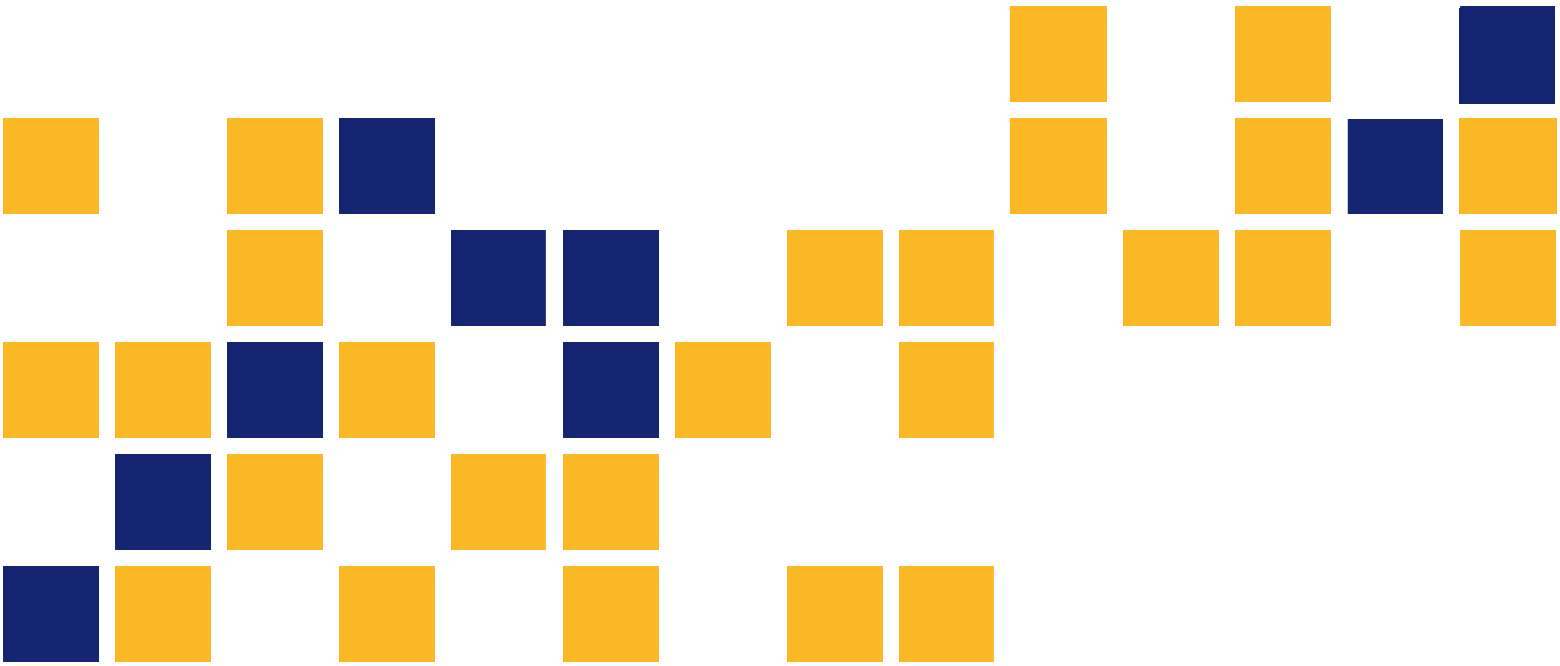
Report No. FHWA-KS-16-13 ▪ FINAL REPORT▪ December 2016

Real Time Laser Scanning of Aggregate Materials in Highway Construction

Warren H. Chesner
Chesner Engineering, P.C.

Nancy J. McMillan
New Mexico State University

A Transportation Pooled Fund Study - TPF-5(278)



Kansas Department of Transportation

1 Report No. FHWA-KS-16-13	**2 Government Accession No.**	**3 Recipient Catalog No.**
4 Title and Subtitle Real Time Laser Scanning of Aggregate Materials in Highway Construction		**5 Report Date** December 2016 **6 Performing Organization Code**
7 Author(s) Warren H. Chesner, Nancy J. McMillan		**7 Performing Organization Report No.**
9 Performing Organization Name and Address Chesner Engineering, P.C. 38 W. Park Avenue, Ste 200 Long Beach, New York 11561		**10 Work Unit No. (TRAIS)** **11 Contract or Grant No.** C1964
12 Sponsoring Agency Name and Address Kansas Department of Transportation Bureau of Research 2300 SW Van Buren Topeka, Kansas 66611-1195		**13 Type of Report and Period Covered** Final Report June 2013–November 2016 **14 Sponsoring Agency Code** RE-0617-01 TPF-5(278)

15 Supplementary Notes
For more information write to address in block 9.
Pooled Fund Study TPF-5(278) sponsored by the following DOTs: Kansas, Oklahoma, New York, Ohio, and Pennsylvania.

The quality and service life of the roadways that make up the highway transportation infrastructure are dependent upon the selection and use of high quality aggregate materials. Five state transportation agencies participated in this Transportation Pooled Fund (TPF) study, which was designed to demonstrate the use of laser scanning as a means to assess, in real-time, the quality of aggregate used in highway construction. Participating states included Kansas, New York, Ohio, Oklahoma, and Pennsylvania.

The referenced technology is based on a process referred to as Laser Induced Breakdown Spectroscopy (LIBS). In this process, a high-powered laser pulse is used to excite atoms that make up the aggregate. This excitation results in the emission of light from a range of unique wavelengths (spectrum) that can be thought of as a "fingerprint" of the material. The development of a database of spectra or fingerprints of many aggregate materials with known engineering properties provides the basis for employing numerical techniques (models), similar to "fingerprint matching," to identify the properties of unknown aggregate material.

Scanning data generated in this demonstration show that the technology can differentiate between approved and unapproved aggregate sources. It has the potential to quantify specific test parameters such as acid insoluble residue (AIR), Micro-Deval loss, and specific gravity, as well as to identify the presence of deleterious materials, such as reactive chert, ASR and ACR, and D-cracking susceptible aggregate. It can be used to identify the aggregate source or sources of a stockpile of unknown material(s).

A total of 113 aggregates supplied by the participating states were laser-scanned using a field prototype system located in a field materials testing laboratory in South Bethlehem, New York. The analyses in this demonstration focused on specific gravity (bulk and SSD) and absorption, D-cracking, acid insoluble residue, Micro-Deval, and Los Angeles (LA) Abrasion Loss. The results show that laser scanning can successfully predict the properties of aggregate, opening up a whole new way of analyzing aggregate materials. Based on the results presented, recommended future work is outlined, some of which has been initiated and presented herein to refine the scanning and modeling process to enhance data quality.

17 Key Words Laser Induced Breakdown Spectroscopy, Laser Scanning, Sample Laser Targeting System, Aggregate Quality		**18 Distribution Statement** No restrictions. This document is available to the public through the National Technical Information Service www.ntis.gov.	
19 Security Classification (of this report) Unclassified	**20 Security Classification (of this page)** Unclassified	**21 No. of pages** 62	**22 Price**

Form DOT F 1700.7 (8-72)

This page intentionally left blank.

Real Time Laser Scanning of Aggregate Materials in Highway Construction

Final Report

Prepared by

Warren H. Chesner

Chesner Engineering, P.C.

Nancy J. McMillan

New Mexico State University

A Report on Research Sponsored by

THE KANSAS DEPARTMENT OF TRANSPORTATION
TOPEKA, KANSAS

December 2016

NOTICE

The authors and the state of Kansas do not endorse products or manufacturers. Trade and manufacturers names appear herein solely because they are considered essential to the object of this report.

This information is available in alternative accessible formats. To obtain an alternative format, contact the Office of Public Affairs, Kansas Department of Transportation, 700 SW Harrison, 2nd Floor – West Wing, Topeka, Kansas 66603-3745 or phone (785) 296-3585 (Voice) (TDD).

DISCLAIMER

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 views or the policies of the state of Kansas. This report does not constitute a standard, specification or regulation.

Abstract

The quality and service life of the roadways that make up the highway transportation infrastructure are dependent upon the selection and use of high quality aggregate materials. Five state transportation agencies participated in this Transportation Pooled Fund (TPF) study, which was designed to demonstrate the use of laser scanning as a means to assess, in real-time, the quality of aggregate used in highway construction. Participating states included Kansas, New York, Ohio, Oklahoma, and Pennsylvania.

The referenced technology is based on a process referred to as Laser Induced Breakdown Spectroscopy (LIBS). In this process, a high-powered laser pulse is used to excite atoms that make up the aggregate. This excitation results in the emission of light from a range of unique wavelengths (spectrum) that can be thought of as a "fingerprint" of the material. The development of a database of spectra or fingerprints of many aggregate materials with known engineering properties provides the basis for employing numerical techniques (models), similar to "fingerprint matching," to identify the properties of unknown aggregate material.

Scanning data generated in this demonstration show that the technology can differentiate between approved and unapproved aggregate sources. It has the potential to quantify specific test parameters such as acid insoluble residue (AIR), Micro-Deval loss, and specific gravity, as well as to identify the presence of deleterious materials, such as reactive chert, ASR and ACR, and D-cracking susceptible aggregate. It can be used to identify the aggregate source or sources of a stockpile of unknown material(s).

A total of 113 aggregates supplied by the participating states were laser-scanned using a field prototype system located in a field materials testing laboratory in South Bethlehem, New York. The analyses in this demonstration focused on specific gravity (bulk and SSD) and absorption, D-cracking, acid insoluble residue, Micro-Deval, and Los Angeles (LA) Abrasion Loss. The results show that laser scanning can successfully predict the properties of aggregate, opening up a whole new way of analyzing aggregate materials. Based on the results presented, recommended future work is outlined, some of which has been initiated and presented herein to refine the scanning and modeling process to enhance data quality.

Acknowledgements

Primary support for the work presented in this report was provided by five state transportation agencies that participated in Transportation Pooled Fund Agreement TPF-5(278):

1. Kansas Department of Transportation (KDOT; Lead State Agency),
2. Oklahoma Department of Transportation (OKDOT),
3. New York State Department of Transportation (NYSDOT),
4. Ohio Department of Transportation (OHDOT), and
5. Pennsylvania Department of Transportation (PennDOT).

The authors gratefully acknowledge the technical input and project guidance provided by Randy Billinger and Rodney Montney (KDOT), Scott Seiter (OKDOT), Marilyn Bradley, Nina Kraev, Bryce Nelson, William Skerritt, Randell Wells and Gary Frederick (NYSDOT), Mickey Cronin and Jeff Wigdahl (OHDOT), Michael Matthews (PennDOT), and Richard Meininger (FHWA).

Supplementary support for the development, design, and fabrication of the laser scanning system prototype employed in this TPF effort was provided by the Transportation Research Board (NCHRP 168) and the New York State Energy Research and Development Authority (Agreement 25533).

Old Castle Materials and Callanan Industries, Inc., generously provided facilities to house and operate the laser scanning system, including facilities for sample storage. Special recognition and appreciation is due to Cindy Lafleur (Old Castle), Sheila Barkevich, Steve Aker, and Derek Berner (Callanan Industries) for their extra level of effort to ensure a successful project.

Special thanks are extended to the TPF project coordination staff at Fort Hays State University, and in particular, the guidance and support of Dr. Kenneth Neuhauser and Rachel Depenbusch for their efforts in ensuring an efficient administrative process.

Finally, this project could not have been undertaken without the assistance of primary research team participants from New Mexico State University (NMSU) and Chesner Engineering, P.C. From NMSU, these participants included Carlos Montoya, John Curry, and Trent Haskell; and from Chesner Engineering, P.C., Henry Justus, Matteo Forgione, Chris Stein, and Carlos Quiles.

Table of Contents

Abstract v
Acknowledgements vi
Table of Contents vii
List of Tables ix
List of Figures x
Chapter 1: Introduction 1
1.1 Overview 1
1.2 Objectives and Scope 3
Chapter 2: Methodology 4
2.1 Hardware and Process Development 4
2.2 Aggregate Spectra 6
2.3 SLT Modeling 7
Chapter 3: Results 10
3.1 Test Samples and Known Engineering Properties 10
3.2 Measurement of Engineering Properties 12
3.2.1 Bulk, Saturated Surface Dry (SSD), and Apparent Specific Gravity Results 12
3.2.2 D-Cracking Results 17
3.2.3 Acid Insoluble Residue (AIR) Results 19
3.2.4 LA Loss Testing 22
3.2.5 Absorption Testing 24
3.2.6 Micro-Deval Loss Results 26
3.2.7 Presence and Quantification of Deleterious Materials: Chert 29
3.2.8 Production Samples Analysis 30
Chapter 4: Discussion of Results 34
4.1 Overview 34
4.2 Mineralogical Heterogeneity 34
4.3 Orientation Heterogeneity 35
4.3.1 Kansas Specific Gravity Results Using Filtered Data 39
4.4 Lab Data Heterogeneity 42

4.5 Laser Power and Signal Reception Heterogeneity 42
4.6 Addressing the Issues 43
Chapter 5: Findings and Conclusions 44
Chapter 6: Recommendations 46
References 47

List of Tables

Table 3.1: Summary of Number of Samples and Test Data Provided by States 10
Table 3.2: Distribution of Sample Types and Sources from Five States 11
Table 3.3: Results of D-Cracking Pass-Fail Models with KDOT Samples 19
Table 3.4: Measured AIR Values for NYSDOT Samples 20
Table 3.5: Test Set Validation for LA Loss (%) using Nine Samples from KS, OH, OK, and PA 24
Table 4.1: The Number and Percent of Spectra as a Function of Total Intensity 39
Table 4.2: Values of Saturated Specific Gravity Calculated for Test Set Validation Samples 42

List of Figures

Figure 2.1: Photograph of SLT at Callanan Industries, Inc., Materials Testing Lab in South Bethlehem, NY 4
Figure 2.2: Operator Introducing Aggregate into SLT for Scanning 5
Figure 2.3: SLT LIBS Process Equipment Concept 6
Figure 2.4: Oklahoma Limestone Spectral Pattern 6
Figure 3.1: Kansas Bulk Specific Gravity Calibration (Black Diamonds) and Test-Set Validation (Green Circles) 13
Figure 3.2: Kansas SSD Specific Gravity Calibration (Black Diamonds) and Test-Set Validation (Green Circles) 14
Figure 3.3: Kansas Apparent Specific Gravity Calibration (Black Diamonds) and Test-Set Validation (Green Circles) 14
Figure 3.4: Pennsylvania Calibration Model for Apparent Specific Gravity 15
Figure 3.5: Pennsylvania Cross-Validation Model for Apparent Specific Gravity 15
Figure 3.6: Combined States (KS, OK, and OH) Bulk Specific Gravity Calibration and Test-Set Validation 16
Figure 3.7: Combined States (KS, OK, and OH) SSD Specific Gravity Calibration and Test-Set Validation 17
Figure 3.8: Kansas Calibration Results for First D-Cracking Pass-Fail Model (90% Success Rate) 18
Figure 3.9: Kansas Calibration Results for Second D-Cracking Pass/Fail Model (90% Success Rate) 19
Figure 3.10: New York AIR Test Set Validation (1) Calibration (Diamonds) and Test Set Validation (Circles) 21
Figure 3.11: New York AIR Test Set Validation (2) Calibration (Diamonds) and Test Set Validation (Circles) 21
Figure 3.12: Ohio Sample Calibration for LA Loss 22
Figure 3.13: Ohio Cross-Validation for LA Loss 23
Figure 3.14: Calibration for LA Loss using Data from Kansas, Ohio, Oklahoma, and Pennsylvania 23

Figure 3.15: Oklahoma Calibration for Absorption (Four Outliers Removed) 25
Figure 3.16: Oklahoma Test-Set Validation for Absorption 25
Figure 3.17: Combined States (KS, OK, and Ohio) VAD Model for Absorption..... 26
Figure 3.18: Oklahoma Calibration for Micro-Deval Loss (Outliers Removed) 27
Figure 3.19: Oklahoma Cross-Validation for Micro-Deval Loss (Outliers Removed) 27
Figure 3.20: Pennsylvania Calibration Model for Micro-Deval..... 28
Figure 3.21: Pennsylvania Cross-Validation Model for Micro-Deval 28
Figure 3.22: Ohio Calibration for Chert Content..... 29
Figure 3.23: Ohio Cross-Validation for Chert Content 29
Figure 3.24: Principal Component Analysis Score Diagram for Kansas Samples..... 31
Figure 3.25: PCA Score Plot for Production Blended Samples KS-41 through KS-47 Compared to Known Samples KS-33 through KS-35 and KS-48 through KS-51..... 33
Figure 4.1: Target Orifice (Outer Circle) and Laser Shot (Red Dot Inside Circle) – Magnified Photo of ~½ Inch Diameter Aggregate..... 35
Figure 4.2: Illustration of Orientation Heterogeneity in the SLT Aggregate Flow Chute..... 37
Figure 4.3: Examples of Spectra of a Kansas Limestone at Different Levels of Intensity..... 38
Figure 4.4: Histogram of Number of Samples as a Function of Total Intensity..... 39
Figure 4.5: Calibration for SSDSG Specific Gravity using the Average of All 5,000 Shots for Each Sample 40
Figure 4.6: Calibration for Saturated Specific Gravity using the Average of all Spectra with Total Intensity Greater than the Average Total Intensity..... 41
Figure 4.7: Calibration of Saturated Specific Gravity Model using the Average of Spectra with Total Intensity Greater than 10,000,000 Counts 41

This page intentionally left blank.

Chapter 1: Introduction

1.1 Overview

It is nearly impossible to produce a mineral aggregate product for use in construction that does not have some risk of contamination. This is particularly true of crushed stone sources comprised of limestones, dolostones, and sandstones that are typically interbedded with low quality layers that cannot be wholly segregated during the extraction process. Gravel beds, created by glacial and/or water transport, are generally found in discrete deposits, but gravels of acceptable quality often occur next to gravels of low quality. Other sources of stone and gravel product contamination include soil overburden, silt/clay, or gravels of low quality underlying the gravel deposit.

As a result, prior to use, construction aggregates are subjected to a series of tests to ensure compliance with appropriate state, local, and federal requirements. Standardized test methods have been developed by agencies such as the American Society of Testing Materials (ASTM) and the American Association of State Highway and Transportation Officials (AASHTO) as a means to evaluate the quality of construction aggregate and to determine whether the materials are in compliance with appropriate specifications. Most of these tests were developed in the early and middle part of the 20th century and are the basis of quality control (QC) and quality assurance (QA) programs currently employed in the industry.

It is generally recognized by most material engineers and material geologists that existing test methods are in need of an upgrading. Many of the current aggregate QC test methods do not adequately characterize aggregate properties, are time consuming, highly inefficient, and have poor reproducibility. Inefficient, time-consuming tests and inaccurate QC test methods negatively impact the service life of our highway infrastructure. They introduce considerable risk to the hot mix asphalt and portland cement concrete production industry, where suppliers generally produce and laydown final pavement products prior to receiving quality control testing results. The ramifications of test failures after the pavement is in-place are financially and administratively burdensome to the producers, contractors, and the transportation agencies charged with ensuring that funds to construct and maintain the highway infrastructure are used efficiently. QC issues in the industry are further aggravated by the poor precision associated with

many of the sampling and acceptance tests employed by the industry and the skill of technicians conducting such tests.

With the advent and rapid development of modern spectroscopy in the mid-20th century and the development of laser technology in the latter part of the 20th century, the means to remotely scan materials in "real-time" began to emerge as a viable commercial technology for many industries. In April 2012, the Transportation Research Board (TRB) published the results of the first documented study in the United States that examined the viability of using laser scanning process to identify and characterize aggregate materials used in highway construction (Chesner & McMillan, 2012). The laser scanning process described in this report is referred to in the scientific literature as Laser-Induced Breakdown Spectroscopy (LIBS). The process makes use of a high-powered laser that couples with and excites atoms present in the microstructure of target aggregate materials. This atomic excitation process induces the release of electromagnetic radiation (spectra) that is unique to the aggregate material targeted. The study demonstrated the feasibility of using the LIBS process along with multivariate (or chemometric) modeling to "fingerprint" aggregates based on the specific wavelength pattern emitted by the laser-induced emission, and to use these fingerprints to identify the type and characteristic of unknown aggregate sources.

The successful deployment of laser scanning technology could have far reaching quality control ramifications. It could provide the means for state transportation agencies to rapidly screen samples that must be monitored to ensure aggregate quality, reducing the laboratory burden facing many agencies. It could provide the means for aggregate producers to monitor their production process at the source in real time, minimizing the potential liability issues associated with off-spec quarried aggregate entering the market place. It could provide to asphalt and concrete producers the assurance that only products specified are used in roadway construction.

Initial interest by several state agencies in this 2012 work led to the development of this Transportation Pooled Fund (TPF) study, spearheaded by the Kansas Department of Transportation (KDOT), to explore in greater detail the practical aspects of employing laser scanning as a QC technique in the highway industry.

1.2 Objectives and Scope

Transportation Pooled Fund Project TPF-5(278) was initiated in late 2013 with five participating agencies. These included Kansas, New York, Ohio, Oklahoma, and Pennsylvania. The primary objective of the effort was to determine whether laser scanning could effectively be employed to predict aggregate types and properties in a commercial environment.

The TPF work plan focused on a series of tasks:

1. Collecting aggregate samples from each of the participating states, and scanning these samples in a field laser prototype developed for this pooled fund effort;
2. Compiling a database (fingerprints) of known materials (calibration models);
3. Validating the models to determine whether the available database (calibration models) could be used to predict the type and characteristics (engineering properties) of unknown samples; and
4. Assessing the operation of the laser scanning hardware and software used in the laser scanning program. A more detailed discussion of the laser scanning prototype is presented elsewhere (Chesner, 2015).

Each participating state provided engineering property data associated with the aggregate samples supplied for scanning. These data were transmitted through an online database developed specifically for this TPF study. The state-supplied engineering property data were generated using traditional laboratory methods and included:

- Specific gravity (bulk, saturated surface dry, and apparent)
- Absorption
- Micro-Deval loss
- Los Angeles (LA) abrasion loss
- Acid insoluble residue (AIR)
- Stability during freeze/thaw (soundness and expansion)
- Percent deleterious materials

Calibration and validation testing of these properties using scanned fingerprints were the primary analytical focus of the investigation. The tasks, findings, and conclusions of this TPF study are described in this report.

Chapter 2: Methodology

2.1 Hardware and Process Development

Initial 2012 aggregate laser scanning research, described in Section 1.1, was conducted on a bench scale (table-top) laser system (Chesner & McMillan, 2012). In a bench scale system, the laser technician targets individual aggregate particles, placed on a target platform, one particle at a time. Such a system was not considered practical for a commercial aggregate scanning operation. Aggregate heterogeneity requires characterization of a large number of particles, which cannot be practically accomplished in a bench scale system in a field environment. The prototype aggregate scanner used in this TPF study is referred to as the Sample Laser Targeting (SLT) System. The SLT system components are contained in a 2 × 3 × 4-ft-high cabinet, shown in Figure 2.1. The cabinet houses a laser, supporting power supply, optical components, material flow equipment, and specialized software, and is capable of scanning thousands of particles introduced into the system in minutes.

Figure 2.1: Photograph of SLT at Callanan Industries, Inc., Materials Testing Lab in South Bethlehem, NY

Figure 2.2 is a photograph of an operator introducing a bulk aggregate sample of approximately 3 gallons into the SLT. The aggregate flows vertically down inside a material flow chute, where a pulsed laser beam is directed at aggregate particles moving past the laser. A conceptual representation of the SLT process concept is shown in Figure 2.3.

Figure 2.2: Operator Introducing Aggregate into SLT for Scanning

The Figure 2.3 schematic shows an Nd:YAG laser emitting a high-powered laser beam at a wavelength of 1,064 nanometers (nm) directed through a focusing lens and a target orifice.[1] The target orifice opens into the material flow chute, through which the aggregate sample migrates. In this arrangement, the laser can be set to fire continuously at the target aggregate as the aggregate flows vertically through the material flow chute. The high power associated with the laser generates a plasma that emits light back through the target orifice. This light, visible and non-visible, is collected and transported through an optical fiber to a spectrometer and charged coupled detector (CCD). The spectrometer resolves the light into its component wavelengths and the CCD transfers the information electronically to a computer for storage and analysis.

[1] The type of laser used in the SLT is an Nd:YAG (neodymium-doped yttrium aluminum garnet) laser; Nd:Y3Al5O12 is a crystal that is used as a lasing medium for select solid-state lasers.

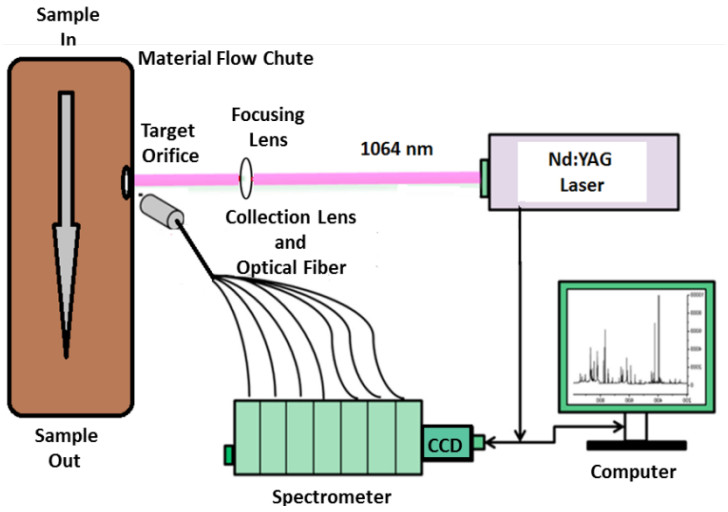


Figure 2.3: SLT LIBS Process Equipment Concept

2.2 Aggregate Spectra

The unique spectral patterns generated in this process can be equated to fingerprints that reveal information about the underlying microstructure of the target aggregate material, and can be used to assess similarities and differences between aggregates. A spectral pattern generated during scanning from an Oklahoma limestone sample is shown in Figure 2.4. The abscissa axis contains the emission wavelengths, and the ordinate axis shows the intensity of light at each corresponding wavelength recorded by the spectrometer.

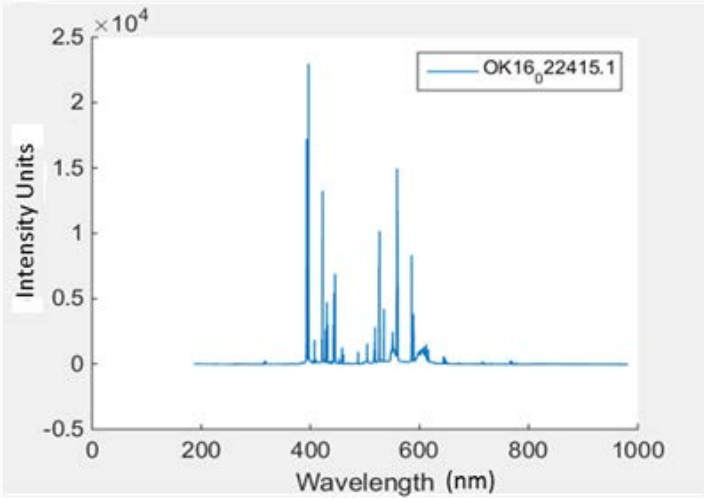


Figure 2.4: Oklahoma Limestone Spectral Pattern

The spectral pattern, shown in Figure 2.4, is comprised of intensity measurements at approximately 14,000 different wavelengths. To discriminate between and among the spectra patterns generated by the variety of aggregates tested, special mathematical modeling techniques (chemometric modeling) are employed. These techniques provide the means to discriminate between observed spectra from known materials and spectra from unknown materials and to predict the engineering properties of the unknown aggregate. Due to the heterogeneity inherent in aggregate particles, suitable characterization requires a sufficient number of laser shots to provide an accurate representation of the spectral pattern associated with the target aggregate. The spectral emission pattern shown in Figure 2.4 is an average of 1,500 laser shots.

2.3 SLT Modeling

As noted above, to analyze spectral data generated by the LIBS process, the SLT employs multivariate analysis, commonly referred to in the scientific literature as chemometric modeling. Two related chemometric modeling techniques were employed in this investigation: (1) Principal Component Analysis (PCA), and (2) Partial Least Squares Regression (PLSR). A detailed description of these techniques is beyond the scope of this report but, in general, PCA is used as a discriminatory model to sort out dissimilar aggregate types from similar types, and PLSR is used as a regression model to predict aggregate properties based on spectral emission data that has been calibrated to specific aggregate types and properties.

The main result of PCA is a PC score plot, which is a two-dimensional diagram that reduces the multi-dimensional nature of the analysis to a more familiar two-dimensional framework that can be visualized. The two dimensions of this plot represent a new coordinate system referred to as principal components (PC). The samples plot on this two-dimensional diagram can be interpreted like any two-dimensional (x vs. y) diagram. Samples that are similar in composition plot near each other and those with very different compositions will plot in a different section of the diagram.

PLSR is similar to PCA but can be used to develop models that project values (dependent variables) for unknown samples based on previously calibrated data generated from an original data set. The dependent variables can be values such as specific gravity or acid insoluble residue,

or any selected engineering property. Like any calibration, once a model is calculated, one can input into the model spectra from an uncharacterized source to predict the values of unknown variables associated with the unknown source. Dependent variables, in addition to engineering material properties, can also be indicator integers. Integers "1" and "0" are used in aggregate analysis to model data in which there are yes or no questions that must be answered (e.g., does the group contain a deleterious material or not).

The efficacy of PLSR models can be evaluated by how well a predicted value correlates with a known input value. This evaluation is typically done using graphical techniques, followed by data precision and accuracy tests. Graphical evaluation techniques are referred to as PLSR validation plots. Both PC score plots and PLSR calibration and validation plots are used to depict the results of the data analysis and are presented in the next section of this report.

The most relevant method to assess the precision and accuracy of SLT scanning measurements and models is to analyze a sample with known engineering values multiple times and calculate the standard deviation (precision) and the absolute difference between the average analyzed value and the known value (accuracy). This was not rigorously done in this TPF project, because the focus was on hardware, software, and database development as much as on sample analysis. The small number of samples available precluded detailed precision and accuracy analyses. Defining the precision and accuracy of SLT measurements is planned in a new TPF project.

Also, it is important to note that the "known" values consist of a single traditional analysis; the precision of these measurements is not known. Thus, the SLT data can be no more accurate than the traditional data used to produce the models.[2] This is an issue that will be considered further in subsequent work.

In this report, analytical deviations (plus and minus range around the predicted answer) are reported for models in which the value of an engineering property (i.e., specific gravity, LA loss, etc.) is predicted using test-set validation, described in greater detail in the next section. They can be applied in the same way as the root mean squared error of prediction (RMSEP) in

[2] Reference here is being made to the standard AASHTO, ASTM, or state transportation agency lab tests, most of which lack rigorous statistical data on both accuracy and precision.

that they provide an estimate of how well the modeled value has been predicted.[3] The deviation values (plus or minus values associated with any prediction) represent the best estimate of the range where 95% of the dependent variables can be expected to occur.

[3] One difference between the deviation values reported here and RMSEP is that the latter is calculated using the absolute errors from all of the measurements. The chemometric deviations used in this analysis do not use absolute errors because the engineering value may not be known for all samples. Instead, the deviation values are derived from a measure of how similar a spectrum is to the other spectra in the model and how well the modeled spectrum compares to the actual spectrum.

Chapter 3: Results

3.1 Test Samples and Known Engineering Properties

To evaluate the performance of the SLT prototype, test aggregate samples were provided to the Research team for scanning by the TPF participating states: Kansas, New York, Ohio, Oklahoma, and Pennsylvania. Each state also provided test data on the respective aggregate samples submitted for scanning. A listing of the number of samples and property test data provided by each respective state is presented in Table 3.1.

Table 3.1: Summary of Number of Samples and Test Data Provided by States

Available Test Data	Kansas	New York	Ohio	Oklahoma	Pennsylvania	All States
Acid Insoluble Residue	36	16	--	8	--	60
Bulk Specific Gravity	36	--	14	33	14	97
SSD Specific Gravity	36	--	14	26	14	90
Apparent Specific Gravity	24	--	--	--	14	38
Absorption	36	--	14	33	14	97
Micro-Deval	--	--	--	27	14	41
LA Loss	36	--	8	33	14	91
Production Samples	12	--	--	--	--	--
D-Cracking	35	--	--	--	--	--
% Chert	--	--	9	--	--	--
% Blast Furnace Slag	--	--	--	--	14	--

Samples provided by each state for scanning were selected from a wide range of sources. A list of the types of samples received from each state and the sample sources are summarized in Table 3.2. It is of note that the wide range of aggregate types and source locations, coupled with the relatively small number of samples that were available for scanning during this effort, are not ideal conditions for calibrating and validating aggregate chemometric models. Additional samples are needed. Nonetheless, as will be shown below, aggregate properties can be defined in most cases, suggesting that additional and more focused sample collection strategies would significantly improve the database.

Each sample received was contained within a 5-gallon bucket, containing approximately 3 to 4 gallons of aggregate material. Scanning was undertaken by introducing the 3 to 4 gallon sample through the SLT, as shown in Figure 2.2, and firing a total of 150 laser shots at the

sample. The spectra associated with these 150 shots were averaged to yield one subsample spectra. This process was repeated 10 times to yield 10 subsample spectral images, each representing the average of 150 laser shots. The 10 subsample results were further averaged to yield one super-averaged spectra to represent the one sample.

Table 3.2: Distribution of Sample Types and Sources from Five States

Source Description	Kansas[a]	New York[b]	Ohio[c]	Oklahoma[d]	Pennsylvania[e]
Limestone	x	x	x	x	
Gravel			x	x	
Sandstone				x	
Dolostone			x	x	
Blast Furnace Slag					x
Steel Slag					x
Recycled Concrete				x	

a. Kansas samples were collected from at least seven different quarries around the state.
b. New York samples were collected from three to four quarries.
c. Ohio samples were collected from at least 15 different locations.
d. Oklahoma samples were collected from at least 26 different locations, and included quarry sites in Arkansas, Texas, and Colorado, as well as Oklahoma.
e. Pennsylvania samples included blast furnace slag and steel slag blends collected from 14 different slag pile locations.

The efficacy of the modeled data was determined using one of two techniques:

1. Test set validation, and
2. Cross-validation.

When a set of data is used to calibrate a model, the most effective approach is to obtain a set of samples for use as a calibration set to calibrate the model, and then use a completely independent sample or validation set to validate the model. This method is referred to as “test set validation.”

When there are an insufficient number of samples available to calibrate the model, in chemometric modeling the “cross-validation” method is commonly used. In cross-validation, the

model is calibrated using all the data, and then in a stepwise fashion, one sample is removed from the model. A new calibration model is then generated without the left out sample. The sample left out is then used as a test sample. This process can be repeated until each sample in the model has been analyzed in this fashion. In general, models were tested using cross-validation, but wherever more than 20 samples were available, test set validation was employed.

The following subsections summarize the procedures, representative results, and findings observed during the laser scanning and modeling efforts. Heterogeneity of various types is thought to affect the quality of the results:

- Inherent mineralogical heterogeneity;
- Heterogeneity of particle orientation in the SLT material flow chute;
- Variations in lab testing results from lab to lab and state to state; and
- Changes in the laser and spectrometer system operations related to temperature, vibration, and other environmental conditions.

These factors are briefly alluded to in this section, where applicable, and discussed more fully in Section 4.3.

3.2 Measurement of Engineering Properties

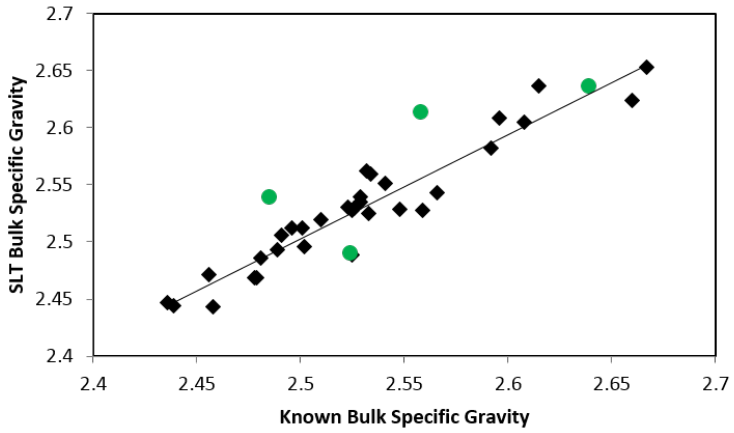
3.2.1 Bulk, Saturated Surface Dry (SSD), and Apparent Specific Gravity Results

Specific gravity models are presented for Kansas (limestones) and Pennsylvania (slag) aggregate types. A third model is also presented using aggregates from all states that submitted specific gravity data (KS, OH, OK, and PA).

Using exclusively the Kansas data, test set validation models were developed for bulk, SSD, and apparent specific gravity. A total of 36 samples were available for model development. Four KS samples (KS-09, KS-19, KS-23, and KS-30) were left out of the three calibration models (bulk, SSD, and apparent specific gravity); each of the three models was calibrated with the remaining 32 samples. The four samples were then inserted into each model to predict the bulk, SSD, and apparent specific gravity values. Test-set validation results for bulk, SSD, and apparent specific gravity are presented in Figure 3.1, Figure 3.2, and Figure 3.3, respectively. Each figure shows the calibration results (diamonds) and the test set validation results (circles).

The actual values and the respective plus/minus deviations and absolute errors between the measured and predicted are included in tabular form in each figure.

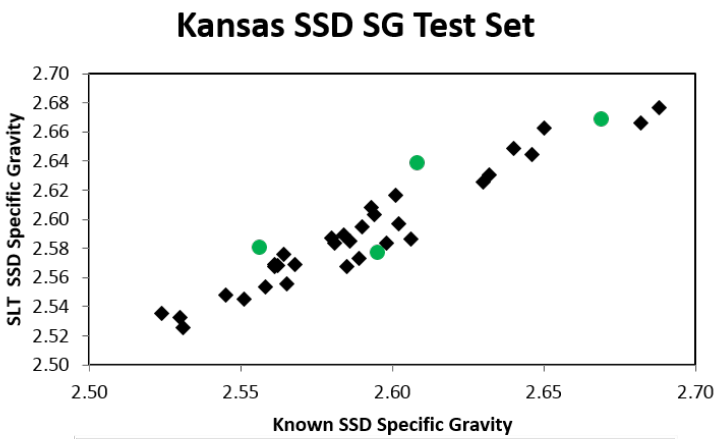
The results demonstrate that the models developed provide impressive predictive capability in a range of +/-0.05 of the expected specific gravity value.



KS Bulk Specific Gravity Test Set			
	known	predicted	absolute error
KS09	2.64	2.64 +/- 0.03	0.00
KS19	2.52	2.49 +/- 0.05	0.03
KS23	2.49	2.54 +/-0.05	-0.05
KS30	2.56	2.61 +/- 0.05	-0.06

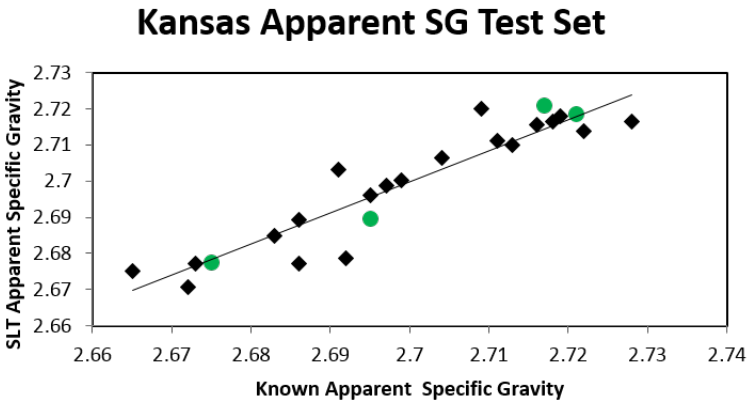
Figure 3.1: Kansas Bulk Specific Gravity Calibration (Black Diamonds) and Test-Set Validation (Green Circles)

The table gives the predicted SLT and known bulk specific gravity results.



KS SSD Specific Gravity Test Set			
	known	predicted	absolute error
KS09	2.67	2.67 +/- 0.02	0.00
KS19	2.60	2.58 +/- 0.03	0.02
KS23	2.56	2.58 +/- 0.04	-0.03
KS30	2.61	2.64 +/- 0.03	-0.03

Figure 3.2: Kansas SSD Specific Gravity Calibration (Black Diamonds) and Test-Set Validation (Green Circles)
The table gives numerical results for test-set samples.



KS Apparent Specific Gravity Test Set			
	known	predicted	absolute error
KS09	2.72	2.72 +/- 0.03	0.00
KS19	2.72	2.72 +/- 0.05	0.00
KS23	2.68	2.68 +/- 0.03	0.00
KS30	2.70	2.69 +/- 0.04	0.01

Figure 3.3: Kansas Apparent Specific Gravity Calibration (Black Diamonds) and Test-Set Validation (Green Circles)
The table gives numerical results for test-set samples.

Pennsylvania aggregate samples are unique in this study because they are all slag aggregates, including both blast furnace slag and steel slag, some originating from decades-old stockpiles and some from fresh sources. The calibration and cross-validation model results for Pennsylvania apparent specific gravity (14 samples) are presented in Figure 3.4 and Figure 3.5, respectively. Given the small sample set and the slag variability encountered, the validation results provide a clear positive trend; however, additional samples from each source of slag are required to fully model specific gravity using slag samples.

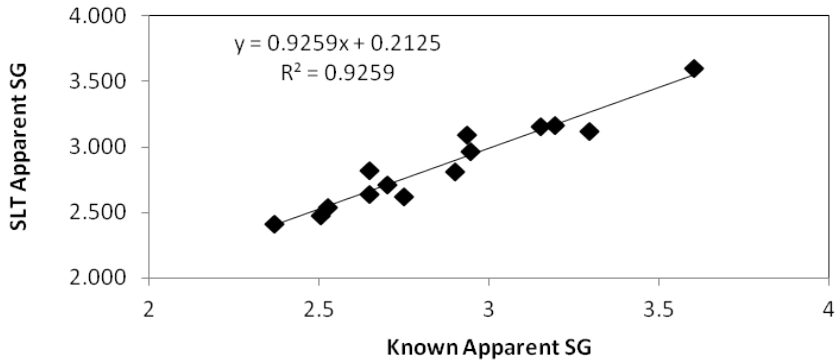


Figure 3.4: Pennsylvania Calibration Model for Apparent Specific Gravity

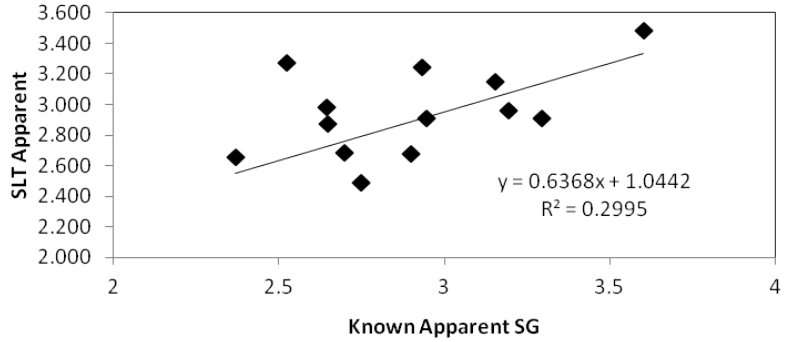
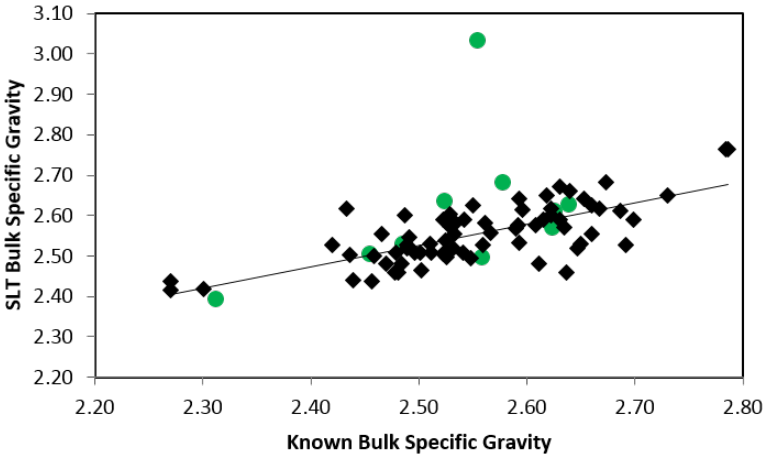


Figure 3.5: Pennsylvania Cross-Validation Model for Apparent Specific Gravity

In the 113 aggregates used in this study, the most common engineering properties available were bulk and SSD specific gravity data. All states, with the exception of New York, reported bulk specific gravity and SSD specific gravity results on all samples. The bulk specific gravity model calibration and test set validation results for 10 test set samples selected from the

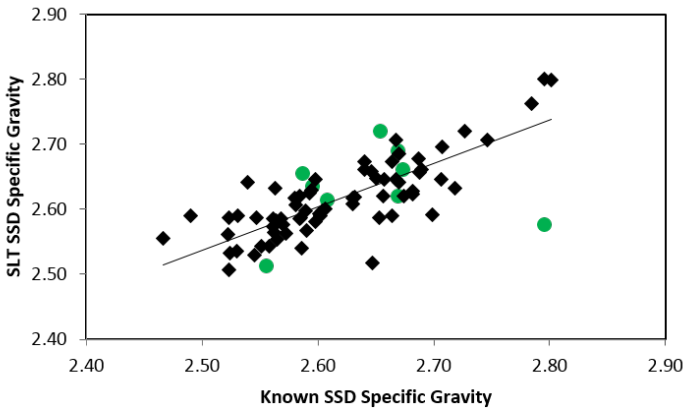
combined KS, OK, and Ohio data were calibrated on 73 samples and are presented in Figure 3.6. Combining data in such a manner introduces all of the heterogeneity issues outlined in Section 3.1, particularly mineral heterogeneity and lab-to-lab heterogeneity associated with lab results from multiple states. Nonetheless, results of the test samples presented in the table below show most test samples to be within + 0.07 of expected values, with one bad outlier (OK-18).



All States Bulk SG			
	known	predicted	absolute error
KS9	2.64	2.63 +/- 0.07	-0.01
KS19	2.52	2.64 +/- 0.05	0.11
KS23	2.49	2.53 +/- 0.08	0.04
KS30	2.56	2.50 +/- 0.08	-0.06
OH1	2.58	2.68 +/- 0.10	0.10
OH8	2.62	2.57 +/- 0.06	-0.05
OK12	2.63	2.61 +/- 0.12	-0.01
OK18	2.55	3.03 +/- 0.62	0.48
OK23	2.31	2.39 +/- 0.07	0.08
OK32	2.45	2.51 +/- 0.13	0.05

Figure 3.6: Combined States (KS, OK, and OH) Bulk Specific Gravity Calibration and Test-Set Validation

The models for SSD specific gravity using the same KS, OH, and OK samples are shown in Figure 3.7. Test-set results exhibited absolute SSD specific gravity values within 0.09 of the expected values, with the exception of the OK-18 outlier.



All States SSD SG			
	known	predicted	absolute error
KS9	2.67	2.69 +/- 0.06	0.02
KS19	2.60	2.64 +/- 0.05	-0.02
KS23	2.56	2.51 +/- 0.07	-0.13
KS30	2.61	2.61 +/- 0.07	0.01
OH1	2.67	2.66 +/- 0.07	0.07
OH8	2.67	2.62 +/- 0.05	0.04
OK12	2.65	2.72 +/- 0.09	0.09
OK18	2.59	2.65 +/- 0.75	0.09
OK23		2.55 +/- 0.08	
OK32	2.80	2.58 +/- 0.16	-0.02

Figure 3.7: Combined States (KS, OK, and OH) SSD Specific Gravity Calibration and Test-Set Validation

3.2.2 D-Cracking Results

A total of 35 samples were used in the Kansas D-cracking studies. The goal in this study was to separate samples into two groups: those that, according to Kansas test data, would pass the D-cracking test and those that would fail. PLS calibration models were generated, using the integer regression variable "1" for passing samples and the integer regression variable "0" for failing samples. Once a model was calibrated by setting passing and failing samples to these values, the predicted variable was calculated for a selected set of test set samples. Samples with calculated variables closer to 1 were assigned to the passing group; those with calculated variables closer to 0 were assigned to the failing group. A value called the VAD (Value of Apparent Distinction) was used to separate the two groups. The VAD is chosen during calibration to give the best fit to the model.

The D-cracking pass/fail model was calculated (calibrated) twice, with the omission of two different sets of test-set samples. Figure 3.8 shows the results of the calibration for the first model in which passing samples KS-22 and KS-41 and failing samples KS-9 and KS-43 were reserved for test set validation; 31 remaining samples were used in the calibration model. Using a VAD of 0.4, the calibration was 90% successful; only three samples were misclassified. The exercise was repeated using four different samples for test-set validation (passing samples KS-29 and KS-11, failing samples KS-8 and KS-37). The results of the calibration are shown in Figure 3.9. This time the model was also 90% successful, with a VAD of 0.4, and only three samples misidentified.

The test-set samples KS-9, KS-22, KS-41, and KS-43 were run through the first model and the predicted variable values were determined. The results, listed in Table 3.3, show that the model correctly predicted whether all four samples would pass or fail. For the test set samples KS-29, KS-11, KS-8, and KS-37, three of the four test-set samples were correctly identified. Sample KS-37 had a calculated value of 0.81, above the VAD of 0.4, and so would be assigned to the group of passing samples, although it is a failing sample. Overall, seven of eight samples were correctly identified, for a success rate of 87.5%.

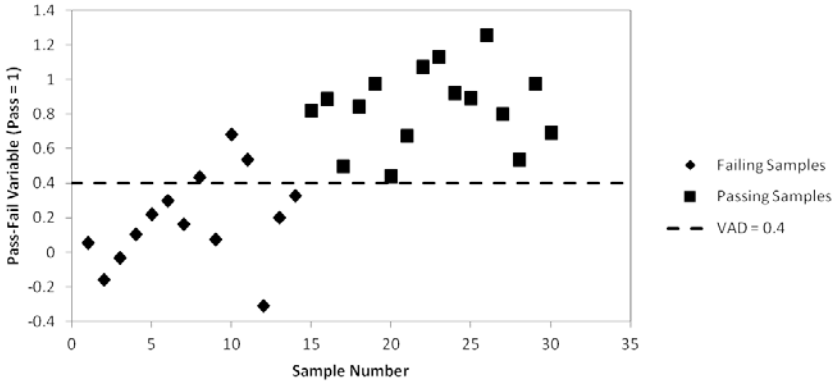


Figure 3.8: Kansas Calibration Results for First D-Cracking Pass-Fail Model (90% Success Rate)

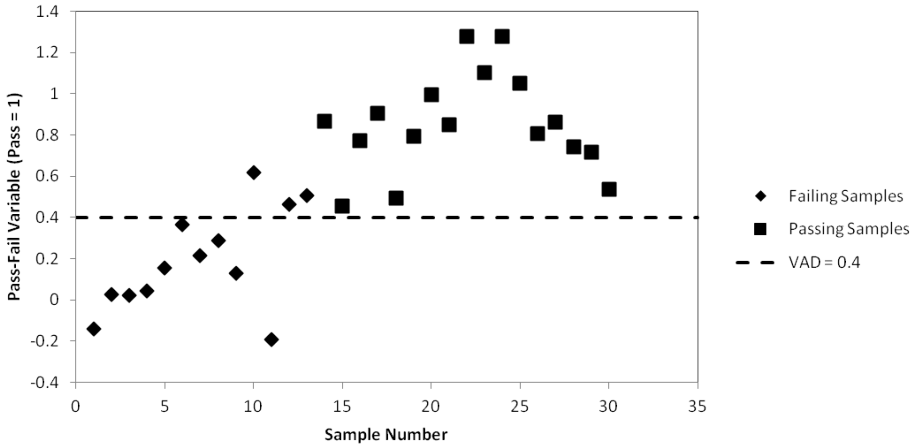


Figure 3.9: Kansas Calibration Results for Second D-Cracking Pass/Fail Model (90% Success Rate)

Table 3.3: Results of D-Cracking Pass-Fail Models with KDOT Samples

Sample	Known to Pass or Fail	Calculated Variable	Correct?
KS-9	Fail	0.21	Yes
KS-43	Fail	0.24	Yes
KS-22	Pass	0.97	Yes
KS-41	Pass	0.45	Yes
KS-8	Fail	-0.36	Yes
KS-37	Fail	0.81	No
KS-29	Pass	0.49	Yes
KS-11	Pass	0.65	Yes

3.2.3 Acid Insoluble Residue (AIR) Results

Kansas and New York submitted known AIR values with their aggregates. Kansas aggregates have low AIR, mostly less than 10%. When calibration samples do not have a good distribution of values, in this case AIR values, modeling over a wide range of values, particularly outside the calibration value range, is extremely difficult. New York provided two samples: New Scotland, with an average AIR value of 54.9%, and Becraft, with an average AIR value of 11.9%. Using prepared blends of New Scotland and Becraft, it was possible to prepare 11 blended samples with expected AIR values ranging from 11.9% to 54.9%. An additional five aggregate sources were supplied by NYSDOT. A list of the aggregates prepared and scanned for AIR analysis is presented in Table 3.4.

Table 3.4: Measured AIR Values for NYSDOT Samples

Sample Name	AIR %	% Becraft / % New Scott
NY-B0	54.9	100/0
NY-B10	52.23	90/10
NY-B20	44.40	80/20
NY-B30	39.40	70/30
NY-B40	31.60	60/40
NY-B50	34.10	50/50
NY-B60	29.40	40/60
NY-B70	24.63	30/70
NY-B80	19.60	20/80
NY-B90	14.00	10/90
NY-B100	11.90	0/100
1.2RDriedScreening	32.70	
100 Becraft	11.90	
100 New Scotland	54.90	
E Kingston	53.60	
LabBlend	29.10	

Test set modeling was used to determine the accuracy of predicting AIR using SLT data. Two test set models were created. The first model removed samples 1.2RDriedScreening and 100 Becraft from the calibration and then used the model to predict AIR in these two samples. Figure 3.10 shows the calibration results (diamonds) and the test set validation results (circles).

The calibration model is very good with an r^2 value equal to 0.989. Test set validation samples modeled relatively close to the known values. The 1.2RDriedScreening (known AIR = 32.7%) was modeled at 39.6 +/- 14.2%, and 100 Becraft (known AIR = 11.9%) was modeled at 15.5 +/- 12%.

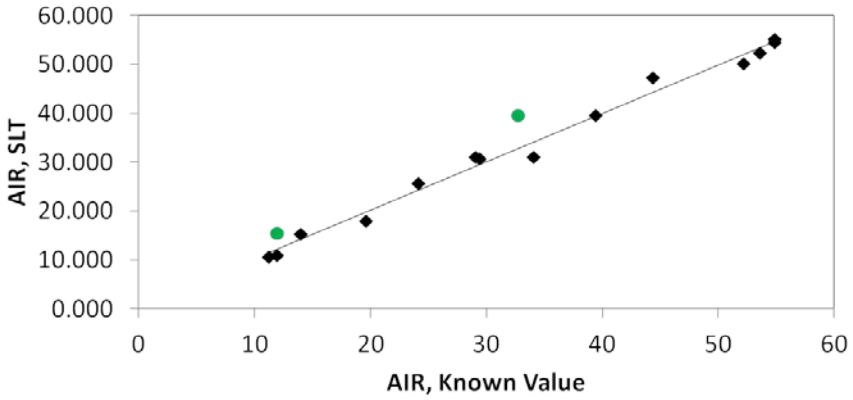


Figure 3.10: New York AIR Test Set Validation (1) Calibration (Diamonds) and Test Set Validation (Circles)

The second test set model was created using samples East Kingston and LabBlend as the test set validation samples. Results are shown in Figure 3.11. The second calibration was also excellent with an r^2 value of 0.966. East Kingston (known AIR = 53.6%) was modeled at 36.2 +/- 6.6%, and LabBlend (known AIR = 29.1%) was modeled at 21.5 +/- 5.2%.

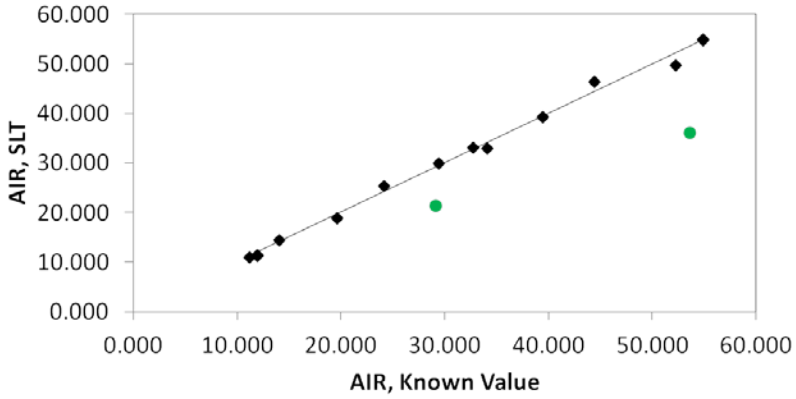


Figure 3.11: New York AIR Test Set Validation (2) Calibration (Diamonds) and Test Set Validation (Circles)

While these results indicate that the SLT data can be modeled to estimate AIR values, the results are not yet as accurate or precise as desired. One reason for this is that the majority of the calibration model is derived from samples blended from the Becraft and New Scotland samples; that is, the calibration is controlled almost entirely by the compositions of two aggregates. If a larger number of aggregates with known AIR values were used to calibrate, it is expected that

the predictions would be more accurate. In addition, only a few samples were available for calibration; the model would be more precise with more samples. Finally, it was found in a separate study that the New York AIR lab data alone exhibited a range of values +/- 8%.[4]

3.2.4 LA Loss Testing

It was originally assumed that since LA Loss is a measure of degradation by abrasion, impact, and grinding, correlating spectral data dominated by chemical information with LA Loss would not be very productive. Our attempts at modeling LA Loss with OHDOT samples suggest, however, that this is not the case. Ohio submitted only eight samples with LA Loss measurements. This was deemed too small a sample set to generate a model with confidence. The calibration of these data is presented in Figure 3.12. The calibration, however, was very good, with $r^2 = 0.984$ and a slope near 1. Validation of the calibration was performed with cross-validation. The results, shown in Figure 3.13, exhibited moderately good correlation, suggesting that this model has promise if a sufficient number of samples are used.

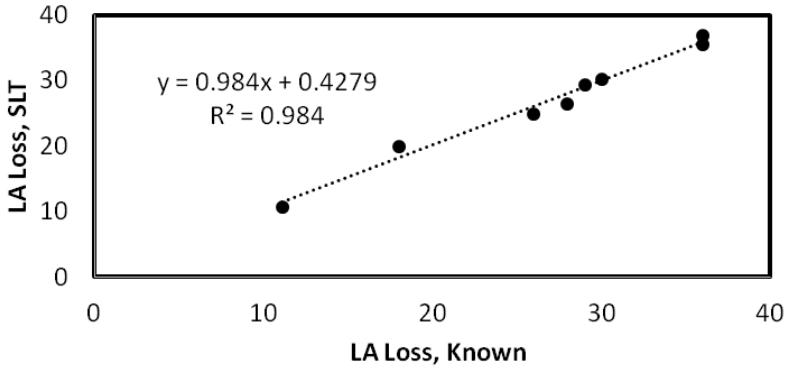


Figure 3.12: Ohio Sample Calibration for LA Loss

[4] Laboratory bias (also referred to herein as lab heterogeneity) is discussed more fully in Section 4.3.

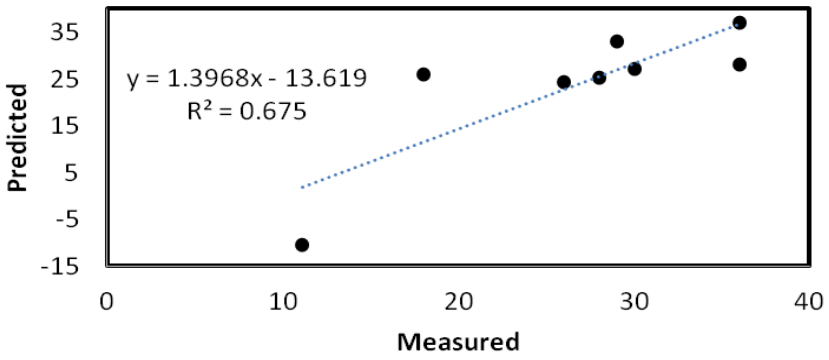


Figure 3.13: Ohio Cross-Validation for LA Loss

An LA Loss model was also calibrated with a total of 82 multistate samples collected from Kansas, Ohio, Oklahoma, and Pennsylvania. Test set validation was performed with nine samples. The calibration is shown in Figure 3.14. Results for the test set validation are presented in Table 3.5.

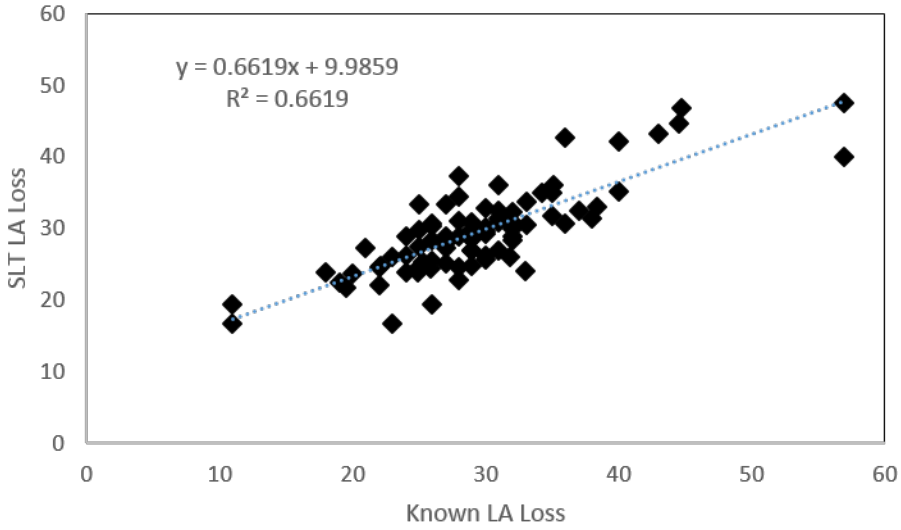


Figure 3.14: Calibration for LA Loss using Data from Kansas, Ohio, Oklahoma, and Pennsylvania

In general, LA Loss predictions tracked the known values. However, the analytical deviation is fairly high (+/- 6 to +/- 24), and the predicted value for OK-18 was very low.[5] As noted previously, mineralogical heterogeneity and state-to-state laboratory bias are thought to have a significant influence on the results.[6]

Table 3.5: Test Set Validation for LA Loss (%) using Nine Samples from KS, OH, OK, and PA

Sample Nos.	Known	Predicted	Absolute Error
KS-09	24	23 +/- 8	-1
KS-19	29	24 +/- 6	-5
KS-23	36	36 +/- 9	0
KS-30	32	37+/- 8	5
OH-01	36	28 +/- 10	-8
OK-12	19	24 +/- 10	5
OK-18	31	8 +/- 24	-23
OK-23	36	43 +/- 9	7
OK-32	41	32 +/- 16	-9

3.2.5 Absorption Testing

Results for the measurement of absorption are presented for Oklahoma aggregates (33 samples) and for the combined aggregates from all states (97 samples). Of the 33 Oklahoma samples analyzed, four samples were found to be significant outliers and were removed from the model.[7] The calibration for the Oklahoma absorption model is presented in Figure 3.15. There is significant scatter, as might be expected for a property like absorption, but the calibration was trending in a positive direction. Results for four test-set samples are plotted in Figure 3.16. Again, the results are not yet ideal, but given the wide range of lithologic and geographic variability, the results are encouraging and suggest that improvements in sample selection would yield improved results.

[5] Note OK-18 has been identified previously (see Section 3.2.1) as an outlier.
[6] See Section 4.3.
[7] Outliers are identified when the selected sample(s) are noticeably unlike the remaining sample data in the calibration model.

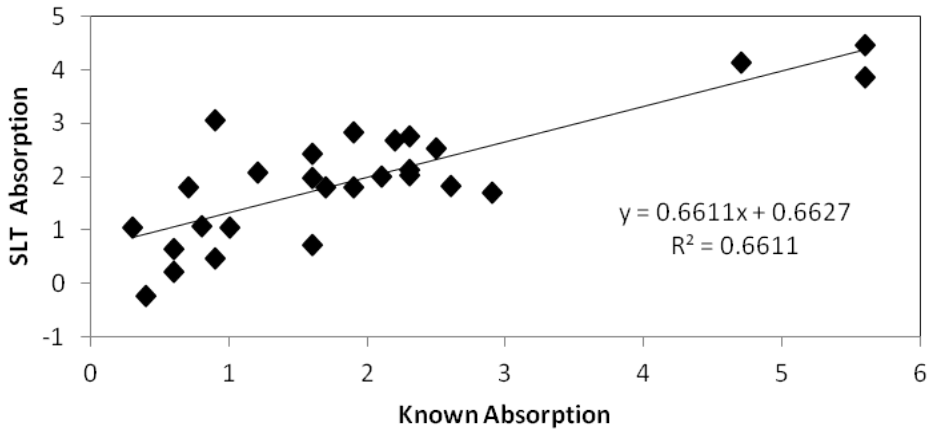


Figure 3.15: Oklahoma Calibration for Absorption (Four Outliers Removed)

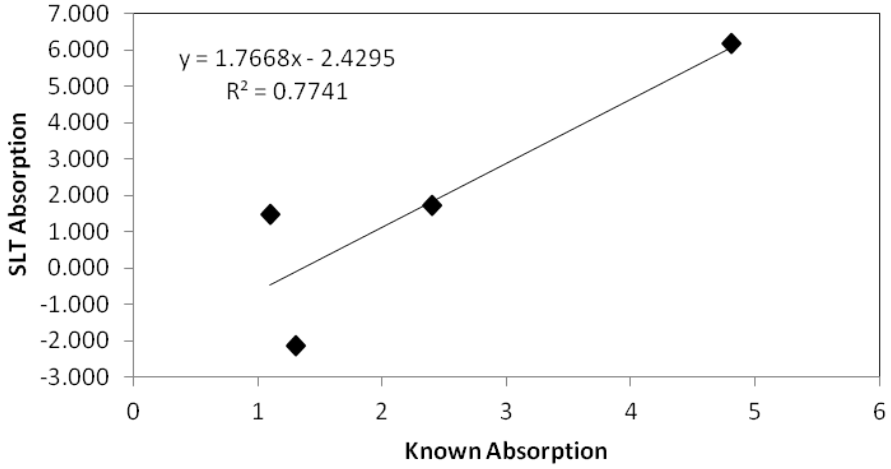


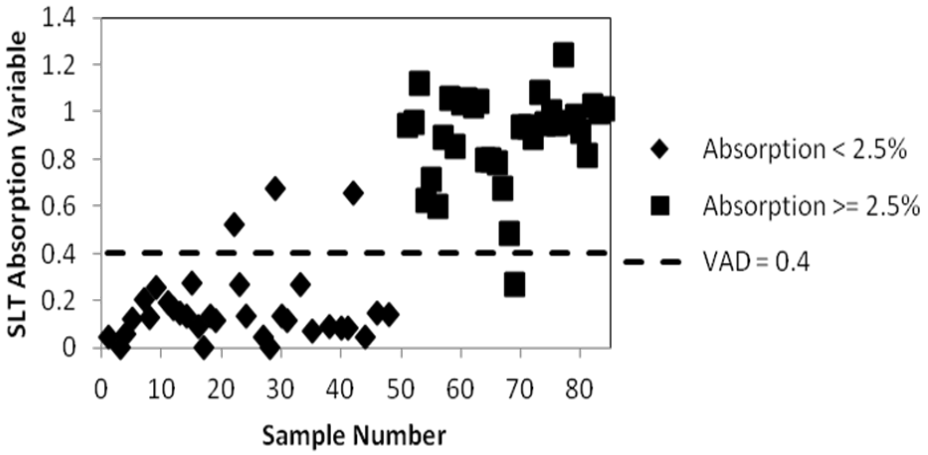
Figure 3.16: Oklahoma Test-Set Validation for Absorption

Ninety-seven samples from Kansas, Ohio, Oklahoma, and Pennsylvania were modeled using a different method. Instead of modeling the actual absorption value for an aggregate, samples were separated into two groups: greater than or equal to 2.5% absorption and less than 2.5% absorption.

If an aggregate exhibited an absorption value A > 2.5%, it was assigned the absorption variable 1; if A < 2.5%, the absorption variable was assigned a value of 0. The combined modeled results for absorption are shown in Figure 3.17. Here, the absorption variables calculated from SLT data for the calibration samples are plotted against sample number.[8] Four samples are misidentified, using a VAD of 0.4, yielding an overall success rate of 95%. The

[8] Sample number is an arbitrary designation used as a simple identifier for each sample in the scan.

results of test-set validation for 10 samples from Kansas, Ohio, and Oklahoma are also presented in tabular form in Figure 3.17. Nine of the 10 samples are correctly classified using the VAD = 0.4 chosen during calibration. The model was 90% successful in predicting whether an aggregate has greater than 2.5% absorption or less than that threshold value.



	known	predicted	Correct?
KS9	< 2.5	0.093	yes
KS30	< 2.5	0.215	yes
OH1	< 2.5	-0.145	yes
OH8	< 2.5	0.069	yes
OK12	< 2.5	0.131	yes
OK18	< 2.5	-5.706	yes
OK32	< 2.5	0.375	yes
KS19	≥ 2.5	0.406	yes
KS23	≥ 2.5	-0.314	no
OK23	≥ 2.5	1.322	yes

Figure 3.17: Combined States (KS, OK, and Ohio) VAD Model for Absorption

3.2.6 Micro-Deval Loss Results

Micro-Deval Loss was modeled for Oklahoma (27 aggregate samples) and Pennsylvania (14 slag samples) aggregates. The calibration for the Oklahoma Micro-Deval model is presented in Figure 3.18; the cross-validation results are presented in Figure 3.19. Due to the heterogeneity of the Oklahoma sample set, only limestone scans were included in the model calibration and validation process.

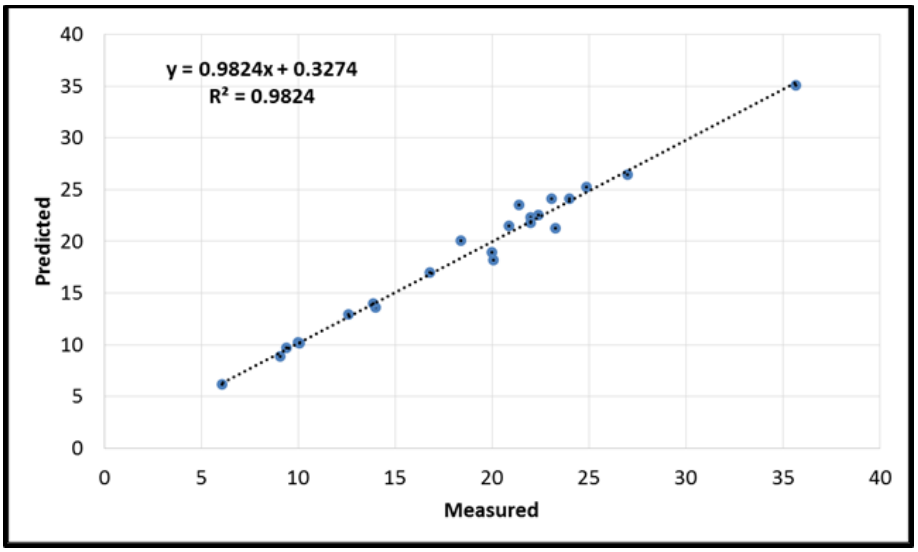


Figure 3.18: Oklahoma Calibration for Micro-Deval Loss (Outliers Removed)

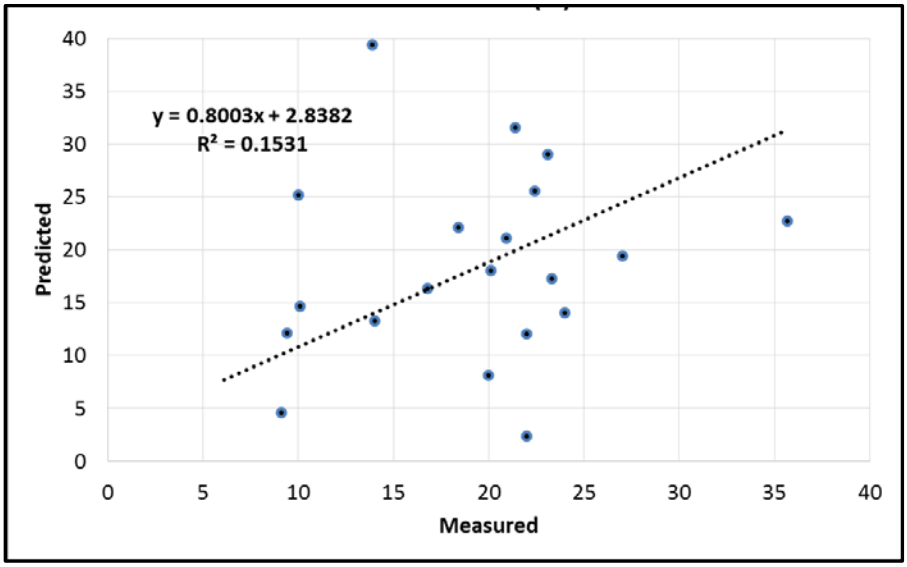


Figure 3.19: Oklahoma Cross-Validation for Micro-Deval Loss (Outliers Removed)

The calibration model developed exhibited very high correlation. As previously noted, multivariate chemometric models have the capability to generate highly correlated calibration models, but this does not necessarily translate into effective predictive models. If the spectral information used to develop the calibration model does not contain sufficient information to define the spectra associated with the samples being tested, then the validation results will not be good. The cross-validation results presented in Figure 3.19 are poorly correlated. It is noteworthy, however, that the cross-validation regression trend line tracks the calibration model

trend quite well. This type of graphical result suggests that the model on average sees the correct trend, but additional calibration data (samples) are needed to provide the information necessary to define the wide heterogeneity associated with the Oklahoma sample set.

The calibration for the Pennsylvania Micro-Deval model is presented as Figure 3.20; the cross-validation results are presented in Figure 3.21. The calibration is strong, with highly correlated results. The cross-validation results are less well correlated with known composition, but are clearly tending in the correct direction.

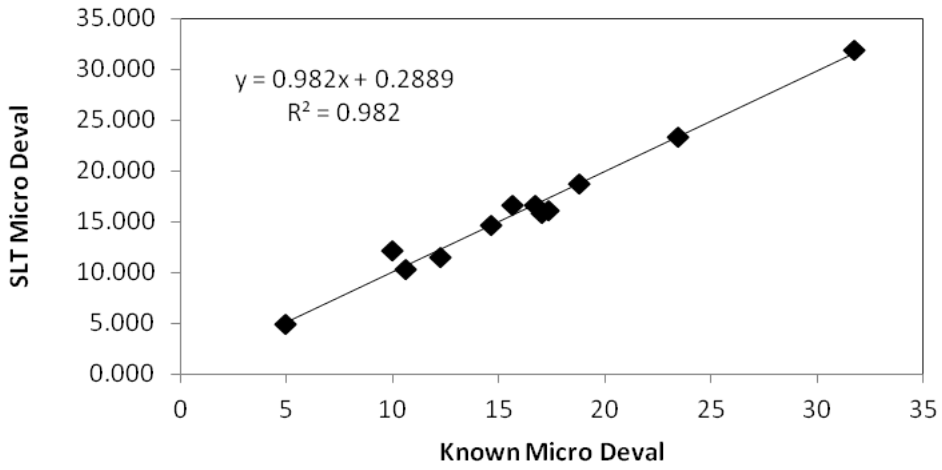


Figure 3.20: Pennsylvania Calibration Model for Micro-Deval

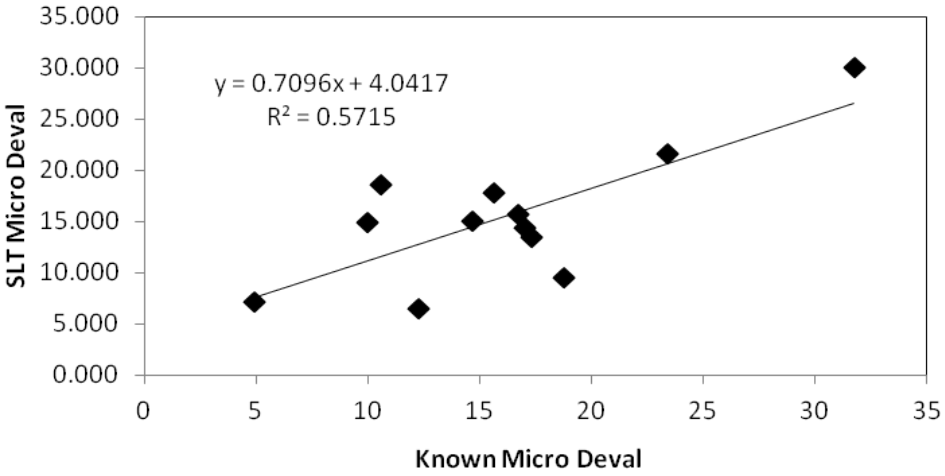


Figure 3.21: Pennsylvania Cross-Validation Model for Micro-Deval

3.2.7 Presence and Quantification of Deleterious Materials: Chert

Nine samples received from the Ohio Department of Transportation were used to develop a chert content model. The calibration model for chert content is shown in Figure 3.22. The calibration is fairly good, with high r^2 and slope near 1, although the slope is strongly controlled by the one sample with more than 10% chert. Once again, cross-validation results, shown in Figure 3.23, do not correlate nearly as well as the calibration model; however, the results trend in the right direction, suggesting that the data should be improved with a larger sample set.

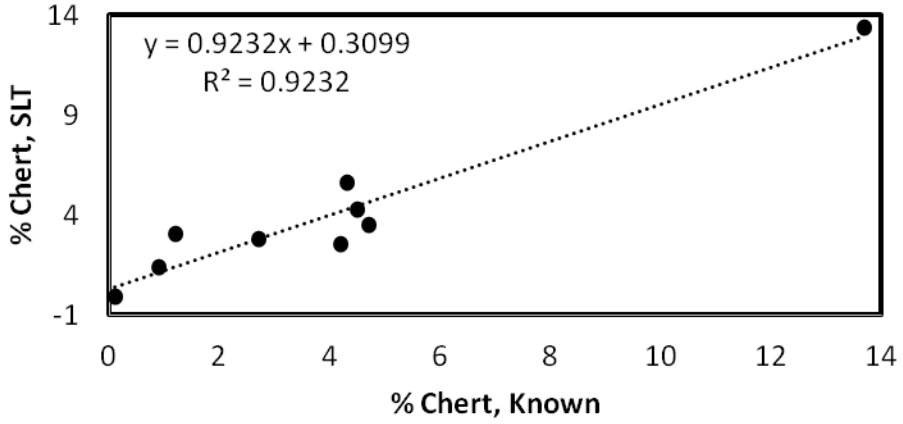


Figure 3.22: Ohio Calibration for Chert Content

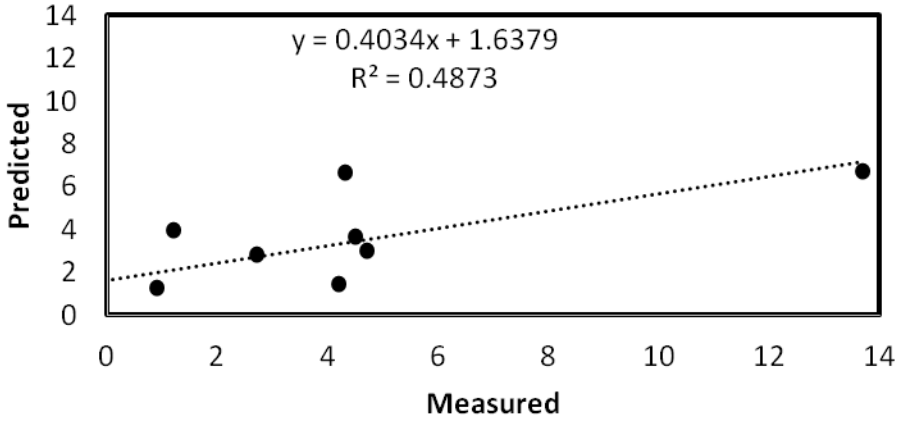


Figure 3.23: Ohio Cross-Validation for Chert Content

3.2.8 Production Samples Analysis

During the TPF scanning and modeling effort, KDOT submitted several samples of a Limestone Member from a Kansas location to the Research team to determine whether laser scanning could provide the answers to the following questions.

Question 1 deals with an issue of possible lithologic changes in a limestone ledge where KDOT was interested in determining whether the aggregate quality in the beds making up the limestone ledge were changing in composition.

The conditions and question are as follows:

- Originally, during the quarrying process, three beds within the ledge were sampled as KS-33, KS-34, and KS-35, respectively.
- As quarrying proceeded, samples KS-48, KS-49, KS-50, and KS-51 were quarried from the same ledge as KS-33, KS-34, and KS-35. The formation here appears to be separated into four beds instead of three.
- The question is: are KS-48, KS-49, KS-50, and KS-51 the same as KS-33, KS-34, and KS-35?

Principal Component Analysis (PCA) was used to answer this question. A Principal Component Analysis (PCA) score plot for all Kansas samples is presented in Figure 3.24. Recall from Section 2.3 that PCA score diagrams plot the scanned samples on a two-dimensional (x vs. y type) diagram, and show the compositional relationship between samples. Score plots are interpreted in the same manner as two-variable X-Y plots. Samples that are similar in composition plot near each other and those with very different compositions will plot in a different section of the diagram. All Kansas samples are plotted in Figure 3.24 in order to see the compositions of the samples in question in the context of all Kansas limestones.

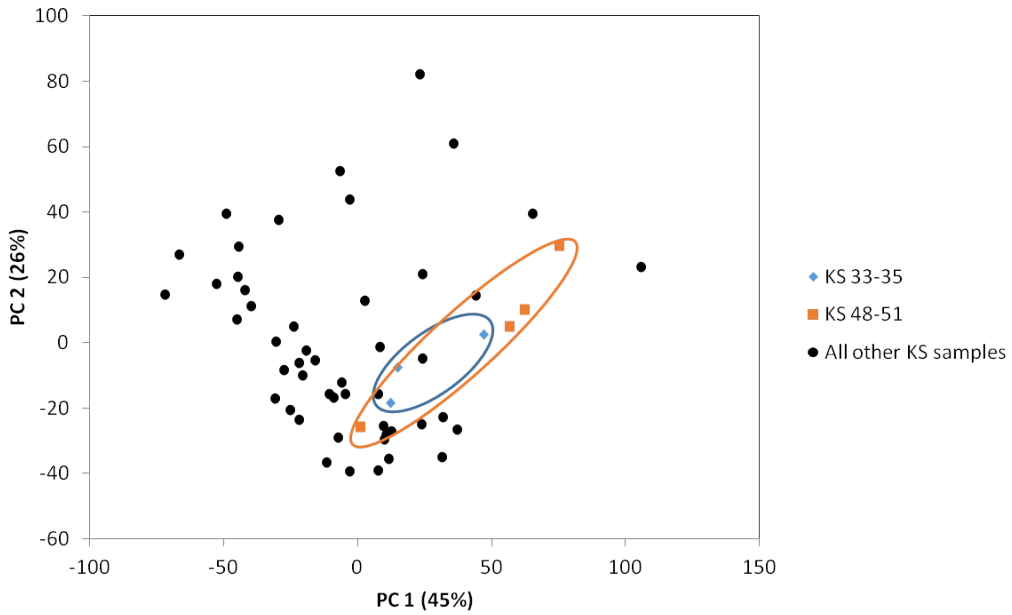


Figure 3.24: Principal Component Analysis Score Diagram for Kansas Samples
Note that limestone samples KS-33 through KS-35 are similar to KS-48 through KS-51.

The three original limestone samples (KS-33, KS-34, and KS-35) are plotted as diamonds and fall in a relatively small compositional area (shown by the small ellipse) compared to the compositions of all Kansas limestones. The samples in question, KS-48, KS-49, KS-50, and KS-51, represented by the orange squares, fall in a larger compositional range (shown by the larger ellipse).

These results indicate that KS-48, KS-49, KS-50, and KS-51 are not identical in composition to the original samples (KS-33, KS-34, and KS-35), but do occupy a similar compositional space. It is difficult to determine unequivocally whether or not the new samples are from the same beds as the original samples; however, it seems likely, based on these data, that the new samples do differ somewhat from the original by changes in composition related to natural facies changes. More samples need to be studied in order to evaluate the natural variability in limestone beds.

Question 2 deals with stockpile verification. KDOT is interested in knowing whether seven aggregate samples collected from blended production piles are made up of aggregate from approved sources.

The conditions and question are as follows:

- KDOT sent in aggregate samples for laser “fingerprinting” from the original three beds represented by KS-33, KS-34, and KS-35 discussed above in Question 1.
- Only samples that pass KDOT D-cracking criteria are permissible in production samples.
- Sample KS-33 failed the KDOT D-cracking criteria test; samples KS-34 and KS-35 passed.
- KDOT provided seven production samples numbered KS-41, KS-42, KS-43, KS-44, KS-45, KS-46, and KS-47 for scanning.
- The question is: Are the production samples (KS-41, KS-42, KS-43, KS-44, KS-45, KS-46, and KS-47) blends of samples from KS-34 and KS-35, or blends of other samples?

Once again, Principal Component Analysis (PCA) was used to answer this question. If the production samples are blends of KS-34 and KS-35, they should fall on a blending line between KS-34 and KS-35, illustrated by the small blue arrow in Figure 3.25. Instead of plotting between KS-34 and KS-35, the production samples KS-41 through KS-47 consistently fall to the left of KS-34 and KS-35. Furthermore, samples KS-48 through KS-51, which are possible correlatives of KS-33 through KS-35, plot to the right of KS-34 and KS-35. It is reasonable to interpret these data to mean that production samples KS-41 through KS-47 are not blends of KS-34 and KS-35. They could be blends of KS-34 or KS-35 and some other unknown aggregate or aggregates with composition(s) in the upper left of the score plot.

These two studies demonstrate that SLT laser scanning data coupled with multivariate analysis can be used for quarry quality and real-time production sample analyses.

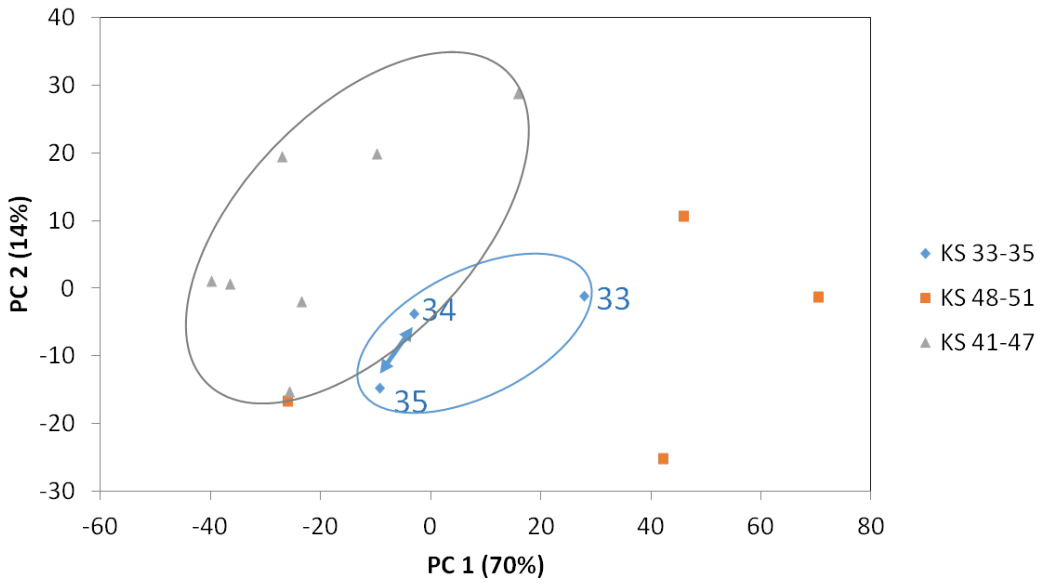


Figure 3.25: PCA Score Plot for Production Blended Samples KS-41 through KS-47 Compared to Known Samples KS-33 through KS-35 and KS-48 through KS-51

Chapter 4: Discussion of Results

4.1 Overview

The primary objective of the TPF-5(278) effort was to determine whether laser scanning could effectively be employed to predict aggregate types and properties in a commercial environment. While the results presented in Chapter 3 showed that the calibrated laser scanning models developed are capable of providing the means to predict aggregate quality, model validation, while impressive, lacked consistent accuracy and precision. Chapter 4 provides an analysis and discussion of the factors identified by the Research team that are believed to be limiting model resolution as well as recommendations designed to mitigate each factor.

Limiting factors identified to date include:

- Mineralogical Heterogeneity
- Particle Orientation Heterogeneity
- Lab Data (Bias) Heterogeneity
- Laser Scanning and Signal Reception Heterogeneity

4.2 Mineralogical Heterogeneity

It is the nature of construction aggregates to exhibit both intra- and inter-particle "mineralogical" heterogeneity. Intra-particle heterogeneity refers to the change in mineralogy along the surface of any one particle. Inter-particle heterogeneity refers to the change in mineralogy between different particles. This intra- and inter-particle mineralogical heterogeneity is illustrated in the limestone sample photograph shown in Figure 4.1.

The "red dot" shown inside the black circle superimposed on the aggregate particle on the right side of Figure 4.1 illustrates the approximate diameter of the laser beam that impacts the aggregate particle during the scanning process. Many laser shots must be directed at the aggregate sample to generate a representative fingerprint of a sample. Due to the inherent mineral heterogeneity of aggregate, it should be apparent that the greater the number of samples scanned and the greater the number of laser shots directed at the aggregate in each sample, the more robust the calibration model. For example, taking 100 laser shots at a 3-gallon sample of aggregate would provide much more information about the nature of the sample than taking only

one or 10 laser shots. The models generated in this study were based on 1,500 laser shots per sample.[9]

Figure 4.1: Target Orifice (Outer Circle) and Laser Shot (Red Dot Inside Circle) – Magnified Photo of ~½ Inch Diameter Aggregate

4.3 Orientation Heterogeneity

In addition to mineralogical heterogeneity, the SLT scanning system introduces into the analysis particle "orientation" heterogeneity. Particle orientation heterogeneity occurs because in the SLT (see Figure 2.3) a column of aggregate is moving past the target orifice through which the laser beam fires. The laser fires at a preset repetition rate as the aggregate passes by.[10] The exact location or point at which the laser couples with an aggregate particle, as well as the orientation of the incoming laser beam and the outgoing spectral emission, are unknown. It is dependent on how the particles orient themselves in the material flow chute column relative to the target orifice. It is random.[11]

The concept of orientation heterogeneity and how it impacts the incoming and outgoing light is shown in Figure 4.2. The most direct laser shot on an aggregate particle is one that hits the particle on a surface that is normal to the laser beam. When this occurs, the induced light

[9] For each aggregate sample scanned, the sample bucket was cycled through the SLT 10 times with 150 laser shots fired each cycle. This generated 10 spectra composed of 150 shots each. The 10 spectra were subsequently averaged to generate one average spectra (fingerprint) for each sample.
[10] The repetition rate during the scanning conducted was initially set at 3 Hz (3 shots per second).
[11] In a bench operation, prior to firing a laser shot, an aggregate particle is placed on a platform, and the laser is focused on a specific location (i.e., point) on the particle. Such a point is shown in Figure 4.1. Orientation heterogeneity is much less of an issue.

emission (spectra) travels directly back along the line of the laser beam. Panel A illustrates the condition where the particle surface is normal to the incoming laser beam. Most particles will be at an angle to the side of the chute, propped up by another particle, as shown in Panel B. In this case, the particle may produce spectral light, but much of the light will not move directly back through the target orifice to the detector. Alternatively, the orifice may be aligned with the space between two particles (Panel C), so that the particles are beyond the focal plane of the laser pulse. In this case, no or very little light will be generated or collected.

The original thinking in the design of the TPF-5(278) program was that "averaging all the data" from many laser shots would take into account the distribution of spectral data resulting from mineralogical and orientation heterogeneity. The SLT operations and software were all designed to run in this manner. The modeled results, presented in Chapter 3, show that modeling highway aggregate properties using LIBS spectra and multivariate models has great promise. This suggested that the "averaging" approach used was the correct one.

Further analysis of the data by the research team, however, found that the spectral information and the models generated in the TPF effort were in general not as accurate or precise as in the lab-based study (Chesner & McMillan, 2012). The research team began to suspect that the spectra generated in the SLT were of lower quality than that generated in the laser bench scale test units, due to the "orientation" heterogeneity introduced by the SLT.

A more detailed examination was made of the individual (non-averaged) laser shots and generated spectra to investigate this issue. The results of this investigation are presented in condensed form in Figure 4.3, which depicts the quality of spectra generated by analyzing intensity ranges of 4,959 individual laser shots. Beginning on the left hand side of the figure and moving downward, the spectra were sorted into 10 total intensity ranges.[12] The high-intensity spectra, presented on the left side of the figure, are similar to those obtained in the 2012 bench scale study.[13] In contrast, the low-intensity spectra (lower than 50% of the maximum spectra) lack many of the peaks present in the high-intensity spectra. The scale of this problem is

[12] Total intensity is the sum of the intensities of all wavelengths for a particular laser scan.
[13] In the lab study, each aggregate particle is placed on a platform by hand, the laser is focused on the target; there are no restrictions between the induced emission light and the optical fiber, which transmits the light to the spectrometer.

illustrated by the data in Figure 4.4 and Table 4.1, which show that very few SLT spectra have high intensity; low-quality, low-intensity spectra are much more common.

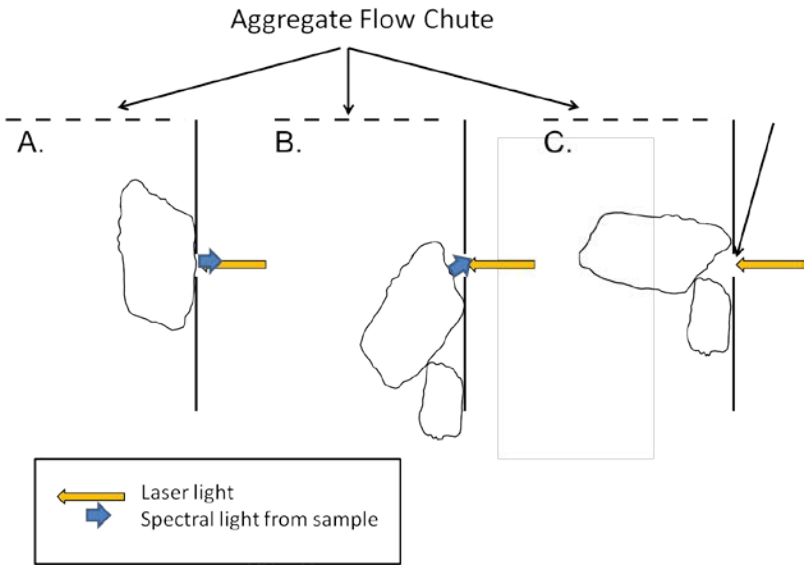


Figure 4.2: Illustration of Orientation Heterogeneity in the SLT Aggregate Flow Chute
A. The aggregate particle is properly oriented with its relatively flat side flush against the side of the chute adjacent to the laser orifice. The light from the sample (short blue arrow) is able to pass back through the orifice and reach the detector.
B. Some particles are at an angle to the laser beam because they are propped up by other particles. Only a small amount of the spectral light can pass back through the orifice and reach the detector.
C. Often, the laser beam interacts with the space between two grains, yielding little or no light.

Over half of the 4,959 spectra acquired from this sample have intensities below 20% of the most intense spectrum (Table 4.1). Since during the TPF effort every spectrum was included in the overall sample average, the high-quality spectra were being diluted by the low-quality spectra data. This can be seen by comparing the last spectra (lower right) in Figure 4.3. Note the signal degradation, and that the average of 4,959 spectra is similar to a low-quality spectrum. Orientation heterogeneity (see Figure 4.2) results in a decreasing number of good quality shots during any laser scan. Averaging the good quality shots with the poor quality shots dilutes the good data.

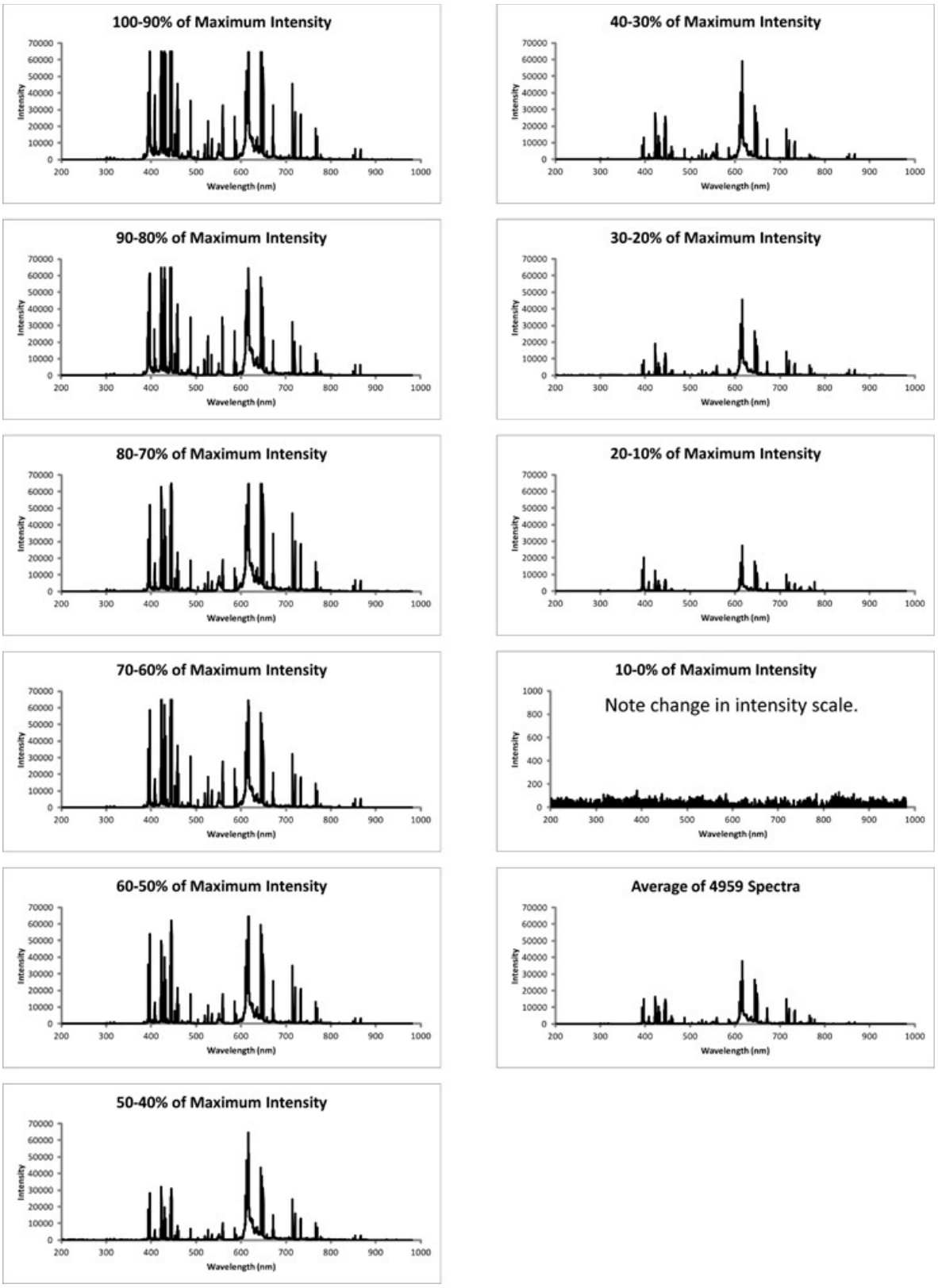


Figure 4.3: Examples of Spectra of a Kansas Limestone at Different Levels of Intensity

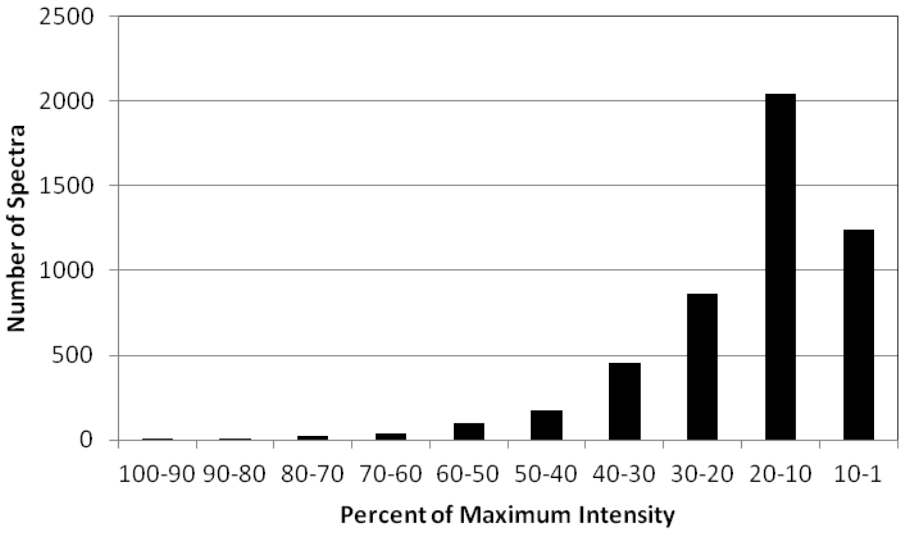


Figure 4.4: Histogram of Number of Samples as a Function of Total Intensity

Table 4.1: The Number and Percent of Spectra as a Function of Total Intensity

% of Maximum Intensity	# Spectra	%
100-90	5	0.1%
90-80	12	0.2%
80-70	23	0.5%
70-60	37	0.7%
60-50	102	2.1%
50-40	175	3.5%
40-30	458	9.2%
30-20	860	17.3%
20-10	2046	41.3%
10-1	1241	25.0%

4.3.1 Kansas Specific Gravity Results Using Filtered Data

To compare the effects of employing numerical techniques to filter out lower quality spectral data from the analysis, 56 samples from Kansas (labeled KS-1 through KS-56) were analyzed with 5,000 individual shots each. Models were calculated for three different filtering levels:

- Average of all 5,000 shots
- Average of all shots with intensity greater than the average intensity
- Average of all shots with intensity greater than 10,000,000 counts[14]

The results of the saturated specific gravity models for each condition are respectively shown in Figure 4.5, Figure 4.6, and Figure 4.7. The correlations (r^2 = 0.80) between the known specific gravity values (provided by KDOT) and the model values are similar for the average of 5,000 (Figure 4.5) and average of spectra with total intensity greater than the average total intensity (Figure 4.6). This is because there are so many low-intensity spectra that even filtering out the lowest half of them does not significantly improve data quality. However, the correlation using the average of spectra with total intensities greater than 10,000,000 (Figure 4.7) is much improved (r^2 = 0.91).

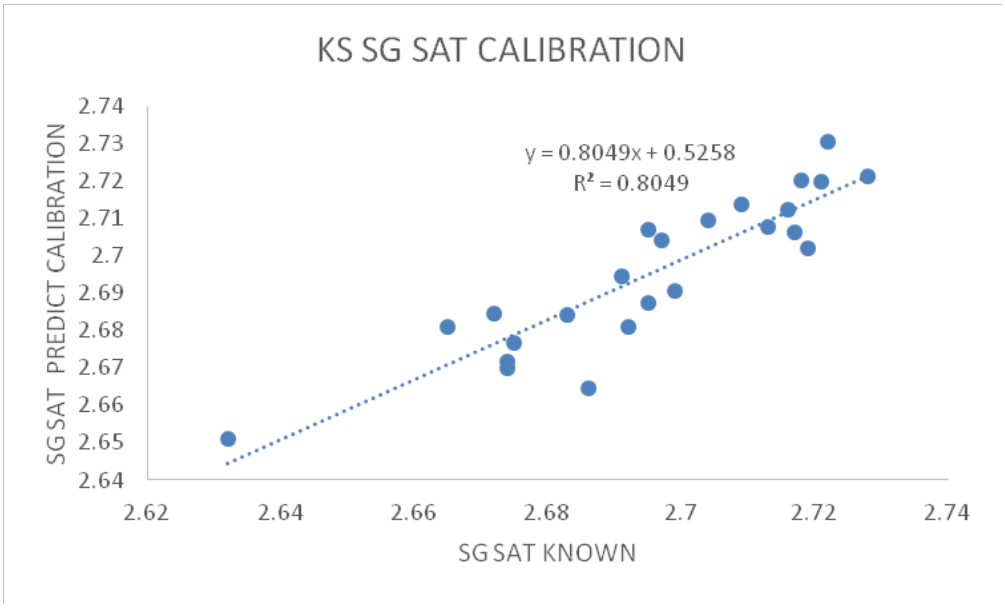


Figure 4.5: Calibration for SSDSG Specific Gravity using the Average of All 5,000 Shots for Each Sample

[14] The limit of 10,000,000 counts is arbitrary and was set at a level for which most samples have at least 100 shots. A large number of shots are necessary to capture the mineralogical heterogeneity in the sample.

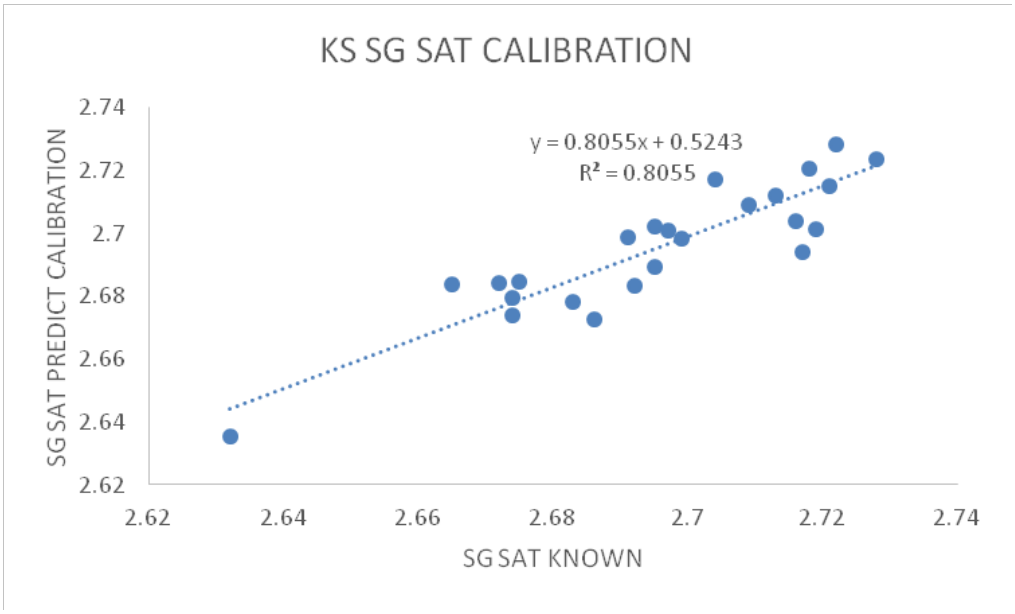


Figure 4.6: Calibration for Saturated Specific Gravity using the Average of all Spectra with Total Intensity Greater than the Average Total Intensity

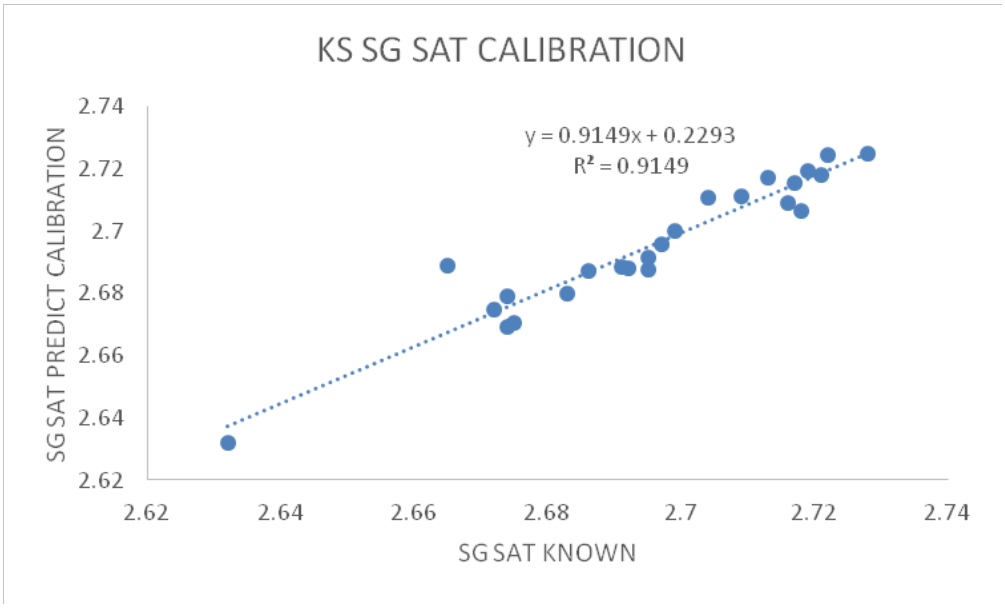


Figure 4.7: Calibration of Saturated Specific Gravity Model using the Average of Spectra with Total Intensity Greater than 10,000,000 Counts

Four Kansas samples (KS-11, KS-21, KS-31, and KS-51) were set aside as a test set to evaluate the effectiveness of the model. The SSD specific gravity values predicted by the model were found to be within 0.04 units of the known values and are listed in Table 4.2. The data

indicate that the filtering method has promise as a workable solution to the orientation heterogeneity issue.

Table 4.2: Values of Saturated Specific Gravity Calculated for Test Set Validation Samples*

Sample	Known SG Saturated	SLT SG Saturated	Analytical Error	Difference
KS-11	2.71	2.75	0.05	0.04
KS-21	2.67	2.69	0.02	0.02
KS-31	2.69	2.66	0.05	0.03
KS-51	2.65	2.61	0.05	0.04

* Laser shots with total average intensity greater than 10,000,000 counts

4.4 Lab Data Heterogeneity

Calibration models are dependent on the use of traditional lab test data as input into the model. For example, if an acid insoluble residue (AIR) calibration model is under development, the AIR value of each sample used in the model must be known. If AIR procedures differ amongst laboratories, the AIR test value of the exact same sample generated in two regions would differ. This would adversely affect the quality of the calibration model. Many of the engineering tests used in this study, such as specific gravity, AIR, and D-cracking, require lengthy tests with some subjective aspects (e.g., Has the aggregate stopped reacting with acid? Is the aggregate appropriately saturated for the various specific gravity measurements?). Results for these properties probably vary from state to state.

4.5 Laser Power and Signal Reception Heterogeneity

In a materials lab environment, laser power output and signal reception can be expected to fluctuate in response to temperature, humidity, system vibrations, and electromagnetic radiation fluctuations induced by surrounding equipment and processes. Laser power and output fluctuations can also affect the quality of the calibration model.

4.6 Addressing the Issues

In the latter part of the first pooled fund study, the research team began work on addressing these types of heterogeneity. A large number of shots are necessary to capture all of the mineralogic heterogeneity in each sample. The study used the average of 1,500 laser shots. However, the recognition of orientation heterogeneity resulted in the collection of 5,000 shots per sample. All shots with total intensity less than 107 are discarded; the high-intensity shots are averaged. The number of high-intensity shots varies from 100 to about 500; this is roughly equivalent to the 150 shots used in the preliminary lab study. This filtering method will be tested in a second pooled fund project.

The variation in known values for engineering properties for aggregates measured in different states imparts an unknown, but potentially large, effect on the chemometric models. It would be interesting to measure this laboratory bias for the engineering properties in a round-robin style study and test the idea that the ensuing SLT models will improve.

Finally, protocols are being developed to monitor lab environmental factors such as temperature, humidity, vibrations, and electromagnetic radiation from surrounding equipment. As the result of these factors on SLT data and models is understood, it will be possible to correct for them.

Chapter 5: Findings and Conclusions

Pooled fund project TPF-5(278) included an evaluation of the promise of laser scanning in a field commercial system. Calibrated models and validation tests were undertaken on aggregate samples received from five states (Kansas, New York, Ohio, Oklahoma, and Pennsylvania) to assess how well the models developed predicted the aggregate quality.

The findings are as follows:

- Multivariate (chemometric) calibration models and validation tests on these models (developed to predict values for AIR, bulk specific gravity, SSD specific gravity, apparent specific gravity, absorption, LA Loss, Micro-Deval, approved materials, D-cracking, and percent chert) yielded very promising results.
- Several types of heterogeneity will affect the quality of the spectra and laser generated fingerprints during laser scanning. These include: (1) mineralogical heterogeneity, a function of the aggregate source; (2) particle orientation heterogeneity, a function of the physical manner in which the aggregate particles orient relative to the laser beam; (3) lab data heterogeneity, a function of differences in test data generated by different laboratories; and (4) laser and reception signal heterogeneity, a function of fluctuating environmental factors.
- The relatively small number of samples tested to date and the large variability in the types of samples tested limited the accuracy of the validation test predictions during this investigation.
- Chemometric models appear to be improved by filtering the spectra and only using the high-intensity, high-quality data, and so new system and data processing improvements that can discriminate between good and poor laser shots, filter out the poor shots, and process a large number of individual (non-averaged) laser shots will improve modeling resolution.

From these findings, we conclude:

- Laser scanning in near real-time of aggregate materials is a technology capable of significantly improving the manner in which aggregate quality control procedures are employed by the industry, altering the paradigm of aggregate monitoring established in the early 20th century.
- A sufficient number of laser shots and a sufficient number of samples must be included in any calibration model to address both the mineralogical and orientation heterogeneity encountered in the test program. The exact number is dependent on the heterogeneity of the samples and can be determined by assessing how the model converges to the correct answer with increasing numbers of samples and laser shots.
- A sufficient number of high quality (filtered) laser shots are needed to mitigate the effects of laser power and signal reception heterogeneity.
- Improvements in data processing and analytical techniques that can handle individual (non-averaged) data and filter out poor laser shots from the analysis are needed to assist in resolving the spectral information generated during the laser scanning process and improving model predictions of aggregate properties.
- Further studies are needed to assess how state-to-state lab testing heterogeneity affects modeled results.

Chapter 6: Recommendations

While the TPF-5(278) findings were very impressive, and while it was concluded that the technology will change the manner in which QC testing occurs in the future, it was also concluded that several technology-related issues remain to be addressed. They include:

- Expanding the number of state agencies participating in the assessment and thereby scanning and modeling of a larger number of aggregates of a broader range of geologic types to strengthen the models and additional engineering properties;
- Modifying the SLT to permit scanning of a larger number of laser shots per sample (5000, 10,000, and 15,000 per sample) to account for mineralogical heterogeneity;
- Redesigning the data processing software to adequately process the large data sets and provide the means to automatically filter out poor laser shots from good ones to account for orientation heterogeneity.
- Scanning all new samples received from participating state agencies and retesting all the aggregate received during the initial TPF to compare the results of the old and upgraded system; and
- Development of an AASHTO Standard of Practice for QC Testing Using Laser Scanning.

The latter recommendation underscores the fact that state participation is a critical component in the development of aggregate laser scanning technology, because without state transportation agency review and acceptance, commercial implementation of the technology will stall.

It is recommended that the above-itemized recommendations be pursued by continuing the effort initiated in the TPF-5(278) effort in an ongoing Phase 2 TPF evaluation. By maintaining the existing state participation in this effort and the current management structure, the project can continue in a seamless manner and provide the means to bring additional states into the project.

References

Chesner, W.H. (2015). Prototype Development: Automated and Continuous Aggregate Sampling and Laser Targeting System (NCHRP 168). Retrieved from http://onlinepubs.trb.org/onlinepubs/IDEA/FinalReports/Highway/NCHRP168_Final_Report.pdf

Chesner, W.H., & McMillan, N.J. (2012). Automated Laser Spectrographic Pattern Matching for Aggregate Identification (NCHRP 150). Retrieved from http://onlinepubs.trb.org/Onlinepubs/IDEA/FinalReports/Highway/NCHRP150_Final_Report.pdf

AD ASTRA PER ASPERA

Kansas

Department of Transportation