o intercept)**	2.90
Normality	pass	**Constant Variance**	pass
		DFFITS	none

Note that several of the major factors in freeze-thaw durability are present: IPI and VSBSG representing pore characteristics, HudA representing pore length, and WBM and NaSO4 for possibly representing tensile strength/elastic accommodation.

The above statistics show that the model met the following criteria: 1) the model has the highest possible adjusted R^2 while best meeting other criteria, 2) the model is significant at the 0.01 level as indicated by the analysis of variance F-statistic, 3) all major effects are significant at the 0.05 level as indicated by the p-values, 4) no problems with multi-collinearity, as indicated by VIF’s being less than 4 to 5 and the CN (without intercept) is below 30, 5) the model passed the test for normality of residuals, 6) the model passed the test for constant variance of residuals, and 7) no single observation (row of data) exerted undue influence on the model, as indicated by each observation's DFFITS being below 2. In Fig. 52 is shown a plot of predicted vs. measured DF values.

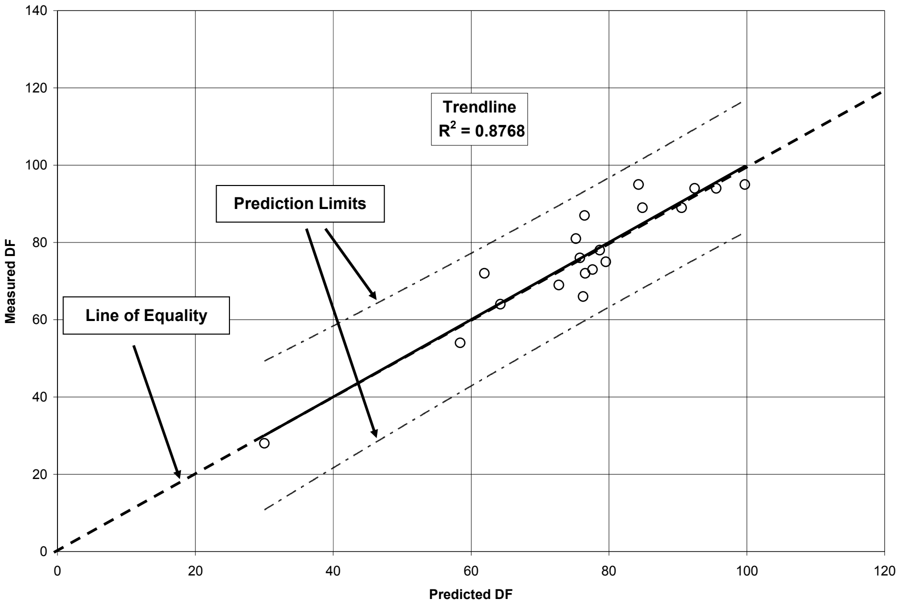


Figure 52: Measured vs. Predicted Durability Factor: Full Aggregate Model (Model 1)

Highest Adjusted R^2 Any Short Duration Aggregate Test Method

Again considering aggregate tests only, the model with the greatest adjusted R^2 (0.785) but not requiring tests that entail long testing times (NaSO4 and WAFT), and meeting all model test criteria was the following. MoDOT currently performs three of these tests, either routinely or for research purposes, but not VSBSG and ACV. VSBSG is just a modified T 85 test, but ACV would require the use of a compression machine (such as for breaking concrete cylinders) with a capacity of approximately 100,000 lb.

$$DF = A + B(MD) + C(WBM) + D(ACV) + E(VSBSG) + F(MIC) \quad (4)$$

Table 16: Statistical Summary: Model 2

Coefficient	**Coefficient**	**p-value**	**VIF**
A (intercept)	-1325.77661	<0.0001	-
B	2.77440	0.0026	4.02
C	1.69482	0.0113	3.02
D	-2.56308	0.0401	2.93
E	554.87969	<0.0001	3.91
F	-0.05081	<0.0001	1.54
R^2	0.845	**F-Statistic**	<0.0001
Adj R^2	0.785	**CN(w/o intercept)**	4.49
Normality	pass	**Constant Variance**	pass
		DFFITS	none

Note that several of the major factors in freeze-thaw durability are present: VSBSG representing pore characteristics, MIC representing pore length, and ACV, MD, and WBM representing tensile strength/elastic accommodation.

The above statistics show that the model met the following criteria: 1) the model has the highest possible adjusted R^2 while meeting other criteria, 2) the model is significant at the 0.01 level as indicated by the analysis of variance F-statistic, 3) all major effects are significant at the 0.05 level as indicated by the p-values, 4) no problems with multi-collinearity, as indicated by VIF's being less than 4 to 5 and the CN (without intercept) is below 30, 5) the model passed the test for normality of residuals, 6) the model passed the test for constant variance of residuals, and 7) no single observation (row of data) exerted undue influence on the model, as indicated by each observation's DFFITS being below 2.

In Fig. 53 is shown a plot of predicted vs. measured DF values.

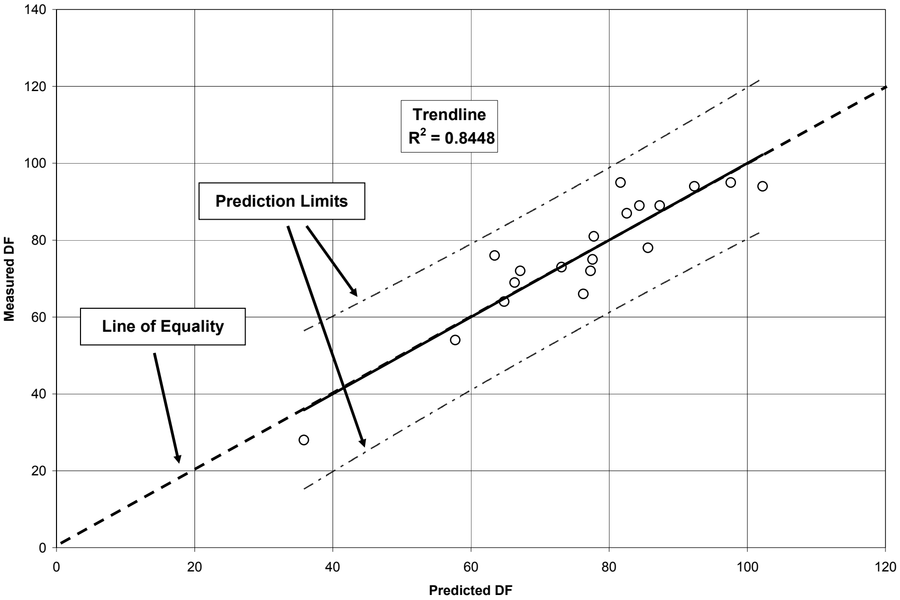


Figure 53: Measured vs. Predicted Durability Factor: Short Duration Aggregate Model (Model 2)

A second model in this category with a slightly lower adjusted R^2 but not requiring ACV, which MoDOT does not currently perform, is as listed below. This suite of tests involves relatively simple methods of short duration.

$$DF = A + B(LogWBM) + C(LogIPI) + D(LogVSBSG) + E(LogHudA) \qquad (5)$$

Table 17: Statistical Summary: Model 3

Coefficient	Coefficient	p-value	VIF
A (intercept)	-1199.77871	<0.0001	-
B	108.80764	0.0004	1.57
C	26.93942	0.0237	1.77
D	2510.84589	<0.0001	1.91
E	129.46521	0.0009	1.14
R^2	0.810	**F-Statistic**	<0.0001
Adj R^2	0.756	**CN(w/o intercept)**	2.48
Normality	pass	**Constant Variance**	pass
		DFFITS	none

Note that several of the major factors in freeze-thaw durability are present: IPI and VSBSG representing pore characteristics, HudA representing pore length, and WBM for representing tensile strength/elastic accommodation.

The above statistics show that the model met the following criteria: 1) the model has the highest possible adjusted R^2 while meeting other criteria, 2) the model is significant at the 0.01 level as indicated by the analysis of variance F-statistic, 3) all major effects are significant at the 0.05 level as indicated by the p-values, 4) no problems with multi-collinearity, as indicated by VIF's being less than 4 to 5 and the CN (without intercept) is below 30, 5) the model passed the test for normality of residuals, 6) the model passed the test for constant variance of residuals, and 7) no single observation (row of data) exerted undue influence on the model, as indicated by each observation's DFFITS being below 2.

In Fig. 54 is shown a plot of predicted vs. measured DF values.

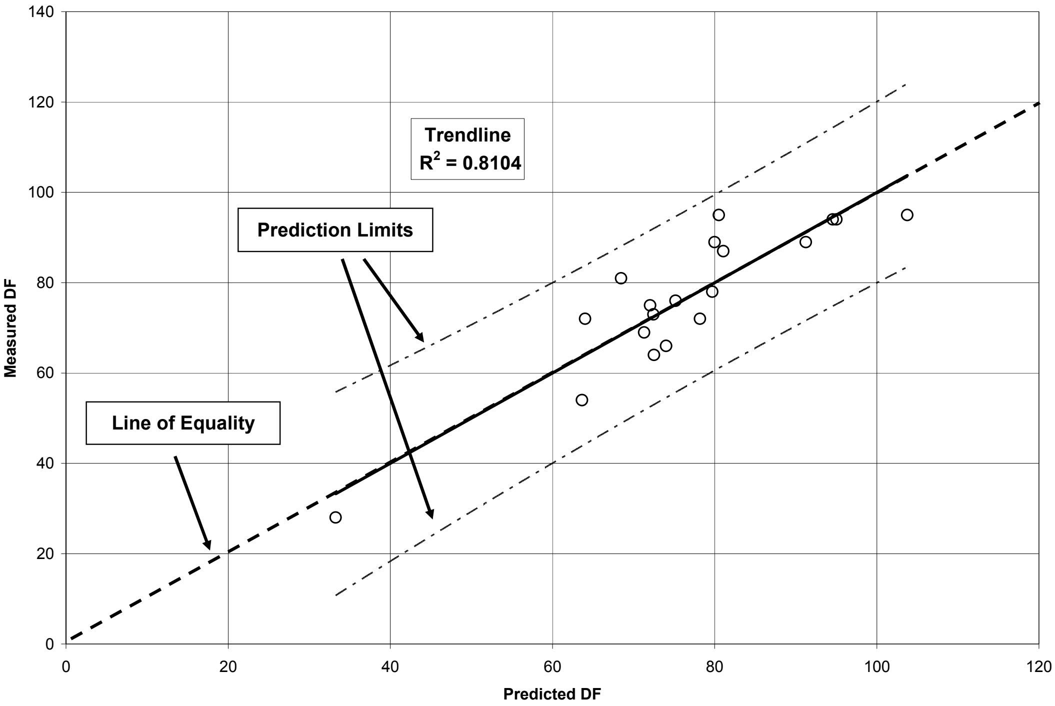


Figure 54: Measured vs. Predicted Durability Factor: Alternate Short Duration Aggregate Model (Model 3)

Highest Adjusted R^2 Short Duration MoDOT Aggregate Test Methods

Again considering aggregate tests only, the model with the greatest adjusted R^2 (0.747) but not requiring tests that entail long testing times (NaSO4 and WAFT) or tests not currently being performed by MoDOT, such as VSBSG, VSAbs, and PLS, and meeting all but one model test criteria was the following:

$$SqRtDF = A + B(LogWBM) + C(LogIPI) + D(LogBSG) + E(LogHudA) \qquad (6)$$

where "SqRtDF" refers to the square root of DF

Table 18: Statistical Summary: Model 4

Coefficient	**Coefficient**	**p-value**	**VIF**
A (intercept)	-62.69700	<0.0001	-
B	4.74903	0.0034	1.24
C	1.85638	0.0163	1.72
D	145.18788	<0.0001	1.56
E	5.07676	0.0245	1.15
R^2	0.804	**F-Statistic**	<0.0001
Adj R^2	0.747	**CN(w/o intercept)**	2.21
Normality	pass	**Constant Variance**	pass
		DFFITS	-2.882

Note that several of the major factors in freeze-thaw durability are present: IPI and BSG representing pore characteristics, HudA representing pore length, and WBM for representing tensile strength/elastic accommodation.

The above statistics show that the model met the following criteria: 1) the model has the highest possible adjusted R^2 while meeting other criteria, 2) the model is significant at the 0.01 level as indicated by the analysis of variance F-statistic, 3) all major effects are significant at the 0.05 level as indicated by the p-values, 4) no problems with multi-collinearity, as indicated by VIF's being less than 4 and the CN (without intercept) is below 30, 5) the model passed the test for normality of residuals, and 6) the model passed the test for constant variance of residuals,. One criterion was not met: the 84SRE203 sample observation had a DFFITS of -2.882, indicating that it may have a somewhat stronger influence on the model than would be preferred. It was decided to leave this model as the choice for this criteria section because the regression model itself would not change significantly by inclusion of the observation in question, and the model was the least problematic of the best models,.

In Fig. 55 is shown a plot of predicted vs. measured DF values.

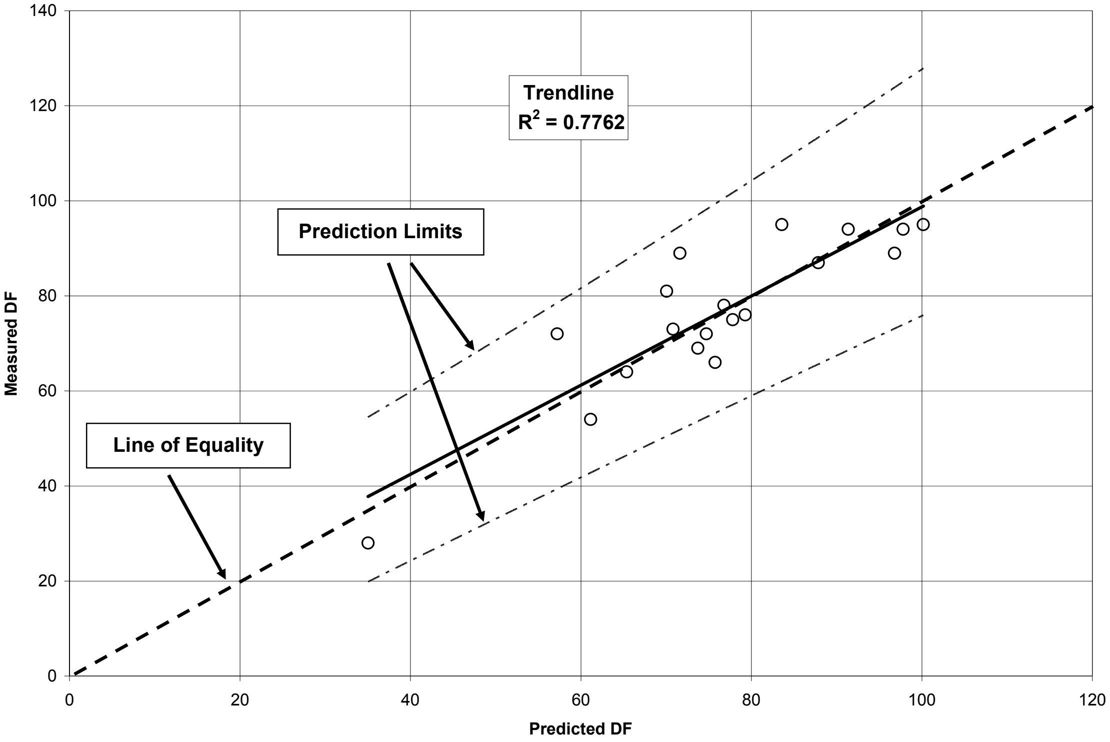


Figure 55: Measured vs. Predicted Durability Factor: Short Duration MoDOT Tests Aggregate Model (Model 4)

Highest Adjusted R^2 Any Aggregate and Strength Test Method

Looking at all aggregate tests and strength tests, the model with the greatest adjusted R^2 (0.954) and meeting all model test criteria was the following. This will be the highest adjusted R^2 model in this study, because of the addition of the paste component and the disregard of test method difficulty. This is what brings the study to the point of offering a hybrid of Option A and Option B, that is, inclusion of a paste characteristic. In all models attempted, 28 day compressive strength was always the emergent strength parameter. However, the model does contain VSBSG and PLS which MoDOT does not currently perform, and Ratio2, which is not trivial to calculate.

$$\mathit{Log\ DF} = A + B(\mathit{LogWBM}) + C(\mathit{LogIPI}) + D(\mathit{LogVSBSG}) + E(\mathit{LogHudA}) + F(\mathit{LogRatio2}) + G(\mathit{LogComp}) + H(\mathit{LogPLS}) \tag{7}$$

Table 19: Statistical Summary: Model 5

Coefficient	Coefficient	p-value	VIF
A (intercept)	-10.57007	<0.0001	-
B	0.61927	0.0004	3.35
C	030775	<0.0001	2.21
D	17.26864	<0.0001	3.59
E	0.55007	0.0023	1.43
F	-0.6094	0.0101	1.27
G	1.07475	0.0004	1.89
H	-.026403	0.0732	1.76
R^2	0.974	**F-Statistic**	<0.0001
Adj R^2	0.954	**CN(w/o intercept)**	4.10
Normality	pass	**Constant Variance**	pass
		DFFITS	2.425

Note that several of the major factors in freeze-thaw durability are present: IPI and VSBSG representing pore characteristics, HudA and Ratio2 representing pore length, PLS and WBM for representing tensile strength/elastic accommodation, and Comp representing concrete strength.

The above statistics show that the model met the following criteria: 1) the model has the highest possible adjusted R^2 while meeting most other criteria, 2) the model is significant at the 0.01 level as indicated by the analysis of variance F-statistic, 3) all major effects are significant at the 0.05 level as indicated by the p-values (except LogPLS which was slightly high at 0.07), 4) no problems with multi-collinearity, as indicated by VIF's being less than 4 to 5 and the CN (without intercept) is below 30, 5) the model passed the test for normality of residuals, 6) the model passed the test for constant variance of residuals, and 7) no single observation (row of data) exerted undue influence on the model, as indicated by each observation's DFFITS being below 2, except for sample 83MA0234 at

2.425. It was decided to leave the observation in the model because, by examination, the regression model itself would not change significantly by the removal of the observation.

In Fig. 56 is shown a plot of predicted vs. measured DF values.

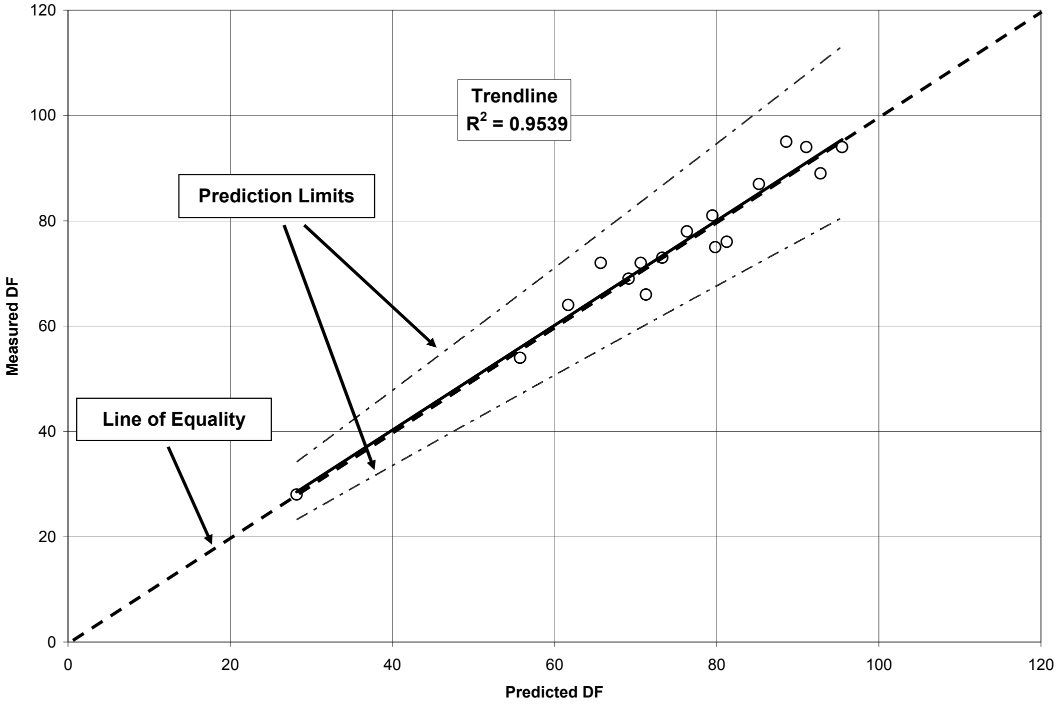


Figure 56: Measured vs. Predicted Durability Factor: Aggregate and Paste Tests Full Model (Model 5)

A second model fitting this category had a somewhat lower adjusted R^2 (0.906) but was much simpler:

$$\mathit{LogDF} = A + B(\mathit{LogWBM}) + C(\mathit{LogIPI}) + D(\mathit{LogVSBSG}) + E(\mathit{LogHudA}) + F(\mathit{LogComp}) \tag{8}$$

Table 20: Statistical Summary: Model 6

Coefficient	**Coefficient**	**p-value**	**VIF**
A (intercept)	-10.28526	<0.0001	-
B	0.62688	0.0011	2.68
C	0.26166	0.0004	1.93
D	16.08879	<0.0001	3.41
E	0.74699	0.0010	1.17
F	1.06713	0.0030	1.88
R^2	0.936	**F-Statistic**	<0.0001
Adj R^2	0.906	**CN(w/o intercept)**	3.71
Normality	pass	**Constant Variance**	pass
		DFFITS	2.258, -2.584

Note that several of the major factors in freeze-thaw durability are present: IPI and VSBSG representing pore characteristics, HudA representing pore length, WBM representing tensile strength/elastic accommodation, and Comp representing concrete strength.

The above statistics show that the model met the following criteria: 1) the model had a high adjusted R^2 while meeting other criteria, 2) the model is significant at the 0.01 level as indicated by the analysis of variance F-statistic, 3) all major effects are significant at the 0.05 level as indicated by the p-values, 4) no problems with multi-collinearity, as indicated by VIF's being less than 4 to 5 and the CN (without intercept) is below 30, 5) the model passed the test for normality of residuals, 6) the model passed the test for constant variance of residuals, and 7) no single observation (row of data) exerted undue influence on the model, as indicated by each observation's DFFITS being below 2,except that 83MA0234 (2.258) and 86L2R021 (-2.584) were somewhat high. It was decided to leave the observations in the model because, by examination, the regression model itself would not change significantly by the removal of the observations.

In Fig. 57 is shown a plot of predicted vs. measured DF values.

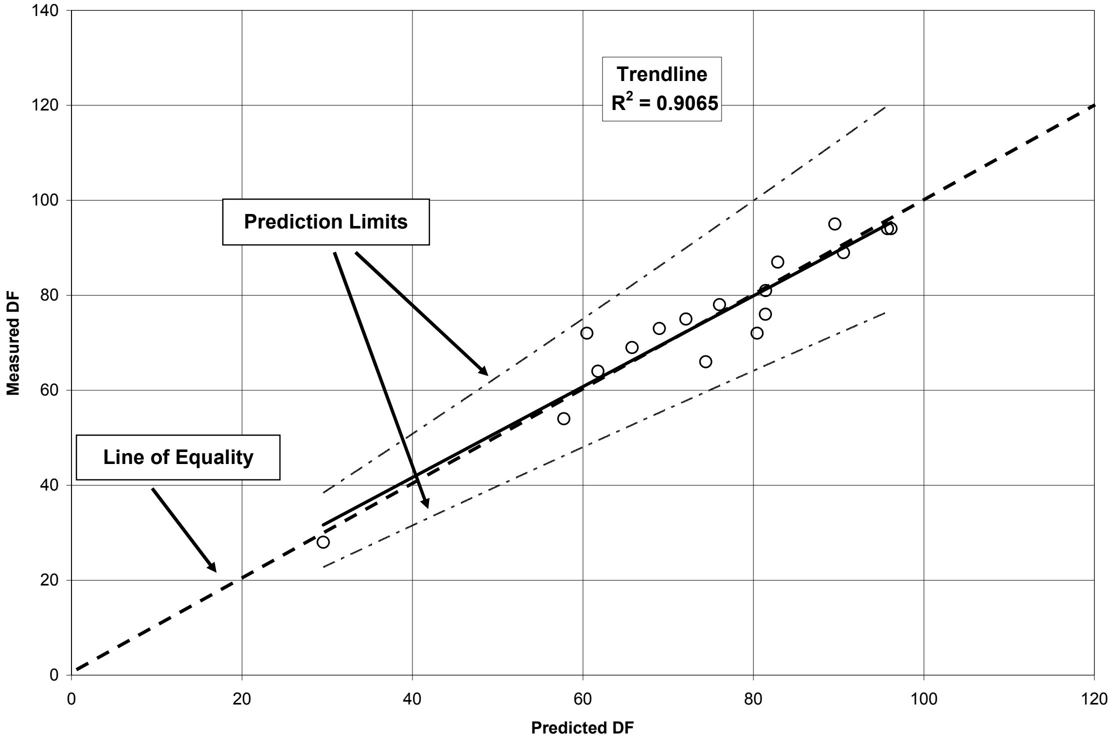


Figure 57: Measured vs. Predicted Durability Factor: Alternate Aggregate and Paste Tests Full Model (Model 6)

Highest Adjusted R^2 Any MoDOT Aggregate and Strength Test Method

The model with the greatest adjusted R^2 (0.792) but not requiring tests that are not currently being performed by MoDOT, such as VSBSG and VSAbs, and meeting all model test criteria, was the following:

$$SqRtDF = A + B(LogWBM) + C(LogIPI) + D(LogBSG) + E(LogHudA) + F(LogComp) \tag{9}$$

Table 21: Statistical Summary: Model 7

Coefficient	Coefficient	p-value	VIF
A (intercept)	-76.51746	<0.0001	-
B	3.09027	0.078	2.02
C	1.60159	0.030	1.83
D	110.61218	0.003	3.29
E	6.12113	0.015	1.15
F	8.14781	0.067	2.34
R^2	0.857	**F-Statistic**	0.0002
Adj R^2	0.792	**CN(w/o intercept)**	3.70
Normality	pass	**Constant Variance**	pass
		DFFITS	-2.141

Note that several of the major factors in freeze-thaw durability are present: IPI and BSG representing pore characteristics, HudA representing pore length, WBM for possibly representing tensile strength/elastic accommodation, and Comp representing concrete strength.

The above statistics show that the model met the following criteria: 1) the model has the highest possible adjusted R^2 while meeting other criteria, 2) the model is significant at the 0.01 level as indicated by the analysis of variance F-statistic, 3) all major effects are significant at the 0.05 level as indicated by the p-values, except for two that were slightly higher than preferred (LogWBM and LogComp), 4) no problems with multi-collinearity, as indicated by VIF's being less than 4 and the CN (without intercept) is below 30, 5) the model passed the test for normality of residuals, 6) the model passed the test for constant variance of residuals, and 7) no single observation (row of data) exerted undue influence on the model, as indicated by each observation's DFFITS being below 2, except for 84SRE203 being slightly high. It was decided to leave the observation in the model because, by examination, the regression model itself would not change significantly by the removal of the observation.

In Fig. 58 is shown a plot of predicted vs. measured DF values.

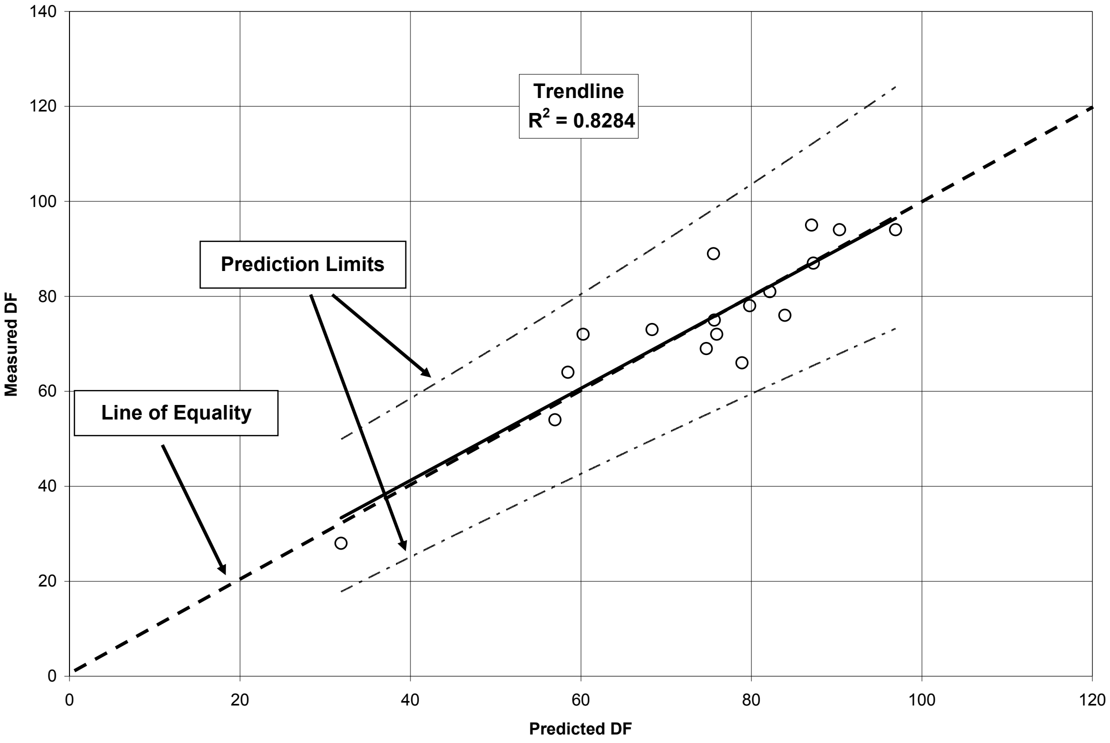


Figure 58: Measured vs. Predicted Durability Factor: Aggregate and Paste MoDOT Tests Full Model (Model 7)

Summary

Seven models have been presented in this section, each offering advantages and disadvantages. There are tradeoffs: usually accuracy is sacrificed by choosing models with simpler tests, tests with shorter duration time, and tests that are more familiar. As mentioned previously, the models need verification. This should be done before any model is implemented.

Table 22 is a summary of the seven models, arranged in order of adjusted R^2.

Table 22: Models in order of adjusted R^2

Model	Adjust R^2	Test Methods
5	0.954	WBM, IPI, VSBSG, HudA, Comp, PLS, Ratio2
6	0.906	WBM, IPI, VSBSG, HudA, Comp
1	0.830	WBM, IPI, VSBSG, HudA, $NaSO_4$
7	0.792	WBM, IPI, BSG, HudA, Comp
2	0.785	WBM, VSBSG, MD, ACV, MIC
3	0.756	WBM, IPI, VSBSG, HudA
4	0.747	WBM, IPI, BSG, HudA

FLOWCHART ACCEPTANCE

Threshold Limit Development

Another approach, besides prediction of DF by regression, is to create a system of threshold limits for several key test methods. Thus, if a given mix exhibits values that exceed the threshold limits, the probability of its DF being greater than 75 would be low. The test methods included in the threshold system are ones which MoDOT currently performs, in some fashion.

Before doing this, the effect of gradation must be taken into consideration. In general, the longer an aggregate particle's pore length, the greater the chance of internal pressure (due to freezing) building up and affecting the aggregate. Typically, NMS is used to easily characterize pore length. However, NMS alone does not take into account the amount of aggregate in the concrete mixture that is possibly above the critical size of the aggregate. Several parameters that would measure the large particle contribution were tried in this study, such as the volume of plus ¾ in. aggregate in a concrete mix, expressed as the percent of total volume. Other parameters attempted were various indices calculated by weighting the amount of aggregate sizes retained on the ½, ¾, and 1 in. sieves. These efforts have been discussed previously.

The other consideration of NMS involves the make-up of the concrete mixtures used in the T 161 tests in this study. Some aggregate types that were tested in the T 161 tests utilized a fine gradation, usually as a ½ in. NMS, with zero material retained on the ¾ in. sieve. Presumably this was done because the aggregate exhibited poor freeze-thaw behavior, and thus was tested with a smaller NMS. This would boost the DF values above what they would have been if tested in a ¾ in. or larger NMS. Thus, aggregate test results may not predict these high DF values sensibly. Consequently, the study data set was divided into two subsets: aggregates with no plus ¾ in. material, and aggregates with some plus ¾ in. material (potentially frost-susceptible).

Looking at the coarser subset, Table 23 shows the actual threshold limits that delineated DF values falling below 75. Both compressive and flexural strengths are in the pre-T 161 testing condition. Plots of various test method results against DF helped delineate where the limits should fall.

Table 23: Actual threshold limits to achieve DF of 75

Test	Actual Threshold
Absorption (%)	2.0 max.
BSG	2.580 min.
IPI	29 max.
MD (%)	25 max.
NaSO4 (%)	11 max.
VSAbs (%)	2.5 max.
VSBSG	2.570 min.
WAFT (%)	12 max
Compressive strength (psi)	5200 min.
Flexural strength (psi)	840 min.
Absorption (%) plus IPI	1.5 max + 19 max

A more generalized set of threshold values was created from the above limits. These are shown in Table 24.

Table 24: Threshold limits

Test Method	Threshold Limit
NaSO4 (%)	11 max.
WAFT (%)	16 max.
Absorption (%)	2.0 max.
IPI	27 max.
LAA (%), if Volume + ¾ in. exceeds 10%	35 max.
Volume + ¾ in. (%)	10 max.
WBM (%)	30 Max.
MD (%)	25 max.
Absorption (%) plus IPI	1.5 max + 19 max

The 11 aggregates in this study that had some plus ¾ in material were used to create these thresholds. Because of the varied nature of each aggregate in regard to T 161 behavior, different tests were needed to exclude different aggregates. Table 25 shows the seven coarser aggregates that were rejected by the system and the tests that were associated which caused rejection. The four coarser aggregates that had DF's greater than 75 (86R3M025, 80MA0051, 89TCR067, 86R3M029) did not have any test values exceeding the above limits.

Consideration should also be given to using flexural or compressive strength limits: suggested minimum values of 800 and 5000 psi for flexural and compressive strength, respectively. However, the issue with using concrete strengths is that the data is based on the particular mixture components utilized in MoDOT's Central Lab T 161 testing program, including proportions, sand source, cement brand and type, and air entraining admixture brand and dosage, with resulting air content. For the same aggregate, changes in the above may give different strengths.

Table 25: Aggregates with + ¾ in. material with associated tests causing rejection

Material (DF)	**$NaSO_4$ 11%**	**WAFT 16%**	**Abs 2.0%**	**IPI 27**	**LAA 35% Vol +3/4 >10 %**	**WBM 30%**	**Volume % + ¾ in >10%**	**Abs 1.5% + IPI 19**	**Comp 5000 psi**	**Flex 800 psi**
85RDP040 (54)	x	x	x	x				x	x	x
86L2R020 (64)	x							x	x	
85DGG007 (66)						x				
86R3M028 (69)	x		x	x				x		
83MA0234 (72)	x		x	x		x		x		x
88MA0024 (73)					x		x			
85DLR012 (75)				x				x		x

Application of the Threshold Limit System

The system would be applied in the following manner: the volume of + ¾ in material in the mix would be calculated, and the above test results analyzed. If any threshold limit was exceeded, the DF could be expected to be equal to or below 75. In cases where DF's below 75 were predicted, and there was still interest in utilizing the aggregate, verification could be done by actually performing the T 161 test.

Threshold Limit System Verification

MoDOT supplied data that was not included in this study (38 samples representing 25 ledges). Running the results through the above system, 30 materials with DFs less than 75 were successfully rejected by this system. However, six materials that should have passed (i.e. DFs > 75) were falsely rejected. Three of the six were tested with little or no + ¾ in. material and one was unknown. Two of the six contained some + ¾ in. material. Testing a gradation with no larger coarse aggregate would cloud the interpretation of the DF levels. If the aggregates would have been tested with a NMS of ¾ in. or larger, those aggregates may have tested with a DF less than 75 and would have been correctly rejected by the system. In any event, if desired, these six could be restored to acceptance by successful T 161 referee testing. Figs. 59 through 64 depict where the data fell. For instance, in Fig. 59, the upper left quadrant contains aggregates that had DF's greater than 75 and were successfully accepted. The lower right quadrant shows aggregates that had DF's less than 75, and were correctly rejected. The lower left quadrant is where the aggregates should have been rejected, but were not. However, by applying other test threshold criteria to these aggregates, most of them would eventually be successfully rejected. The upper right quadrant shows one aggregate that was incorrectly rejected. No system is perfect. Of more concern are the two aggregates that should have been rejected by the system, but were not.

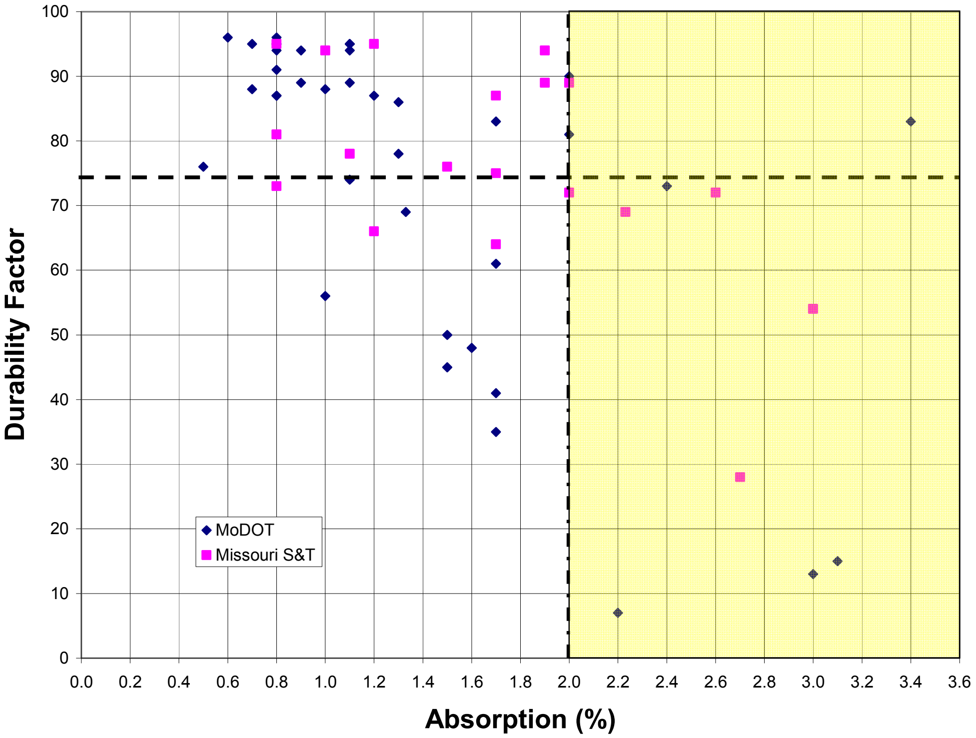


Figure 59: Absorption Threshold Limit

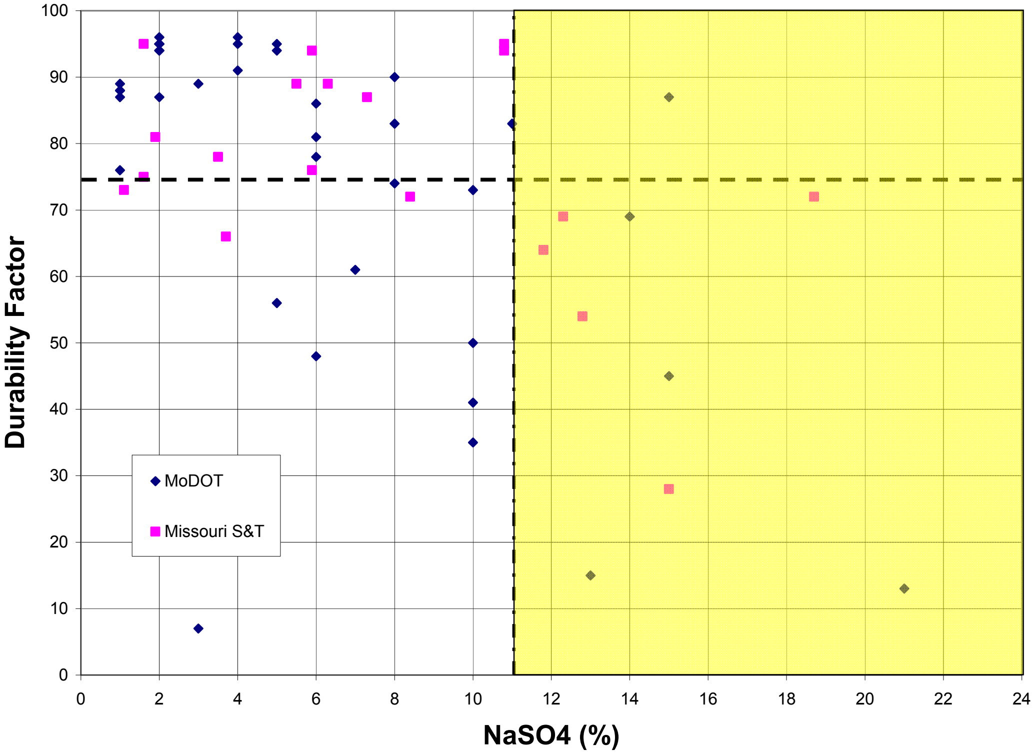


Figure 60: Sulfate Soundness Threshold Limit

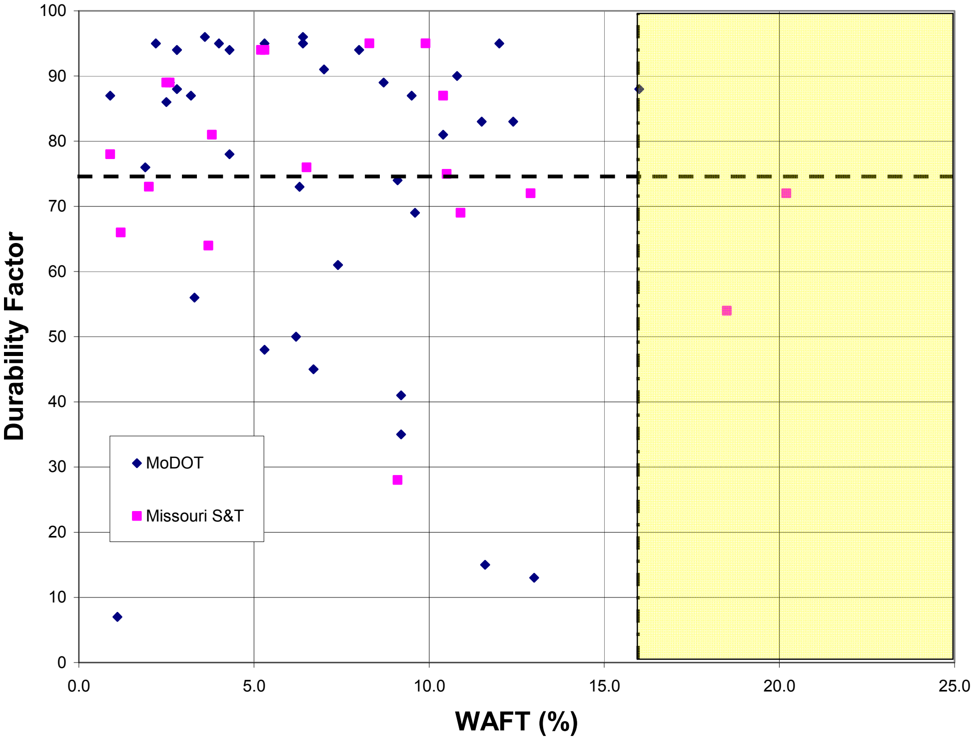


Figure 61: WAFT Threshold Limit

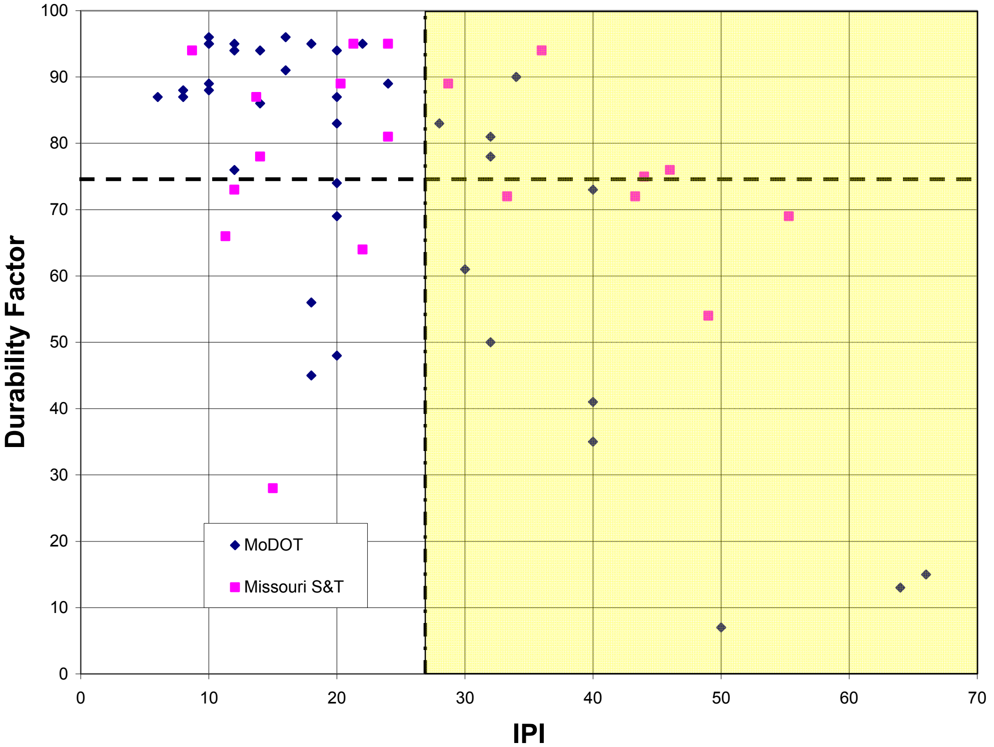


Figure 62: IPI Threshold Limit

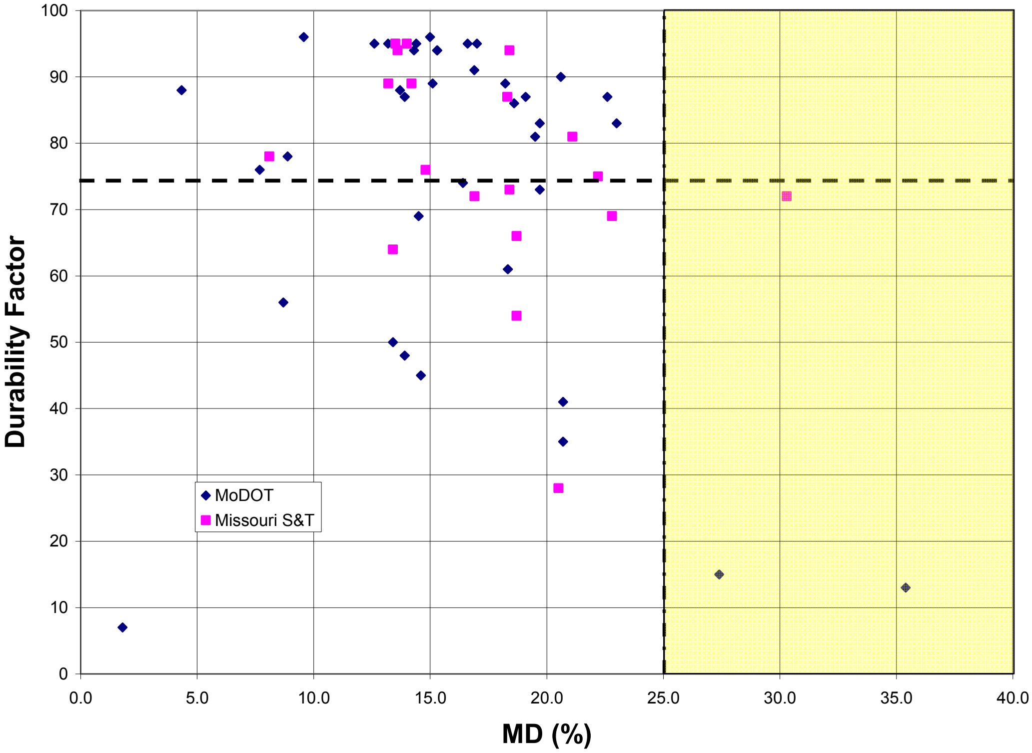


Figure 63: MD Threshold Limit

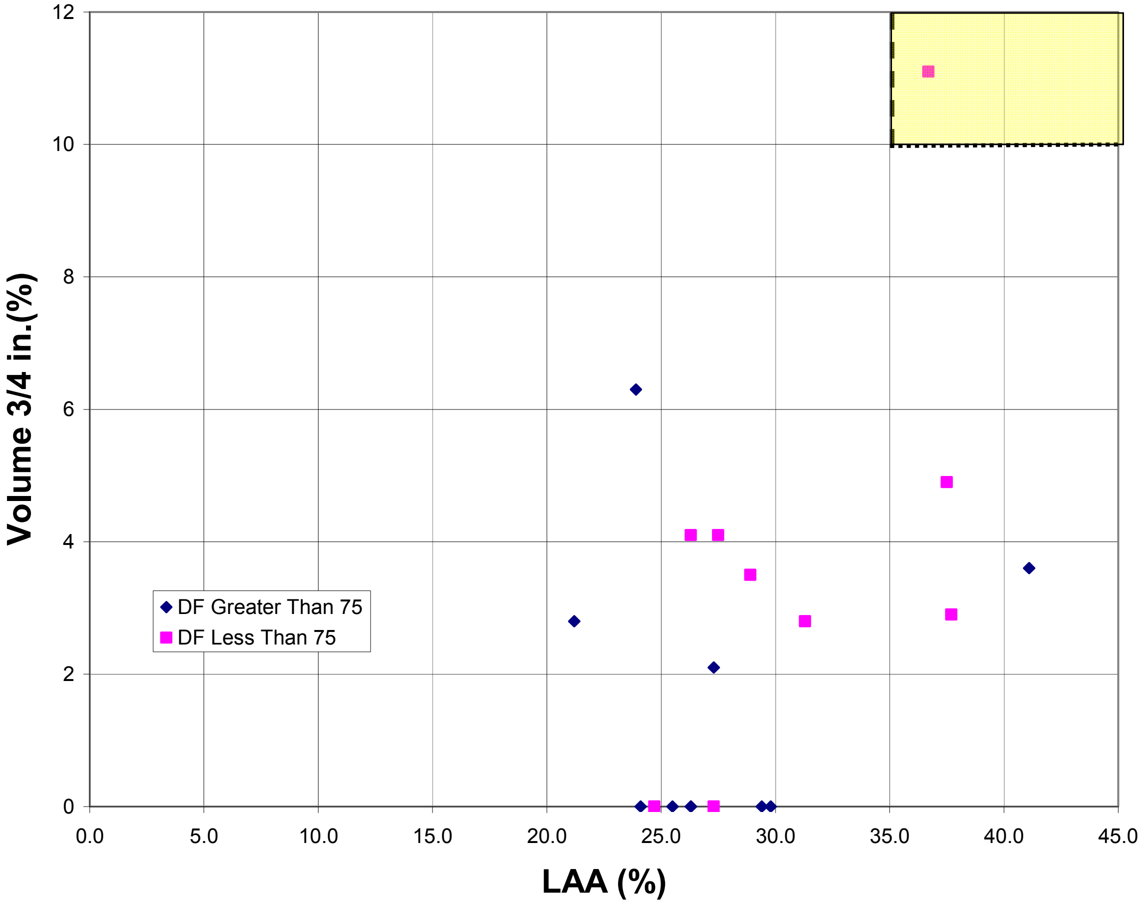


Figure 64: LAA With Volume Plus ¾ in. Threshold Limit

In summary, using all 57 samples, if an aggregate fails three or more of the six aggregate tests, there is almost a 100% chance that a DF of 75 or more cannot be achieved. If an aggregate fails two of the six tests, there is almost an 83% chance that a DF of 75 or more cannot be achieved. If an aggregate fails one of the six tests, there is almost a 50% chance that a DF of 75 or more cannot be achieved. And, if an aggregate fails no tests, there is almost a 10% chance that a DF of 75 or more cannot be achieved.

CORRELATION OF TEST RESULTS, MODELS, AND SERVICE RECORDS

Ideally, an evaluation system should not only be tied to T 161 test results, but also to historical field records of pavement performance. Unfortunately, at the time of writing, a relationship between the materials used in this study and field service records was not available.

CONCLUSIONS

Seven regression models have been developed to predict the T 161 DF. Four models feature only aggregate tests, while three models contain 28 day compressive strength as well. The choice of model depends on the desired ease of testing, familiarity with test methods, equipment cost, sensitivity to test duration, and level of accuracy that is considered acceptable. There is a trade off between accuracy of prediction and the above-listed factors. All models contained some test methods for which MoDOT has no historical data, thus, verification of the models was not possible. Therefore, the models should be considered preliminary until proven.

The most accurate model (adjusted $R^2 = 0.954$) entailed a routine test (gradation), tests that MoDOT is currently evaluating (IPI, WBM), and tests that are not currently being performed (PLS, VSBSG). The model also contains 28 day compressive strength. Thus, total testing time would be on the order of a month.

The second-most accurate model (adjusted $R^2 = 0.906$) was simpler: it contained a routine test (gradation), and tests that MoDOT is currently evaluating (IPI, WBM), and a test that is not currently being performed (VSBSG). The model also contains 28 day compressive strength. Thus, total testing time would be on the order of a month.

The third-most accurate model (adjusted $R^2 = 0.830$) was fairly simple: it contained routine tests (gradation, NaSO4), and tests that MoDOT is currently evaluating (IPI, WBM), and a test that is not currently being performed (VSBSG). However, the model does not entail 28 day compressive strength testing. Thus, total testing time would be on the order of two weeks.

The fourth-most accurate model (adjusted $R^2 = 0.792$) was fairly simple: it contained routine tests (gradation, BSG), and tests that MoDOT is currently evaluating (IPI, WBM). The model also contains 28 day compressive strength. Thus, total testing time would be on the order of a month.

The fifth-most accurate model (adjusted $R^2 = 0.785$) contained routine tests (gradation, MD), and a test that MoDOT is currently evaluating (WBM), and tests that are not currently being performed (VSBSG, ACV). However, the model does not entail 28 day compressive strength. Thus, total testing time would be on the order of several days.

The sixth-most accurate model (adjusted $R^2 = 0.756$) was quite simple: it contained a routine test (gradation), and tests that MoDOT is currently evaluating (IPI, WBM), and a test that is not currently being performed (VSBSG). However, the model does not entail 28 day compressive strength. Thus, total testing time would be on the order of several days.

The least accurate model (adjusted $R^2 = 0.747$) was the simplest: it contained routine tests (gradation, BSG), and tests that MoDOT is currently evaluating (IPI, WBM).The model does not entail 28 day compressive strength. Thus, total testing time would be on the order of several days.

A second system of evaluation of a given mixture entailed the use of a set of threshold limits set on various aggregate and, possibly, strength test method results. The limits were based, in part, on historical MoDOT limits. The test methods include NaSO4, WAFT, absorption, IPI, LAA, gradation, WBM, MD, and possibly concrete compressive and flexural strength. Successfully passing the system would most likely result in a T 161 result greater than 75, in most cases. The limits are to be considered preliminary until proven against a larger data set.

RECOMMENDATIONS – FUTURE RESEARCH

Future research should include verification, or even extension, of the models by performing additional tests to obtain the necessary data. Unfortunately, MoDOT's database (Site Manager) does not include WBM data, and very little IPI data, thus none of the models can be verified. As a start, models 4 and 7 can be verified by performing WBM tests on the 38 member data set used in this study in the Threshold Limits section. By performing VSBSG, models 1, 3, and 6 can be checked. To check model 5, PLS would have to be added, and for model 2, ACV would need to be performed. At any point, the regressions can be run again with a larger data set. And, the threshold system can be fine-tuned by moving the limits to balance acceptance and rejection.

The predictions of the models and threshold system should also be compared to actual field service records as a final check, should the records become available.

The WBM procedure has promise, and needs fine-tuning by standardizing such variables as matching the number of balls to the NMS (like LAA), the number of revolutions, and standard gradations like LAA or MD.

As available aggregate sources dwindle, more marginal aggregates will be forced into service, thus requiring the use of a smaller NMS to avoid D-cracking problems. The IPI procedure needs to be explored as to the effect of aggregate particle size on test results.

IMPLEMENTATION

Several implementation scenarios are envisioned:

1. Aggregate source approval by MoDOT- once a ledge is approved via T 161 testing, from that point, annual source approval testing could utilize one of the models or the threshold system in lieu of T 161 testing.

2. Mix design approval- for contractor designed mixes, aggregate sample test results could be submitted to MoDOT for approval; MoDOT could also perform verification testing on the aggregates for approval. Both sets of results could be used to predict DF, in lieu of T 161 testing.

3. Quality Control- aggregate could be sampled from stockpiles at the concrete batch plant and run through one of the systems presented in this study. If the model entailed concrete testing as well, concrete would be sampled at the plant or behind the paver, cylinders cast and cured, and subsequently tested at 28 days. From this, DF would be approximated. In regard to equipment costs, (i.e. IPI, WBM), the equipment costs for these items are about the same as a concrete compression machine which is currently required for core testing. And, a precedent for the necessity of a fully-equipped mobile QC testing laboratory has been established for MoDOT asphalt (Superpave) paving projects.

NOTATION

AASHTO= American Association of State Highway and Transportation Officials
Abs= T85 absorption
ACV= Aggregate Crushing Value
ADF= Aggregate Durability Factor
Adj R^2= adjusted R^2
AVA= Air Voids Analyzer
ASTM= American Society of Testing and Materials
BSG= T85 bulk specific gravity (dry)
Comp= 28 day compressive strength of concrete
CN= Condition Number
CV= coefficient of variation
DF= T161 Durability Factor
DFFITS= difference of fits
DR= deleterious rock
Flex= 35 day flexural strength of concrete
HudA= Hudsons A
IC= Index of Crushing
IPI= Iowa Pore Index
LAA= Los Angeles Abrasion
MAS= maximum aggregate size
MB= methylene blue
MD= Micro-Deval
MIC= modified Index of Crushing
MoDOT= Missouri Department of Transportation
MR= modulus of rupture
$NaSO_4$= sodium sulfate soundness
NMS= nominal maximum size
NMSDF= nominal maximum size Durability Factor
PCA= Portland Cement Association
PLS= point load strength
QC= quality control
R= correlation coefficient
R^2= coefficient of determination
Ratio2= gradation index tied to a unique value
RetFlex= retained flexural strength after T161 testing
VIF= Variance Inflation Factor
VolPlus3/4= volume of plus ¾ in. aggregate
VSAbs= vacuum saturated absorption
VSBSG= vacuum saturated bulk specific gravity (dry)
WAFT= water alcohol freeze thaw
WBM= wet ball mill
WBMM= wet ball mill modified

REFERENCES

American Association of State Highway and Transportation Officials (AASHTO). (2000). “Standard Method of Test for Resistance of Concrete to Rapid Freezing and Thawing.” AASHTO T 161-00, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). (2007). “Standard Method of Test for Soundness of Aggregates by Freezing and Thawing.” AASHTO T 103-07, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). (2002). “Standard Method of Test for Specific Gravity and Absorption of Coarse Aggregate.” AASHTO T 85-02, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). (2002). “Standard Method of Test for Resistance to Degradation of Small-Size Coarse Aggregate by Abrasion and Impact in the Los Angeles Machine.” AASHTO T 96-02, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). (2006). “Standard Method of Test for Resistance of Coarse Aggregate to Degradation by Abrasion in the Micro-Deval Apparatus.” AASHTO T 327-06, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). (2003). “Standard Method of Test for Soundness of Aggregate by Use of Sodium Sulfate or Magnesium Sulfate.” AASHTO T 104-03, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). (2007). “Standard Method of Test for Methylene Blue Value of Clays, Mineral Fillers, and Fines.” AASHTO T 330-07, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). (2005). “Standard Method of Test for Theoretical Maximum Specific Gravity and Density of Asphalt Paving Mixtures.” AASHTO T 209-05, Washington, D.C.

American Association of State Highway and Transportation Officials (AASHTO). (2003). “Standard Method of Test for Resistance of Compacted Asphalt Mixtures to Moisture-Induced Damage.” AASHTO T 283-03, Washington, D.C.

American Society of Testing and Materials (ASTM). (2007). “Standard Test Method Determination of the Point Load Strength Index of Rock and Application to Rock Strength Classifications.” ASTM D 5731-07, Philadelphia.

American Society of Testing and Materials (ASTM). (2008). “Standard Practice for Dealing with Outlying Observations.” ASTM E 178-08, Philadelphia.

Bieniawski, Z.T. (1975). “The Point-Load Test in Geotechnical Practice.” *Engineering Geology*. Vol. 9, Amsterdam, 1-11.

Bisque, R. E., and Lemish, J. (1958). “Chemical Characteristics of Some Carbonate Aggregate as Related to Durability of Concrete.” *Highway Research Board*. Bulletin 196, Transportation Research Board, Washington, D.C, 29-45.

Bloem, D.L. (1966). “Soundness and Deleterious Substances.” *Significance of Tests and Properties of Concrete and Concrete-Making Materials*. ASTM STP 169a. American Society for Testing and Materials, Philadelphia, 497-512.

Brink, R.H. (1958). “Rapid Tests for Aggregate and Concrete Presented at the Thirty-Seventh Annual Meeting, January 6-10, 1958.” National Research Council (U.S.), Highway Research Board.

British Standards Institution. (BSI). (1990). “Methods for Determination of Aggregate Crushing Value (ACV).” BS 812: Part 110, London.

British Standards Institution. (BSI). (1998). “Tests for Mechanical and Physical Properties of Aregates. Methods for the Determination of Resistance to Fragmentation.” BS 1097: Part 2, London.

Buth, E., Ivey, D. L., and Hirsch, T. J. (1964). “Correlation of Concrete Properties with Tests for Clay Content of Aggregate.” Texas Transportation Institute, Texas A & M University, College Station, TX.

Buth, E., Ivey, D. L., Hirsch, T. J. (1967). “Clay, Aggregate, and Concrete.” Texas Transportation Institute, Texas A & M University, College Station, TX.

Cady, P.D. (1984). “Soundness Testing and Durability of Coarse Aggregates in Concrete.” *Cement, Concrete, and Aggregate (CCAGDP)*. 6(1), 43-46.

Chamberlin, W. (1981). “Aggregate Porosity and Concrete Freeze-Thaw Resistance.” *FHWA/NY/SR-81/71*, Washington, D.C.

Clifton, A.W., Klassen, M.J., and Watters, B.R. (1987). “Production and Testing of Ballast,” *Transportation Research Record*. 1131, Transportation Research Board, Washington, D.C, 26-34.

Clifton, A.W., M.J. Klassen, and Watters, B.R. (1987). “Track Evaluation and Ballast Performance Specifications,” *Transportation Research Record*. 1131, Transportation Research Board, Washington, D.C, 35-44.

Cramer, S., and Carpenter, A. (1999). “Influence of Total Aggregate Gradation on Freeze-Thaw Durability and Other Performance Measures of Paving Concrete.” *Transportation Research Record*. 1668, Transportation Research Board, Washington, D.C, 1-8.

Cuelho, E., Mokwa, R., and Obert, K. (2007). “Comparative Analysis of Coarse Surfacing Aggregate Using Micro-Deval, L.A. Abrasion and Sodium Sulfate Soundness Tests.” *Final Project Report*. Western Transportation Institute, College of Engineering, Montana State University, Bozeman, MT.

Dolch, W.L. (1959). “Studies of Limestone Aggregates by Fluid-Flow Methods.” *American Society for Testing and Materials (ASTM)*. Proceedings, Vol. 59, 1204-1215.

Dolch, W.L. (1966). “Porosity.” *Significance of Tests and Properties of Concrete and Concrete-Making Materials*. ASTM STP 169a. American Society for Testing and Materials, Philadelphia, 443-461.

Domaschuk, L., and Garychuk, G. (1988). “Pore Stucture Analyses of Aggregates Susceptible to D-cracking.” *Annual Conference Proceedings of the Transportation Association of Canada*. Vol. 1, B25-B46.

Dubberke, W. (1983). “ Factors Relating to Aggregate Durability in Portland Cement Concrete.” *Interim Report for Project HR-2022*. Iowa Dept. of Transportation.

Dunn, J., and Hudec, P. (1965). “The Influence of Clays on Water and Ice in Rock Pores. Phase 4 of Physical Research Project No. 4 ‘Acceptability Tests for Coarse Aggregates.’” New York Dept. of Public Works, Bureau of Physical Research.

Dunn, J., and Hudec, P. (1966). "Water, Clay and Rock Soundness." *The Ohio Journal of Science*. Vol. 66(2), 153-168.

Dunn, J., and Hudec, P. (1972). "Frost and Sorption Effects in Agrillaceous Rock." *Highway Research Record*. 393, Transportation Research Board, Washington, D.C., 65-78.

Hardin, Bobby O. (1985). “Crushing of Soil Particles.” *Journal of Geotechnical Engineering*. 111(10), 1177-1192.

Harman, J., Cady, P., and Bolling, N. (1970). “Slow-Cooling Tests for Frost Susceptibility of Pennsylvania Aggregates.” *Highway Research Record*. 328, Transportation Research Board, Washington, D.C, 26-37.

Hiltrop, C. L., and Lemish, J. (1960). “Relationship of Pore-Size Distribution and Other Rock Properties to Serviceability of Some Carbonate Aggregates.” *Highway Research Board*. Bulletin 239, Washington D.C, 1-23.

Hossain, M., Lane, D., and Schmidt, B. (2007). “Use of the Micro-Deval Test for Assessing the Durability of Virginia Aggregates.” Virginia Transportation Research Council, Charlottesville, VA.

Hudson, S.B. and Waller, H.F. (1969). “Evaluation of Construction Control Procedures: Aggregate Gradation Variations and Effects”, Highway Research Board, Washington, D.C., 1-58.

Iowa Department of Transportation (2000). “Chemical Testing to Determine the Salt-Susceptibility of an Aggregate: A Guide for Materials Personnel.” Iowa Department of Transportation, Materials Division.

Iowa Department of Transportation. (2000). “Method of Test for Determining the Pore Index of Aggregates.” Iowa 219-D.

Iowa Department of Transportation. (2004). “Vacuum-Saturated Specific Gravity & Absorption of Combined or Individual Aggregate Sources” Materials IM 380.

Iowa Department of Transportation. (2007). “Source Approvals for Aggregates” Materials IM 409.

Janssen, D. and Snyder, M. (1994). “Resistance of Concrete to Freezing and Thawing,” *SHRP-C-391*, Washington, D.C.

Jayawickrama, P. W., Phillips, W., and Hossain, M. (2001). “Comparative Analysis of the Micro-Deval and Magnesium Sulfate Soundness Tests.” Texas Tech University, Center for Multidisciplinary Research in Transportation, Lubbock, TX.

Kandhal, P., and Parker, F. (1998). *Aggregate Tests Related to Asphalt Concrete Performance in Pavements*. National Academy Press, Washington, D.C.

Kaneuji, M. (1978). “Correlation between Pore Size Distribution and Freeze Thaw Durability of Coarse Aggregate in Concrete,” West Lafayette: Purdue University.

Klieger, P., Monfore, G., Stark, D., and Teske W. (1974). “D-Cracking of Concrete Pavements in Ohio.” Portland Cement Association. Skokie, Ill.

Klieger, P., Stark, D., and Teske, W. (1978). “The Influence of Environment and Materials on D-Cracking.” *Final Report, Portland Cement Association*. Skokie, IL.

Koubaa, A. and Snyder M. (2001). "Assessing Frost Resistance of Concrete Aggregates in Minnesota." *Journal of Cold Regions Engineering*. 15(4), 187-210.

Koubaa, A., and Snyder, M. (1996). "Evaluation of Frost Resistance Tests for Carbonate Aggregates." *Transportation Research Record*. 1547, Transportation Research Board, Washington, D.C., 35-45.

Lade, P., Yamamuro, J., and Bopp, P. (1996). “Significance of Particle Crushing in Granular Materials.” *Journal of Geotechnical Engineering*. 122(4), 309-361.

Larson, T. D., Boettcher A., Cady, P., Franzen, M., and Reed, J. (1965). “Identification of Concrete Aggregates Exhibiting Frost Susceptibility.” Highway Research Board of the Division of Engineering and Industrial Research, National Academy of Sciences-National Research Council, Washington, D.C.

Larson, T.D., and Cady, P.D. (1969). “Identification of Frost-Susceptible Particles in Concrete Aggregates.” *NCHRP 66, Highway Research Board*. Transportation Research Board, Washington, D.C.

Lewis, D.W., Dolch, W.L, and Woods, K.B. (1953). “Porosity Determinations and the Significance of Pore Characteristics of Aggregates.” *American Society for Testing and Materials (ASTM)*. Proceedings. Vol. 53, 949-958.

Litvan, G.G. (1972). “Phase Transitions of Adsorbates: IV, Mechanisms of Frost Action in Hardened Cement Paste.” *American Ceramic Society Journal*. 55(1).

Litvan, G.G. (1973). “Frost Action in Cement Paste.” *Materials and Structures*. 6(34).

Litvan, G.G. (1975). “Phase Transitions of Adsorbates: VI, Effects of deicing agents on the freezing of cement paste.” *American Ceramic Society Journal*. 58(12).

Litvan, G.G. (1978). “Adsorptive Systems at Temperatures below the Freezing Point of the Adsorptive.” *Advances in Colloid and Interface Science*. Vol. 9.

Lobo-Guerrero, S., and Vallejo, L. E. (2006). “Application of Weibull Statistics to the Tensile Strength of Rock Aggregates.” *Journal of Geotechnical and Geoenvironmental Engineering*. 132(6), 786-790.

Long, J. (1987). “Optimization of Richardson’s Shale Durability Rating System”, M.S. Thesis, Department of Civil Engineering, University of Missouri-Rolla.

Lusher, S.M. (2004). “Prediction of the Resilient Modulus of Unbound Granular Base and Subbase Materials Based on the California Bearing Ratio and Other Test Data”, M.S. Thesis, Department of Civil, Architectural, and Environmental Engineering, University of Missouri-Rolla.

Marks, V. and Dubberke, W. (1982). “Durability of Concrete and the Iowa Pore Index Test,” *Transportation Research Record*. 853, Transportation Research Board, Washington, D.C, 25-30.

Marks, V.J., and Dubberke, W. (1982). “Durability of Concrete and the Iowa Pore Index Method.” *Tranportation Research Record*. 853, Transportation Research Board, Washington, D.C., 25-30.

McDowell, G.R., and Bolton, M.D. (1998). “On the Micromechanics of Crushable Aggregates.” *Géotechnique*. 48(5), 667-679.

Meininger, R.C. (1978). “Abrasion Resistance, Strength, Toughness, and Related Properties.” *Significance of Tests and Properties of Concrete and Concrete-Making Materials*. ASTM STP 169b. American Society for Testing and Materials, Philadelphia, 657-694.

Mindess, S., Young, J., and Darwin, D. (2003). *Concrete*. 2nd ed. Prentice Hall, Upper Saddle River, NJ.

Missouri Department of Transportation. (1993). “Evaluation of Concrete and Aggregate Durability Tests”, *Project No. MCHRP 86-1*. MoDOT.

Missouri Department of Transportation. (2000). “Test Method TM-67: Preparation of Concrete for Use in AASHTO T 161 Freeze and Thaw Durability”, Missouri Department of Transportation, Materials Division. Jefferson City, Missouri.

Missouri Department of Transportation. (2006). “Deleterious Content of Aggregate.” Engineering Policy Guide. MoDOT T 71, Missouri Department of Transportation, Materials Division. Jefferson City, Missouri.

Missouri Department of Transportation. (2007). “Soundness Test of Coarse Aggregate – Water-Alcohol Freeze Method.” *Engineering Policy Guide*. MoDOT TM-14, Missouri Department of Transportation, Materials Division. Jefferson City, Missouri.

Myers, J., and Dubberke, W. (1980). “Iowa Pore Index Test.” *Interim Report, Iowa Department of Transportation*. Iowa Department of Transportation.

Paxton, John T. (1982). “Ohio Aggregate and Concrete Testing to Determine D-Cracking Susceptibility.” *Transportation Research Record*. 853. Transportation Research Board, Washington, D.C., 20-24.

Pigeon, M., and Pleau, R. (1995). *Durability of Concrete in Cold Climates*. E. & F.N. Spon, London.

Portland Cement Association, and Stark, D. (1976). “Characteristics and Utilization of Coarse Aggregates Associated with D-Cracking.” *Portland Cement Association Research and Development Bulletin (RD)*, 047-01. Skokie, Ill.

Portland Cement Association. (2008). "Aggregate." *Cement and Concrete Basics*. http://www.cement.org/basics/concretebasics_aggregate.asp (22 Jan. 2008).

Powers, T. C. (1945). “A Working Hypothesis for Further Studies of Frost Resistance of Concrete.” *Proceedings of the Highway Research Board*. Bulletin 5, Vol. 41, Washington, D.C.

Powers, T.C. (1949). “The Air Requirement of Frost-Resistant Concrete.” *Proceedings of the Highway Research Board*. Bulletin 33, Vol. 29, Washington, D.C.

Powers, T.C. (1955). “Basic Considerations Pertaining to Freezing and Thawing Tests.” *American Society for Testing and Materials (ASTM)*. Proceedings. Vol. 55, Philadelphia, PA, 1132-1155.

Powers, T.C. (1975). “Freezing Effects in Concrete.” *ACI Special Publication, SP-47*. American Concrete Institute, Detroit, MI.

Powers, T.C., and Helmuth, R.A. (1953). “Theory of Volume Changes in Hardened Portland Cement Pastes During Freezing.” *Proceedings of the Highway Research Board*, Vol. 32, Transportation Research Board, Washington, D.C, 285-297.

Richardson, D. N. (1989). “Point-Load Test for Estimating Compressive Strength.” Journal of the American Concrete Institute: *ACI Materials Journal*. 86(4), 409-416.

Richardson, D.N. (1984). “Relative Durability of Shale”, Ph.D. Dissertation, Department of Civil Engineering, University of Missouri-Rolla, 316 p.

Richardson, D.N. (1985). “Relative Durability of Shale-A Suggested Rating System”, *Proceedings of the 36th Annual Highway Geology Symposium*, Clarksville, Indiana.

Richardson, D.N. and Long, J. (1987), “The Sieved Slake Durability Test”, *Bulletin of the Association of Engineering Geologists*, 24(2), 247-258.

Richardson, D.N. and Wiles, T. (1990), “A Shale Durability System Based on Loss of Shear Strength”, *ASCE Journal of Geotechnical Engineering*, 116(12), 18641880.

Richardson, D.N., Morrison, W.J., Kremer, P.A., and Hubbard, K.M. (1996). “Determination of AASHTO Drainage Coefficients”, Study 90-4, Missouri Highway and Transportation Dept., Rolla MO, 194.

Rogers, C. A. (1989). “Development of an Unconfined Freeze-Thaw Test for Coarse Aggregates.” *Annual Conference Proceedings of the Transportation Association of Canada*. Vol. 1, B47-B67.

Saeed, A., Hall, J., and Barker, W. (2000). “Performance Related Tests of Aggregates for Use in Unbound Pavement Layers Final Report.” National Highway Research Program, Transportation Research Board, National Research Council, Washington, D.C.

Schulman, R.S.(1992). *Statistics in Plain English*, Van Nostrand Rheinhold, New York, New York.

Selig, E. T., and Boucher, D. L. (1990). “Abrasion Tests for Railroad Ballast,” *Geotechnical Testing Journal*. 13(4), 301-311.

Shakoor, A., and Scholer, C.F. (1985). “Comparison of Aggregate Pore Characteristics as Measured by Mercury Intrusion Porosimetry and Iowa Pore Index Tests.” *Proceedings of the Journal of the American Concrete Institute*. 82(4), 453-458.

Stark, D. (1976). “Characteristics and Utilization of Coarse Aggregates Associated with D-Cracking.” Portland Cement Association, Skokie, Ill.

Stark, D. and Klieger, P. (1973). “Effect of Maximum Size of Coarse Aggregate on D-cracking in Concrete Pavements,” *Highway Research Record*. 441, Transportation Research Board, Washington, D.C, 33-43.

Surdahl, R.W. (1990). “A New Aggregate Gradation Modulus”, M.S. Thesis, Department of Civil Engineering, Montana State University, Bozeman, MT, 121.

Swenson, E.G., and Chaly, V. (1956). “Basis for Classifying Deleterious Characteristics of Concrete Aggregate Materials.” *Journal of the American Concrete Institute*. 27(9), Proceedings Vol. 52, 987-1002.

Texas Department of Transportation. (2000). “Ball Mill Method for Determining the Disintegration of Flexible Base Material.” TEX-116-E.

Thompson, S., Olsen, M., and Dempsey, B. (1980). “D-cracking in Portland Cement Concrete Pavements,” Illinois Department of Transportation, Synthesis Report, FHWA/IL/UI-187, Washington, D.C.

Traylor, M. “Efforts to Eliminate D-Cracking in Illinois.” *Transportation Research Record*. 853. Transportation Research Board, Washington, D.C, 9-14.

Union Pacific and Burlington Northern Santa Fe Railroads. (2001). “Specifications for Main, Branch, and Yard Track Ballast.”

Verbeck, G., and Landgren, R. (1960). “Influence of Physical Characteristics of Aggregates on Frost Resistance of Concrete.” *American Society for Testing and Materials (ASTM)*. Proceedings. Vol. 60, Philadelphia, PA, 1063-1079.

Wiles, T. (1988). “A Shale Durability Rating System Based on Shear Strength”, M.S. Thesis, Department of Civil Engineering, University of Missouri-Rolla.

Williamson, G., Weyers, R., Mokarem, D., Lane, D., and Reid, D. (2007). "Vacuum Saturated Absorption As Aggregate Durability Indicator." *American Concrete Institute (ACI) Materials Journal*. 104(3). 307-312.

Winslow, D. N., Lindgren, M. K., and Dolch, W. L. (1982). “Relation Between Pavement D-Cracking and Coarse Aggregate Pore Structure.” *Transportation Research Record*. 853, Transportation Research Board, Washington, D.C., 17-20.

Winslow, D.N. (1987). “The Rate of Absorption of Aggregates.” *Cement, Concrete, and Aggregates (CCAGDP)*. 9(2), 154-158.

Woolf, D.O. (1966). “Toughness, Hardness, Abrasion, Strength, and Elastic Properties.” *Significance of Tests and Properties of Concrete and Concrete-Making Materials*. ASTM STP 169a. American Society for Testing and Materials, Philadelphia, 462-475.

Wu, Y., Parker, F., and Kandhal, P. (1998). "Aggregate Toughness/Abrasion Resistance and Durability/Soundness Tests Related to Asphalt Concrete Performance in Pavements." *Transportation Research Record*. 1638, Transportation Research Board, Washington, D.C, 85-93.

Wu, Y., Parker, F., and Kandhal, P. (1998). “Aggregate Toughness/Abrasion Resistance and Durability/Soundness Tests Related to Asphalt Concrete Performance in Pavements.” *NCAT Report 98-04*.

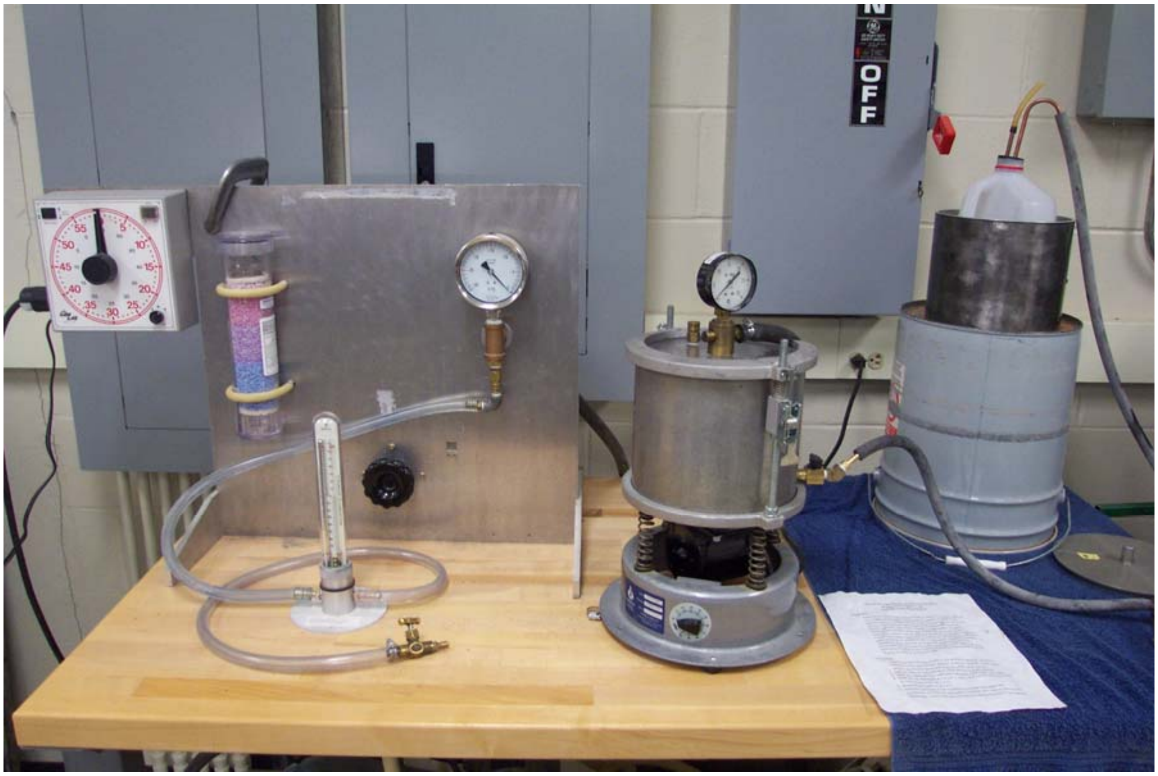
Yool, A.I.G., Lees, T.P., and Fried, A. (1998). “Improvements to the Methylene Blue Dye Test for Harmful Clay in Aggregates for Concrete and Mortar.” *Cement and Concrete Research*. 28(10), 1417-1428.

APPENDICES

APPENDIX A
Vacuum Saturated Specific Gravity and Absorption
Modified from AASHTO T 85
Durability Factor Study Method
Revised 7-31-08

Equipment

Equipment includes a 4500 ml pycnometer modified to introduce water, a non-modified pycnometer, a vacuum pump capable of sustaining a vacuum pressure of at least 25 mm of mercury (absolute pressure) with mercury manometer with appropriate ancillary equipment such as a vacuum regulator, and a towel. A weigh-in-water station should be available that includes a water bath suitable for immersion of the suspended container with its saturated specimen, an overflow outlet for maintaining a default water level, a method for controlling or monitoring water temperature, a balance with a weigh-below capability (nearest 0.1 g readability), and some type of suspended platform on which the pycnometer can be supported while submerged in the water bath. The platform and rod/wires that connect the platform to the balance should displace a minimum amount of water.



Procedure

1. Obtain a sample size appropriate for the gradation. MoDOT 1005 Gradation F requires a 2000 gram sample. MoDOT 1005 Gradation D and

1007 Type 5 requires a 3000 gram sample, and MoDOT 1005 Gradation B requires a 4000 gram sample.

2. Split the material over a #4 sieve. Work with the plus #4 material.
3. Wash the aggregate over a #4 sieve repeatedly until water runs clear.
4. Oven-dry the aggregate at 110 ± 5 °C (230 ± 9 °F) for 24 ± 6 hours. Cool to room temperature (25 ± 5 °C).
5. Bring the test water to 25 ± 5 °C.
6. Place the specimen in the modified pycnometer and attach the pycnometer to a vacuum apparatus. Close the vent valve on the mercury manometer. Turn on the vacuum pump by setting the timer to an arbitrary value such as 45 minutes then switching on the timer (the pump is connected to the timer). Gradually increase the vacuum to 27.5 ± 2.5 mm of mercury absolute pressure as measured by the mercury manometer. As soon as this level is reached, reset the timer to 30 minutes. Allow the aggregate to sit under vacuum for 5.0 minutes ± 15 sec from the time the timer was set to 30 minutes.
7. After the 5.0 minute period, while the vacuum pump is still running, turn the valve that is connected to the water slowly to the open position. Allow the pycnometer to fill with water until at least one inch of water is over the top of the aggregate. Then shut off the valve to the water. Start the mechanical agitator and turn the setting to 8.
8. When the timer goes off, the vacuum pump will automatically stop (30 minutes at 27.5 ± 2.5 mm of mercury absolute pressure will have been achieved). Stop the mechanical agitation.
9. Using the vent valve on the mercury manometer, slowly release the vacuum at a rate not to exceed 2.36 inches mercury gage per second as displayed on the vacuum gage on the lid of the pycnometer.
10. Remove the lid from the pycnometer.
11. Place an empty pan in a water bath. Without exposing the aggregate to the air, carefully submerge the pycnometer in the water and, while underwater, empty the contents of the pycnometer into the pan. Be sure that the pycnometer and aggregate are completely submerged when transferring the aggregate. Avoid loss of material.
12. Remove the pycnometer from the bath.
13. Carefully remove the pan filled with aggregate from the water bath. Decant some of the water, but leave about 2 inches of water above the surface of the aggregate and allow the specimen to sit for 24 ± 2 hours.
14. After 24 ± 2 hours, drain water from aggregate and roll aggregate in a pre-dampened towel to obtain SSD state. The SSD state is reached when the sheen on the aggregate just barely disappears.
15. Once the sheen on the aggregate surface has disappeared, immediately remove the sample from the towel into a pan and weigh in air. Record the weight (W_{SSD}).
16. Place the aggregate into a container such as the non-modified pycnometer.
17. The water bath should be maintained at 25 ± 1 °C.

18. Tare the weigh-in-water system.
19. Suspend the pycnometer containing the specimen in the water bath. Stir the aggregate to release air bubbles. Allow the scale to stabilize and record the weight of the pycnometer with sample (W_w) underwater when no more fluctuations on the scale's display occur.
20. Remove, drain, and completely empty the pycnometer into a pan.
21. Reset the weigh-in water system and immediately weigh the empty pycnometer under water and record the weight (W_t).
22. Oven-dry the aggregate at 110 ± 5°C (230 ± 9°F) for 24 ± 6 hours.
23. Remove the specimen from the oven and allow to cool to room temperature.
24. Weigh and record the oven-dried weight (W_{OD}).
25. Calculate the vacuum saturated specific gravity and the vacuum saturated absorption as follows:

VS Gsa = $W_{OD} / (W_{OD} - W_{SW})$
VS Gsb = $W_{OD} / (W_{SSD} - W_{SW})$

VS Abs = $(W_{SSD} - W_{OD}) / W_{OD}$

PRECISION: Single operator, 3 replicates

	1s	**d2s**
G_{sa}	0.007	0.020
$G_{sb, od}$	0.009	0.025
Abs, %	0.088	0.25

APPENDIX B
Methylene Blue
AASHTO T 330-07
Durability Factor Study
Revised 7-31-08

Equipment

Equipment includes a 500 ml Griffin beaker, one magnetic mixing plate with stir bar, one amber-colored burette of at least 50 ml capacity with 0.1 ml graduations, one glass rod, and Whatman No. 2 filter paper. A 200 ml capacity volumetric flask and a balance capable of reading to 1.00 grams. Methylene Blue reagent should be stored for no more than 4 months in a brown bottle wrapped in foil inside of a dark cabinet at lab temperature. Sieves: #200

Solution Mixing Procedures

1. Place 1 gram of methylene blue dye into a 200 ml volumetric flask.
2. Add distilled water at lab temperature to the cylinder until the 200 ml mark is reached.
3. Pour the solution into a beaker.
4. Mix thoroughly.

Procedure

1. Save all material passing the #12 sieve from the LA Abrasion test and further sieve over a #200 sieve.
2. Oven-dry the minus #200 material at 110 ±5°C (230 °F) for 24 ± 6 hours.
3. Weigh a 10.00 ± 0.05 gram sample of minus #200 material [W].
4. Flush the burette with methylene blue solution by filling it with 25 ml of solution, then opening the valve and allowing the entire 25 ml of solution to drain into a beaker. This is done to ensure that water in the burette is not left to dilute the solution. Flushing removes water. Discard the solution once drained.
5. Fill the burette with the methylene blue solution.
6. Place the sample in the Griffin beaker and add 30 ml of distilled water.
7. Place the Griffin beaker on a magnetic mixing plate and insert the magnet into the beaker.
8. Mix the sample and water to create a slurry.
9. With the slurry mixing, add 0.5 ml of the solution. Allow to mix for 1 minute.
10. If a sample has been previously tested: it is permissible to add more than 0.5 ml, up to 2.0 ml less than what has been required for previous samples to reach titration. (Example: if a previous sample has taken 16 ml to reach titration, it is permissible to immediately add 14 ml on the first dose of a subsequent specimen. If the halo appears on this first round, then the sample must be discarded and the first dose must be lessened by

at least 2.0 ml). This step is allowed to reduce the amount of time required to run replicate tests.

11. Using the glass rod, place a single drop of the slurry on a filter paper to check for a blue halo (signifying a fulfilled cation exchange capacity).
12. Continue to add 0.5 ml increments, mix, and check until the blue halo appears. Larger increments can be used, especially at the beginning of the test, if it is known that the sample's adsorption of the dye is high.
13. Once the halo is achieved, mix the solution for an additional 5 minutes and place another drop on filter paper to ensure that the exchange capacity is met.
14. Record volume of solution required for titration [V] and calculate the Methylene Blue Value [MBV].
15. Rinse the burette with distilled water; allow to drain by leaving the petcock open. Calculate MBV to the nearest 0.1 mg/g.

Formula: [MBV] = CV/W

Where: C = concentration of methylene blue dye in the solution (mg/ml). "C" is equal to 5 if the solution is made as directed in the instructions.
V = Amount of solution required for titration to occur (ml)
W = Original oven-dry sample weight

APPENDIX C
Aggregate Crushing Value
Modified from BS EN 812:110
Durability Factor Study
Revised 7-31-08; 10-1-08

Equipment

Equipment includes a heavy steel cylinder with an internal diameter of 154 mm, a solid steel plunger 152 mm in diameter, a metal slump rod (16 mm diameter, 600 mm long), a metal scoop, and a balance capable of reading to 1.0 grams. Sieves: 14 mm, 10 mm, and 2.36 mm (#8) sieves

Procedure

1. Starting with an air-dry sample, dry sieve the aggregate over a #4 sieve.
2. Obtain material that passes the 13.2 mm sieve and is retained on the 3/8 in. (9.5 mm) sieve. For finer materials lacking this size, build the specimen using 6.3 to 9.5 mm.
3. Oven-dry the aggregate at 110 °C (230 °F) for 24 ± 6 hours.
4. Add enough aggregate to fill about 33 mm (1.3 inches) of the mold; visually this should be 1/3 of the height from the bottom of the mold up to the 100 mm mark.
5. Tamp the aggregate 25 times with a slump rod: tamping consists of dropping the rod from a height of 50 mm (1 inch) from the surface of the specimen, evenly distributing the strokes over the entire surface of the specimen
6. Continue filling the test cylinder up to the 100 mm mark in two more equal lifts. Be sure to tamp each layer 25 times.
7. After tamping the last layer, level the top layer and be sure that the top of the layer is just at the 100 mm mark.

8. Remove the sample from the mold and into a pan. Be sure to remove all materials from the mold and retain all material in the pan.
9. Weigh the sample and record it (M_1). Divide the sample into thirds and add the sample to the steel cylinder again in three layers, rodding each layer 25 times. Level the top.
10. Insert the plunger into the top of the mold and place the entire mold with sample in a compression load frame (Tinius Olsen). Rotate the plunger slightly (about 1/3 of a turn) to ensure that the plunger is not stuck and to further level the sample.
11. Load the sample at a constant rate such that 89,924 lbs (400kN) of force is achieved in 10 minutes (a load rate of about 150 lbs/sec (40kN/minute) is preferable if controls are available).
12. Remove the load once 89,224 lbs (400 kN) has been achieved.
13. Remove the plunger. While removing it, slide the bottom of the plunger against the top of the wall of the cylinder to scrape off any aggregate stuck on to the bottom of the plunger.
14. Unbolt the mold from the base plate and remove it. Turn the cylindrical mold over and place into a pan. Use a small, wood 2x4 piece and a hammer to break loose the compacted aggregate. It has worked well to place the 2x4 along the interior edges of the mold and lightly tap it with a hammer. Once loose, the entire aggregate sample should be easily pushed through the entire mold.
15. Empty the entire sample into a pan and dry sieve the material over a #8 (2.36 mm) sieve. Be careful to apply only sufficient sieving action to accomplish the separation. Do not overwork the specimen. It is recommended to use a nested set of sieves: 3/8”, #4, #8, and a pan. Place the set of sieves filled with the aggregate in a shaker for 5 minutes.
16. After 5 minutes of shaking, empty the contents of all material retained on the #8 and larger sieves into a pan.
17. Record the mass passing (M_2) and the mass retained (M_3) on the #8 (2.36 mm) sieve to determine % loss. If ($M_2 + M_3$) differs by more than 10 g from M_1, discard the sample. Calculate to the nearest 0.1%.

    $ACV = (M_2 / M_1) * 100$

APPENDIX D
Point Load Index
ASTM D 5731-07
Durability Factor Study
10-6-08

Equipment

Equipment includes calipers and the MATEST point load testing machine.

Procedure (Unsoaked)

1. Oven-dry the aggregate at 110 ±5°C (230 °F) for 24 ± 6 hours.
2. Obtain a 20 piece sample. Each piece must be approximately 30 mm or greater in size. Square and rectangular-shaped aggregates are preferred
3. Insert the aggregate specimen in between the two platens. Close the platens to make contact with the specimen. Measure the initial specimen dimension with a calipers (D).
4. Load the specimen at 0.1 in./minute and record the load at which it breaks. The aggregate is required to fail within 10 to 60 seconds of loading. Adjust the load rate as needed to ensure aggregate failure within the specified time range.
5. Check to ensure that the aggregate ruptured at or very near one platen all the way through the aggregate piece to or very near to the other platen. If the specimen fails before a measurable load can be applied, record the load as zero.
6. Check to ensure that a tensile break has occurred (make sure no crushing occurred where the platens touched the aggregate).
7. Measure the dimensions of the aggregate at the fracture point to the nearest mm: W is the smallest dimension of the aggregate, and is perpendicular to the loading direction, D' is distance between the platens when the aggregate ruptures. L is the distance from the platens to the nearest free end. See the diagram below for a visual reference of the dimensions and how they are measured.

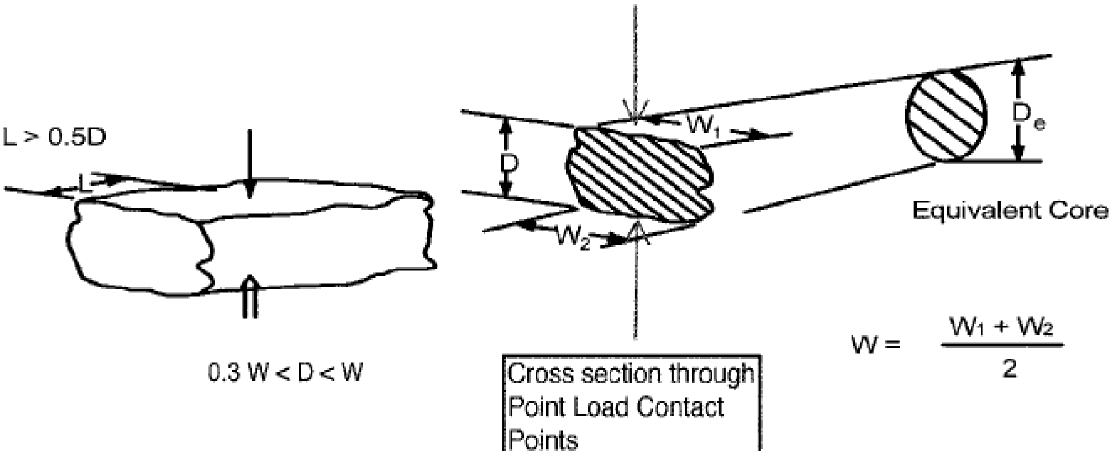


Figure from ASTM D 5731

8. Measure the distance, D’, and record.
9. Run the calculations as shown in the datasheet.
10. After calculating $I_{S(50)}$, delete the smallest two and largest two values from the dataset.
11. Calculate the average $I_{S(50)}$ value.

Equations

$$W = \frac{W_1 + W_2}{2}$$

where

W_1 = Width at the top of the aggregate piece (see figure), mm

W_2 = Width at the bottom of the aggregate piece (see figure), mm

$$A = W \times D'$$

$$D_e^2 = \frac{4 \times A}{\pi}$$

where

D_e = Equivalent diameter, converting irregular shapes to circular, mm

$$I_s = \frac{P}{D_e^2}$$

where

P = Load, N [the SBEL gage reads directly in kN; if a regular pressure gage is used (readable in psi), multiply the psi reading by the ram area 2.236 in^2 to obtain lbs; convert this to N]

$$F = \left(\frac{D_e}{50}\right)^{0.45}$$

where

F = Size correction factor used to compare samples of all sizes

$$I_{s(50)} = F \times I_s$$

where

$I_{s(50)}$ = Point load index adjusted with the size correction factor, MPa

APPENDIX E

Ratio2

Similar to Hudson's $\bar{A}$, Ratio2 is a gradation index. However, unlike Hudson's $\bar{A}$, Ratio2 is based on two characteristics of a particular gradation (Lusher, 2004). The idea grew out of work involving a third gradation index, R-modulus (Surdahl, 1990). R-modulus is calculated as follows:

$$\text{R-Modulus} = \frac{1}{\sum\left(\frac{1}{P_i}\right)} \tag{10}$$

Where: P_i = Total % passing each standard sieve

The standard sieves are the same as in the determination of $\bar{A}$ except that the top size to be included in the calculation is the smallest sieve through which 100% of the material passes.

Like $\bar{A}$, R-Modulus is non-unique in its characterization of a particular gradation; i.e. there are many gradations that can result in the same $\bar{A}$ or R-modulus values. To improve this situation, Surdahl utilized the 0.45 power gradation chart and a specific definition for determining the maximum density line (MDL) of a particular gradation then calculated the difference between the R-Modulus of the actual gradation and the R-Modulus of the gradation that would fall along the MDL. The result was that a positive difference indicated a fine gradation relative to the MDL and a negative difference indicated a coarse gradation relative to the MDL. The benefit of this method is that two characteristics of a particular gradation are used to quantify the particle size distribution thus increasing the uniqueness of the index.

To avoid using independent variables in regression analyses that could be negative or positive, Lusher proposed using the ratio of the MDL R-Modulus (MDL_{RMod}) to the actual gradation R-Modulus. Ratio2, therefore, has the benefit of utilizing two characteristics of a particular gradation but is always a positive value.

MDL_{RMod} is calculated by first determining the effective maximum size of the gradation. MDL_{RMod} is derived graphically by extending a straight line from the origin on the 0.45 power gradation chart through the point where the actual gradation crosses the 95% passing line and onto the 100% passing line. The interpolation equation is given by

$$x^{0.45} = SS_L^{0.45} - \frac{\left(SS_L^{0.45} - SS_S^{0.45}\right)\left(PP_L - 95\right)}{\left(PP_L - PP_S\right)} \tag{11}$$

Where: x = Size of particle at the 95% passing intersection
SS_L = Larger discrete sieve size
SS_S = Smaller discrete sieve size
PP_L = Total % passing the larger sieve (not always 100%)
PP_S = Total % passing the smaller sieve

To obtain the size of the particle at the 100% line, multiply the value generated by Eq. 11 by the ratio of $^{100}/_{95}$ as follows:

$$EMS^{0.45} = \left(x^{0.45}\right)\left(\frac{20}{19}\right) \tag{12}$$

Where: EMS = Effective maximum size

Substituting Eq. 11 into Eq. 12 gives the following:

$$EMS^{0.45} = SS_L^{0.45} - \frac{\left(SS_L^{0.45} - SS_S^{0.45}\right)\left(PP_L - 95\right)}{\left(PP_L - PP_S\right)} \times \left(\frac{20}{19}\right) \tag{13}$$

Surdahl developed a relationship through regression analysis that expressed the MDL_{RMod} as a function of the EMS and presented it as follows:

$$\frac{1}{MDL_{RMod}} = 0.5155\left(EMS^{0.45}\right) - 0.0274 \tag{14}$$

Therefore by substitution, Ratio2 can be calculated as follows:

$$\text{Ratio2} = \frac{1}{\left[0.5155\left(EMS^{0.45}\right)\text{-}0.0274\right]\left(\text{R-Modulus}\right)} \tag{15}$$

APPENDIX F

CORRELATION MATRIX

	DR	Chert	LAA	MD	WBM	ACV	IPI	MB	NaSO4	WAFT	PLS	VSBSG	VSSA	BSG	ABS	WBMM	Vol (+3/4")	NMS,DF	WP0.5	HudA	MIC	Rmod	Ratio2	Flex	Comp	%RetFlex	DF
DR	1																										
Chert	0.25869	1																									
LAA	0.460012	0.699416	1																								
MD	0.46316	0.132788	0.548932	1																							
WBM	0.351086	0.152826	0.490649	0.400468	1																						
ACV	0.02753	0.317224	0.602287	0.253505	0.762381	1																					
IPI	0.242322	-0.237242	-0.236725	0.361928	-0.242207	-0.480572	1																				
MB	0.213319	-0.085175	-0.266678	0.179334	0.029565	-0.190528	0.569534	1																			
NaSO4	0.339066	0.019312	0.116939	0.446417	0.206225	0.176109	0.31299	0.355102	1																		
WAFT	0.39499	0.103416	0.065961	0.414209	0.103472	-0.248697	0.613619	0.52785	0.463181	1																	
PLS	0.124465	-0.201019	-0.429775	-0.383173	-0.562028	-0.780049	0.310822	-0.026083	-0.050383	0.27769	1																
VSBSG	-0.506501	0.172631	-0.140108	-0.759685	-0.333897	-0.050242	-0.500181	-0.289813	-0.628502	-0.4616	0.102399	1															
VSSA	0.556601	0.085058	0.180983	0.637059	0.086271	-0.155513	0.626058	0.42506	0.705891	0.644434	0.079995	-0.805675	1														
BSG	-0.387164	0.231054	-0.111768	-0.734478	-0.169462	0.027694	-0.485526	-0.157791	-0.726131	-0.495732	0.002258	0.895789	-0.775919	1													
ABS	0.429547	-0.051408	0.05287	0.45816	-0.035205	-0.214855	0.530417	0.364376	0.791172	0.579955	0.194665	-0.710039	0.918027	-0.780246	1												
WBMM	0.238744	0.288936	0.561125	0.575655	0.881104	0.758989	-0.179731	-0.016851	0.354994	0.166861	-0.591834	-0.429853	0.245815	-0.323037	0.081338	1											
Vol (+3/4")	-0.028931	0.237828	0.399374	0.131883	-0.098692	0.04313	-0.187662	-0.332019	-0.168865	-0.12552	-0.00982	0.276419	-0.205399	0.149512	-0.208557	-0.1009	1										
NMS,DF	0.20701	0.293761	0.314381	0.042987	-0.189725	-0.166899	-0.095049	-0.41006	-0.010446	-0.044971	0.246081	0.229124	0.033712	0.052473	0.028354	-0.110436	0.735329	1									
WP0.5	0.130579	0.314059	0.411873	0.275339	-0.217216	-0.206299	-0.089074	-0.436213	-0.135669	0.173626	0.185251	0.087628	0.071508	-0.102334	0.025582	-0.098994	0.764849	0.753379	1								
HudA	-0.005021	-0.364363	-0.470264	-0.128567	-0.093765	-0.10986	0.329355	0.520554	0.232832	-0.127203	0.02503	-0.181713	0.147199	-0.042139	0.209835	-0.154317	-0.515641	-0.411387	-0.77152	1							
MIC	0.079802	0.352118	0.473634	0.245317	-0.067659	-0.023977	-0.200767	-0.524272	-0.156012	0.155984	0.072594	0.13782	-0.027092	-0.06162	-0.095167	0.051989	0.708015	0.680887	0.956518	-0.897469	1						
Rmod	0.114195	-0.160558	-0.068535	0.27243	-0.083258	-0.042135	0.313741	0.095208	0.316702	0.050972	0.049696	-0.238743	0.280035	-0.305294	0.257344	0.053974	0.094726	0.299532	-0.008198	0.51776	-0.127197	1					
Ratio2	0.133376	-0.076191	-0.167344	-0.059326	0.078322	-0.10879	0.128538	0.494451	-0.003929	0.255082	-0.046723	-0.044458	-0.028029	-0.009338	0.034333	-0.151236	-0.317491	-0.514679	-0.275332	-0.014709	-0.249317	-0.631143	1				
Flex	-0.3281	-0.085311	-0.350835	-0.589439	-0.069172	-0.012509	-0.37998	-0.212392	-0.62009	-0.409137	0.178443	0.473344	-0.650205	0.543628	-0.686713	-0.12313	-0.251786	-0.274775	-0.322934	0.066016	-0.220582	-0.140501	-0.008101	1			
Comp	-0.078239	0.22344	0.191195	-0.132935	0.328779	0.30787	-0.334342	-0.189523	-0.680095	-0.307699	-0.182779	0.380274	-0.605323	0.557723	-0.770864	0.260694	-0.064959	-0.201308	-0.169076	-0.141453	-0.044506	-0.246369	0.085677	0.666734	1		
%RetFlex	-0.440195	-0.014453	-0.133954	-0.19265	0.211122	0.212095	-0.142248	0.272591	-0.447571	-0.215147	-0.316791	0.413302	-0.426977	0.513471	-0.494758	0.191963	-0.137337	-0.347	-0.448988	0.250616	-0.42009	-0.088873	0.135075	0.472899	0.63549	1	
DF	-0.329928	0.059353	-0.134786	-0.352178	0.141369	0.154599	-0.12515	0.254633	-0.428867	-0.267061	-0.158833	0.524996	-0.529882	0.66363	-0.566761	0.054033	-0.093392	-0.205968	-0.46864	0.376353	-0.453089	0.021901	0.054483	0.476899	0.698007	0.918678	1