

COMPARATIVE ANALYSIS OF COARSE SURFACING AGGREGATE USING MICRO- DEVAL, L.A. ABRASION AND SODIUM SULFATE SOUNDNESS TESTS

FHWA/MT-06-016/8117-27

Final Report

prepared for
THE STATE OF MONTANA
DEPARTMENT OF TRANSPORTATION

in cooperation with
THE U.S. DEPARTMENT OF TRANSPORTATION
FEDERAL HIGHWAY ADMINISTRATION

February 2007

prepared by
Eli Cuelho
Robert Mokwa
Keely Obert

Western Transportation Institute
College of Engineering
Montana State University - Bozeman



RESEARCH PROGRAMS

Montana Department of Transportation
MDOT
serving you with pride

You are free to copy, distribute, display, and perform the work; make derivative works; make commercial use of the work under the condition that you give the original author and sponsor credit. For any reuse or distribution, you must make clear to others the license terms of this work. Any of these conditions can be waived if you get permission from the sponsor. Your fair use and other rights are in no way affected by the above.

COMPARATIVE ANALYSIS OF COARSE SURFACING AGGREGATE USING MICRO-DEVAL, L.A. ABRASION AND SODIUM SULFATE SOUNDNESS TESTS

Final Project Report

by

Eli Cuelho, Robert Mokwa and Keely Obert

of the

Western Transportation Institute
College of Engineering
Montana State University – Bozeman

prepared for the

State of Montana
Department of Transportation
Research Programs

in cooperation with the

U.S. Department of Transportation
Federal Highway Administration

January 2007

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA/MT-06-016/8117-27	2. Government Access No.	3. Recipient's Catalog No.	
4. Title and Subtitle Comparative Analysis of Coarse Surfacing Aggregate Using Micro-Deval, L.A. Abrasion and Sodium Sulfate Soundness Tests		5. Report Date January 2007	
		6. Performing Organization Code	
7. Author(s) Eli Cuelho, Robert Mokwa and Keely Obert		8. Performing Organization Report Code	
9. Performing Organization Name and Address Western Transportation Institute PO Box 174250 Montana State University – Bozeman Bozeman, Montana 59717-4250		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. MSU G&C #4W0753 MDT Project #8117-27	
12. Sponsoring Agency Name and Address Research Programs Montana Department of Transportation 2701 Prospect Avenue Helena, Montana 59620-1001		13. Type of Report and Period Covered Final Report September 2005 – January 2007	
		14. Sponsoring Agency Code 5401	
15. Supplementary Notes Research performed in cooperation with the Montana Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration. This report can be found at http://www.mdt.mt.gov/research/projects/mat/coarse_aggregate.shtml .			
16. Abstract Aggregates used in the construction of roads must be durable, abrasion resistant, and freeze-thaw resistant in order to perform well in pavement or as base course. The objective of this study was to investigate whether the Micro-Deval test will provide better, timelier, and more repeatable information about the quality of an aggregate than the Sodium Sulfate Soundness test. This objective was met by 1) conducting a thorough literature review, 2) testing a variety of soils from across Montana, and 3) analyzing these results to reveal potential trends. The literature review indicated that the Micro-Deval test has good repeatability characteristics, and that it correlates well to field performance. Laboratory tests (Micro-Deval, L.A. Abrasion, and Sodium Sulfate) were conducted on a variety of soil types to examine the repeatability of each test method, and to examine how well the methods correlate with each other in terms of predicting aggregate durability and degradation. Test results were normalized to facilitate direct comparisons between the three methods. Linear regression of the data points and corresponding confidence intervals were plotted to qualitatively assess agreement or disagreement between test methods. The authors concluded that the Micro-Deval test is a suitable replacement for the Sodium Sulfate test as the primary method for evaluating aggregate durability, with limitations. Because there were some inconsistent durability determinations between test methodologies, the authors recommend that the Micro-Deval test results be further supported by a second aggregate durability test whenever the Micro-Deval results fall between the cutoff value and plus 30% of the cutoff value. In other words, when the Micro-Deval test result for an aggregate is between 18 and 24 percent loss, a second test using an alternate method is recommended before any conclusions are made regarding the durability or quality of an aggregate.			
17. Key Words Micro-Deval, L.A. Abrasion, Sodium Sulfate, synthesis, comparative analysis, aggregate durability		18. Distribution Statement No restrictions. This document is available through NTIS, Springfield, Virginia 22161.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 47	22. Price

DISCLAIMER

This document is disseminated under the sponsorship of the Montana Department of Transportation and the United States Department of Transportation in the interest of information exchange. The State of Montana and the United States Government assume no liability of its contents or use thereof.

The contents of this report reflect the views of the authors, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official policies of the Montana Department of Transportation or the United States Department of Transportation.

This report does not constitute a standard, specification, or regulation.

ALTERNATIVE FORMAT STATEMENT

MDT attempts to provide accommodations for any known disability that may interfere with a person participating in any service, program, or activity of the Department. Alternative accessible formats of this information will be provided upon request. For further information, call (406) 444-7693, TTY (800) 335-7592, or Montana Relay at 711.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the efforts of the two undergraduate students (Keely Obert and James Robinson) who conducted the laboratory tests on the materials considered in this study. Acknowledgement of financial support and matching dollars for this research is extended to the Montana Department of Transportation, as well as the Research and Innovative Technology Administration (RITA) at the United States Department of Transportation through the Western Transportation Institute at Montana State University.

EXECUTIVE SUMMARY

Aggregates used in the construction of roads must be durable, abrasion resistant, and freeze-thaw resistant in order to perform well in pavement or as base course. Montana Department of Transportation (MDT) currently uses the Sodium Sulfate Soundness test and the Los Angeles Abrasion and Impact test (L.A. Abrasion test) to determine aggregate quality. The Sodium Sulfate test's validity and repeatability has been questioned by transportation officials and researchers across the country, mostly due to its known poor repeatability. As such, several other state departments of transportation have recently investigated the use of the Micro-Deval Abrasion test as an alternative for determining aggregate durability. The objective of this study was to investigate whether the Micro-Deval test will provide better, timelier, and more repeatable information about the quality of an aggregate than the Sodium Sulfate Soundness test.

A literature review conducted as part of this study focused on research that has attempted to correlate or quantify relationships between various soil durability/toughness tests. The literature review focused on the Micro-Deval, L.A. Abrasion and Sodium Sulfate tests. Generally, these tests are geared toward providing information related to an aggregate's ability to resist degradation during construction and under traffic loading or to determining how well aggregates withstand environmental changes (i.e., freezing and thawing, wetting and drying). Sixteen research studies were critically reviewed to qualitatively examine the relationships between different durability tests. Although results from the literature were at times inconsistent, some trends emerged. Generally, however, most of the authors were able to obtain favorable or useful results using the Micro-Deval test. The Micro-Deval test generally was considered repeatable and it reportedly correlated well to field performance. In contrast, several studies indicated that the L.A. Abrasion test does not accurately predict field performance. The Sodium Sulfate test commonly received poor ratings in terms of repeatability and correlation to field performance. In the two studies that investigated how field performance relates to the Magnesium and Sodium Sulfate Soundness tests, the Magnesium Sulfate test rated higher overall than the Sodium Sulfate test based on raveling (Brandes and Robinson, 2006) and raveling, popouts and potholes (Wu et al., 1998). Reported research that attempted to correlate results between various test methods yielded inconsistent conclusions. Some authors suggested that a more reliable assessment of aggregate durability could be obtained by employing multiple test methods. However, specific practical implementation recommendations were not provided.

Comparison laboratory testing was conducted using the Micro-Deval, Sodium Sulfate, and L.A. Abrasion tests to examine the repeatability of each test and to compare the results in terms of aggregate durability and degradation. The soils examined in this study were primarily granular and cohesionless, typical of material that would be excavated and processed for use as aggregate for plant mix asphalt or crushed base course. Soil samples were obtained from gravel pits throughout the state of Montana.

Micro-Deval tests in this study were conducted according to standard test method AASHTO T327 (ASTM D6928), L.A. Abrasion tests were conducted according to standard test method AASHTO T96 (ASTM C131), and Sodium Sulfate tests were conducted by Montana Department of Transportation according to standard test method AASHTO T104 (ASTM C88). For the purposes of this study, the following percent loss pass-fail standards were used for each test:

- Micro-Deval: passing (i.e., durable), if % loss $\leq 18\%$;
- L.A. Abrasion: passing (i.e., durable), if % loss $\leq 40\%$; and
- Sodium Sulfate: passing (i.e., durable), if % loss $\leq 12\%$.

The testing program provided data for evaluating the repeatability of multiple tests conducted on the same material (five repeats for the Micro-Deval test and three or more repeated tests for the L.A. Abrasion test) and for examining correlations between the three tests. The scope of this study did not include a repeatability evaluation of the Sodium Sulfate test; consequently, the Sodium Sulfate percent loss values were obtained from single tests on each aggregate. This single-lab repeatability study indicated that both the Micro-Deval and L.A. Abrasion tests have good repeatability with no statistically significant difference between the two tests. The Sodium Sulfate test has a reported coefficient of variation (COV) for single laboratory testing that is significantly higher than the COVs measured in this study for Micro-Deval and L.A. Abrasion tests.

Results from the suite of laboratory tests were normalized to facilitate direct comparisons between the three test methods. Normalized test results were obtained by taking the average percent loss for a particular soil and dividing it by the pass-fail standard for that test. Plots were generated to make direct comparisons between test results: 1) Micro-Deval to Sodium Sulfate, 2) Micro-Deval to L.A. Abrasion, and 3) L.A. Abrasion to Sodium Sulfate. Data points within these plots fell within one of two regions: regions where the test results agreed (i.e., pass-pass and fail-fail regions) or regions where the test results disagreed (i.e., pass-fail and fail-pass regions). Linear regression of the data points and its corresponding confidence interval were plotted in addition to the points to qualitatively assess the degree of positive correlation between test methods.

The majority of the aggregate samples passed all three tests, indicating the majority of aggregates examined in this study would be characterized as durable. Based on the materials tested in this study, the Micro-Deval and L.A. Abrasion tests appear to correlate well for aggregates that have relatively low percent loss values. The Sodium Sulfate test appeared to have the poorest correlation of the three test methods examined in this study. Discontinuities between all three tests are more apparent with materials that have percent losses near or slightly greater than the cutoff values. Because of the scarcity of borderline and failing tests, the statistical significance of this observation could not be quantified.

The Micro-Deval test appears to be a suitable substitute for the Sodium Sulfate test for aggregates that are not borderline pass/fail in terms of percent loss durability. The Micro-Deval test has good repeatability characteristics, and it correlates reasonably well with results from L.A. Abrasion and Sodium Sulfate tests. Furthermore, the Micro-Deval test is less time-consuming and easier to perform than the Sodium Sulfate test.

Because there were some inconsistent durability determinations between test methodologies, the authors recommend that the Micro-Deval test results be further supported by a second aggregate durability test whenever the Micro-Deval results fall between the cutoff value and plus 30% of the cutoff value. In other words, when the Micro-Deval test result for an aggregate is between 18 and 24 percent loss, a second test should be conducted before any conclusions are made regarding the durability or quality of an aggregate. Suggested alternatives include recognized methods such as the Sodium Sulfate Soundness test, the Magnesium Sulfate Soundness test, or the L.A. Abrasion test.

TABLE OF CONTENTS

1	Introduction.....	1
2	Literature Review.....	3
2.1	Arm (2003)	3
2.2	Bjarnason et al. (2002)	3
2.3	Brandes and Robinson (2006)	4
2.4	Cooley et al. (2002)	5
2.5	Fowler et al. (2006)	6
2.6	Hunt (2001)	7
2.7	Jayawickrama et al. (2006).....	8
2.8	Kline et al. (2006 – in press)	8
2.9	Lim (2004).....	9
2.10	Prowell et al. (2005)	9
2.11	Rangaraju et al. (2005)	10
2.12	Rismantojo (2002).....	10
2.13	Rogers et al. (1991)	11
2.14	Senior and Rogers (1991).....	12
2.14.1	Granular Base Course Aggregates	13
2.14.2	Portland Cement Concrete	13
2.14.3	Surface Course Asphalt Concrete	13
2.15	Tarefder et al. (2003).....	14
2.16	Wu et al. (1998).....	14
2.17	Summary of Literature Review	16
3	Results and Analysis.....	20
3.1	Soil Samples	20
3.2	Laboratory Test Results.....	22
3.3	Repeatability.....	24
3.4	Comparison of Three Test Methods.....	25
3.4.1	Micro-Deval versus Sodium Sulfate	28
3.4.2	Micro-Deval versus L.A. Abrasion.....	28
3.4.3	L.A. Abrasion versus Sodium Sulfate.....	30
3.5	Summary of Comparisons	31
4	Conclusions and Recommendations	34
5	References.....	37

LIST OF TABLES

Table 1. Test methods used to measure aggregate durability (from Wu et al., 1998)	2
Table 2. Correlation (r) between Micro-Deval, L.A. Abrasion, and Soundness Tests (Brandes and Robinson, 2006).....	4
Table 3. Correlation (r) between Aggregate Tests and HMA Pavement Performance (Brandes and Robinson, 2006).....	5
Table 4. Aggregate Field Performance Evaluation Criteria Used by Fowler et al. (2006).....	7
Table 5. Statistical Correlations between Micro-Deval, L.A. Abrasion, Magnesium Sulfate Soundness and Water Absorption Tests (Rismantojo, 2002).....	11
Table 6. Differences between Standard Micro-Deval and Modified Test for Fine Aggregates	12
Table 7. Aggregate Field Performance Evaluation Criteria Used by Senior and Rogers (1991).....	13
Table 8. Linear Correlations (R^2) between Micro-Deval and Other Tests (Tarefder et al., 2003)	14
Table 9. Aggregate Field Performance Evaluation Criteria Used by Wu et al. (1998)	15
Table 10. Summary survey findings (Wu et al., 1998).....	15
Table 11. Summary of literature review conducted by Wu et al. (1998).....	16
Table 12. Literature Review Summary Table.....	18
Table 13. Summary of Materials Examined in This Study.....	21
Table 14. Summary of Test Results.....	23
Table 15. Summary of All Data from Soils Tested in This Study	33
Table 16. Summary Comparison of Test Methods	34

LIST OF FIGURES

Figure 1. Soil sample locations in Montana.....	20
Figure 2. COV for the Micro-Deval and L.A. Abrasion tests: a) compilation of all test data and b) data displayed by soil number.....	25
Figure 3. Normalized comparison between Micro-Deval and Sodium Sulfate tests.....	27
Figure 4. Normalized comparison between Micro-Deval and L.A. Abrasion tests.....	29
Figure 5. Normalized comparison between L.A. Abrasion and Sodium Sulfate tests.....	30
Figure 6. Three-dimensional representation of the test data.....	32

1 INTRODUCTION

In the construction of roads, aggregate quality is critical. The aggregate must be durable, abrasion resistant, and freeze-thaw resistant in order to perform well in pavement or as sub-base. Several abrasion and durability tests are available to quantify the durability of aggregates. Montana Department of Transportation (MDT) currently uses the Sodium Sulfate Soundness test and the Los Angeles Abrasion and Impact test (L.A. Abrasion test) to determine aggregate quality. The Sodium Sulfate test is a method of durability testing in which the aggregate is subjected to simulated freeze-thaw cycles. After five cycles the loss in each gradation size is measured. The L.A. Abrasion test is an abrasion test in which loss in mass is caused by the impact of steel balls dropping on the aggregate in a rotating drum.

The Sodium Sulfate test's validity and repeatability has been questioned by transportation officials and researchers across the country. The Sodium Sulfate test has poor repeatability between laboratories and even between samples in the same laboratory (Kline et al., 2006). When samples fail to meet minimum criteria, contractors may try to contest the results because of the variability of the Sodium Sulfate test. In order to accurately describe the durability of the aggregate, minimize litigation, and costs on transportation projects, MDT is looking for a more repeatable durability test. Several other state departments of transportation have started using the Micro-Deval Abrasion test as an alternative for determining aggregate durability. The objective of this study is to investigate whether the Micro-Deval test will provide better, timelier, and more repeatable information about the quality of an aggregate than the Sodium Sulfate Soundness test.

Estimating the durability or toughness of aggregate can be a challenging task. There are a variety of test methods used to quantify durability of aggregates, as shown in Table 1. These tests use different mechanisms to determine durability, but they all give information about how the aggregate will degrade during construction and throughout its life in a pavement section. Another method that is used is a petrographic analysis. This method uses geology to define the characteristics of the aggregate to determine its durability. However, the analysis is subjectively dependent on the geologist that performs the test, so it may not be a reliable test for construction quality control.

Table 1. Test methods used to measure aggregate durability (from Wu et al., 1998)

Test Method	Test Type
L.A. Abrasion	Abrasion (dry)
Micro- Deval Abrasion	Abrasion (wet)
Nordic Ball Mill	Abrasion (wet)
Magnesium Sulfate Soundness	Simulated freeze-thaw
Sodium Sulfate Soundness	Simulated freeze-thaw
Freeze-Thaw Soundness	Freeze-thaw
Canadian Freeze-Thaw	Freeze-thaw plus sodium chloride solution
Aggregate Impact Value	Impact
Aggregate Crush Value	Compressive load
Degradation in the SHRP* Gyratory Compactor	Compaction
Petrographic	Geologic analysis

*SHRP stands for Strategic Highway Research Program

2 LITERATURE REVIEW

Coarse aggregates make up the majority of the constituents of asphalt pavements; therefore, their ability to perform under varying traffic and environmental conditions is important. Several tests exist to determine the strength, durability and toughness of aggregates; however, information from these tests has been known to conflict. This literature review has focused on research that has attempted to correlate or quantify relationships between various soil durability/toughness tests. Numerous durability tests have been proposed over the years; this review focused on three tests germane to this study: the Micro-Deval, L.A. Abrasion and Sodium Sulfate tests. Generally, tests are geared toward providing information related to an aggregate's ability to resist degradation during construction and under traffic loading or to determining how well aggregates withstand environmental changes (i.e., freezing and thawing, wetting and drying).

2.1 Arm (2003)

Research was conducted in Stockholm for the Swedish Geotechnical Institute to investigate the mechanical properties of recycled aggregates and industrial byproducts. As part of this study, the results from durability tests conducted on these materials were compared to similar tests conducted on standard construction aggregates including sand, gravel and crushed rock. Tests conducted in Arm's (2003) study to assess aggregate durability included: L.A. Abrasion, Micro-Deval, and Freeze-Thaw tests. L.A. Abrasion tests were run according to ASTM¹ C131 and C535 (ASTM, 2003), and to the European standard EN² 1097-2. Micro-Deval and freeze-thaw tests were conducted according to the European standards only: EN 1097-1 and EN 1367-1, respectively. Results indicated that:

- the Micro-Deval and both L.A. Abrasion tests ranked the materials similarly,
- the test results were not significantly effected by the different versions of the L.A. Abrasion tests, and
- the Micro-Deval test offered the best differentiation between the various material types.

2.2 Bjarnason et al. (2002)

This study investigated 20 sources of Icelandic aggregates (primarily basaltic) prepared as base course material with grain sizes between 0 and 25 mm. A total of 17 test methods were conducted including L.A. Abrasion, Micro-Deval, Magnesium Sulfate Soundness, and Freeze-

¹ ASTM stands for American Society of Testing and Materials

² EN stands for European Standard

Thaw. The test methods were classified as fragmentation (e.g., L.A. Abrasion), weathering (e.g., Magnesium Sulfate Soundness and Freeze-Thaw) or abrasion (e.g., Micro-Deval) tests.

Good correlation was found within each category of tests. Correlations were presented in terms of a correlation coefficient, r , which is simply the square root of R^2 determined from regression analysis, where values near zero indicate poor correlation and those closer to 1 or -1 are associated with high correlation. The correlation coefficient (r) between the fragmentation tests ranged from 0.73 to 0.96; the range of r between the weathering tests was slightly less at 0.63 to 0.94. The single correlation value between the two abrasion tests was calculated to be 0.87. The primary difference between the two abrasion tests is that the rotating drum in the Nordic studded tire test has three ribs, while a smooth drum is used in the Micro-Deval test. Because the tests within each of the categories correlated well with one another, Bjarnason et al. (2002) concluded that any of the tests within a particular group (i.e., fragmentation, weathering, and abrasion) can be used to assess aggregate quality for road construction.

2.3 Brandes and Robinson (2006)

This study analyzed parameters from chemical and mechanical tests on aggregates from twelve quarries in Hawaii. Correlation was computed and reported in terms of the correlation coefficient (r). Correlations among the tests were deemed reasonable in most cases, as summarized in Table 2. The two tests with the greatest correlation to the Micro-Deval were the Sulfate Soundness tests (correlation coefficient of 0.87 for sodium and 0.77 for magnesium). However, the correlation between Micro-Deval and L.A. Abrasion was very low, at -0.12.

Table 2. Correlation (r) between Micro-Deval, L.A. Abrasion, and Soundness Tests (Brandes and Robinson, 2006)

	Micro-Deval	L.A. Abrasion	Sodium Sulfate Soundness	Magnesium Sulfate Soundness
Micro-Deval	1.00	-0.12	0.87	0.77
L.A. Abrasion		1.00	-0.08	-0.30
Sodium Sulfate Soundness			1.00	0.80
Magnesium Sulfate Soundness				1.00

The relationships of aggregate tests to pavement performance were also made using raveling distress data. The results for Micro-Deval, L.A. Abrasion, and Sulfate Soundness are reproduced

in Table 3. Hawaii DOT requires a 9% maximum loss of Sodium Sulfate (NaSO_4) Soundness and 30% maximum loss from L.A. Abrasion tests; they do not have requirements for Magnesium Sulfate (MgSO_4) Soundness and Micro-Deval tests. Unexpectedly, high values of L.A. Abrasion loss were associated with better pavement performance. The authors attribute this to the possibility that tough aggregate with low abrasion loss may be less flexible and yield a more brittle asphalt mixture. The correlation of pavement performance to Micro-Deval aggregate loss was not high enough to suggest that Micro-Deval tests should replace L.A. Abrasion tests. Instead, the authors suggested the Aggregate Durability Index test, which was performed according to the AASHTO T210 test standard (AASHTO, 2005), may be a more appropriate mechanical test even though its correlation to pavement performance (r) was only calculated to be -0.49. The best correlation to pavement performance was seen in the Magnesium Sulfate Soundness test, which far exceeded the Sodium Sulfate Soundness (second best).

Table 3. Correlation (r) between Aggregate Tests and HMA Pavement Performance (Brandes and Robinson, 2006)

Test	Correlation (r)
Micro-Deval	-0.39
L.A. Abrasion	0.33
NaSO_4 Soundness	-0.66
MgSO_4 Soundness	-0.89
Aggregate Durability Index	-0.49

2.4 Cooley et al. (2002)

Researchers at the National Center for Asphalt Technology (NCAT) investigated the existing methods of testing aggregate durability in the Southeastern region of the United States. This study's main objective was to characterize aggregate quality throughout the region using the Micro-Deval test. The results were compared to data obtained from the L.A. Abrasion, and Magnesium and Sodium Sulfate Soundness tests.

The research examined 72 aggregates from eight states, with at least five aggregates from each state. Aggregate samples were assigned a performance rank of either good, fair, or poor based on its field performance. Micro-Deval, L.A. Abrasion, and Sodium Sulfate tests were conducted on each aggregate. The tests were performed at the respective state departments of transportation, and the results were compiled by the NCAT researchers. Most of the aggregates tested were categorized as: gravel, limestone, sandstone, or chert.

Statistical relationships between testing methods were examined using p-values. P-values range between zero and one, with p-values closer to zero indicating that the parameters being compared are statistically less similar to one another and p-values closer to one indicating that the parameters being compared are statistically more similar to one another. Of the eight states

that submitted samples, three found a significant relationship between the Micro-Deval test and the Sodium Sulfate test based on the p-value. However, only two of the states, Alabama and Georgia, found that the AASHTO specification of 18% loss of material adequately distinguished between their good and poor aggregates. Only the Alabama samples showed any correlation between the Micro-Deval and the L.A. Abrasion tests. The testing also revealed that gravels categorized as poor in several states passed the Micro-Deval test with under 6 % loss.

The coefficient of variation (COV) was also determined for each sample, and averaged for the L.A. Abrasion, Micro-Deval, and Sodium Sulfate test methods. The Sodium Sulfate test had the highest COV with an average value between 114.3% and 141.8%. The L.A. Abrasion test had the least variation with an average COV value between 29.8% and 39.3%, for all states. The Micro-Deval test had a COV between 57.9% and 65.3%. All of the tests showed significant variation.

The study concluded that aggregate type may affect the correlation between the Micro-Deval test and the performance rating of the aggregate. The study suggested that no significant relationship exists between the Micro-Deval and either the L.A. Abrasion or the Sodium Sulfate Soundness test. For Micro-Deval testing, Cooley et al. (2002) recommended that maximum loss values may need to be established for different aggregate types in order to obtain a better correlation with established performance history. The significance of the relationship was also drawn from the historical performance ranking of the aggregate. Performance rankings varied based on how individual state agencies perform these assessments; therefore, a cross examination of all aggregates and states was not feasible.

2.5 Fowler et al. (2006)

The International Center for Aggregate Research Project 507 was a comprehensive evaluation of 117 sources of aggregates in the U.S. and bordering Canadian provinces. Fowler et al. (2006) examined correlations between 10 different aggregate durability tests, including: Micro-Deval, L.A. Abrasion, Magnesium Sulfate Soundness, British Aggregate Crushing Value (original and saturated-surface dry), Canadian Freeze-Thaw, Water Absorption, and Specific Gravity (bulk, saturated-surface dry, and apparent). In general, very little correlation was observed, even when outliers were removed from the data sets. For example, the Micro-Deval and Magnesium Sulfate Soundness tests had R^2 values less than 0.6 for all correlations with other test methods. The highest correlation for the L.A. Abrasion test was to the British aggregate crushing value in which an $R^2 = 0.650$ was obtained using an exponential regression.

In addition to statistical correlations, the authors compared durability test results to field performance ratings of the aggregates' sources. Field performance information was obtained using a telephone survey of personnel from DOTs and Canadian Ministries of Transportation. Aggregates were rated as good, fair or poor, which is based on the three-tiered performance

evaluation criteria conducted by Senior and Rogers (1991) and Wu et al. (1998) (Table 4). According to Fowler et al. (2006), those failures referred to in Table 4 as ‘non-chemical’ may include “pop-outs, excessive degradation, and production of plastic fines, rutting, D-cracking, and any other failure relating only to the physical properties of the aggregate.” Most of the sources were rated as “good” (only 5 of the 117 aggregates were not classified in this category), which limited the analysis. The Micro-Deval test was consistent and repeatable and the best indicator of field performance, but test results did not correlate well with other aggregate tests studied. The predictions were improved when the Micro-Deval test was used in conjunction with results from Canadian Freeze-Thaw, Magnesium Sulfate Soundness, or Specific Gravity tests; however, it was not clear how the tests results would be used in conjunction with one another.

Table 4. Aggregate Field Performance Evaluation Criteria Used by Fowler et al. (2006)

Evaluation	Description
Good	Used for 10 or more years with no reported non-chemically related failures
Fair	Used at least once where minor non-chemically related failures require repair, but life extends beyond 10 years
Poor	Used at least once with severe degradation of failure occurring within two years of service or during construction which severely inhibits or prevents the use of the application

2.6 Hunt (2001)

Prompted in part by studded tire use in Oregon, the Oregon Department of Transportation (ODOT) sponsored a research project to evaluate alternative means of measuring aggregate durability. The project considered the Micro-Deval and Nordic Ball Mill tests as viable alternatives to the L.A. Abrasion test.

Initial testing to familiarize the technicians with the apparatus and procedures indicated that results from multiple tests conducted on a single aggregate were repeatable and there were no statistically significant differences between the two technicians’ results. Further testing was conducted on additional aggregates (44 samples from 22 sources) that had varying degrees of wear resistance, as assessed through field measurements. The results indicated there was little difference in the abilities of the Micro-Deval and L.A. Abrasion tests to predict field wear performance. In fact, all but one aggregate passed the Micro-Deval test and all passed the L.A. Abrasion test regardless of their field performance. In contrast, the few samples tested using the Nordic Ball Mill did indicate its ability to differentiate aggregate quality.

2.7 Jayawickrama et al. (2006)

This study was conducted by Texas Tech University in cooperation with the Texas Department of Transportation. The viability of the Micro-Deval test was investigated as a method for evaluating the durability of bituminous aggregate, and the results were compared with Texas' current method, the Magnesium Sulfate Soundness test.

Jayawickrama et al. (2006) examined 52 sources that were used in bituminous pavements in Texas. Three repeat samples of each test were conducted. Mineralogical and petrographic analyses were also performed on each sample. Most of the aggregates tested contained calcite and dolomite minerals. Regression analyses were performed on the results to determine the relationship between the Micro-Deval and Magnesium Sulfate tests. Two statistical models were developed from this comparison. The first was a direct empirical relationship between the tests. The second model included absorption as an additional independent variable.

The study concluded that a correlation existed between the Magnesium Sulfate test and the Micro-Deval test. This study also confirmed the results obtained through the study by Rogers et al. (1991) that the Micro-Deval test is more repeatable than the Magnesium Sulfate Soundness test. In addition to these results, the study found that both tests had an inherent dependence on the absorption rate of the aggregate. Jayawickrama et al. (2006) recommended that the Micro-Deval test be used as a "quality control tool" because it is more repeatable than the Magnesium Sulfate test and is quicker to perform.

2.8 Kline et al. (2006 – in press)

The objective of this study was to investigate if the Sodium Sulfate was correctly quantifying the durability of dolomite aggregates in Arkansas. The study achieved this by comparing the Sodium Sulfate results to several physical features of the aggregate to determine if a reliable relationship exists. The Sodium Sulfate was compared to the following physical properties: insoluble residue, x-ray diffraction, rock grain size, absorption, lamination and other visible physical features.

The study included 50 Sodium Sulfate tests conducted on a prevalent source of dolomite obtained from two quarries in Arkansas. The study found that the results from Sodium Sulfate tests were not related to any physical properties of the aggregate. Kline et al. (2006) conducted an extensive literature review and determined the Sodium Sulfate test had a reputation for poor repeatability. They suggested that hand sieving between each cycle may have caused the variation. The study recommended that the Sodium Sulfates test's "poor precision and lack of correlation to the actual weathering stresses mean that this test should not be used as a basis for rejection of the aggregate," but suggested that the performance history of an aggregate should be the final judgment of whether the aggregate is adequate for use in pavement (Kline et al., 2006).

2.9 Lim (2004)

Research was conducted at the University of Nottingham to investigate the correlation between simple railroad ballast index tests and box tests that simulated field loading conditions. Relative to this project, Wet Attrition Value, L.A. Abrasion and Micro-Deval tests were conducted on six ballast materials. Based on the latest European specifications for railway ballast (BS¹ EN² 13450-02), five properties are required to define appropriate ballast materials: ballast grading, L.A. Abrasion, Micro-Deval, flakiness index, and particle length. The Micro-Deval tests conducted by Lim (2004) were performed according to the European standard test method (EN 1097-1-96), and included modifications specified in BS EN 13450-02. Results from Lim's study showed that the simulated box tests correlated well with the Wet Attrition Value, L.A. Abrasion, and Micro-Deval test results in identifying ballast field performance.

2.10 Prowell et al. (2005)

This was an extensive study conducted by the National Cooperative Highway Research Program (NCHRP – Project 9-35) to synthesize information related to the performance of pavements designed using the Superpave method, and the relationship between pavement performance and specific aggregate properties. The study included a literature review to establish the state-of-the-practice, a survey to determine state agency specifications, a review of performance data from field test sections, and full-scale accelerated testing.

The literature review considered (among other things) methods of characterizing aggregates based on: coarse and fine aggregate angularity, particle shape (flat and elongated), particle texture, susceptibility to moisture damage, and durability. The L.A. Abrasion, Sodium Sulfate, and Micro-Deval durability tests were considered in the review. These tests were selected for review since they specifically addressed aggregate breakdown during construction, abrasion or weathering during use, and freeze-thaw durability.

Results from the extensive review generally indicated that coarse aggregate angularity may be a useful parameter for predicting rutting potential. Several main conclusions were presented regarding aggregate particle shape:

1. handling of aggregates changes the amount of flat and elongated particles,
2. aggregates that have a higher percentage of flat and elongated particles tend to break down easier during compaction,
3. aggregates that have a higher percentage of flat and elongated particles generally have higher voids within the aggregate, and
4. the ASTM Standard Test Method for Flat Particles, Elongated Particles, or Flat and Elongated Particles in Coarse Aggregate specification, designated as ASTM D4791

¹ BS stands for British Standard

² EN stands for European Standard

(ASTM, 2003), is highly variable and may be the reason that relationships with field performance are difficult to establish.

Prowell et al. (2005) recommended that the Micro-Deval test should be used instead of Sulfate Soundness tests to measure aggregate's resistance to abrasion. The authors recommend that for states that experience freeze-thaw cycles, a freeze-thaw test in addition to the Micro-Deval test may be beneficial. The authors conclude that additional research is needed to further evaluate specification limits for different aggregate types.

2.11 Rangaraju et al. (2005)

The South Carolina Department of Transportation (SCDOT) sponsored a project to compare the Micro-Deval and Magnesium Sulfate Soundness tests to the L.A. Abrasion and Sodium Sulfate tests. Correlation of the four tests to field performance was undertaken, as well as the effect of aggregate gradation on loss from L.A. Abrasion and Micro-Deval testing. Nineteen of the twenty-three sources of aggregate were classified as granite, two as marine limestone, and one each as granite-gneiss and marble-schist. Correlations between the Micro-Deval, the Sodium Sulfate, and the L.A. Abrasion tests were statistically insignificant. The Sodium and Magnesium Sulfate Soundness tests were found to have a high correlation with one another ($R^2 = 0.906$). Field performance ratings of good, fair, and poor were provided by SCDOT for each of the aggregates. These ratings were based on the occurrence of raveling, popouts and potholes. The results showed that a Micro-Deval loss limit of 17% accurately identified marginal aggregates, but an L.A. Abrasion loss limit of 55% did not accurately identify marginal aggregates. Aggregate gradation appeared to influence the results of the Micro-Deval test, in that smaller aggregate sizes showed greater loss. The authors recommended the Micro-Deval test over the other durability tests examined in their study.

2.12 Rismantojo (2002)

The research conducted by Rismantojo (2002) focused on aggregate tests related to permanent deformation and moisture susceptibility of Hot Mix Asphalt (HMA) pavements. Relationships between HMA pavement performance and individual aggregate properties were assessed based on the results of a number of durability tests, including: the Micro-Deval, L.A. Abrasion, Magnesium Sulfate, and Water Absorption tests. Five course aggregates and six fine aggregates were tested; and the test results from NCHRP Project 4-19 (Kandhal and Parker, 1998), were also included in the analyses. Test sections were constructed in the Indiana Department of Transportation Purdue Accelerated Pavement Testing facility located in West Lafayette, Indiana to evaluate field performance of the aggregates. A correlation analysis was also conducted between the various aggregate properties.

Rismantojo (2002) reported strong correlations between the Micro Deval, L.A. Abrasion, Magnesium Sulfate Soundness, and Water Absorption tests. Rismantojo's statistical correlation results are summarized in Table 5 (statistical measures r and p are described in Sections 2.2 and 2.4 of this report, respectively). Rismantojo (2002) reported instances where water aided in the breakdown of aggregates in the Micro-Deval test, resulting in increased loss in the Micro-Deval when compared to the L.A. Abrasion. Neither the Micro-Deval nor the L.A. Abrasion test correlated well with rutting parameters in terms of predicting aggregate field performance. The author concluded that the Micro Deval test is significantly correlated with the Magnesium Sulfate Soundness test, but recommended the Micro-Deval as a preferable alternative since it is quicker to perform.

Table 5. Statistical Correlations between Micro-Deval, L.A. Abrasion, Magnesium Sulfate Soundness and Water Absorption Tests (Rismantojo, 2002)

	Micro-Deval	L.A. Abrasion	Magnesium Sulfate Soundness	Water Absorption
Micro-Deval	$r = 1.00$	$r = 0.802$ $p = 0.055$	$r = 0.863$ $p = 0.027$	$r = 0.961$ $p = 0.002$
L.A. Abrasion		$r = 1.00$	$r = 0.427$ $p = 0.398$	$r = 0.684$ $p = 0.134$
Magnesium Sulfate Soundness			$r = 1.00$	$r = 0.894$ $p = 0.016$
Water Absorption				$r = 1.00$

2.13 Rogers et al. (1991)

Rogers et al. (1991) evaluated alternative tests to measure soundness of fine aggregates for use in Portland cement and asphalt concretes. The standard Micro-Deval test was modified to more appropriately evaluate fine aggregates, as described in Table 6. Results were compared between several other test methods: the Attrition Test (ASTM C1137-90), the MTO (Ministry of Transportation Ontario) Attrition Test, Gradation, Absorption, Relative Density and Magnesium Sulfate Soundness. Rogers et al. (1991) concluded that the Magnesium Sulfate Soundness test is not very precise, is not always suitable for testing fine aggregates, and is time consuming. They reported that the Micro-Deval test is comparatively quick and simple, and correlates well with Magnesium Sulfate Soundness test results. Variability between different laboratories that used the modified Micro-Deval test was low, and the test was reportedly well suited for identifying fine aggregates derived from weak and poor rocks.

Table 6. Differences between Standard Micro-Deval and Modified Test for Fine Aggregates

	Standard Micro-Deval*	Modified Micro-Deval for Fine Aggregates (Rogers et al., 1991)
Mass of soil	500 g	500 g**
Mass of steel charge	5000 g	1250 g
Volume of water	2.0 L	750 ml
Rotation	100 rpm	100 rpm
Rotation time	2 hours	15 minutes
Soak time	1+ hours	24 hours
Loss calculation	Loss through 1.18 mm sieve	Loss through 75 µm sieve

*Performed according to ASTM D6928 (ASTM, 2003)

** Taken from a 700 g sample that was washed over a 75 µm sieve and oven dried

2.14 Senior and Rogers (1991)

The objective of Senior and Rogers' (1991) research was "to investigate alternative methods of measuring the performance of coarse aggregates using tests that are relatively quick and simple to perform, yet give repeatable and reproducible results." The three-tiered field performance evaluation criteria used by Senior and Rogers (1991) are summarized in Table 7. One of the expected benefits from this work was to identify aggregate tests that would more accurately help engineers reject or accept certain aggregates for use as granular base, Portland cement concrete, and asphalt pavements. Ultimately, these tests should be reliable, precise, quick, simple and inexpensive. Tests considered in this study included (standard method follows in parentheses):

- Unconfined Freeze Thaw Test for Coarse Aggregate (MTO¹ LS-614),
- Micro-Deval Test (BNQ² 2560-070/82),
- Aggregate Impact Value Test (BS³ 812),
- Polished Stone Value Test (BS 812), and
- Aggregate Abrasion Value (BS 812).

These relatively quick and simple tests were compared with the following more traditional test methods (standard method follows in parentheses):

- L.A. Abrasion Test (MTO LS-603, ASTM C131),
- Magnesium Sulfate Soundness Test (MTO LS-606, ASTM C88),

¹ MTO stands for Ministry of Transportation Ontario

² BNQ stands for Bureau of Normalization of Quebec

³ BS stands for British Standard

- 24-Hour Water Absorption Test (MTO LS-604, ASTM C127), and
- Petrographic Evaluation (MTO LS-609).

Table 7. Aggregate Field Performance Evaluation Criteria Used by Senior and Rogers (1991)

Evaluation	Description
Good	Used for many years with no reported failures, popouts, or other signs of poor durability
Fair	Used at least once where popouts or some reduced service life had resulted, but pavement or structure life extended for over 10 years
Poor	Used once with noticeable disintegration of pavement after one winter, severely restricting pavement life

2.14.1 Granular Base Course Aggregates

Results from the study indicate that the Micro-Deval test, when used in conjunction with the Petrographic evaluation, was able to distinguish between good and bad aggregates. Even though the authors suggest using both tests to better distinguish aggregate quality due to limitations within each test, they do not offer a procedure on how to practically combine data from both tests to make a definitive conclusion on aggregate quality. This is true for other aggregate test combinations suggested by Senior and Rogers (1991). On its own, the Micro-Deval test was able to reasonably distinguish between marginal and good aggregates. Additionally, when the Micro-Deval test was used with the 24-Hour Water Absorption test, it was able to identify potentially poor aggregates.

2.14.2 Portland Cement Concrete

The Micro-Deval test was able to distinguish between marginal and poor aggregates when used in conjunction with the Unconfined Freeze Thaw test. In addition, when the 24-Hour Water Absorption test was used in conjunction with either the Micro-Deval or Unconfined Freeze Thaw test, it was able to identify poor aggregates.

2.14.3 Surface Course Asphalt Concrete

The Unconfined Freeze Thaw test was reportedly better than the Magnesium Sulfate Soundness test when used with the Petrographic examination to assess aggregate quality. The Aggregate Abrasion Value test provides a good measure of wear resistance; however, the Micro-Deval test is simpler. The Polished Stone Value test, when used with the Micro-Deval test, was reportedly able to identify good aggregates.

2.15 Tarefder et al. (2003)

The objective of this study was to compare Micro-Deval test results to L.A. Abrasion, Freeze-Thaw, Aggregate Durability Index, Specific Gravity, and Water Absorption tests for 18 aggregates from the state of Oklahoma. The study also investigated test repeatability and the ability of the Micro-Deval test to predict aggregate field performance. Fourteen of the aggregates were predominately limestone; three were sandstone and one was rhyolite. When compared to the other tests, the Micro-Deval results for the sandstone aggregate exhibited much better correlation than tests conducted on the limestone aggregate, as shown in Table 8. Note that there was a limited sample of sandstone aggregates when compared to the limestone.

Table 8. Linear Correlations (R^2) between Micro-Deval and Other Tests (Tarefder et al., 2003)

	Freeze-Thaw	L.A. Abrasion	Aggregate Durability Index	Specific Gravity	Water Absorption
Sandstone	0.657	0.633 (all agg.)	0.939	0.884	0.937
Limestone	0.389 (log fit)		0.045	0.021	0.039

Micro-Deval tests were found to be highly repeatable based on ten tests conducted on each of four aggregates. The greatest average coefficient of variation was calculated to be 5.4 percent. All four average coefficients of variation were below specified acceptable maximums recommended in AASHTO TP58 (AASHTO, 2005).

Aggregates having a wide range of field performance were considered in this investigation. Of the 18 aggregates sources considered, the ratings were as follows: 3 excellent, 8 good, 5 fair and 2 poor. Field performance was evaluated by the Oklahoma Department of Transportation, and evaluation measures were not specified in the literature. The L.A. Abrasion test did not accurately predict field performance except for two aggregates classified as poor (losses greater than 32%). In contrast, a clear trend was evident with the Micro-Deval test results. Results showed that losses less than 10% indicated excellent performance, losses between 10 and 18% indicated good performance, losses between 18 and 25% indicated fair performance, and losses greater than 25% indicated poor performance. Based on these results, the authors provisionally suggested that a maximum Micro-Deval loss of 25% be used to identify poor performers.

2.16 Wu et al. (1998)

The objective of this study was to identify and evaluate various laboratory tests used to determine durability, toughness, soundness, or abrasion resistance in terms of their ability to predict aggregate field performance. The three-tiered field performance evaluation criteria used by Wu et al. (1998) are summarized in Table 9. A survey conducted as part of this study

generally showed that most states in the late 1990s required that L.A. Abrasion tests be used to determine aggregate suitability, as shown in Table 10. Durability tests considered in their literature review and laboratory evaluation included: L.A. Abrasion, Aggregate Impact Value (British), Aggregate Crushing Value (British), Micro-Deval, Degradation in the SHRP Gyratory Compactor, Sodium and Magnesium Sulfate Soundness, Freeze-Thaw Soundness, Aggregate Durability Index, and the Canadian Freeze-Thaw Test. Results from their literature review are summarized in Table 11.

Table 9. Aggregate Field Performance Evaluation Criteria Used by Wu et al. (1998)

Evaluation	Description
Good	Used for many years with no significant degradation problem during construction and no significant popouts, raveling or potholes during service life
Fair	Used at least once where some degradation occurred during construction and/or some popouts, raveling, and potholes developed, but pavement life extended for over 8 years
Poor	Used at least once where raveling, popouts, or combinations developed during the first two years, severely restricting pavement

Table 10. Summary survey findings (Wu et al., 1998)

Laboratory Index Test	Extent of State Use	Parameters
Sodium Sulfate Soundness	53%	Max. allowable loss: 5-25%, 14% avg.
Magnesium Sulfate Soundness	19%	Max. allowable loss: 10-30%, 16% avg.
L.A. Abrasion	94%	Max. allowable loss: 40-45%
Durability Index Test	2% (one state)	
Freeze-Thaw Test	10%	
No requirements	16%	

Table 11. Summary of literature review conducted by Wu et al. (1998)

Author(s)	Result(s)
Woolf (1937)	Good correlation between L.A. Abrasion test and field performance
Shelburne (1940)	
Shergold (1948)	
Bullas and West (1991)	Aggregate Impact Value not suitable, but Aggregate Crushing Value is suitable for predicting field performance of aggregates
Fookes et al. (1988)	Recommends that a combination of index be used to determine suitability of aggregates
Senior and Rogers (1991)	Recommends using the Micro-Deval to determine aggregate quality because it has good correlation with performance
Paul (1932)	Sulfate soundness tests are good for predicting aggregate performance
Garrity and Kriege (1935)	Sulfate soundness tests have poor correlation with performance
Gandhi and Lytton (1984)	Mixed review with respect to the ability of sulfate soundness tests to predict field performance of aggregates
Papalentiou et al. (1987)	
Hasan et al. (1991)	
Rogers et al. (1991)	Canadian Freeze-Thaw marginally better than $MgSO_4$ in predicting aggregate field performance
Senior and Rogers (1991)	

Sixteen aggregates sources were tested using the L.A. Abrasion, Micro-Deval, Sodium Sulfate and Magnesium Sulfate Soundness tests to evaluate their ability to predict field performance. Three repeat trials were conducted on each soil type, which were obtained from 12 U.S. states. Results from a statistical analysis of the data found that the L.A. Abrasion and Sodium Sulfate tests did not predict pavement performance nearly as well as the Micro-Deval and the Magnesium Sulfate Soundness tests. Results from the analysis also suggested that a loss of approximately 18% was a good cutoff value for the Micro-Deval and Magnesium Sulfate Soundness tests.

2.17 Summary of Literature Review

Out of the sixteen studies reviewed, fifteen considered the Micro-Deval test within their research or analysis. Twelve considered the L.A. Abrasion test, eleven considered the Magnesium Sulfate Soundness test and five considered the Sodium Sulfate test. Results from the literature were mixed, but some trends emerged. A tabulated summary of pertinent results,

conclusions, or recommendations from the various studies described in this chapter is provided in Table 12. Generally, most of the authors were able to obtain favorable or useful results using the Micro-Deval test. Additionally, the Micro-Deval test was considered repeatable by those who studied its repeatability (Jayawickrama et al., 2006; Tarefder et al., 2003; Hunt, 2001). Four studies specifically indicated that the Micro-Deval test related well with field performance (Fowler et al., 2006; Rangaraju et al., 2005; Tarefder et al., 2003; Wu et al., 1998); however, one study indicated that it did not (Brandes and Robinson, 2006). Several studies indicated that the L.A. Abrasion test does not accurately predict field performance (Brandes and Robinson, 2006; Tarefder et al., 2003; Wu et al., 1998). Studies that examined the Sodium Sulfate test indicated that it was not a very good test, it wasn't very repeatable, and it appeared to have only a limited relationship to field performance or to aggregate properties (Kline et al., 2006; Prowell et al., 2005; Wu et al., 1998). In contrast, Brandes and Robinson (2006) indicated that the Sodium Sulfate test had a fairly high correlation with field performance. The Magnesium Sulfate Soundness test generally rated higher overall than the Sodium Sulfate test, especially in terms of field performance (Brandes and Robinson, 2006; Wu et al., 1998). Rogers et al. (1991) reported that the Magnesium Sulfate Soundness test was less reliable when fine aggregates were used.

Several studies correlated the results from different durability test methods. Good correlations were generally found when the Micro-Deval test results were compared to Magnesium Sulfate Soundness test results (Brandes and Robinson, 2006; Fowler et al., 2006; Jayawickrama et al., 2006; Rismantojo, 2002; Rogers et al., 1991). However, in contrast, Cooley et al. (2002) found poor correlations between the Micro-Deval test and other durability tests. Correlations between the Micro-Deval and the Sodium Sulfate tests were also mixed: Brandes and Robinson (2006) reported good a correlation, while and Cooley et al. (2002) reported a relatively poor correlation. Rismantojo (2002) reported a good correlation between the Micro-Deval and the Water Absorption tests. Correlations between the Micro-Deval and the L.A. Abrasion tests were mostly poor (Brandes and Robinson, 2006; Fowler et al., 2006; Rangaraju et al., 2005; Cooley et al., 2002), however, Arm (2003) reported a relatively good correlation.

Results from a number of studies suggest that evaluating results from more than one type of test on a single aggregate may provide a more reliable assessment of aggregate durability. Based on their results, two studies suggested using the Micro-Deval test with the Freeze-Thaw test (Fowler et al., 2006 and Prowell et al., 2005). Fowler et al. (2006) also suggested using the Micro-Deval test with the Magnesium Sulfate Soundness test or specific gravity test to better predict field performance. Senior and Rogers (1991) suggested using the Micro-Deval test with Petrographic analysis and the Water Absorption test to better assess the durability of base course aggregates. Although several authors suggest using combinations of various tests to better distinguish aggregate quality, partly due to limitations within individual tests, analytical procedures on how to practically merge data from these suggested combinations are generally not specified.

Table 12. Literature Review Summary Table

Author(s)	Tests Considered	Aggregate Information	Research Components	Brief Summary/Recommendations
Arm, 2003	L.A., M-D, F-T		mechanical properties of recycled aggregates	M-D and L.A. ranked materials similarly. M-D offered the best differentiation between the various material types.
Bjarnason et al., 2002	L.A., M-D, MgSO ₄ , F-T	20 sources	correlations between test results	Tests were broken into three categories: 1) fragmentation (e.g., L.A.), 2) weathering or durability (e.g., MgSO ₄ , F-T) and 3) abrasion (e.g., M-D). Tests within each category correlated well to one another, so using any test within a particular category will help assess aggregate quality.
Brandes and Robinson, 2006	M-D, L.A., MgSO ₄ , NaSO ₄	12 quarries	mechanical and chemical tests on aggregates, statistical correlations between results, correlations with pavement performance	Correlation between M-D and L.A. was low. M-D and sulfate soundness tests correlated well. MgSO ₄ and NaSO ₄ correlated well to one another. L.A. and M-D had poor correlations with pavement performance. MgSO ₄ had best correlation to pavement performance. NaSO ₄ had second best correlation to pavement performance.
Cooley et al., 2002	M-D, L.A., MgSO ₄ , NaSO ₄	72 aggregates from 8 states	statistical correlations between the various methods and between repeated tests	No correlation between M-D and L.A., M-D and MgSO ₄ or M-D and NaSO ₄ . Aggregate type may affect correlations between M-D and performance ratings.
Fowler et al., 2006	M-D, L.A., MgSO ₄ , ACV, F-T, H ₂ O, G _s	117 sources within US and Can.	correlations between test results, correlations with field performance	Very little correlation between tests, best being the L.A. to British ACV ($R^2 = 0.650$), M-D to MgSO ₄ ($R^2 = 0.600$). M-D works well to predict field performance, works even better when used in conjunction with Canadian F-T, MgSO ₄ or G _s .
Hunt, 2001	M-D, NBM, L.A.	44 samples from 22 sources	investigating alternative means of measuring aggregate durability	M-D results repeatable between two technicians. Little differences between the M-D and L.A. in predicting field performance. NBM was able to identify aggregate quality.
Jayawickrama et al., 2006	M-D, MgSO ₄ , Pet	52 aggregate sources	statistical analysis, correlations between repeated tests	Correlation between M-D and MgSO ₄ ; M-D more repeatable than MgSO ₄ , recommends using M-D as a quality control tool.
Kline et al., 2006	NaSO ₄	50 tests	literature review, comparisons of NaSO ₄ to various physical properties	NaSO ₄ not related to other aggregate material properties, literature review stated low repeatability is evident in NaSO ₄ , recommended not to use the NaSO ₄ to determine whether a particular aggregate should be used.
Lim, 2004	WAV, L.A., M-D	ballast material	investigate correlations between a simulated box test and other aggregate durability tests	M-D, WAV and L.A. all worked well to identify ballast field performance.
Prowell et al., 2005	L.A., M-D, MgSO ₄ , NaSO ₄	N/A	survey, literature review, performance data review	Replace Sulfate Soundness tests with Micro-Deval, use freeze-thaw test in conjunction with Micro-Deval to improve results.
Rangaraju et al., 2005	M-D, L.A., MgSO ₄ , NaSO ₄	23 sources (19 classified as granite) in South Carolina	correlations between test results, correlations with field performance	No correlation between M-D and L.A., MgSO ₄ and NaSO ₄ had very good correlation. M-D better than L.A. to identify marginal aggregates based on field performance, aggregate size affects outcome of M-D test – smaller aggregates show greater loss.
Rismantojo, 2002	L.A., M-D, MgSO ₄	5 course, 6 fine	statistical correlations of test results, test sections	Good correlation between M-D and MgSO ₄ , correlation between M-D and water absorption.

Author(s)	Tests Considered	Aggregate Information	Research Components	Brief Summary/Recommendations
Rogers et al., 1991	MgSO ₄ , M-D,	fine aggregates	several factors compared and test methods	MgSO ₄ not very precise for fine aggregates and is time consuming; M-D correlates well with MgSO ₄ , M-D variability is low and well-suited to identify fine aggregates that are derived from weak and poor rocks.
Senior and Rogers, 1991	F-T, M-D, AIT, PSV, AAV, L.A., MgSO ₄ , H ₂ O, Pet	granular base courses, Portland cement concrete, surface course asphalt concrete	statistical correlations of test results, correlations with field performance	Granular bases: M-D+Pet works well to distinguish between good and bad aggregates, M-D alone able to distinguish between marginal and good aggregates, M-D+H ₂ O able to identify poor aggregates; PCC: M-D+F-T able to distinguish between marginal and poor aggregates, H ₂ O+M-D or H ₂ O+F-T able to identify poor aggregates; Surface Course Asphalt Concrete: M-D+PSV able to identify good aggregates.
Tarefder et al., 2003	M-D, L.A., F-T, ADI, G _s , H ₂ O	18 aggregates	compare results between various tests, evaluate repeatability, correlations with field performance	M-D highly repeatable, M-D correlates well with other tests for sandstone and not as well for limestone, L.A. did not accurately predict field performance, M-D did accurately predict field performance of aggregates.
Wu et al., 1998	L.A., AIV, ACV, M-D, Gyr, NaSO ₄ , MgSO ₄ , F-T, DIT, F-T (Can.)	16 sources from 12 states	survey, literature review, performance data statistical review	L.A. and NaSO ₄ did not predict pavement performance as well as M-D and MgSO ₄ .

Notes:

AAV = Aggregate Abrasion Value
 ACV = Aggregate Crushing Value
 ADI = Aggregate Durability Index
 AIT = Aggregate Impact Test
 AIV = Aggregate Impact Value
 F-T = Freeze-Thaw
 G_s = Specific Gravity
 Gyr = Gyrotory Compactor
 H₂O = Water Absorption
 L.A. = L.A. Abrasion
 M-D = Micro-Deval
 MgSO₄ = Magnesium Sulfate Soundness
 NaSO₄ = Sodium Sulfate Soundness
 NBM = Nordic Ball Mill
 Pet = Petrographic Analysis
 PSV = Polished Stone Value
 WAV = Wet Attrition Value

3 RESULTS AND ANALYSIS

Comparison laboratory testing was conducted using the Micro-Deval, Sodium Sulfate, and L.A. Abrasion tests to examine the repeatability of each test and to compare the results in terms of aggregate durability and degradation.

3.1 Soil Samples

The soils examined in this study were primarily granular and cohesionless, typical of material that would be excavated and processed for use as aggregate for plant mix asphalt or crushed base course. A complete list of the soils and the quantity of tests conducted on each soil are provided in Table 13. Micro-Deval tests for soils 1 – 20 were conducted in the Montana State University (MSU) soils laboratory. L.A. Abrasion tests for soils 1 – 20 were conducted in either the MSU soils laboratory or the Montana Department of Transportation (MDT) Helena materials laboratory. All of the Sodium Sulfate tests, and the L.A. Abrasion and Micro-Deval tests for soils 21 – 32 were conducted in the MDT Helena materials laboratory. Soil samples were obtained by MDT district personnel from borrow areas and gravel pits throughout the state of Montana, as shown graphically in Figure 1. Specific source and location designations as well as MDT lab test numbers are provided in Table 13.

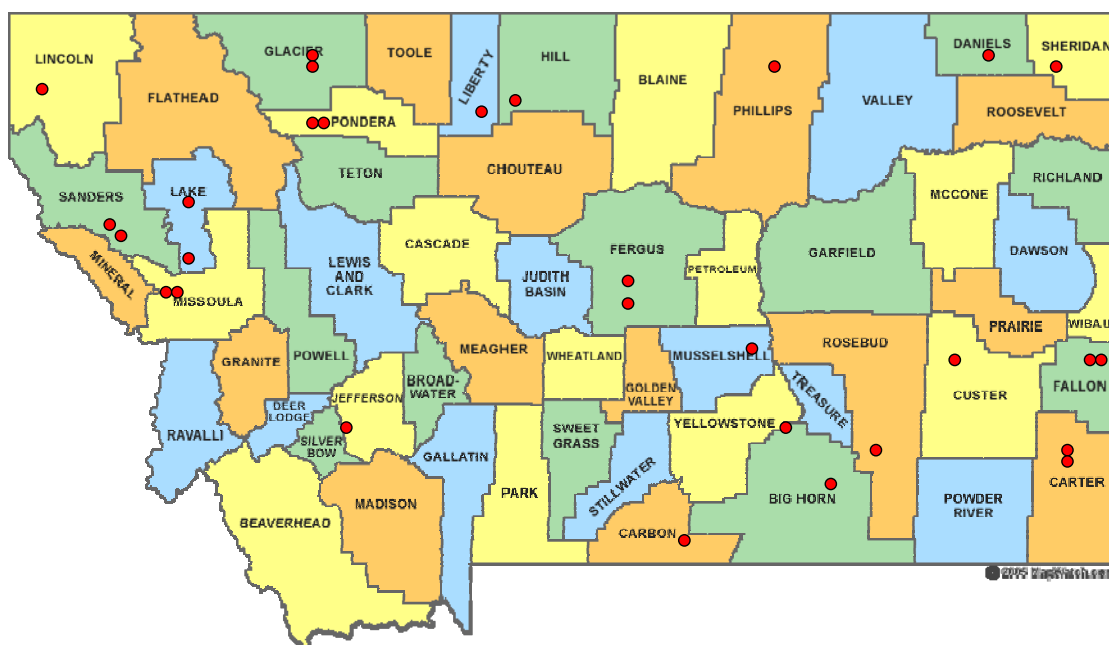


Figure 1. Soil sample locations in Montana.

Table 13. Summary of Materials Examined in This Study

Soil No. ¹	MDT Lab No.	Source or Location ²	Total Number of Tests Conducted		
			L.A. Abrasion	Micro-Deval	Sodium Sulfate Soundness
1	861553	John Depuydt, Whitewater	3	5	1
2	861554	Reece Puckett, Flaxville	7	5	1
3	861555	State of Montana Pit	4	5	1
4	861556	John Montgomery, Hilger	5	5	1
5	861557	DNRC/John Montgomery, Hilger	3	5	1
6	861678	Summers Pit, Ekalaka	5	5	1
7	861871	Newman Source, Melstone	10	5	1
8	861872	Newman Source, Melstone	0	5	1
9	863381	Pugsley Ranches Inc., Chester	11	5	1
10	863382	Marsha D. Horel, Rudyard	6	5	1
11	863383	Pioneer Concrete/Rich Deihl	8	5	1
12	863384	Veland Ranch Source	3	5	1
13	864332	Peter Andersen, Lewistown	5	5	1
14	864941	Singer, N. Dunmore, Crow Agency	15	5	1
15	864944	Art Peaslee, Miles City	4	5	1
16	864945	Barone Ranch Inc., Savage	13	5	1
17	864946	Donald Fried, Baker	8	5	1
18	864947	Allan and Ruth Wang, Baker	11	3	1
19	865006	Lawyer Nursery, Plains	6	5	1
20	865937		4	5	1
<hr/>					
21	NC25647	Summers Pit, Ekalaka	0	3	1
22	NC25648	Summers Pit, Ekalaka	0	3	1
23	NC25649	Brewer Pit, Rosebud County	1	3	1
24	854276	KH Concrete	1	1	0
25	856725	Krug and Sons	1	1	1
26	855919	S and T Redi-Mix, Belfry	1	1	1
27	854578	JTL Inc., Bozeman	1	1	1
28	855264	Missoula Ready Mix, Arlee	1	1	1
29	856443	Jensen Contracting, Ronan	1	1	0
30	857696	NH-NHTSA-1-3(36)234F[0594B]	0	1	1
31	858053	Baxbaum Pit	1	1	0
32	857540	Butte-Spangler-Gilman	1	1	0

1. Tests for soil sample numbers 21 – 32 were conducted at the MDT Helena lab during the summer of 2005

2. Locations and sources are all in Montana

3.2 Laboratory Test Results

The Micro-Deval, L.A. Abrasion, and Sodium Sulfate tests are used to characterize the durability of aggregate based on the percent loss of material that occurs during the test. Micro-Deval tests in this study were conducted according to standard test method AASHTO T327 (ASTM D6928), L.A. Abrasion tests were conducted according to standard test method AASHTO T96 (ASTM C131), and Sodium Sulfate tests were conducted by Montana Department of Transportation according to standard test method AASHTO T104 (ASTM C88) (AASHTO, 2005 and ASTM, 2003). Each method differs in their treatment of the aggregate during testing; and consequently, each method produces a unique value of percent loss, which is used to distinguish between durable and non-durable aggregates. For the purposes of this study, the following percent loss pass-fail standards were used for each test:

- Micro-Deval: passing (i.e., durable), if % loss $\leq 18\%$;
- L.A. Abrasion: passing (i.e., durable), if % loss $\leq 40\%$; and
- Sodium Sulfate: passing (i.e., durable), if % loss $\leq 12\%$.

The testing program provided data for evaluating the repeatability of multiple tests conducted on the same material (5 repeats for the Micro-Deval test and 3 or more repeats for the L.A. Abrasion test) and for examining correlations between the three tests. Test results are summarized in Table 14 in terms of average percent loss values and coefficients of variation (COV) for the Micro-Deval and L.A. Abrasion tests. The scope of this study did not include a repeatability evaluation of the Sodium Sulfate test; consequently, the Sodium Sulfate percent loss values presented in Table 14 were obtained from single tests on each aggregate. As can be observed in Table 14, 78% of the aggregate samples passed all three tests, indicating the majority of aggregates examined in this study would be characterized as durable. Normalized values of percent loss were used in many of the comparisons provided in the following sections, because of the different pass-fail criteria used for each test.

The L.A. Abrasion results presented in this report, and summarized in Table 14, exclude tests that were determined to contain errors. A change in undergraduate testing personnel occurred during the later phases of the project. While training a new student for the project, it was observed that the previous student had made a minor deviation from the standard L.A. Abrasion test procedures. Because the deviation was minor, it was physically difficult to detect, and consequently it was not observed in time by the principal investigators. Unfortunately, the testing error influenced the numeric results of the majority of L.A. Abrasion tests conducted by this student at the MSU lab. The mistake involved an incorrect placement of a catch pan beneath the ball mill drum. The rectangular-shaped catch pan inadvertently was oriented 90 deg from the proper direction. Because of the incorrect orientation of the pan, a small quantity of material was not captured at the end of each test. The amount of material missed by the pan varied from test

to test. Consequently, after considerable study, it was determined that the earlier test results could not be corrected with any degree of certainty and were therefore excluded from the study.

Table 14. Summary of Test Results

Soil No. ¹	MDT Lab No.	Micro-Deval ²		L.A. Abrasion ²		Sodium Sulfate ²
		Average (% Loss)	COV (%)	Average (% Loss)	COV (%)	Single Test (% Loss)
1	861553	2.1	28.0	17.0	15.6	1.0
2	861554	3.9	5.9	16.3	5.8	0.0
3	861555	5.7	4.0	20.5	2.8	0.0
4	861556	16.5	6.9	34.4	1.6	4.0
5	861557	15.8	9.4	36.7	3.1	8.0
6	861678	40.3	2.3	45.2	0.9	11.0
7	861871	6.0	10.4	22.0	12.7	1.0
8	861872	13.7	4.9	---	---	1.0
9	863381	7.4	6.9	19.6	4.6	1.0
10	863382	17.5	9.0	28.4	4.6	6.0
11	863383	6.6	5.3	11.9	10.9	1.0
12	863384	23.4	2.3	36.3	1.7	16.0
13	864332	36.4	5.9	34.2	7.0	14.0
14	864941	5.7	3.9	22.3	5.4	0.0
15	864944	21.3	3.0	28.4	2.5	6.0
16	864945	6.0	15.0	20.5	9.3	3.0
17	864946	41.0	1.1	36.0	7.5	38.0
18	864947	32.1	3.2	35.5	4.8	24.0
19	865006	11.4	4.9	29.6	5.3	2.0
20	865937	16.4	8.8	24.0	3.3	8.0
21	NC25647	61.3	2.6	---	---	30.0
22	NC25648	59.1	1.9	---	---	17.0
23	NC25649	3.5	10.8	---	---	0.87
24	854276	5.6	19.6	22.0	---	---
25	856725	6.6	2.3	25.0	---	0.94
26	855919	7.4	0.9	30.0	---	1.0
27	854578	5.7	19.0	16.0	---	0.2
28	855264	10.1	0.0	16.0	---	0.26
29	856443	8.9	4.5	15.0	---	---
30	857696	8.4	---	---	---	1.0
31	858053	4.1	---	20.0	---	---
32	857540	6.6	1.5	14.0	---	---

1. Tests for soil sample numbers 21 – 32 were conducted at the MDT Helena lab during the summer of 2005

2. Bolded numbers indicate failing values

3.3 Repeatability

Results from the materials tested in this study were examined to investigate the repeatability of multiple tests conducted on the same material. The coefficient of variation was calculated for each material source using data from the Micro-Deval and the L.A. Abrasion tests. The Sodium Sulfate test was not included in this repeatability evaluation because only one Sodium Sulfate test was conducted on most of the materials.

The coefficient of variation (COV) is mathematically the standard deviation of a set of data divided by the mean and multiplied by one hundred to obtain a percent. This value provides information about the variability of the test procedure. A higher value of COV indicates greater variability between repeated tests. The average COV for the Micro-Deval tests was 6.8% and the average COV for the L.A. Abrasion test was 6.7%. If the major outlier for each test is eliminated (28% for the Micro-Deval and 23.6% for the L.A. Abrasion), the average COV values are 6.1% for the Micro-Deval test and 5.8% for the L.A. Abrasion test. These correspond to average standard deviations of 0.7 and 1.7 percent loss for the Micro-Deval and L.A. Abrasion tests, respectively. This single-lab repeatability study indicates both the Micro-Deval and L.A. Abrasion tests have good repeatability with no statistically significant difference between the two tests. Similar conclusions regarding the repeatability of the Micro-Deval test were reported in the literature by Jayawickrama et al. (2006), Tarefder et al. (2003), and Hunt (2001).

The majority of COV values fall within a range of 3% to 15% for both tests, as shown in Figure 2a. No particular trends were observed in the variability of the COV between material types as shown graphically in Figure 2b. The large COV value of 28% for Micro-Deval test sample number 861553 is attributed to the very small average test result for this sample, which was 2.1% loss. Larger standard deviations have a larger affect on low mean values since COV is the standard deviation divided by the mean. If the standard deviation increases and the mean decreases, values for COV can be large.

The Sodium Sulfate test has a reported coefficient of variation of 41% for multi-laboratory testing and 24% for a single laboratory testing, as stated in the AASHTO T104 test standard (AASHTO, 2005).

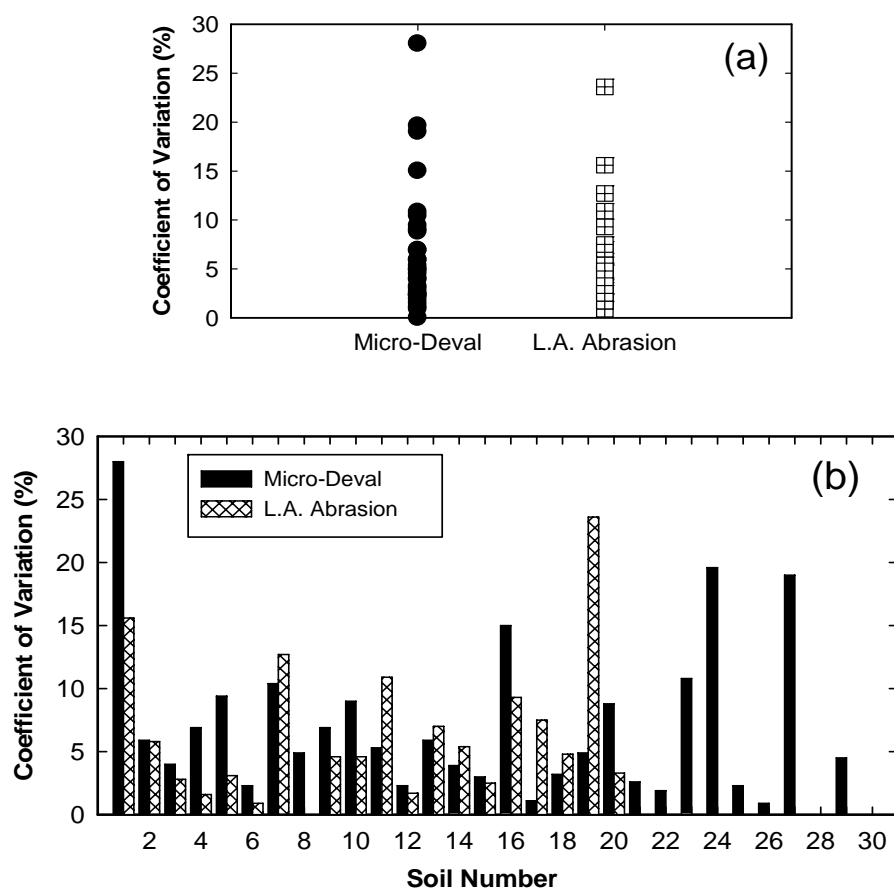


Figure 2. COV for the Micro-Deval and L.A. Abrasion tests: a) compilation of all test data and b) data displayed by soil number.

3.4 Comparison of Three Test Methods

The Micro-Deval, Sodium Sulfate, and L.A. Abrasion tests use a percentage of specific sizes of aggregate lost to differentiate between durable aggregate and non-durable aggregate. Materials that have a percent loss greater than the respective cutoff are considered non-durable; and likewise, materials that have a percent loss that is less than the respective cutoff are considered durable. The percent loss (or cutoff) value used for each test is as follows:

- Micro-Deval = 18%,
- Sodium Sulfate = 12%, and
- L.A. Abrasion = 40%.

Results from the suite of laboratory tests were normalized to facilitate direct comparisons between the three test methods. Normalized test results were obtained by taking the average percent loss for a particular soil and dividing it by the cutoff for that test, as shown in Equation 1.

$$\text{normalized value} = \frac{\text{average percent loss (\%)}}{\text{cutoff (\%)}} \quad (1)$$

For example, five separate Micro-Deval tests were conducted on MDT Sample No. 861556, with an average percent loss of 16.5%. The normalized result for that series of tests is thus calculated as:

$$\text{normalized value} = \frac{16.5\%}{18\%} = 0.917 \quad (2)$$

The normalized value of 0.917 is less than 1.0, indicating the average of the five samples passed the test criteria of 18% maximum loss, and this material would be classified as durable. A normalized value greater than 1.0 is considered failing; that is, the material would be classified as non-durable. This normalized approach was used to make direct comparisons of results obtained using the Micro-Deval, Sodium Sulfate, and L.A. Abrasion tests. The normalized values were plotted together to obtain a graphical representation of the relationship between the tests. An example of the head-to-head comparisons examined in this study is shown in Figure 3 for the normalized Micro-Deval and Sodium Sulfate test results. In general, the plot is broken into four quadrants, explained as follows:

1. top right quadrant (NE), sample failed both the x-axis test (Micro-Deval) and y-axis test (Sodium Sulfate);
2. bottom left quadrant (SW), sample passed both the x-axis test (Micro-Deval) and y-axis test (Sodium Sulfate);
3. top left quadrant (NW), sample passed the x-axis test (Micro-Deval) and failed the y-axis test (Sodium Sulfate); and
4. bottom right quadrant (SE), sample failed the x-axis test (Micro-Deval) and passed the y-axis test (Sodium Sulfate).

Data points that plot in either the NE or SW quadrants indicate that both tests provided consistent results, either pass/pass (SW quadrant) or fail/fail (NE quadrant). Data points that plot in the cross-hatched quadrants are indicative of a discontinuity or disagreement between test results. In other words, one test characterized the material as durable aggregate, while the other test characterized the material as non-durable aggregate, thus signifying poor agreement between tests.

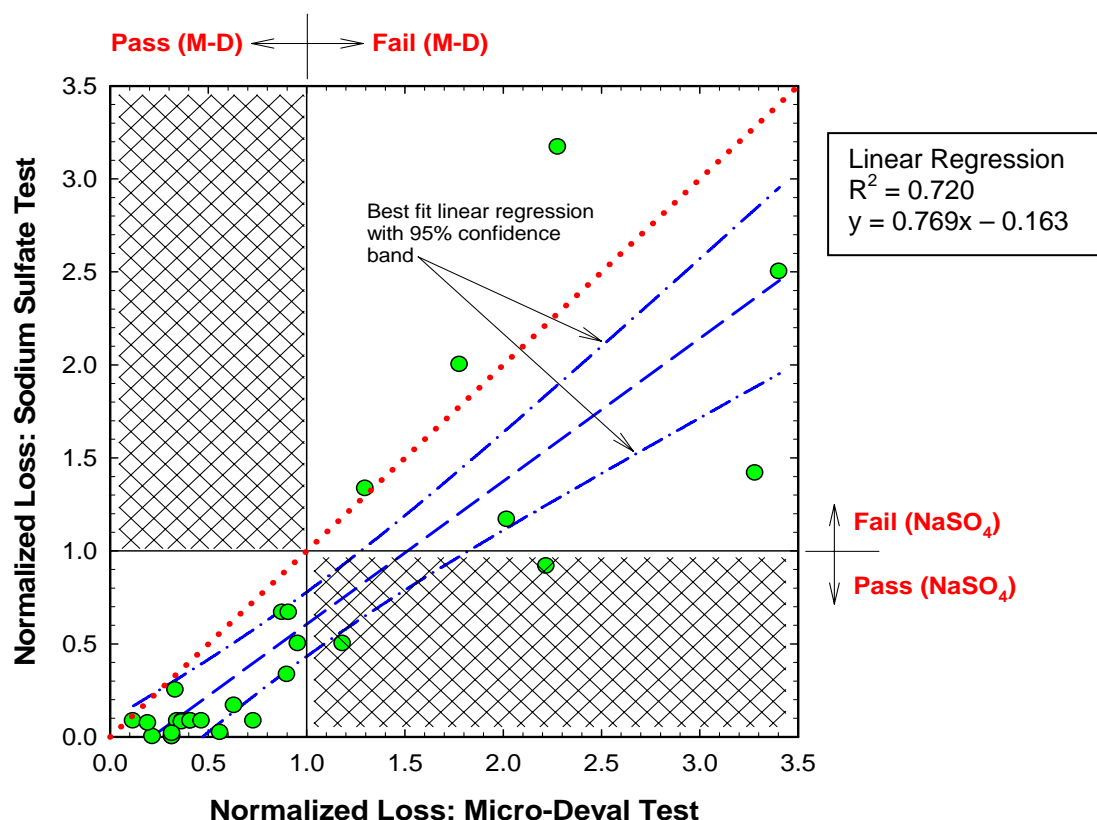


Figure 3. Normalized comparison between Micro-Deval and Sodium Sulfate tests.

A best fit linear regression equation and R^2 value are shown in the plots for each relationship. The R^2 value is a measure of how well the linear regression correlated with the data. Values of R^2 close to 1.0 imply that the data correlates well to a linear best fit line, while values close to 0.0 imply a poor linear correlation between tests. Irrespective of the linearity of the data, a perfect correlation between the two tests would result in a best fit line that passes through the origin and slopes upward at a 45 degree angle measured from horizontal (dotted line in Figure 3). Qualitatively, data points that plot close to this line indicate the two test methods correlated well, while data points that plot relatively far from the 45 degree line indicate poor correlation between tests. Results that are of most concern are those that fall into one of the two cross-hatched zones and are far removed from the 45 degree line. The 45 degree line would be quantified with an R^2 value of 1.0, a slope of 1.0, and a y-intercept of 0.0. The authors believe that the closeness of the data points to this 45 degree line provides the most informative approach for qualitatively evaluating a correlation between test methods.

Data scatter associated with typical regression analysis will also indicate the likelihood that points could fall into *areas* of agreement or disagreement, adjacent to the best fit line. Even in the case of perfect *linear* agreement, data scatter may result in data points that fall within areas of disagreement (i.e., pass/fail or fail/pass zones). Therefore, the standard deviation of data points

about the regression line was used to generate a 95% confidence band about the best fit regression line, as indicated by the dashed-dotted lines in Figure 3. This confidence band provides width to the regression line and takes into account the variation of scatter across the spectrum of normalized results.

3.4.1 Micro-Deval versus Sodium Sulfate

Comparisons between Micro-Deval and Sodium Sulfate tests were conducted on twenty-eight different materials, obtained from different sources. The normalized Micro-Deval results shown in Figure 3 are based on averages obtained from 3 to 5 tests on each material. All of the Sodium Sulfate results are based on a single test. Table 13 contains an itemized list of the number of tests conducted on each material.

The dotted line drawn at a 45 degree angle represents a perfect correlation between test methods. Data points close to the dotted line represent very good correlation between tests, while data points that plot relatively far from the 45 degree line indicate poor correlation between tests. The most concerning results are those that fall into one of the two cross-hatched zones and are far removed from the 45 degree line. In this comparison, results for two of the materials (7.1% of the materials tested) would be considered problematic because the Micro-Deval test indicated the material was non-durable (failure) but the Sodium Sulfate test results indicated the material was durable (passing). These two data points are located in the SE cross-hatched zone in Figure 3. Overall, the majority of the data points, 89.3%, are located on the right side of the 45 deg line indicating the Micro-Deval test provided more “conservative” results than the Sodium Sulfate test for the materials examined in this study. In this context, the term “conservative” means the test results imply a higher degree of non-durability.

Best fit linear correlation between normalized results for the Micro-Deval and Sodium Sulfate tests is somewhat significant, with an R^2 value of 0.72 (shown as a dashed line in Figure 3). The two test methods were in agreement regarding an overall pass/fail determination for 26 out of 28 of the materials, or 92.9%. Both tests had failing results (indicating non-durable aggregate) for six of the materials (7.1%). The 95% confidence band passes through the top left corner of the SE cross-hatched zone. This further highlights the possibility that a failing result could be obtained for a sample using the Micro-Deval test while a conflicting passing result could be obtained from the Sodium Sulfate test.

3.4.2 Micro-Deval versus L.A. Abrasion

Comparisons between Micro-Deval and L.A. Abrasion tests were conducted on twenty-seven different materials, obtained from different sources. The normalized Micro-Deval results shown in Figure 4 are based on an average of multiple tests for each material, typically 3 to 5. The L.A. Abrasion results are based on 3 to 15 tests on each material. Table 13 contains an itemized list of the number of tests conducted on each material.

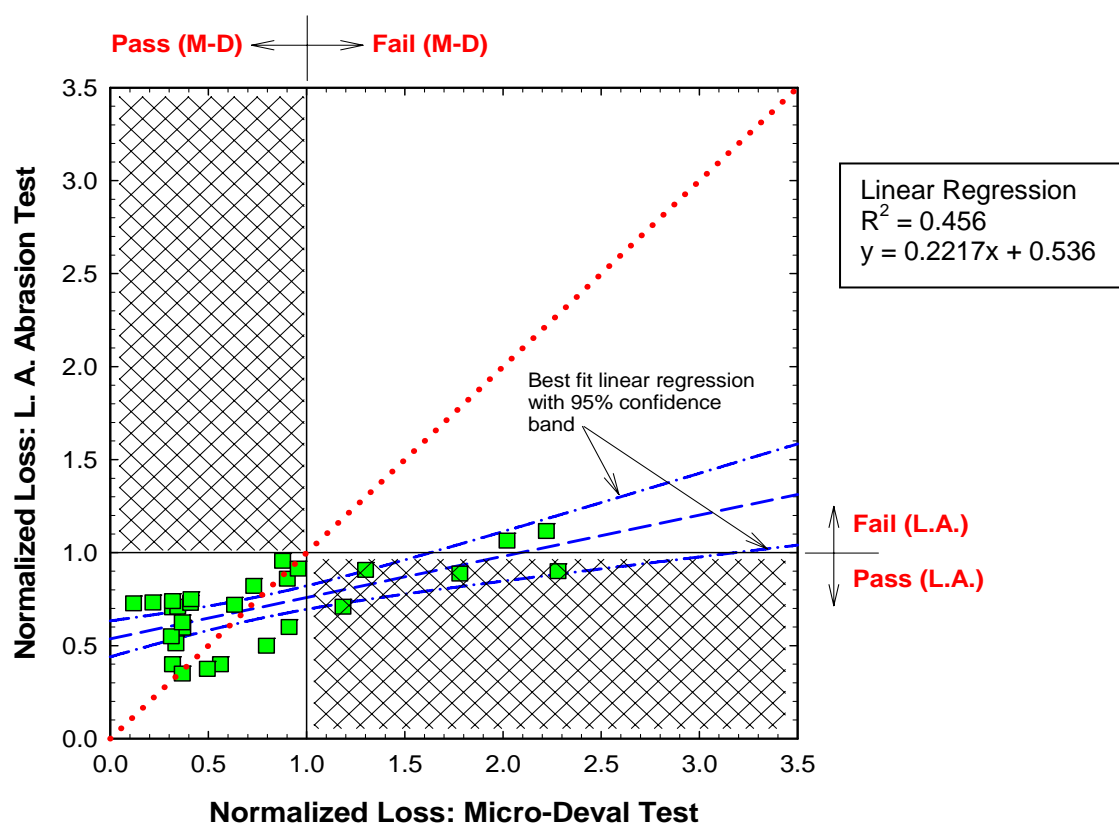


Figure 4. Normalized comparison between Micro-Deval and L.A. Abrasion tests.

In this comparison, four of the materials (14.8% of the materials tested) would be considered problematic because the Micro-Deval test indicated the material was non-durable (failure) but the L.A. Abrasion test results indicated the material was durable (passing). These four data points are located in the SE cross-hatched zone in Figure 4. The four tests that failed the Micro-Deval test were relatively close to the failure cutoff for the L.A. Abrasion test as well.

Overall, the normalized comparison between these two tests exhibited a significant amount of scatter on both sides of the 45 degree line, with about an equal number of data points above and below the 45 degree line. The Micro-Deval test tended to provide more “conservative” results than the L.A. Abrasion test for the materials examined in this study. In this context, the term “conservative” means the test results imply a higher degree of non-durability.

Best fit linear correlation between normalized results for the Micro-Deval and L.A. Abrasion tests yielded an R^2 value of 0.46. The best fit linear regression is shown as a dashed line in Figure 4. The two test methods were in agreement regarding an overall pass/fail determination for 23 out of 27 of the materials, or 85.2%. Both tests had failing results (indicating non-durable aggregate) for 2 of the materials (14.8%). Similar to Figure 3, the 95% confidence band passes through the top left corner of the SE cross-hatched zone. This indicates

the possibility that a failing result could be obtained for a sample using the Micro-Deval test while a conflicting passing result could be obtained from the L. A. Abrasion test.

3.4.3 L.A. Abrasion versus Sodium Sulfate

Comparisons between L.A. Abrasion and Sodium Sulfate tests were conducted on twenty-five different materials, obtained from different sources. The normalized L.A. Abrasion results shown in Figure 5 are based on averages obtained from 3 to 15 tests on each material, while Sodium Sulfate results are based on a single test. Table 13 contains an itemized list of the number of tests conducted on each material.

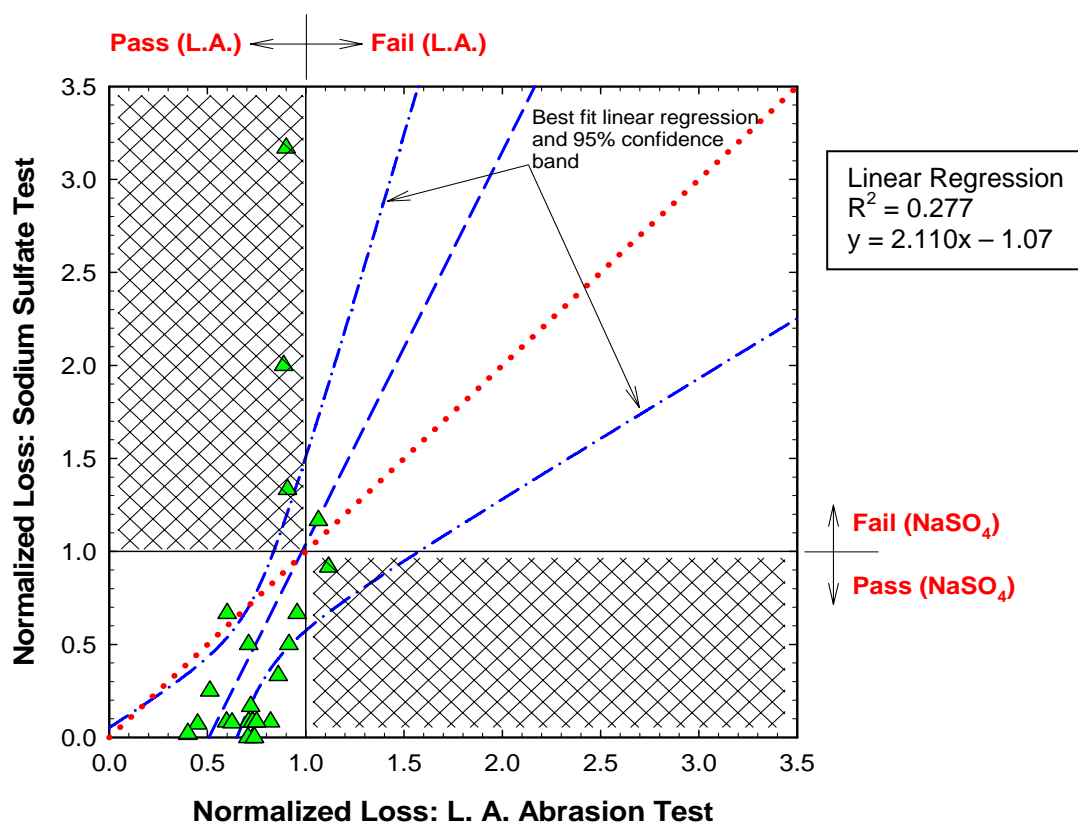


Figure 5. Normalized comparison between L.A. Abrasion and Sodium Sulfate tests.

Statistical correlation between L.A. Abrasion and Sodium Sulfate was rather low based on an R^2 value of 0.28. The best fit linear regression is shown as a dashed line in Figure 5. The two test methods were in agreement regarding an overall pass/fail determination for 21 out of 25 of the materials, or 84.0%. Both tests had failing results (indicating non-durable aggregate) for only one of the materials.

In this comparison, four of the results (16.0% of the materials tested) would be considered problematic because the Sodium Sulfate test indicated the material was non-durable (failure) but the L.A. Abrasion test indicated the material was durable (passing). Three of these “problematic” data points are located in the NW cross-hatched zone and one is located in the SE

cross-hatched zone, in Figure 5. Overall, the majority of the data points, 80.0%, are located on the right side of the 45 degree line in Figure 5, indicating the L.A. Abrasion test provided more “conservative” results than the Sodium Sulfate test for the materials examined in this study. However, the three data points in the NW cross-hatched zone represent important exceptions to this observation. The materials represented by these data points failed the Sodium Sulfate test, but passed the L.A. Abrasion test. Most notably, two of the materials failed the Sodium Sulfate test by large margins (factors of 2 and 3.2), but passed the L.A. Abrasion test. These materials were from MDT sample numbers 864947 and 864946, respectively. These two materials are the primary culprits for the poor correlation observed between the L.A. Abrasion and Sodium Sulfate tests. The 95% confidence band for these tests is hour-glassed shaped and nearly symmetrical about the linear regression line. The narrowest portion of the confidence band occurs near coordinate (0.7, 0.5), indicating the tests correlate reasonably well for aggregate that experiences only minor losses during the tests.

3.5 Summary of Comparisons

Results using the three test methods were available for twenty-four of the materials examined in this study. Figure 6 shows this data graphically in a three-dimensional plot. The majority of data points in this plot are clustered between the origin and the 3-D point (coordinates: 1, 1, 1), indicating a significant number of the materials were characterized as durable aggregates. A complete compilation of laboratory test results for the soils examined during this research project are provided in Table 15.

Based on the laboratory tests conducted in this study, the Micro-Deval, L.A. Abrasion, and Sodium Sulfate tests appear to correlate well for aggregate that has a relatively low percent loss value. Discontinuities between the three tests begin to appear with materials that have percent losses near the cutoff values. Because of the scarcity of borderline and failing tests, the statistical significance of this observation could not be quantified. The following observations are presented based on a qualitative review of the results, and the 95% confidence bands that were created through a statistical evaluation of the data.

1. The largest scatter of data occurred in the comparison between the L.A. Abrasion and Sodium Sulfate tests.
2. Of the three tests, the Sodium Sulfate appears to be the most difficult and time-consuming test to perform. This test also has the poorest record for repeatability and the poorest correlation to field durability.
3. The Micro-Deval test tended to provide more “conservative” results than the L.A. Abrasion and Sodium Sulfate tests.
4. Aggregates that pass the Micro-Deval test will likely also pass the L.A. Abrasion and Sodium Sulfate tests.

5. Based on the 95% statistical confidence bands, the authors suggest that the greatest likelihood of pass-fail conflicts will occur when the percent loss of a sample is slightly greater than the Micro-Deval cutoff criteria.
6. If the Micro-Deval test is selected by MDT as the primary test for evaluating aggregate durability, the authors recommend that an alternate test be implemented whenever an aggregate percent loss is within about 1.0 to 1.3 times the Micro-Deval failure cutoff. In other words, a second verification test is suggested when the Micro-Deval percent loss is between 18 to 24%.

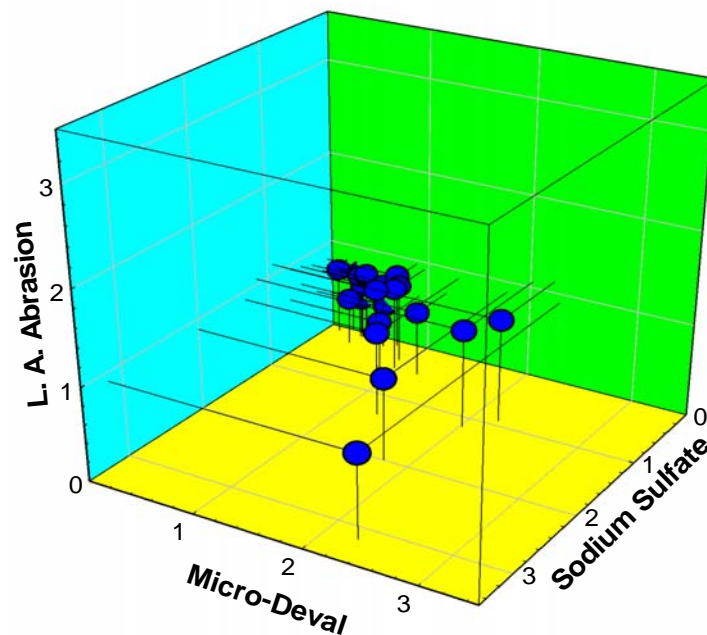


Figure 6. Three-dimensional representation of the test data.

Table 15. Summary of All Data from Soils Tested in This Study

MDT Lab No.	Micro-Deval (% Loss)	L.A. Abrasion (% Loss)	Sodium Sulfate (% Loss)
861553	1.4, 2.0, 2.9, 1.9, 2.5	16.0, 20.0, 15.0	1.0
861554	4.3, 3.8, 3.8, 3.7, 4.0	15.0, 15.0, 17.0, 17.0, 16.0, 17.0, 17.0	0.0
861555	5.9, 5.7, 5.6, 6.0, 5.4	21.0, 21.0, 20.0, 20.0	0.0
861556	16.1, 17.9, 15.1, 15.8, 17.3	35.0, 34.0, 34.0, 34.0, 35.0	4.0
861557	14.9, 15.6, 14.9, 15.2, 18.4	38.0, 36.0, 36.0	8.0
861678	40.0, 40.3, 39.1, 40.6, 41.6	45.0, 45.0, 46.0, 45.0, 45.0	11.0
861871	5.7, 6.8, 6.2, 6.1, 5.1	20.0, 21.0, 20.0, 20.0, 19.0, 21.0, 25.0, 28.0, 22.0, 24.0	1.0
861872	14.3, 13.2, 12.5, 12.9, 12.9		1.0
863381	6.6, 7.1, 7.7, 7.7, 7.7	19.0, 19.0, 20.0, 18.0, 19.0, 19.5, 21.0, 19.2, 20.3, 20.4, 20.2	1.0
863382	17.7, 15.7, 17.1, 18.5, 18.6	30.0, 27.0, 29.2, 26.9, 29.4, 28.1	6.0
863383	7.3, 6.4, 6.5, 6.5, 6.6	10.8, 11.0, 11.0, 10.7, 11.5, 14.0, 13.0, 13.0	1.0
863384	23.8, 23.1, 24.6, 24.1, 23.8	37.0, 36.0, 36.0	16.0
864332	39.9, 36.7, 35.0, 35.7, 34.6	32.0, 35.0, 33.0, 38.0, 33.0	14.0
864941	5.5, 5.7, 5.7, 5.5, 6.1	25.0, 24.0, 22.0, 23.0, 22.0, 23.0, 22.0, 20.0, 22.0, 21.0, 22.0, 22.0, 23.0, 21.0, 22.0	0.0
864944	21.4, 21.5, 20.4, 22.1, 21.3		6.0
864945	6.1, 5.4, 5.1, 7.4, 6.1	18.0, 19.0, 20.0, 21.0, 20.0, 20.0, 24.0, 18.0, 20.0, 24.0, 20.0, 22.0, 20.0	3.0
864946	41.3, 41.5, 40.3, 41.0, 40.7	38.0, 38.0, 36.0, 36.0, 38.0, 38.0, 32.0, 31.6	38.0
864947	32.6, 32.7, 30.9	40.0, 35.0, 35.0, 34.0, 35.0, 34.0, 35.0, 35.0, 37.0, 35.0, 35.0	24.0
865006	10.8, 11.1, 11.1, 11.6, 12.3	17.6, 18.8, 23.0, 22.0, 27.0, 23.0	2.0
865937	17.9, 14.9, 15.9, 15.4, 18.0	24.0, 23.0, 25.0, 24.0	8.0
NC25647	63.0, 59.8, 61.2		30.0
NC25648	58.0, 60.2, 59.0		17.0
NC25649	3.2, 3.9, 3.3	15	0.87
854276	5.0, 5.0, 6.9	22	
856725	6.6, 6.7, 6.4	25	0.94
855919	7.4, 7.3	30	1.0
854578	4.5, 6.6, 6.0	16	0.2
855264	10.1, 10.1	16	0.26
856443	8.5, 9.3, 8.9	15	
857696	8.4		1.0
858053	4.1	20	
857540	6.5, 6.7, 6.6		

4 CONCLUSIONS AND RECOMMENDATIONS

The objective of this study was to evaluate the suitability of the Micro-Deval test using aggregates typically encountered on Montana highway projects, and to determine if the Micro-Deval test provides better, timelier, and more repeatable information about the quality of an aggregate than the Sodium Sulfate test. The laboratory testing program was structured to examine how well three aggregate durability test methods correlate for a sampling of Montana soils. Aggregate durability tests were conducted on 32 different soils using the Micro-Deval, L.A. Abrasion, and Sodium Sulfate tests. Multiple Micro-Deval and L.A. Abrasion tests were conducted on some of the soil samples to investigate the same-lab repeatability of the test methods.

The methods differ in their treatment of the aggregate during testing; and consequently, each method produces a unique value of percent loss, which is used to distinguish between durable aggregate and non-durable aggregate. For the purposes of this study, the following percent loss pass-fail standards were used for each test:

- Micro-Deval: passing (durable), if % loss $\leq 18\%$;
- L.A. Abrasion: passing (durable), if % loss $\leq 40\%$; and
- Sodium Sulfate: passing (durable), if % loss $\leq 12\%$.

Because of the differences in percent loss criteria for each method, results from the suite of laboratory tests were normalized to facilitate direct comparisons between the three methods. Normalized results were obtained by taking the average percent loss for a particular soil and dividing it by the cutoff for that test. Table 16 summarizes the comparisons between each test using data collected during this study. Based on the metrics identified in the table, the Micro-Deval and Sodium Sulfate tests had the best correlation, while the Micro-Deval/L.A. Abrasion and the L.A. Abrasion/Sodium Sulfate correlations were significant, but not as strong.

Table 16. Summary Comparison of Test Methods

Test Methods	R^2	Pass/Fail Agreement (%)	Inconsistent Durability Determination* (%)
M-D versus NaSO ₄	0.72	92.9	7.1
M-D versus L.A.	0.46	85.2	14.8
L.A. versus NaSO ₄	0.28	84.0	16.0
Perfect Correlation	1.0	100.0	0.0

*Note: Column 4 refers to the percentage of samples that passed one of the tests but failed the other test. This inconsistency is identified as a data point that plots in one of the cross-hatched zones identified in Figures 3, 4, and 5.

The percentages of inconsistent durability determinations (pass or fail inconsistencies) listed in column 4 of Table 16 are indicative of a discontinuity between tests and are probably the most important metric for the comparison study. In this case, one test characterized the material as durable aggregate, while the other test characterized the same material as non-durable aggregate. Qualitatively, the authors believe that an excellent correlation between two test methods is obtained when the percentage of inconsistent results is less than about 5%, values between 5 to 10% signify a good correlation, values between 10 to 20% signify a fair to poor correlation, and values above 20% signify a poor or unreliable correlation between tests.

Multiple tests conducted on samples obtained from the same sources indicate similar values of same-lab repeatability for both the Micro-Deval and L.A. Abrasion tests. The coefficients of variation for the multiple tests were less than 10% for both methods. Considering the natural variability that occurs within an aggregate source, the measured variations were low, indicating good repeatability of the test methods. This conclusion has also been supported by others (Jayawickrama et al., 2006; Tarefder et al., 2003; and Hunt, 2001). Repeatability of the Sodium Sulfate test was not examined in this study.

The relationship between Micro-Deval test results and field performance was not examined in this study; however, evaluations by Fowler et al. (2006), Rangaraju et al. (2005), Tarefder et al. (2003) and Wu et al. (1998) indicate that Micro-Deval test results relate well with field performance. An excellent correlation between rutting performance and Micro-Deval test results were observed by White et al. (2006). They suggest that a maximum cutoff value of 15 percent loss for the Micro-Deval test would limit rutting in hot mix asphalt to 12.5 mm at 20,000 wheel passes.

The Micro-Deval test has good repeatability characteristics and it correlates well with results from the Sodium Sulfate test. Furthermore, the Micro-Deval test is less time-consuming and easier to perform than the Sodium Sulfate test. Consequently, the Micro-Deval test appears to be a suitable substitute for the Sodium Sulfate test for aggregates that are not borderline pass/fail in terms of percent loss durability.

The authors of this study conclude that the Micro-Deval test is a suitable replacement for the Sodium Sulfate test as the primary test for evaluating aggregate durability. However, because there were some inconsistent durability determinations between test methodologies, the authors recommend that the Micro-Deval test results be further supported by a second aggregate durability test whenever the Micro-Deval results *slightly exceed* the cutoff value of 18 percent loss. We define *slightly exceed* as a test result that falls between the cutoff value and plus 30% of the cutoff value. In other words, when the Micro-Deval test result for an aggregate is between 18 and 24 percent loss, a second test should be conducted before any conclusions are made regarding the durability or quality of an aggregate. Suggested alternative tests include

recognized methods such as the Sodium (or Magnesium) Sulfate Soundness test or the L.A. Abrasion test.

5 REFERENCES

- AASHTO. (2005). Standard Specifications for Transportation Materials and Methods of Sampling and Testing, 22nd Edition, Washington, D.C.
- Arm, M. (2003) "Mechanical Properties of Residues as Unbound Road Materials." Ph.D. Thesis. The Royal Institute of Technology, Stockholm, Sweden.
- ASTM. (2003). Annual Book of ASTM Standards, West Conshohocken, Pennsylvania.
- Bjarnason, G., Petursson, P. and Erlingsson, S. (2002) "Quality Assessment of Aggregates for Road Construction: Fragmentation, Weathering, and Abrasion." Presented at Ráðstefna norrænna steinefnaframleiðenda, Public Roads Administration, Reykjavik, Iceland September 12, 2002.
- Brandes, H.G. and Robinson, C.E. (2006). "Correlation of Aggregate Test Parameters to Hot Mix Asphalt Pavement Performance in Hawaii." *Journal of Transportation Engineering*, Vol 132, Issue 1.
- Bullas, J.C. and West, G. (1991). "Specifying Clean, Hard and Durable Aggregate for Bitumen Macadam Roadbase." *Research Report 284*, Transport and Road Research Laboratory, Department of Transport (British).
- Cooley, Allen L., Huner, M.S. and James, R.H. (2002). "Micro Deval Testing of Aggregates in the Southeast." NCAT Report No. 02-09. National Center for Asphalt Technology, Auburn University, Auburn, Alabama.
- Fookes, P.G. Gourley, C.S. and Ohkere, C. (1988). "Rock Weathering in Engineering Time." *The Quarterly Journal of Engineering Geology (British)*, Vol. 21.
- Fowler, D.W., Allen, J.J., Lange, A. and Range, P. (2006). "The Prediction of Coarse Aggregate Performance by Micro-Deval and Other Aggregate Tests." The University of Texas at Austin. International Center for Aggregate Research ICAR Report 507-1F.
- Gandhi, P.M and Lytton, R.L. (1984). "Evaluation of Aggregates for Acceptance in Asphalt Paving Mixtures." *Journal of the Association of Asphalt Paving Technologists*, Vol. 53.
- Garrity, L.V. and Kriege, H.F. (1935). "Studies of Accelerated Soundness Tests." *Proceedings: Highway Research Board*, Vol. 15.
- Hasan, S., Meyer, A.H. and Fowler, D.W. (1991). "Establishment of Acceptance Limits for 4-Cycle MSS and Modified Wet Ball Mill Tests for Aggregates Used in Seal Coats and HMAC Surfaces," *Research Report 12222-1F*, Project 3-9-90/1-1222, Center for Transportation Research, University of Texas, Austin, Texas.
- Hunt, E. (2001). "Micro-Deval Coarse Aggregate Test Evaluation." Oregon Department of Transportation.
- Jayawickrama, P., Hossain, S.M. and Phillips, F. (2006). "Evaluation of Aggregate Durability Using Micro-Deval Test." *Proceedings: TRB 2006 Annual Meeting*, Transportation Research Board, Washington D.C.

- Kandhal, P.S. and Parker Jr., F. (1998). "Aggregate Tests Related to Asphalt Concrete Performance in Pavements." NCHRP Report 405, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Kline, S.W., Phiukhao, W., Griffin, M.L. and Miller, J.W. (2006 – in press). "Evaluation of the Sodium Sulfate Soundness Test for Qualifying Dolomites of Northern Arkansas for Construction Aggregate." Proceeding: Forum on the Geology of Industrial Minerals, Indiana University.
- Lim, W.L. (2004). "Mechanics of Railway Ballast Behavior." Ph.D. Thesis. The University of Nottingham, Nottingham, United Kingdom.
- Papalentiou, C.G., Meyer, A.H. and Fowler, D.W. (1987). "Evaluation of the 4-cycle Magnesium Sulfate Soundness Test." *Research Report 438-1F*, Project 3-9-85-438, Center for Transportation Research, University of Texas, Austin, Texas.
- Paul, I. (1932). "Magnesium Sulfate Accelerated Soundness Test on Coarse Aggregates." *Proceedings: Highway Research Board*, Vol. 12, p. 319.
- Prowell, B.D., Zhang, J. and Brown, E. R. (2005). "Aggregate Properties and the Performance of Superpave-Designed Hot Mix Asphalt," NCHRP Report 539, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Rangaraju, P.R., Edlinski, J. and Amikhanian, S. (2005). "Evaluation of South Carolina Aggregate Durability Properties." South Carolina Department of Transportation. Report No. FHWA-SC-05-01. Final Report, January 31, 2005.
- Rismantojo, E. (2002). "Permanent Deformation and Moisture Susceptibility Related Aggregate Tests for use in Hot-Mix Asphalt Pavements." Ph.D. Thesis. Purdue University, West Lafayette, Indiana.
- Rogers, C.A., Bailey, M.L. and Price, B. (1991). "Micro-Deval Test for Evaluating the Quality of Fine Aggregate for Concrete and Asphalt." *Transportation Research Record 1301*.
- Senior, S.A and Rogers, C.A. (1991). "Laboratory Tests for Predicting Coarse Aggregate Performance in Ontario," *Transportation Research Board 1301*.
- Shelburne, T.E. (1940). "Surface Treatment Studies," Proceedings: Association of Asphalt Paving Technologist, Vol. 11, pp. 44-58.
- Shergold, F.A. (1948). "A Review of Available Information on the Significance of Roadstone Tests." Road Research Technical Paper 10. DSIR, H.M.S.O., London.
- Tarefder, R.A., Zaman, M. and Hobson, K. (2003). "Micro-Deval Test for Evaluating Properties of Roadway Aggregate." *International Journal of Pavements: maintenance and rehabilitation of pavements and technological control*, Vol. 2, Issue 1-2.
- White, T.D, Haddock, J.E. and Rismantojo, E. (2006). "Aggregate Tests for Hot-Mix Asphalt Mixtures Used in Pavements." NCHRP Report No. 557, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Woolf, D.O. (1937). "The Relation Between Los Angeles Abrasion Test Results and the Service Records of Coarse Aggregates," *Proceedings, Highway Research Board*, p. 350.

- Wu, Y., Parker, F. and Kandhal, K. (1998). "Aggregate Toughness/Abrasion Resistance and Durability/Soundness Tests Related to Asphalt Concrete Performance in Pavements." NCAT Report N. 98-4, National Center for Asphalt Technology, Auburn University, Auburn, Alabama.

175 copies of this public document were produced at an estimated cost of \$1.55 each, for a total cost of \$271.14. This includes \$119.25 for postage and \$151.89 for printing.