2015

Optimizing the value of soil compaction testing quality assurance and control using stochastic life cycle cost, comparative and statistical analysis

Kevin Wade McLain
Iowa State University

Optimizing the value of soil compaction testing quality assurance and control using stochastic life cycle cost, comparative and statistical analysis

by

Kevin Wade McLain

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Civil Engineering (Construction Engineering and Management)

Program of Study Committee:
Douglas D. Gransberg, Major Professor
Charles Jahren
Hyung Seok "David" Jeong
David White
Iris Rivero
Daniel Bumblauskas

Iowa State University

Ames, Iowa

2015

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

AIAG	Automotive Industry Action Group
DB	Design-Build
DBB	Design-Bid-Build
DCP	Dynamic Cone Penetrometer
EDG	Electrical Density Gauge
FHWA	Federal Highway Administration
GBR	Geotechnical Baseline Report
MnDOT	Minnesota Department of Transportation
MoDOT	Missouri Department of Transportation
LWD	Light Weight Deflectometer
NDG	Nuclear Density Gauge
R&R	Repeatability and Reproducibility
SDG	Soil Density Gauge

ABSTRACT

The nuclear density gauge has been the standard soil compaction acceptance method for the Missouri Department of Transportation (MoDOT) for several decades. However, the cost of licensing, security, transport protocol and training imposed by the federal government have caused MoDot to question whether it remains a cost effective testing technology.. Nuclear density testing's rapidity and accuracy has been crucial in enabling MoDOT inspectors to keep contractor grading processes on schedule. But, in the last two years MoDOT's Quality Management program has shifted the bulk of testing requirements to the contractor, reducing the need for MoDOT inspection on grading projects. As a result,, MoDOT is investigating compaction testing alternatives to the nuclear density gauge which can provide the necessary results at a lower life cycle cost. The investigation comprised a comprehensive review of previous research into compaction testing alternative as well as key findings and gaps in research. This led to the purchase of XX pieces of alternative test equipment which were employed simultaneously alongside the nuclear density gauge on four large structural fill projects. The field testing yielded a set of comparable test results taken at the same time, in roughly the same location, and under the same environmental conditions, and arguably making this research the most comprehensive study of compaction testing technology on record.

The dissertation discusses MoDOT's Quality Management program's development and links to its origin in Design-Build project best practices, which provided the motivation to seek alternatives to the nuclear density gauge. Life Cycle Cost Analysis and Cost Index theory was utilized in comparing the compaction testing alternatives and presenting present cost per compaction test for the Department. For MoDOT project sites, linear and multiple regression

analyses were developed to determine if correlations existed between soil density and associated modulus or Clegg Impact Values. Lastly, an assessment of the repeatability and reproducibility of the light weight deflectometer and the dynamic cone penetrometer on a project site was completed with three distinct statistical analytical methods. The data presented herein can be integral elements in MoDOT's decision to eliminate or keep the nuclear density gauge.

ACKNOWLEDGEMENTS

I would like to thank Dr. Douglas Gransberg for his tremendous guidance, mentoring and support throughout the whole dissertation process. Dr. Gransberg has changed my life forever both academically and professionally. I will always be thankful and appreciative.

I would also like to thank the members of my committee, Dr. Jeong, Dr. Jahren, Dr. Rivero, Dr. White, and Dr. Bumblauskas, for the time they took out of their busy schedules to share their experiences with me.

Additional thanks are reserved for Dave Ahlvers, The State Construction Engineer, and Mike Fritz, former Geotechnical Director at the Missouri Department of Transportation, for their sponsorship of this Dissertation Research. I also appreciate the time and effort my coworkers at MoDOT put into this project. They include Nicole Scott, Kenny Barnett, Trevor Libbert, Ricardo Todd, Paul Hilchen, Cullen Hudson, Rylan Ellis, Tyler Lacy, and Robert Massman. Special appreciation is noted for the assistance provided to me by Renee McHenry of the Research Section in obtaining literature that made the research possible. I would also like to acknowledge John Donahue from the Pavement Section for his advice and support during the research process.

Finally I would very much thank my family--Mom, Dad and Brother--for their support and forgiveness for my infrequent visits home during this research. Special and loving thanks to my wife, Ellen, for her support not only through the Doctoral Dissertation process but through my previous academic endeavors and professional career.

CHAPTER 1. INTRODUCTION

Design-Build (DB) project delivery alters the traditional public highway project delivery system by awarding both the design and the construction to a single entity in a single contract. The literature details many different advantages and disadvantages that a public highway agency must consider when selecting an alternative project delivery method (FHWA 2006)that evolve from the change to DB, but none are more important than the change that occurs when the project's engineer-of-record is moved from being directly contracted with the owner to holding privity of contract with the design-builder, Due to the requirement to furnish performance bonds, this entity is typically a general contractor. With that shift comes a commensurate shift in project quality management responsibilities.

"In the traditional Design-Bid-Build (DBB) system, quality is fixed through the plans and specifications. Thus, in DBB, with schedule and quality fixed, the cost of construction is a factor in which the owner seeks competition. Conversely, in DB, with cost and schedule fixed, the scope and hence the level of quality is the main element of competition.(Gransberg and Molenaar 2004) Once the DB project has been awarded, the follow-on quality management system shifts many of the owner's quality assurance responsibilities to the design-builder because it owns the engineer-of-record (Kraft and Molenaar 2013). Therefore, the owner no longer plays as active a role in personally performing many of the construction quality assurance tasks such as verifying contractor quality control compaction tests and as a result no longer requires the same testing and inspection staff with its inventory of field testing equipment.

The Missouri Department of Transportation (MoDOT) has learned several lessons from initial state DB projects and that of other DOT DB projects. One is requiring the MoDOT

Geotechnical Engineering Section to perform geotechnical exploration borings at probable structure locations and areas of concern for settlement, liquefaction or landslide. It then proceeds to produce a Geotechnical Baseline Report (GBR) or Geotechnical Data Report (GDR) (Gransberg and Loulakis 2012), which allocates the geotechnical risk and allows prospective DB team to increase project scope on other elements of the project rather than include a large contingency for subsurface conditions. MoDOT places its project team beside the DB team to facilitate the development of non-standard design solutions and ensure quality management in construction.

MoDOT assigns the DB team responsibility for quality control inspection and itself for quality assurance at designated points in each project element during construction. MoDOT so satisfied with the quality it received on its DB projects (Ahlvers et al. 2013).that the Department instituted DB Quality management practices on select DBB in 2012 and fully implemented the new system for all projects in 2013. The new system was named Quality Management.

Prior to 2012 on DBB projects, MoDOT performed all required compaction testing. The nuclear density gauge (NDG) provided a quick moisture and density verification that allowed the contractor to proceed at a fast-moving pace on embankment and pavement construction. On DB projects, the contractor performed the majority of compaction tests with MoDOT inspectors testing at designated points in the project. With the implementation of Quality Management in 2013, compaction testing by NDG completed by MoDOT resident engineer offices dropped significantly. This led the MoDOT Construction Division and Geotechnical Engineering Office to question the value of retaining its inventory of NDGs and to investigate the life cycle cost (LCC) of the NDG and other compaction testing devices. The cost per NDG test completed by

MoDOT inspectors was also assessed. These costs were assessed against the speed, repeatability, and reproducibility of the NDG versus less costly testing alternatives.

This dissertation contains a collection of 5 journal articles arranged in sequence to match the direction and purpose of the research as described above. Chapter 2 will provide background and the reasons for the research and Chapter 3 will detail the methodology used in the research.

The first stage of the research was to conduct a review of preceding investigations that assessed and compared differing soil compaction testing devices. The review concentrated at identifying areas that had little or no previous research and the outcome is discussed in Chapter 4.The second topic covered is State DOT best practices in reducing geotechnical risk/uncertainty on DB projects and is found in Chapter 5.On MoDOT DB projects, inspection and quality management protocols led to the majority of inspection being performed by the DB team with specific MoDOT assurance testing. To quantify the impact of the new Quality Management practices on all DBB projects, the reduction in MoDOT compaction testing and associated testing costs for the NDG and its alternatives in terms of LCC were evaluated and reported in Chapter 6.

Next MoDOT’s Quality Management evolution is examined in conjunction with correlation of the standard NDG to modulus and Clegg impact values (CIV) which has implications in the revised QA and QC processes Chapter 7 contains the output from that analysis. Lastly the repeatability and reproducibility of the lightweight deflectometer and dynamic cone penetration test, two alternatives of particular interest to MoDOT are examined using contrasting evaluation methods in Chapter 8.

CHAPTER 2. RESEARCH BACKGROUND AND MOTIVATIONS

Background

The Federally required training, licensing and security for use of the NDG have become a barrier both financially and for efficiency. As a result, MoDOT has questioned the utility of maintaining the NDG and is actively seeking more cost-effective alternatives to fulfill its compaction testing needs. The research presented in Chapters 3 through 9, represent a convergence of two MoDOT initiatives. The first, being able to translate DB QC/QA processes and methods to routine DBB projects. The Quality Management process reduces the proportion of compaction tests conducted by MoDOT staff. Secondly, the Department questioned the cost of maintaining, providing security for and licensing and radiation safety training for NDGs that see vastly reduced usage. Recently, a third initiative was approved by senior management at MoDOT, to implement Intelligent Compaction (IC), which is based on a modulus testing methodology, and is expected to further influence the future usage of the NDG which measures density and moisture rather than modulus. Quality Management and the future implementation of IC have combined to make an impact on the assessment of replacing the NDG with a differing compaction method or device. The research described in this dissertation is a combination of comparative field studies, life cycle cost analysis of alternate compaction test technologies, and an examination of the repeatability and reproducibility of some of the most promising compaction testing alternates. The life cycle cost analysis is adapted to supply cost indices that allow direct comparisons on a life cycle cost per test basis. The research imparts both technical and economic information needed to make an informed decision on whether or not to replace the NDG and a concurrent “apples to apples” comparison of potential alternatives.

Field testing of the NDG and all the alternatives in the same locations on the same projects was utilized to validate the decision framework. Field testing was completed on four active construction projects to permit the experiment to be applied to differing soils with differing contractors. At this writing the work described in this thesis may constitute the most complete assessment of compaction testing alternatives under uniform field conditions to date. Table 2-1 summarizes the details of the four test sites.

Table 2-1 Testing Locations

Route/Location	County/City	Main Soil Types	Comments
Route 50	Osage/Linn	Lean Clay	12 mile DB Project 5 differing proctor areas
Route 364 Phase 3	St. Louis/ O'Fallon	Lean Clay	9 Mile DB Project proctor testing frequency and locations were designated by the design-build team
SAMS Construction Site	Cole/Jefferson City	Manufactured Sand	Private construction and grading
Discovery Parkway	Boone/Columbia	Lean Clay	Private construction and grading 3 Differing Proctor Areas

Additional testing was also conducted at the Jefferson City MoDOT Main Maintenance facility on three test beds constructed with the assistance of maintenance forces. The three different test beds consisted of sand, lean clay and Type 5 base.

The initial testing plan consisted of the following 8 alternate test methods/ devices with the nuclear density gauge performing as the reference device:

Density and Moisture

1. Electrical Density Gauge - EDG
2. Soil Density Gauge – SDG 200

3. Sand Cone - SC
4. Density Drive Sampler -DDS

Modulus/Stiffness/Clegg Impact Value (CIV)

1. Light Weight Deflectometer - LWD
2. Dynamic Cone Penetrometer – DCP
3. GeoGauge- GG
4. 10 kg Clegg Impact Hammer

The Electrical Density Gauge was dropped after use on the first two MoDOT test sites listed in Table 2-1, due to the extreme difficulty in building a satisfactory soil model in the calibration process. The Geogauge was also dropped in the initial testing stages due to problem with repeatability in field conditions.

Motivation

There were several motivations in conducting this research. The primary being that the MoDOT Construction and Materials Division required a thorough analysis of compaction testing options to the NDG on both technical and financial levels. The Division was willing to give time and personnel to the research as well as financial support in buying or renting needed testing equipment.

Secondly, further motivation was found as the researcher conducted the require literature review for this dissertation and discovered gaps in the body of knowledge as well as previously unrecognized important discoveries that appeared to be promising in the context of implementing Quality Management on all projects within the MoDOT construction program. (See Chapter 4). First, MoDOT required sound financial justification to replace the NDGs it already owned by some other technology. The majority of the literature reviewed concerned itself with density and moisture content reading comparisons, not cost comparisons. Cho et al. (2011) calculated a limited life cycle cost analysis between a TransTech SDG 200, Soil Density Gauge and a NDG. Thus, the financial impact of exclusive NDG use compared to possible alternatives was virtually unresearched.

Secondly, previous research was primarily devoted to laboratory testing with no regard for the construction production impact of each alternative. Specifically, no authoritative time and motion studies of field moisture measurements were found on which to base personnel costs for input to the LCCA Berney and Kyzar (2012), and Berney et al. (2012 compared differing moisture measurement devices to a laboratory oven, but only focused on numerical deviation in the readings and not the time needed to perform the test. Time to conduct the moisture test in keeping the contractor on schedule is critical to maintaining as-bid production rates. Field moisture was found to be critical (Abu-Farsakh et al. 2004) for modulus/stiffness measurement devices. Stiffness variation due to moisture is much larger than density variation and maximum stiffness occurs before optimum moisture, which is vastly different from density measurements base on proctor results.

Problem Statement

The research will provide applicable data regarding life cycle costs, correlation potential, reliability and accuracy, to MoDOT and its industry partner the Missouri Association of General Contractors (AGC) Grading Division, to make informed decisions about the future of construction compaction field testing in Missouri. The research question that the work will answer is:

Should MoDOT replace its current density specifications with modulus/stiffness measurements?

To answer this question other key related questions must be answered and include the following:

1. How has MoDOT Quality Management changed the frequency and number of compaction tests conducted by construction personnel and how many test devices are needed for a Resident Engineer's project office?
2. If IC is adopted for majority of projects, will the moisture-density technology used by the NDG become obsolete and can the results of this research contribute to making that decision?
3. Can the contractor use a different compaction testing device than the MoDOT grading inspectors?
4. How reliable and precise are the reviewed compaction measuring devices and how do they compare to the NDG?
5. What are the costs associated with the purchase and long term use of the testing device(s) in the field?

CHAPTER 3. METHODOLOGY AND VALIDATION

The research steps research mechanisms that comprise the methodology are shown in Figure 3-1 and are explained in detail in the methodology sections of Chapters 4,5,6,7,8, and 9.

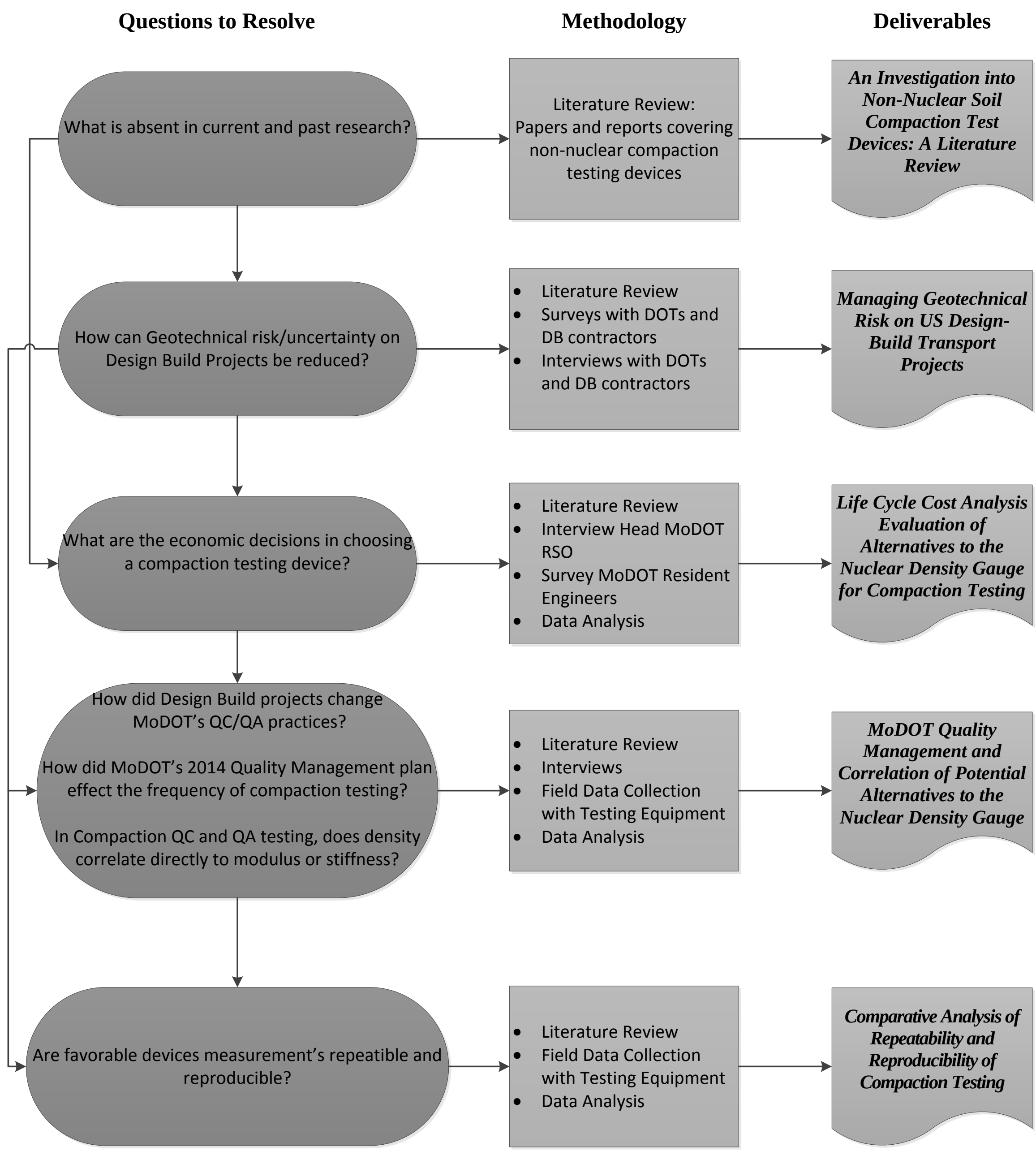


Figure 3-1 Research Methodology

Comparing compaction test devices followed the protocol shown in Figure 3-2.

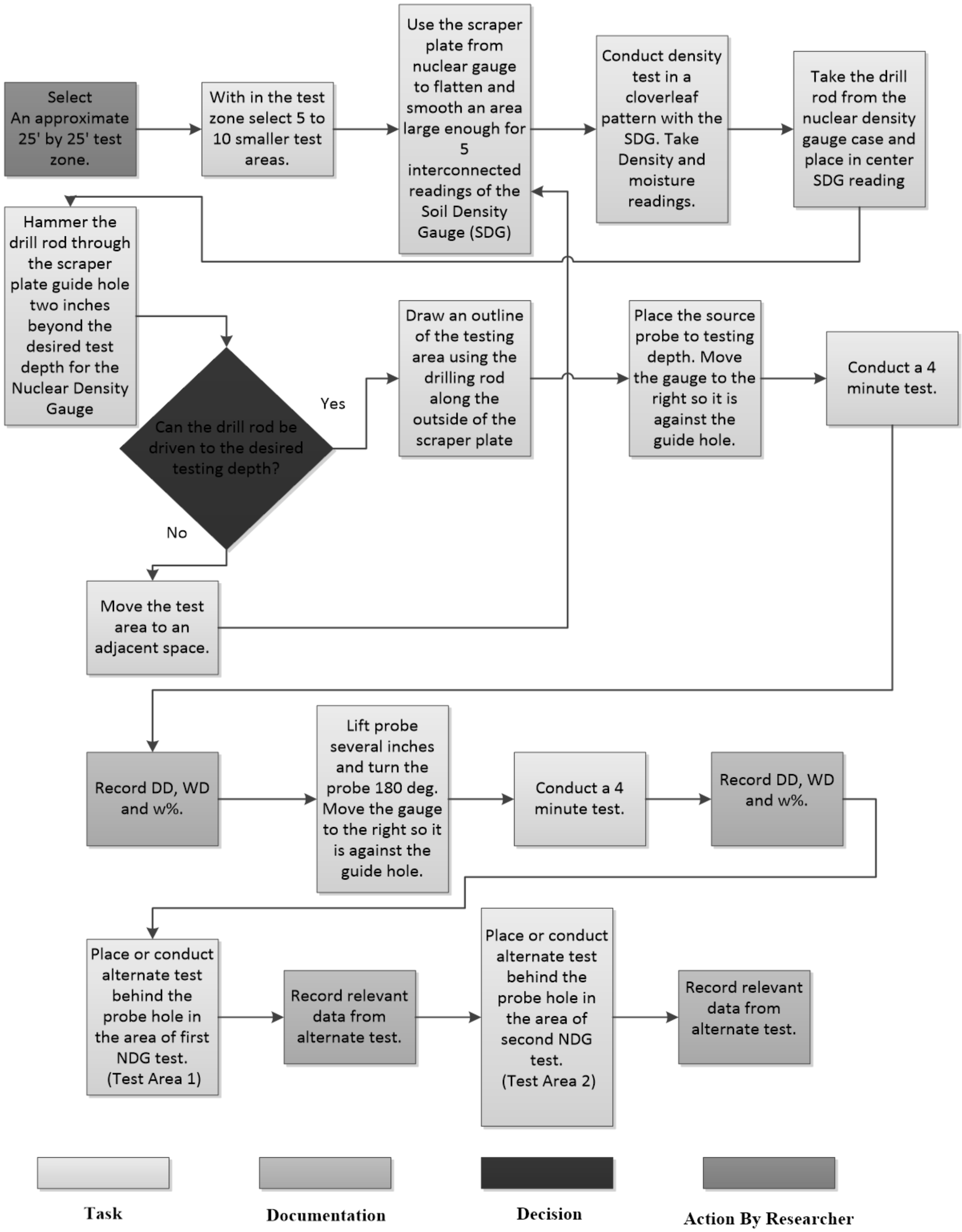


Figure 3-2 Comparison Compaction Protocol

The comparative testing protocol is shown in a generalized testing arrangement shown in Figure 3-3. This arrangement allowed for the testing of four differing compaction test devices.

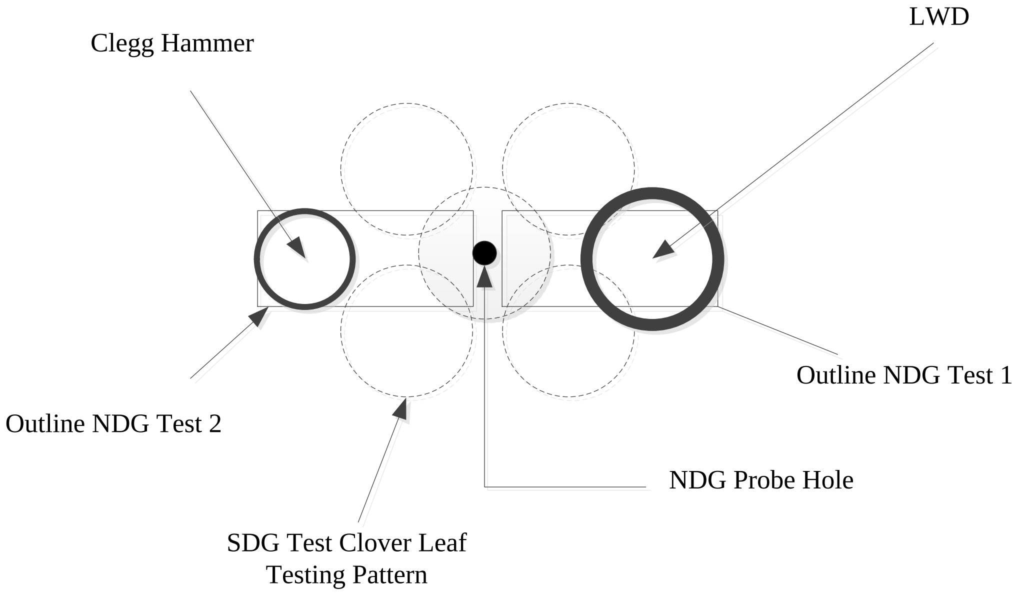


Figure 3-3 General Comparative Testing Arrangement

Validation

During the development and progression of the research, both formal and informal meetings have been held with Dave Ahlvers, MoDOT State Construction and Materials Engineer, John Donahue, Construction Liaison Engineer (Pavements) , Dennis Brucks, Construction Liaison Engineer (Grading) and William Stone, Research Director, to discuss and examine the research direction, status, and preliminary findings.

Research direction and findings were presented at the 2012, 2013 and 2014 annual joint MoDOT/ Missouri Association of General Contractors meeting. Points and questions from the presentation participants were noted and were integrated into upcoming research segments and methodology.

CHAPTER 4. AN INVESTIGATION INTO NON-NUCLEAR SOIL COMPACTION TEST DEVICES: A CRITICAL ANALYSIS OF THE LITERATURE

McLain, K.W., and Gransberg, D.D. *An Investigation into Non-Nuclear Soil Compaction Test Devices: A Critical Analysis of the Literature.* (Submitted to *ASTM Geotechnical Journal*, October 2015)

This chapter discusses completed comprehensive review of the literature on soil compaction testing alternatives to the nuclear density gauge. The literature review had three aims, gather needed background information, determine what research in this area has been completed and to Identify gaps in the body of knowledge alternative compaction testing that require further investigation.

Abstract

State Departments of Transportation as well as other large construction organizations are looking for a compaction testing alternative to the nuclear density gauge. The Federally required training, licensing and security mandates have become a hindrance due to the annual costs of operating and maintaining the devices. Research units of these organizations in partnership with universities have been tasked to investigate alternatives that might replace the nuclear density gauge for routine compaction testing at a lower life cycle cost. This paper examines the evolution of research in the non-nuclear quality assurance and quality control of soil and base compaction. The literature review focused on three comparators: difference in approaches between tested devices and methods, the setting of targets for modulus stiffness devices, and the examination of accuracy and reliability. The paper also summarizes tested devices, how well the

devices were reported to have performed, and how they ranked if compared to other devices for studied research projects. The paper finds that previous research either sparingly covers or fails to include the information on field test performance parameters such as repeatability and time to complete a field test for most alternatives. This may be due to an inability to conduct testing over the broad set of conditions with which practicing geotechnical engineers are faced. It concludes that as a starting point, research that includes a broad set of testing methods and equipment on the same set of soil conditions and that includes both density and stiffness test methods is needed to determine whether or not the advantages of the nuclear density gauge outweigh its disadvantages. The paper's primary contribution is to consolidate the current literature on the subject of compaction testing in a single document and provide a critical analysis of the same and as such the paper serves to benchmark the state-of-the-practice in this area.

Background

Achieving minimum densities in structural embankments, as well as in the subgrade, subbase and base for pavement structures are essential for long term performance for roadways (Schaefer et al. 2008). For many state Departments of Transportation (DOT), this is achieved by testing soil and aggregate layers for a target density and moisture level. Commonly, the contractor conducts field quality control testing and the DOT conducts quality assurance testing (Berney and Kyzar 2012). In most cases both the density and moisture of a compacted course is found using a nuclear density gauge (NDG). The NDG's major advantage is its ability to rapidly measure density and moisture content, typically between 1 and 4 minutes testing time (ASTM 2010). While the NDG is reliable and efficient, it also entails a considerable expenditure of time and expense to meet the statutory requires of the U.S. Nuclear Regulatory Commission and state emergency management agency for technician safety training, radiation and certification classes, licenses and storage facilities (Rathje et al. 2006). This time and expense is borne not only by the state DOTs but also construction contractors and materials testing consultants.

Some state DOTs and their university research partners have undertaken research to examine compaction testing alternatives to the NDG. Studies have assessed differing methods and tools, tested materials, and evaluation methodologies. The literature shows that researchers conduct two types of studies. The first type evaluates two or more different testing technologies on a comparative performance basis, and the second type simply concentrates on an in-depth performance analysis of a single compaction testing technology At this writing, no comprehensive research was found that evaluates all the possible alternatives to the NDG in speed, reliability, repeatability and acceptance with field inspectors from both DOTs and contractors. The alternate testing methods and tools that could theoretically supplant the NDG can be placed in three categories (Berney and Kyzar 2012).

1. Electrical Density and Moisture Gauges
    a. Electrical Density Gauge (EDG)
    b. Soil Density Gauge (SDG)
2. Volume Replacement/Volume Measurement
    a. Balloon (RB)
    b. Sand Cone (SC)
    c. Density Drive Sampler
3. Stiffness/Modulus Measurement
    a. Light Weight Deflectometer (LWD)
    b. Dynamic Cone Penetrometer (DCP)

c. Clegg Soil Impact Tester

d. GeoGauge (GG)

Other methods such as the Soil Compaction Supervisor (SCS) , the Panda Dynamic Cone Penetrometer (Farrag et al. 2005) , Steel Shot, the now defunct Moisture Density Indicator (M+DI) (Berney and Kyzar 2012) and the Portable Seismic Property Analyzer (PSPA) (Rathje et al. 2006) have also been investigated. A number of research projects investigated more than one alternative and compared them to one another and to a baseline method such as the NDG, sand cone, rubber balloon or drive sampler. Research teams have investigated compaction testing tools in each of the three device classes previously mentioned. Some researchers halted investigation of devices in the initial stages, while some did so after operating the devices/methods for some time. Conversely, Berney and Kyzar (2012) dropped stiffness and modulus measurement devices from their investigation because they believed that there was little correlation between proctor densities/moistures and stiffness or modulus values. Further NDG comparison efforts by this team were limited to volume replacement devices, as well as electrical density and moisture devices. Rathje et al. (2006) took an opposite direction dropping two electrical density and moisture devices, the EDG and the SDG due to calibration problems encountered during testing.

While examining several compaction testing devices in a single project requires more initial work and time to procure the test devices, not to mention becoming familiar with established American Society for Testing and Materials (ASTM) test specifications and device manuals shown in Table 4-1, it allows the investigation team to compare each device on a pairwise basis against the baseline compaction measurement device (NDG). Hence, further research then can be focused on the alternatives that performed the best and hold the most promise in adequately replacing the NDG.

Table 4-1 ASTM Compaction Test Device Test Methods

Compaction Test Device	ASTM Test Method
Nuclear Density Gauge	ASTM D6938
Balloon	ASTM D2167
Sand Cone	ASTM D1556
Density Drive Sampler	ASTM D2937
Light Weight Deflectometer	ASTM E2835
Dynamic Cone Penetrometer	ASTM D6951/D6951M
Clegg Soil Impact Tester	ASTM D5874
GeoGauge	ASTM D6758

Table 4-2 is a summary of those studies found in the literature that compared more than a single compaction testing device. One can see that many studies actually examined one of each category of technologies.

Table 4-3 summarizes those research projects have that concentrated on a single alternative method. This research allows the researchers to focus in on one specific alternative method (usually in the same category or class of alternative methods). This usually gave a more in depth examination and background on the non–nuclear alternative testing device with the base line device generally being the NDG. The drawback to investigating limited alternatives is that the researchers and readers must be careful to become myopic or biased towards the compaction alternative investigated.

Table 4-2 Research Projects that Examined Multiple Compaction Test Devices

Author	Nuclear and Electrical Devices Moisture and Density	Volume Replacement/ Traditional Devices	Stiffness/Modulus Devices
Farrag et al. 2005	NDG	SC (top layer-evaluate NDG results)	Utility DCP; DCP; CIegg; Panda; GeoGauge
Berney &Kyzar 2012	NDG; SDG; EDG; M+DI (initial)	SC; WB; SS	Only initial investigation then dropped: Clegg; GeoGauge; LWD; DCP
Cho et al. 2011	NDG; EDG; M+DI(initial)	Density Drive Sampler	LWD
Kim, Prezzi and Salgado 2010	NDG	SC	Clegg; DCP; LWD Summary; GeoGauge Summary
Rathje et al. 2006	NDG; SDG(SQI) (initial); EDG (initial); MDI	RB	Clegg; DCP; Panda; PSPA
Siekmeier et al. 2009	None	SC	DCP; LWD
Mooney et al. 2008	NDG; EDG summary	None	DCP; LWD; Clegg; GeoGauge
Brown 2007	M+DI; EDG; NDG	None	None
Meehan and Hertz 2011	NDG; EDG	SC; Density Drive Cylinder	None
Meehan et al. 2012	NDG	None	LWD; DCP; GeoGauge
Abu-Farsakh et al. 2004	NDG	None	LWD; DCP; GeoGauge; FWD; Plate load Test
White et al. 2013	None	None	LWD; DCP; FWD*; CMV and MDP**
***FWD = Falling Weight Deflectometer; ** Roller Impact Compaction Monitoring Technologies: CMV = Compaction Meter Value; MDP = Machine Drive Power**			

Table 4-3 Research Projects that Investigated Single Compaction Test Devices

Author	LWD	EDG	GeoGauge	Clegg
Tehrani and Meehan 2010	X			
Meehan and Hertz 2013		X		
Ooi and Pu 2003		X		
Lenke et al. 2003			X	
Vennapusa and White 2009	X			
Maher et al. 2002			X	
Erchul and Meade 1990				X
Erchul 1999				X
Vanden Berge 2003				X

Methodology

Comparison Analysis

A diverse set of means and methods were used to conduct the comparative analyses presented in this paper. These include direct comparison, linear regression, statistical analysis, and the setting of targets. Direct comparison is the most straightforward and usually used to compare testing method results that are in the same unit of measure. An example of this is comparing density from an electrical density gauge and the density measured with a NDG. This is usually displayed in tabular or graphical form. Figure 4-1 illustrates typical output for this type of comparison. Meehan and Hertz (2013) use root mean square error to quantify the difference between values estimated by the researcher with true values of the quantity being estimated.

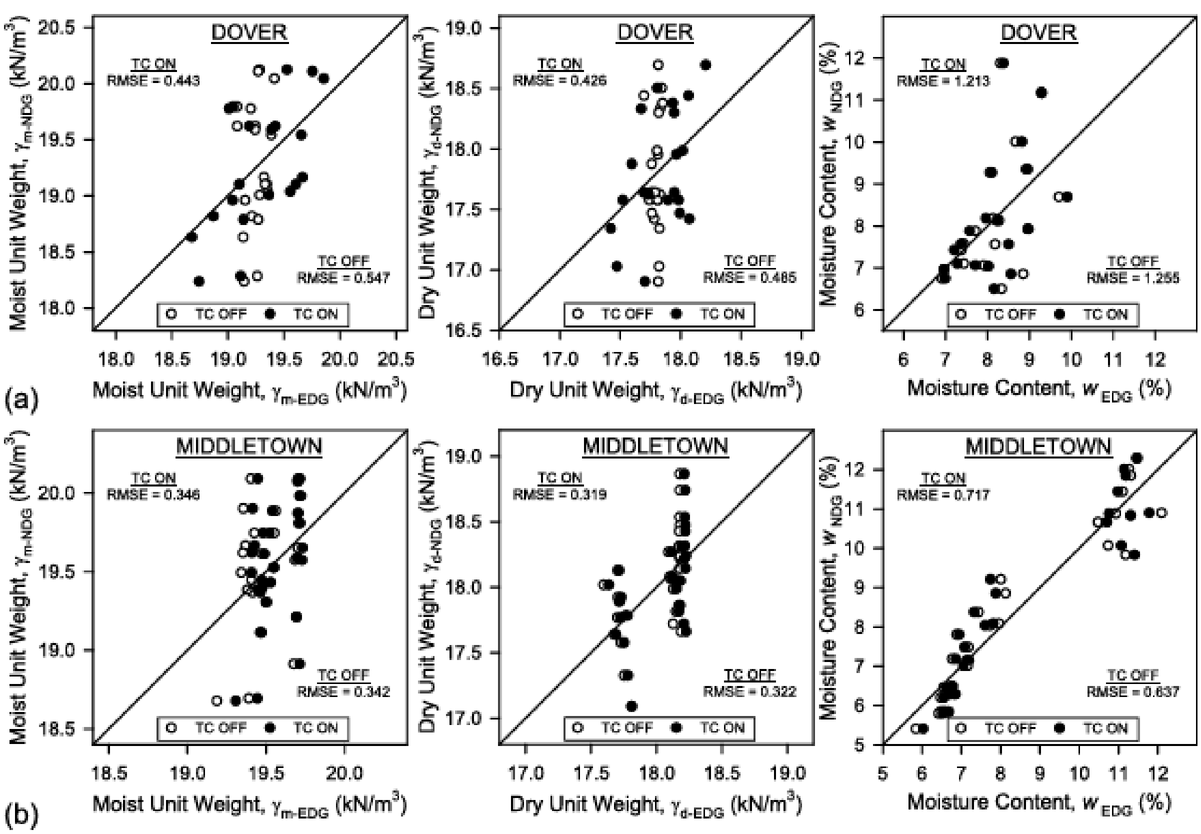


Figure 4-1 Example of Direct Comparison (Meehan and Hertz 2013)

Linear regression is also a standard method to compare testing method results that are not the same units of measure. According to (Yale 1997), “linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable.” Linear regression is used to fit a predictive model to data set of y and x. With linear regression the coefficient of determination is also calculated. This is seen when comparing density to stiffness or modulus. R^2 or the coefficient of determination is a statistical function that provides data about the exactitude of fit of a model. A coefficient of determination of 1 indicates that the calculated regression line fits perfectly to the data.

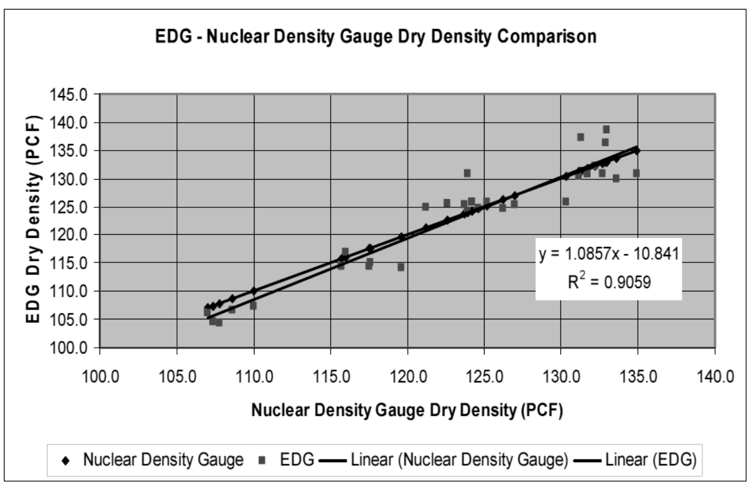


Figure 4-2 Linear Regression with Coefficient of Determination (Brown 2007)

One of the more complex statistical comparisons of compaction measurement devices is that presented in Berney and Kyzar 2012. The resulting analysis evaluated both accuracy and precision for devices. The first step in the analysis was to establish how much the density measured by the alternative device deviated from that measured by the NDG for different soil types tested (Figure 4-3).

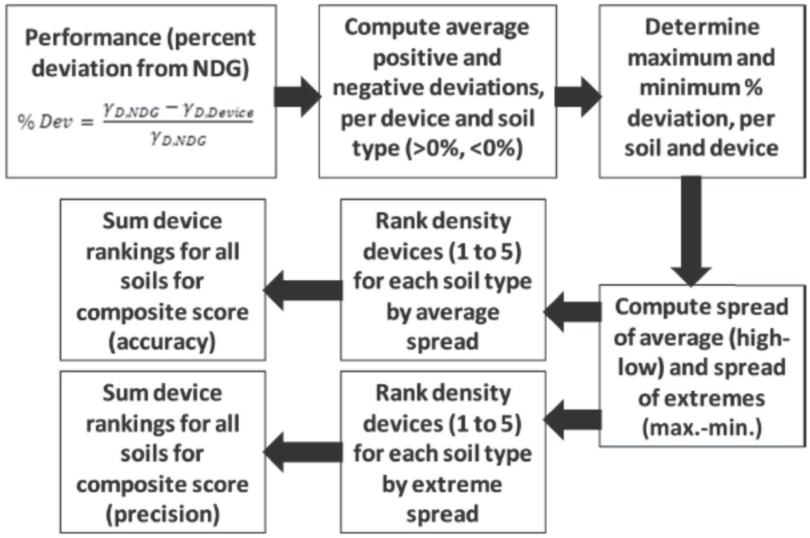
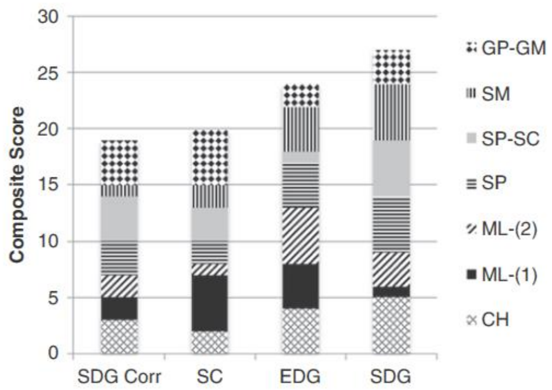
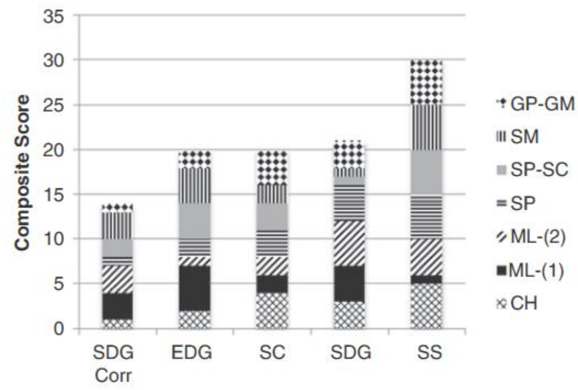


Figure 4-3 Measurement of Compaction Method (Berney and Kyzar 2012)

The process then calculates two device ratings as shown in Figure 4-4. When comparing the compaction test device the problem of accuracy versus precision arises. *The Form and Style for ASTM Standards* (2013), defines precision as the "closeness of agreement between test results obtained under prescribed conditions." Accuracy, on the other hand, is the ability of a measurement to match the actual value of the quantity being measured. These two definitions match what many researchers are examining during the investigation process for NDG alternatives.



Average dry density spread/ Accuracy



Max- Min dry density spread/ Precision

Figure 4-4 Combined Graph of Accuracy and Precision (Berney and Kyzar 2012)

The setting of targets is also a common practice for the researchers to show the strength, modulus, or stiffness values need to be obtained to reach certain or target percentage proctor value. Siekmeier et al. (2009) used target values for implementation of the use of LWD and DCP in quality control and assurance practices for both granular and fine grained soils. The target values were presented in tabular form for fine grained soils as shown in Table 4-4 and for course grained soils as shown in Table 4-5.

Table 4-4 Target Table for Zorn LWD – Fine Grained Soils. (Siekmeier et al. 2009)

Plastic Limit	Estimated Optimum Moisture	Field Moisture as a Percent of Optimum Moisture	DCP Target DPI at Field Moisture	Zorn Deflection Target at Field Moisture (minimum)	Zorn Deflection target at Field Moisture (maximum)
(%)	(%)	(%)	(mm/drop)	(mm)	(mm)

Table 4-5 Target Table for DCP and Zorn LWD – Course Grained Soils. (Siekmeier et al. 2009)

Grading Number	Moisture Content	Target DPI	Target DPI Modulus CSIR	Target LWD Modulus Dynatest	Target Modulus Zorn	Target LWD Deflection Zorn
GN	(mm/drop)	(mm/drop)	(MPa)	(MPa)	(MPa)	(mm)
CSIR - The Council for Scientific and Industrial Research (CSIR) in South Africa where the DCP was developed. DPI - DCP penetration index, penetration distance per drop. Dynatest- Manufacturer of LWD Units Grading Number is equal to percent passing (1” sieve + ¾” sieve + 3/8 inch sieve + #4 sieve)/(100) Zorn - Manufacturer of LWD Units						

Targets are also commonly presented in graphical form (Farrag et al. 2005; Mooney et al. 2008). Mooney et al. (2008) went into considerable depth in describing the steps and calculations for setting of target values. Figure 4-5 is a graph for target value for LWD meeting 95 percent compaction.

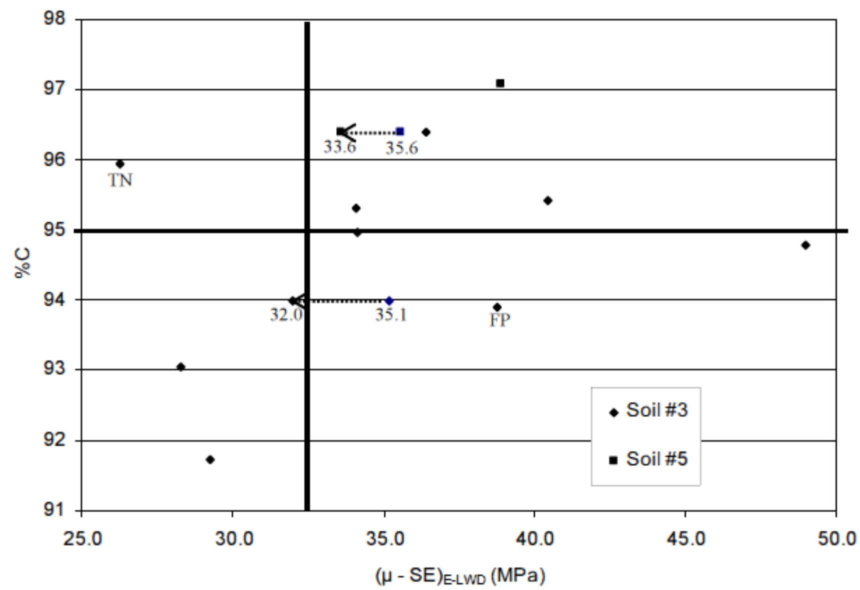


Figure 4-5 Target Values for LWD (Mooney et al. 2008)

Precision and Repeatability of Testing Devices

The paper by Mooney et al. (2008) set itself apart from other reviewed research by examining repeatability of data from investigated compaction testing systems by using two approaches. The first approach performed initial tests for each device, and then 5 to 10 tests were performed without removing the compaction test device. In the second approach the compaction test device is removed and replaced before starting a new test. This procedure is important because it quantifies precision for investigated testing devices. For a device to be considered a viable option, it must also display repeatability (Fig.4-6).

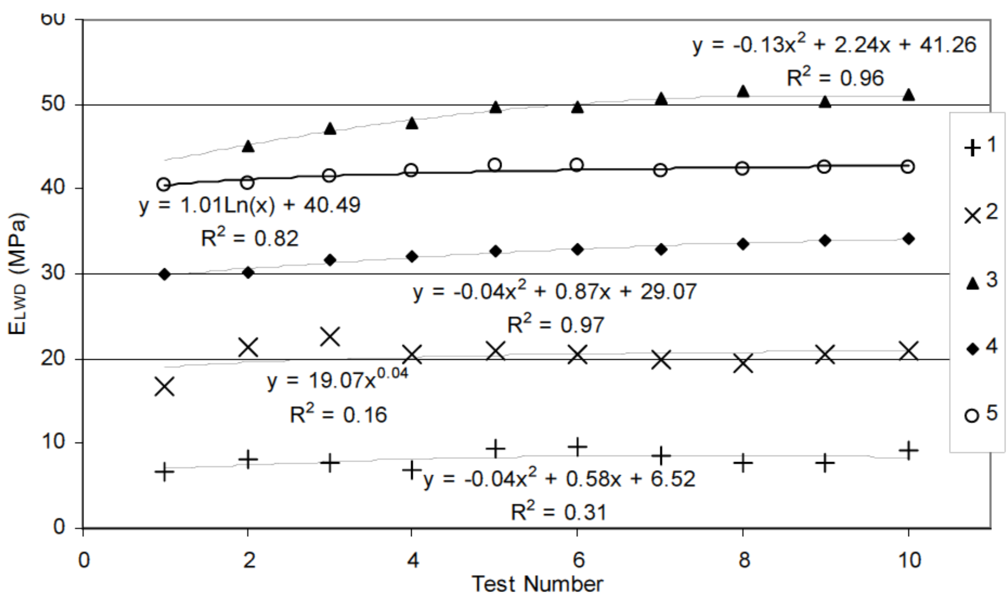


Figure 4-6 Testing Repeatability in Various Soils Using LWD (Mooney et al. 2008)

Maher et al. (2002) examined the repeatability of the Humboldt Stiffness Gauge (also known as the GeoGauge) by testing at the same depth three times in the row and then repeated again for a total of six measurements in a large soil bin (8 ft. deep and 15 ft. in diameter). The six measurements were completed at five to seven different depths for four soils. The first three readings were averaged and a standard deviation calculated this was also completed for the second set of three readings. Berney and Kyzar (2012) performed a study of precision that compared directly to the NDG as shown in Figures 4-3 and 4-4.

Analysis and Results

The researchers who investigated modulus/ stiffness devices noted that when comparing modulus or stiffness to proctor density. Maximum density and maximum modulus/stiffness did not occur concurrently at the same moisture content. As shown in the Figure 4-7 from Lenke et al. (2003), maximum stiffness occurs before maximum density and at lower moisture content.

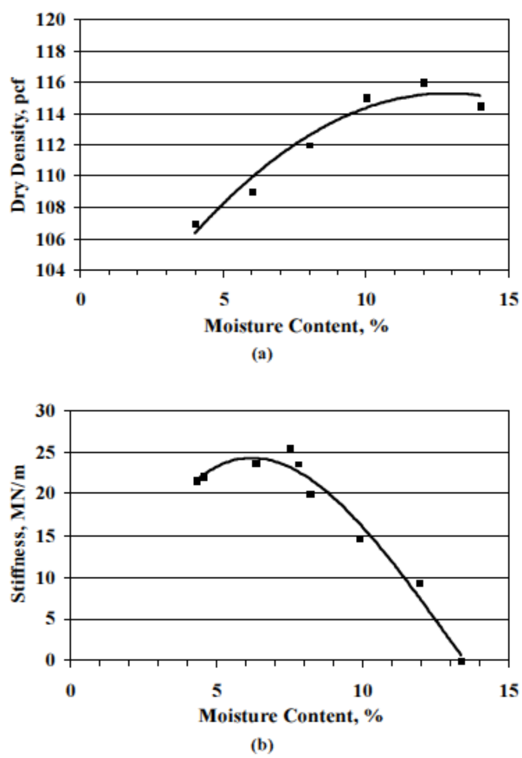


Figure 4-7 Lag of Maximum Stiffness to Maximum Density- Cohesive Soil (a) Density (b) GeoGauge (Lenke et al. 2003)

Another significant discovery, noted in Abu-Farsakh et al. 2004, GeoGauge testing, (modulus/stiffness device) showed that the variation of stiffness within the plus or minus two percent of maximum density is much larger than the variation of density with in the plus or minus two percent range. This implies that the using stiffness/modulus devices for quality control/assurance may be challenging because of sensitivity to moisture content.

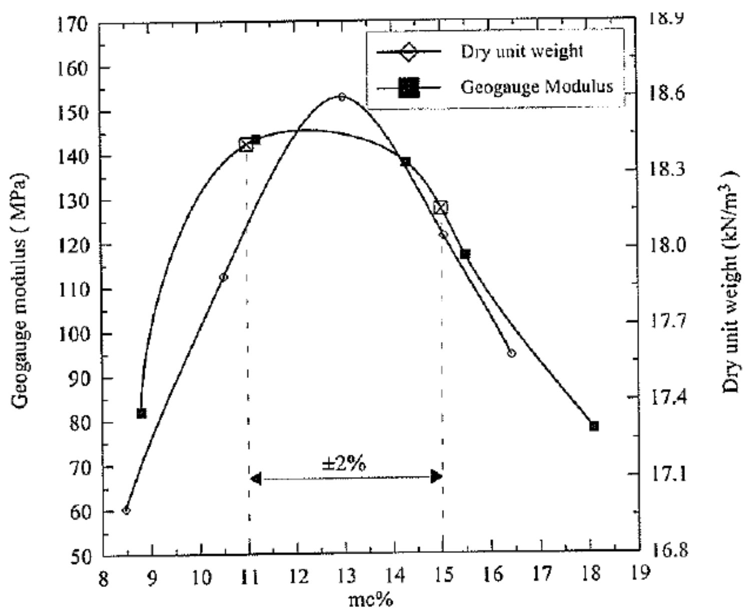


Figure 4-8 GeoGauge Modulus Readings Compared to Corresponding Density and Moisture Content (Abu-Farsakh et al. 2004)

The variation of moisture content influences stiffness/modulus and Clegg Impact Values (CIV) measurements. This is shown in material within a small range of relative compaction (Fig. 4-9).

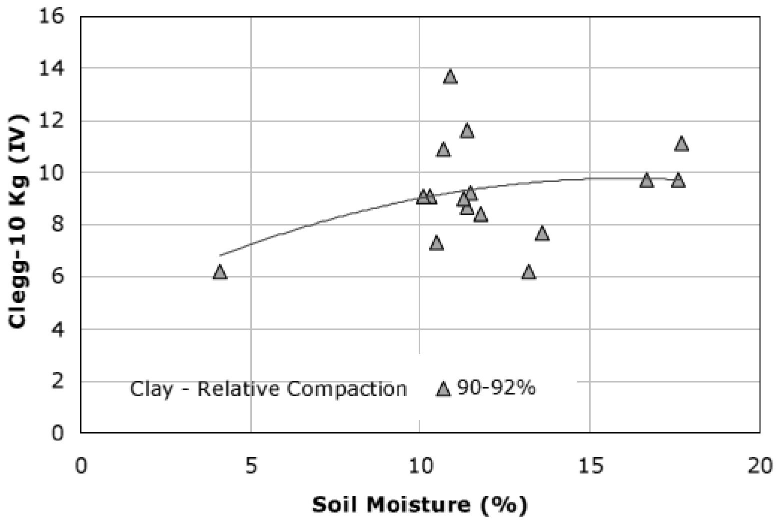


Figure 4-9 Change in 10 Kg Clegg Impact Values VS. Moisture Content. (Farrag et al. 2005)

Farrag et al. (2005) also observed that stiffness and modulus DCP blow count and CIV increased with increasing moisture content to optimum moisture content (Fig. 4-10). This moisture content did not correspond to the optimum moisture from the Modified Proctor Test.

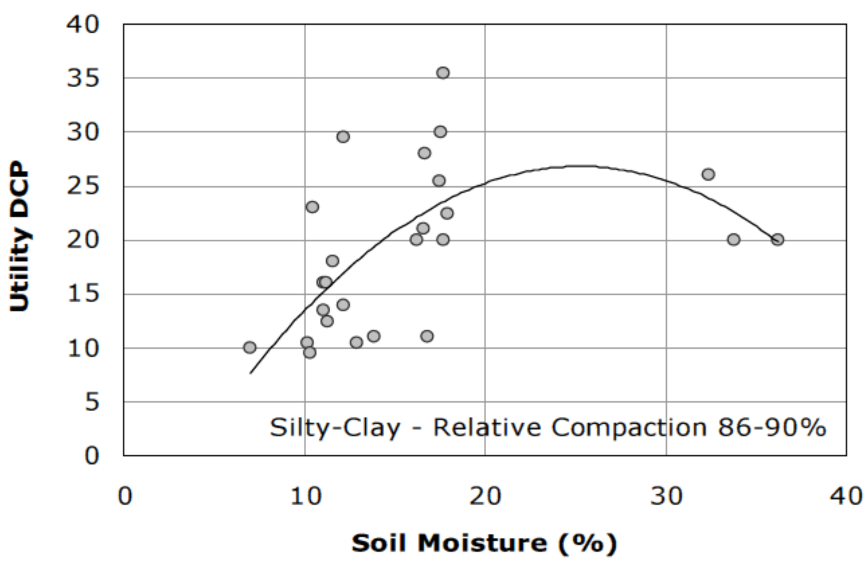


Figure 4-10 Change in Blow Numbers and Moisture Content of Soil for Utility DCP (Farrag et al. 2005)

For the research projects noted in Table 4-1 that consisted of testing various compaction QA/QC devices, summaries are noted below in Tables 4-6a, 4-6b, 4-6c and 4-6d. The summaries focus on test preparation, either field and/or laboratory tests, tested soils, and comparison methods that researchers followed. Also included in the tables are the research recommendations and devices to use and devices to eliminate from consideration for QC/QA compaction testing.

Individual or dual compaction testing devices research was noted in Table 4-3. Further information on those projects is presented in Tables 4-7a and 4-7b.

Information Not Found in the Literature

To thoughtfully evaluate whether or not to replace the NDG, DOT practitioners need not only the engineering and statistical performance of each alternative, but also information on each option's performance during the field testing. The output from laboratory testing is valuable and informative but because the test conditions are highly controlled, not necessarily reflective of how a given alternative will performance under the challenges of the field environment and when results are produced by a number of different technicians. The literature generally recognizes that for some testing methods, variability increases and repeatability decreases in the field due to the scale of the material tested and the fact that different testers rather than the same laboratory team will eventually be performing the tests. As was found by one research team, "the loss rate of soil moisture messages in field trials was disappointing, because in laboratory trials with the same hardware and software the delivery rate was close to 100%. The loss rates in field trials were time related, with significant changes in reliability during different time intervals." (Cardell-Oliver et al. 2004). Additionally, information on each alternative's capital and life cost, as well as the requirements for training and certification are also required. Lastly, since field testing operations are often conducted from the back of a vehicle, on rough ground, and in all kinds of weather, the ergonomics associated with each option must be evaluated to determine if it can be safely and effectively employed outside the lab. The above analysis found that the following information was not adequately or authoritatively covered by the literature reviewed for this study.

- <u>Field Measurement of Moisture:</u> Many research analyses that investigated stiffness/ modulus devices noted that obtaining moisture contents is important during the compaction process, but did not include comprehensive moisture measurement test device comparison into the studies. Berney and Kyzar (2012), and Berney et al. (2012) described research on comparing moisture measurement devices. However, the projects focused only on measurement deviation from the standard laboratory oven for the compared moisture measurement devices. The authors did not investigate the time to obtain moisture readings which is crucial in keeping roadway projects moving and on schedule. The work by Cardell-Oliver et al. (2004) was specifically focused on soil moisture measurement and as quoted above was unable to replicate the excellent results achieved in the lab with similar results in the field.
- <u>Cost Analysis:</u> The majority of the literature reviewed concerned itself with density and moisture content reading comparisons, not cost comparisons. Cho et al. (2011) completed a 15-year (life of source capsule integrity for NDG) economic analysis comparing initial and yearly costs for a nuclear gauge and a TransTech SDG 200 with a calculated breakeven point for the SDG 200 of approximately 4 years. Farrag et al. (2005), Mooney et al. (2008), and Rathje et al. (2006) reported initial costs for studied and/or tested

devices. Again costs for keeping and using NDGs are not only carried by DOTs but also by the contractors and consultants that construct the roadways and structures. The reported NDG cost information has been restricted to initial costs, cost of licenses, leak tests and training costs. The cost of security, the expense of adapting a building to properly store and secure NDGs, and annual costs for sending personnel to training and processing paperwork have not seriously been investigated and totaled up in compaction test device comparison study.

- Training and Ergonomics: Farrag et al. (2005) reported minimum descriptions of device testing ergonomics, calibration and training needed for each tested compaction device in the research program. But none of the reviewed research projects reported a detailed examination of human motion required to load and unload the device, move it around the jobsite and conduct a number of tests during a typical workday. Ergonomics is now a an important issue for employers due to lost time and medical costs incurred from workplace injuries caused by harmful and repetitive movements. Complexity of testing and training is also an important issue for the practitioners because as test methods become more complex the time and expense for training increases to ensure that errors in conducting and determining the results of compaction tests in the field are minimized.

Thus, these three areas constitute gaps in the recorded body of knowledge and are recommended as areas requiring further research by agencies wishing to evaluate replacing the NDG with another alternative.

Table 4-6a Multiple QC/QA Compaction Devices Test Research Projects Summary

Source	Test Preparation/Comparison Methods	Recommended Devices/Methods	Devices /Method Not Recommended or Have Complications
Farrag et al. 2005	• SC to NDG: Direct Comparison • SCS to NDG: Pass or Fail • Utility DCP to NDG: Target • GeoGauge to NDG: Target • Clegg to NDG: Target • DCP to NDG: Pass or Fail • Panda DCP to NDG: Acceptance or Refusal- based on soil type	The Utility DCP and 10 kg Clegg had top overall performance of compaction QC devices.	• The GeoGauge had deficient readings in sand and stone backfills. • 20 kg Clegg Hammer was found to be unwieldly due to its weight.
Berney & Kyzar 2012	• 50 ft. by 12 ft. test sections. • Test sections used the following materials: clay, loess, concrete sand, silty sand, clayey gravel, silty gravel and crushed limestone. • The SDG, SC, EDG, WB, and M+DI were directly compared to corresponding NDG readings.	• The corrected SDG using a linear offset factor had the least variability. • The sand cone was found to the next best device.	• The EDG required an extensive calibration routine to establish accuracy. • The uncorrected SDG had more variable readings than the EDG or sand cone. • Steel Shot had the most variability. • The GeoGauge, LWD, DCP, and Clegg Hammer were dropped from this study - no clear correlation between modulus//stiffness and density . • The M+DI and RB had greater than 25 percent null readings and also dropped.
Cho et al. 2011	• Researchers conducted tests at two sites containing loessial soils. • Drive sampler tests were the reference test in comparing NDG and EDG. • The reference for moisture was a laboratory oven. • Drive Sampler and NDG/EDG: linear regression equations and standard deviation calculations. • Drive Sampler and LWD: MNDOT Targets (no other information given).	• The NDG correlated with Drive Sampler 72.81% of time • The EDG correlated with Drive Sampler 39.80% of time • The LWD correlated with Drive Sampler 54.37% of time • EDG and LWD took much less time to record measurements. • Initial cost is higher, but there is a high return of investment.	N/A

Table 4-6b Multiple QC/QA Compaction Test Devices Research Projects Summary

Source	Test Preparation/Comparison Methods	Recommended Devices/Methods	Devices /Method Not Recommended or Have Complications
Kim, Prezzi and Salgado 2010	• DCP and Clegg tests along with SC in test pit composed of clay. • Field Test: Tests were run on 3 INDOT construction projects composed of sandy soils. • DCP blow count data was placed in histograms vs. frequency. • Clegg Impact Values were plotted against relative compaction.	Developed targeted relations based on Coefficient of Uniformity C_u for sandy (AASHTO A-3) soils for the required blow counts for depth intervals DCP.	The Clegg Impact Value compared against relative compaction was inconsistent.
Rathje et al. 2006	Field Tests 1: devices were first run on compacted field test pads on constructed of clay, lean clay and 3 structural wall backfills. • Clegg to NDG: Target • Panda DCP to NDG: Acceptance or Refusal- based on soil type. • NDG to DCP: MNDOT pass or fail criteria. Field Test 2: focused on EDG , M+D and rubber balloon method. Tested in clay, lean clay and sandy clay.	The M+DI and the SDG have promise and are based on good theoretical basis, but the manufactures need to improve the device by developing a clear-cut calibration process.	None of the evaluated test devices were found feasible to replace the NDG.
Siekmeier et al. 2009	• Granular material tested in the bottom half of 55 gallon barrel • Fine grained soil tested in 23” x 23” x 15” steel containers. • Two differing LWDs (Zorn & Dynatest) and DCP were compared on graphs of modulus VS DPI and DCP drops per 4 “. • Granular and fined grained tabular target values were developed for both the LWD and DCP. The granular target values were based on grading number and moisture content. • The target values for fine grained soil based on plastic limit and estimated optimum moisture.	• LWDs and DCPs should be implemented more widely by MnDOT. Recommended targets should be verified. • Specific LWDs should be specified and used by the contractor and MNDOT.	N/A

Table 4-6c Multiple QC/QA Compaction Test Research Projects Summary

Source	Test Preparation/Comparison Methods	Recommended Devices/Methods	Devices /Method Not Recommended or Have Complications
Mooney et al. 2008	Performed tests on structural backfill for Mechanically Stabilized Earth (MSE) walls and bridge approaches. Target values were established for the LWD, Clegg, and DCP against 95 percent proctor requirements.	The LWD and Clegg hammer were deemed suitable QA devices for structural backfill used on MSE walls & bridge end approaches. Recommended for Colorado DOT usage.	The DCP readings were found to be sensitive to moisture readings and would give false readings when penetrating geogrid or hitting reinforcement behind a MSE wall.
Brown 2007	• Tested M+DI, EDG, and NDG on gravel subbase, granular backfill and sandy borrow. • Tests for the M+DI and EDG were conducted in the foot print of NDG test conducted 180 deg. from one another. • NDG to EDG : linear regression • NDG to M+DI: linear regression	N/A	• The EDG and M+DI had considerable time in calibration, setting up and running. • Spikes for the M+DI tended to bend in coarse and very stiff soils. • The EDG calibration process involved the use of a NDG so NDG not fully replaced.
Meehan and Hertz 2011	• Comparison tests conducted using sandy silt in 5 ft. x 3 ft. x 1 ft. box. • Moisture and dry density were compared with the following comparisons with root mean square error calculated: ○ NDG-SC ○ Drive Cylinder – NDG ○ Drive Cylinder – SC.	• The NDG and drive cylinder had good comparison. • Drawback with the drive cylinder is obtaining moistures so the contractor is not held up.	• The EDG displayed higher root mean square error and relative error than the other tested density and moisture devices. • EDG readings have the chance to improve with better or alternate calibration procedures.
Meehan et al. 2012	• Comparative tests on a 200 ft. x 20 ft. embankment constructed with coarse grained fill. • Linear regression comparisons were conducted for the following comparisons: ○ NDG-GeoGauge ○ NDG-LWD 300 ○ NDG-LWD 200 ○ NDG-DCP(average) ○ NDG-DCP(weighted mean) • LWD readings were also compared against moisture readings.	N/A	• Modulus based tests had poor correlation with NDG dry density measurements • The modulus based devices did not agree with one another.

Table 4-6d Multiple QC/QA Compaction Test Research Projects Summary

Source	Test Preparation/Comparison Methods	Recommended Devices/Methods	Devices /Method Not Recommended or Have Complications
Abu-Farsakh et al. 2004	• Conducted a testing in both lab and field. • GeoGauge, LWD and DCP were examined individually in tabular form for laboratory tests. • GeoGauge, LWD and DCP were compared with number of passes for differing sections from a compactor. • GeoGauge was compared graphically to dry unit weight and moisture content for clayey silt and sandy lean clay soil • The LWD was also compared to Plate Load Test using linear regression correlations.	• Found GeoGauge, LWD and DCP were dependable devices for stiffness/modulus measurements of embankment as well as subgrades and base layers. • Recommended DCP for QA and QC on materials not acceptable to moisture content.	GeoGauge readings were affected when testing compacted lime and cement modified soil due to shrinkage cracks.
White et al. 2013	• Sixteen differing sections tested. • 15 sections contained 6“. crushed limestone subbase. • One section contained 7” geocell filled with crushed limestone. • Various subgrades/suubbases contained geotextiles, geogrids • Other subgrades were stabilized with Portland cement and fly ash. ○ LWD to FWD: linear regression ○ FWD to CMV: linear regression ○ LWD to CMV: linear regression ○ FWD to MDP: linear regression ○ LWD to MDP: linear regression	• FWD correlates to the LWD. • RICM values (CMV & DCP) provide a continuous record of stiffness values. • CMV values correlate better with LWD & FWD values	• In this program and Iowa roadway projects -the NDG displayed problems with test reproducibility, limited test frequency on roadway projects. • NDG readings are not in direct correlation to strength or stiffness measurements..

Table 4-7a Individual QC/QA Compaction Test Devices Research Projects Summary

Source	Device	Testing Methods	Findings/Conclusions
Tehrani and Meehan 2010	LWD	• Two models of Zorn LWDs • (plate diameters of 200 mm and 300mm. • 19 silty sand locations tested. Regression analysis was performed on two different LWDS. • Calculated coefficients of determination R^2on the recorded data.	• Univariate regression analysis displayed a relationship between LWD modulus and moisture content readings. • Recommended that for DOTs implementing the LWD for compaction QC/QA should use time limit for testing compaction after completion of passes.
Meehan and Hertz 2013	EDG	• Built soil models using two different soils from corresponding NDG tests. • Conducted 12 calibration tests in large proctor mold at varying densities. • Conducted assessment of the EDG using a field box. EDG was compared to test results from the NDG, SC, and the Drive Cylinder.	• Three different comparison methods root mean square error (RSME) were calculated between the EDG and compared compaction measuring devices. • Lower values of RSME indicate better correlation or prediction. ○ EDG unit weight readings are more scattered than NDG and Drive Cylinder but showed better results than the SC. ○ EDG moisture content readings tended to be more scattered than standard moisture QA/QC tests.
Ooi and Pu 2003	GeoGauge	• Conducted tests with the Humboldt GeoGauge on silt soil compacted at three different blows per lift in a 150 mm (6 in) diameter inverted molds. • Second set of tests were at varying blows and dry unit weights but with six constant water contents. • Third set of tests, samples were compacted in molds.	• Concluded that the maximum stiffness occurs dry of optimum moisture. • Also concluded there is no direct relationship between stiffness and dry density. • Stiffness increases with increasing density at low moisture. • Stiffness decreases upon wetting this is more significant for soils dry of optimum than wet of optimum.
Lenke et al 2003	GeoGauge	• Tested the GeoGauge on both and silty sand soils. • The sand was compacted in a container (28 in W X 30 in L X 24 in D). • Test in sand was to determine that GeoGauge measurements corresponded with both theoretical and practical soil mechanics. • The silty sand was compacted in a 1.56 ft^3 and in 6 in proctor molds and tested with varying compactive effort and moisture content.	• GeoGauge tests in silty sand revealed that stiffness measurements varied with moisture content of the soil. • Discovered and graphed that the optimum moisture for maximum stiffness does not coincide with the optimum moisture for maximum density. • Obtaining target values with 6 in. proctor molds were not successful due to boundary effects.
Vennapusa and White 2009	LWD	• Examined factors that effected LWD modulus readings which included the diameter of the loading plate, plate contact stress, LWD transducer type and location, plate rigidity, buffer stiffness plate rigidity and other factors. • Compared three devices that included the Zorn, Dynatest and Keros at a number of pavement construction project sites.	Determined from their literature review and test results the following abridged findings: ○ The Keros modulus readings on average were 1.75 and 2.16 times greater than the Zorn with 200 mm plate and 300 mm plate respectively. ○ LWD devices that use accelerometers are expected to have higher deflection readings than those devices which use geophones that measure deflections at the ground. • LWD modulus readings increase with decreasing plate diameters.

Table 4-7b Individual QC/QA Compaction Test Devices Research Projects Summary

Source	Device	Testing Methods	Findings/Conclusions
Maher et al. 2002	GeoGauge	• Performed an extensive literature search. Tested GeoGauge in both the lab and in the field. • Lab tests used a 55 pound drum lined with Styrofoam and a large soil bin 15 ft. in diam. and 8 ft. deep. • Soil bin tests were to determine if the GeoGauge readings would change over buried utility pipes. • The laboratory tests used 4 differing soils. • GeoGauge reading was taken every lift (3 inches) and rubber balloon density test taken every 6 inches. ○ Field test studies were ran on two constructed test embankments. ○ NDG and GeoGauge measurements were taken. ○ NDG moistures were compared with oven dried moistures.	• Convert the GeoGauge stiffness measurements to dry density. • Used several regression analysis equations from three differing sources for sandy soils. • Examined oven dried moistures content versus NDG moistures. • Compared calculated dry densities versus actual densities. • Established that the GeoGauge provided repeatable results. Found that they had to develop regression analyses for each specific soil to convert stiffness readings to dry density. • Concluded that the presences of small objects like small diameter utility pipes have little effect on stiffness readings.
Erchul and Meade 1990	Clegg	• Study to refine the use of the Clegg Hammer to confirm compaction in trench backfills. • Conducted field tests on four sites with Clegg Hammer and NDG. • Performed Standard Proctor tests in conjunction with the field tests.	• Developed acceptance charts by with two acceptance zones using empirical methods. • First zone -90 percent max. density. • Second zone- 95 percent max density. • Zones placed on chart of depth of penetration vs impact values. • Found from eight tests conducted the risk of accepting a failed test was about 10 percent. • Recommended procedure does not evaluate density directly and cannot be used to enforce specifications based on density.
Erchul 1999	Clegg	• Identified a relationship between the impact value and depth of penetration from fourth blow of the hammer. • Determined 90 and 95 percent proctor zone on graph of depth of penetration versus impact values.	The 90 percent proctor zone acceptance criterion was a better comparison to the NDG (97 percent) agreement in cohesive soils.
Vanden Berge 2003	Clegg	• Gathered data in sand using standard method of using only the maximum hammer acceleration to predict compaction. • Investigated compaction prediction from secondary acceleration peaks. • The net velocity change by integration of output signals recorded by a modified hammer.	• The Clegg Hammer with additional analytical methods and knowledge of hammer dynamics will allow the device to be used for compaction QC and QA on construction sites. • Concluded that the impact test can be conducted independent of moisture content to estimate soil density.

Conclusions and Recommendations

The literature review shows that researchers have investigated compaction testing devices and methods as an aggregate of differing categories of compaction testing devices or have focused on a limited number of devices to investigate. Comparisons of device measurement vary from direct comparison to calculating linear or second order regression equations. To implement the use of stiffness and modulus devices for compaction testing, researchers are using the process of setting targets or pass-fail criteria to obtain equivalent compaction levels determined by the NDG or other standard optimum moisture- maximum density method.

The majority of the researchers working with stiffness and modulus devices have discovered in their comparison with the proctor test and other density-moisture measurements, that maximum modulus or stiffness lags behind max density at moisture dry of optimum. This occurrence demonstrates to modulus/stiffness devices users the necessity of obtaining moisture readings to prevent incidents of false pass readings.

The future direction of research for DOTs will be driven by the use of mechanistic design for pavements championed by American Association of State Highway and Transportation Officials as well the emerging compaction means and methods such as “intelligent compaction equipment,” both of which rely on the use of modulus and stiffness of subgrade and base in design and quality control during construction. These new technologies are directing compaction research towards modulus and stiffness devices such as the LWD and the DCP. Implementation and routine use of these devices has already occurred for the Minnesota and Indiana DOTs (Cho et al. 2011; Siddiki et al. 2015).

The analysis of the literature reported in this paper found that previous research was highly focused on a limited comparison of compaction testing methods. Much of the work glosses over or ignores the fundamental geotechnical principles and testing approaches that will be encountered in the field, devoting itself to the highly controlled environment of the laboratory. This leads to the inference that an inability to conduct testing over the broad set of conditions that faces practicing geotechnical engineers led these research teams to limit their investigations to the lab.. As a result, it is concluded that research that includes a broad set of testing methods and equipment on the same set of soil conditions and that includes both density and stiffness test methods is required to furnish the technical, financial, and practical ergonomic aspects of each alternative before an authoritative determination can be made as to whether the advantages of the NDG outweigh its life cycle disadvantages.

CHAPTER 5. MANAGING GEOTECHNICALRISK ON US DESIGN-BUILD TRANSPORT PROJECTS

McLain, K., Gransberg, D., and Loulakis M. (2014). "Managing Geotechnical Risk on US Design-Build Transport Projects." *Australasian Journal of Construction Economics and Building,*Vol. 14(1), pp. 1-19.

This Chapter examines effective practices in use by State Departments of Transportation to manage and mitigate geotechnical risk on design-build (DB) projects. The purpose of this paper is to make the connection between compaction testing methods and the challenges that are found in a DB project where the contractor will typically perform not only the compaction testing but also develop the geotechnical design, specifications, and standards. The chapter discusses the elements of DB contract administration that are linked to the quality assurance and quality control practices that have been incorporated into MoDOT's Quality Management System (QMS). QMS has been implemented on all MoDOT projects regardless of project delivery method. The details of that policy are that are covered in Chapters 6 and 7. The shift from agency control of compaction testing to contractor control will have a huge influence on the choices for possible replacement of the nuclear density gauge.

Abstract

Awarding design-build (DB) contracts before a complete subsurface investigation is completed,

makes mitigating the risk of differing site conditions difficult, if not impossible. The purpose of the study was to identify effective practices for managing geotechnical risk in DB projects, and it reports the results of a survey that included responses from 42 of 50 US state departments of transportation and a content analysis of DB requests for proposals from 26 states to gauge the client's perspective, as well as 11 structured interviews with DB contractors to obtain the perspective from the other side of the DB contract. A suite of DB geotechnical risk manage tools is presented based on the results of the analysis. Effective practices were found in three areas: enhancing communications on geotechnical issues before final proposals are submitted; the use of project-specific differing site conditions clauses; and expediting geotechnical design reviews after award. The major finding is that contract verbiage alone is not sufficient to transfer the risk of changed site conditions. The agency must actively communicate all the geotechnical information on hand at the time of the DB procurement and develop a contract strategy that reduces/retires the risk of geotechnical uncertainty as expeditiously as possible after award.

Background

"Geotechnical engineering is fundamentally about managing risk" (Ho et al., 2000). Managing the risk of geotechnical site conditions is never simple; however, when a DB (also termed "design and build" in many countries) contract is awarded before a complete subsurface investigation is complete; it becomes even more difficult (Perkins, 2009). In the US, recent government pressure to expedite the delivery of highway construction projects to address the current infrastructure deterioration crisis has created a procurement environment where DB projects are being awarded as soon as environmental consents can be obtained without regard to the potential impact of failing to quantify and mitigate geotechnical risk on post-award project cost and schedule (Mendez, 2010; Hatem, 2011; Federal, 2013).

The US Federal Highway Administration's (FHWA) Special Experimental Projects No. 14 –Alternative Contracting (SEP-14) was introduced in 1990 and by 2009 had authorized over 400 DB highway projects (Federal, 2006). A decade later, the FHWA announced its "Every Day Counts" (EDC) initiative to address the rapid renewal of the nation's rapidly deteriorating infrastructure. The program is designed to accelerate the implementation of innovative practices that are immediately available as described by the current FHWA Administrator, Victor Mendez. "Our society and our industry face an unprecedented list of [infrastructure] challenges. Because of our economy, we need to work more efficiently... But it's not enough to simply address those challenges. *We need to do it with a new sense of urgency.* It's that quality— *urgency—that* I've tried to capture in our initiative, Every Day Counts." (Mendez, 2010; *italics added).*

Replacing traditional design-bid-build (DBB) with DB project delivery is one of the tools being specifically encouraged by the EDC program. A report to the US Congress on the effectiveness of DB in highway construction reviewed every project authorized under SEP-14 and found that on average DB "reduced the overall duration of their projects by 14 per cent, reduced the total cost of the project by 3 per cent, and maintained the same level of quality as compared to DBB"(Federal, 2006). The same report also concluded that clients select DB primarily as a mea ns to accelerate a project's schedule, validating a trend reported nearly a decade earlier by Songer and Molenaar (1996). Higbee (2004) found that the major hurdle to achieving an accelerated schedule in DB is the client's approval to release the design for construction. The geotechnical investigation and subsequent foundation design is often the first design package that must be released. Since geotechnical uncertainty is often high at the time of DB contract award, the design-builder's geotechnical designers are under pressure to complete their work as quickly as practical to allow foundation and other subsurface construction to begin. The 2006 Report to Congress found that less than 3 per cent of total highway projects were delivered using DB (Federal, 2006) and, because DB transport projects could only be delivered after obtaining FHWA permission via the SEP-14 application process, the overall impact of managing geotechnical risk has been low on a nation-wide, programmatic basis. DB contracting could not be described as a "routine" method to deliver construction projects. That changed in 2007 when the FHWA DB contracting "final rule" was ratified by the US Congress (Federal, 2007), making it fully eligible for delivering federally-funded projects. The impact was profound with the number of states authorized by their own state legislatures to use DB growing from 33, at the time of the 2006 Report to Congress, to 47 today (Design 2013). The growth in DB usage was further spurred in 2012 by the passage of Public Law 112-141, the Moving Ahead for

Progress in the 21st Century Act (MAP-21), which reduced the state funding share of federal-aid highway projects delivered using DB from 10 per cent to 5 per cent (Federal 2013). Thus, a substantial financial incentive has been provided to state transport agencies that have previously chosen not to implement DB. The political pressure to implement DB is further increased by the FHWA EDC program objective of increasing the number of projects delivered using DB by 50% by 2014 (Mendez, 2010).

From the public agency geotechnical engineer’s perspective, the net result will be to increase the number of DB projects awarded before subsurface investigations are complete. From the design-builder’s perspective, increased geotechnical risk will translate into higher contingencies included in proposed prices to mitigate those risks from the contractor’s perspective (Christensen and Meeker, 2002). Higher prices translate to higher potential that the agencies will ultimately not have sufficient budget to actually award the DB project once the proposals have been opened. Therefore, successfully managing the geotechnical risk during the procurement phase of a DB project becomes essential to being able to award the given project within its budget (Clark and Borst, 2002). Accordingly, the researchers address the following questions:

- What measures can a public transport agency take to manage geotechnical risk during the DB procurement process?
- What can be done after the award of the DB contract to mitigate and retire geotechnical risk in an expeditious manner?

The answers to these questions found in the study are synthesized into a set of geotechnical risk management tools that can be used by public clients to better allocate geotechnical risk among DB project stakeholders. Additionally, the authors provide construction procurement researchers a platform from which to support future research on optimizing DB procurement risk with the need to accelerate project delivery of transport projects.

The current emphasis on accelerated project delivery in the US creates an environment where public engineers may be forced to focus on expediting the procurement process rather than fully developing the project's geotechnical requirements. This includes evaluating how much of the geotechnical investigation should be done by the design-builder after contract award. The geotechnical investigation decision has a number of ramifications, including the level of liability for the underground conditions that can be transferred along with the geotechnical investigation and design responsibility for the foundation/subsurface design.

Risk-based Geotechnical Design

The FHWA introduced risk-based geotechnical design in 1987 when it published the *Geotechnical Risk Analysis User's Guide* developed by G.B. Baecher. This document moved geotechnical design on federally-funded highway projects away from a set of "conservative factors of safety" and toward modeling uncertainty on a project-specific basis via a statistically determined reliability index. This shift was necessary because a "fixed factor of safety implies a different likelihood of failure" in each project and creates a situation where "the overall factor of safety in a design is unknown" (Baecher, 1987). Baecher's work assumed that the project would be delivered using DBB project delivery and the data used as input for the risk-based design would spring from a thorough program of geotechnical investigation, testing, and analysis. An Australian study of the implications of inadequate site investigations agreed with Baecher's assertion regarding communicating geotechnical uncertainty and recommended that geotechnical uncertainty be expressed using statistical measures such as confidence limits. Jaksa (2002) argues that doing so permits "any other engineer utilizing these values, as well as the client, to appreciate the uncertainty associated with the parameters and, hence, appropriately account for them in the design process." The ability to understand the amount of as-designed geotechnical risk is one key to effectively managing that risk after award. A critical discourse on the subject of quantifying geotechnical risk in the design asserted that "designers sometimes wishfully classify those factors which they cannot confidently characterize as being of minor importance, or hope that such imponderables would be compensated by conservatism built in the system elsewhere" (Ho et al., 2000). This notion also agrees Baecher's findings regarding the weakness of using fixed factors of safety. Ho et al. go on to advocate the use of quantitative risk assessment in conjunction with traditional deterministic methods to better communicate geotechnical risk throughout the project's design and construction process. Van Straveren (2000) builds on the quantitative risk analysis theme and extends the argument to actively managing geotechnical risk across a project's entire life cycle with a focus on articulating risk during procurement and cited a "1:10 cost-benefit ratio…as a result of better contracting practices by improved risk allocation."

Oberguggenberger and Fellin (2002) take an opposing view of the value of statistically-based geotechnical risk analysis. Their opinion is founded in theoretical mathematics rather than geotechnical engineering design. While they recognize the role of failure probabilities and safety factors in comparative analysis of design options, they state that "these numerical values do not make quantitative assertions about reality… the failure probability cannot be interpreted as a frequency of failure." While their proof of the superiority of fuzzy sets over probability-based risk analysis appears to be eloquent, it is also arcane requiring an understanding of mathematics at a level beyond the academic preparation of most practicing geotechnical engineers, an argument made by Ho et al. (2000). Those authors contended that resistance to the increased knowledge required to move from deterministic to probabilistic geotechnical risk analysis, much less fuzzy set theory, "is by no means easy to resolve... [requiring] appropriate grounding of the basic concepts [of statistical risk analysis] in university education and focused professional training…" Thus, while Oberguggenberger and Fellin's approach may produce a better means to model geotechnical uncertainty, it must be regarded as impractical due to the need to educate a significant number of engineers, an idea supported by Baynes (2010).

Baynes plays heavily on the human factor and finds that total geotechnical risk is a combination of the technical conditions and the competence of the project staff. In fact, "the project staff may actually be the largest source [of risk]." Baynes emphasizes the need to educate and train project staff to "manage and mitigate the geotechnical risks, rather than generate them." He, like van Straveren (2000), also stresses the need to manage geotechnical risk throughout the project's life cycle, specifically identifying the procurement phase as a point where "inadequate understanding of the importance of ground conditions results in poor acquisition… [that] leads to claims based on contractually unforeseen ground conditions" (Baynes, 2010). To summarize the literature, geotechnical risk management is more than the use of sophisticated statistical models to quantify the risk in probabilistic terms. It must be continually evaluated as an integral part of the project development decision-making process (Baecher, 1987; Ho et al., 2000; van Staveren, 2000; Baynes, 2010). All of the above literature was written in the DBB context where the procurement was based on a completed geotechnical design that was in turn based on subsurface investigation. The issue of subsurface risk becomes weightier when the geotechnical investigation moves from being a condition precedent to DBB construction contract award to a deliverable required after award of a DB contract.

Design-build Contracting

DB project delivery has proven itself to be one method to accelerate the construction, reconstruction, and rehabilitation of aging, structurally deficient infrastructure (Federal, 2006). DB also allows the public transport agency to shift some of the responsibility for completing the geotechnical investigations necessary to support the geotechnical design to the design-builder after the award of the DB contract. This creates a different risk profile than when the project client has full responsibility for design (and hence geotechnical investigations) in a traditional DBB project (Loulakis et al., 1995).

The FHWA mandates the use of a Differing Site Conditions (DSC) clause for DBB project on federal aid highway projects, unless the use of such a clause is contrary to state law (Loulakis etal., 1995). The DSC clause provides broad relief to a contractor for physical conditions that materially differ from what is anticipated by the contract. FHWA does not, however, have the same mandate for DB projects. Instead, FHWA encourages state DOTs to use these clauses when appropriate for the risk and responsibilities that are shared with the design-builder.On DBB projects, the risk of differing site conditions is almost always the responsibility of the client (Tufenkjian, 2007), based on the contract's DSC clause and prevailing case law (Higbee,2004). Diekmann, et al. (1987) confirmed this point specifically for infrastructure projects. On DB projects, the risk of differing site conditions is not as clear (Clark and Borst, 2002). The DB contract can be awarded before a full geotechnical site investigation is made by either the client or the winning design-builder (Smith, 2008). This leads to a question of how to identify an appropriate reference point for implementing the DSC clause if one is included in the contract (Hatem, 2011). There is also a policy question for the agency as to how much information it should furnish about the geotechnical site conditions (Blanchard, 2007; Dwyre et al., 2010). The more information that is provided, the more likely it is that the design-builder can submit a competitive price proposal since it is able to reduce the contingencies contained in the price proposal for geotechnical uncertainty (Christensen and Meeker, 2002). Additionally, this will enable the agency to have a better sense of its program and expected costs. However, because the DB delivery method has proven to be an effective means of compressing project delivery periods to their shortest states (FHWA, 2006), there is frequently an incentive for the agency to start the procurement process before a robust geotechnical program has been performed (Higbee, 2004; Kim et al., 2009).

Pre-award Geotechnical Risk Distribution

Given the above, an agency should first address whether or not a given project is a good candidate for DB project delivery in the context of the geotechnical conditions' impact on the preliminary design, price, and time. Table 5-1 is a synopsis of the risk profiles for DBB and DB found in Koch et al. (2010) and adapted for geotechnical risks. One can see that the major change in the risk profile is due to the shift in design responsibility to the design-builder. The client's new DB risks result in many cases from failing to relinquish the design responsibility to the design-builder. The client's DB scope risk for geotechnical design review comments and/or directives is an example of this. Direct and tacit approval of constructive changes to the geotechnical design during construction is another example.

Table 5-1 DBB versus DB risk profiles

	Contractor/Design-Builder	Client
	Geotechnical Scope Risk	
DBB	• Warranties and Guarantees • Latent Defects - Workmanship • Competent Geotechnical Construction Personnel Available	• Design Error and Omissions • Latent Defects - Design • Direct & Tacit Approval of Constructive Changes to Design
DB	• Design Errors & Omissions • Warranties & Guarantees • Latent Defects - Design & • Workmanship • Competent Geotechnical Design Personnel Available	• Clear Geotechnical Scope Definition • Direct & Tacit Approval of Constructive Changes to • Geotechnical Design • Geotechnical Design Review Comments & Directives • Technical Review Capability
	Geotechnical Cost Risk	
DBB	• Rework • Subcontractor Default • Market Fluctuation after Award	• Redesign and Resultant Rework • Construction Contract Amount • Market Fluctuation During Design - Material & Labor
DB	• Redesign & Rework • Subcontractor Default • Market Fluctuation During Design – Material & Labor	• Design-Build Contract Amount • Prompt Payment • Design-Builder Default
	Geotechnical Schedule Risk	
DBB	• Contract Completion Date • Liquidated Damages	• Timely Design Completion • Client Furnished Property Delivery
DB	• Delivery on Approved Schedule • Fast-Track Geotechnical Rework • Liquidated Damages	• Unrealistic Schedule • Timely Geotechnical Design Approvals Client Furnished Property Delivery

Hatem (2011) maintains that DB geotechnical uncertainty "is always high until the post-award site investigation and geotechnical design report can be completed." The geotechnical/site engineering is the first major design package that must be released to get construction started (Higbee 2004) and competing design-builders must base their schedule estimates on expeditiously completing this key design task (Centennial, 2004). Given the criticality of the geotechnical investigation and design to DB project success, the inclusion of proposal evaluation criteria specifically addressing the competing proposers' approach to project geotechnical issues, ensures that the competing design-builders will focus on those aspects of the project in the proposal because if they do not, their proposal will be found to be nonresponsive (Higbee, 2004).

The Washington State DOT (WSDOT) *Guidebook for Design-Build Highway Project Development* (2004) maintains that the agency is "responsible for establishing the scope, project definition, design criteria, performance measurements, and existing conditions of the site (initial geotechnical investigation, subsurface conditions)." The responsibilities listed in this passage form a foundation for determining what specific data should be included in the DB RFP. This agency agrees with Hung et al. (2009) and goes on to elaborate that "it is necessary for WSDOT to establish a baseline for design-builders to develop their technical and price proposals" and that "preliminary geotechnical investigations will be conducted by WSDOT with data provided to Proposers."

WSDOT is consciously creating an environment of open communication regarding geotechnical uncertainty and allocating differing site conditions risk. In fact, the document states:

"Ultimately, WSDOT will own responsibility for Changed and Differing Site Conditions." Since the geotechnical portion of a DB contract is the combination of information contained in the RFP and the winning proposal (Koch et al., 2010), the amount of geotechnical information contained in the RFP effectively creates the baseline from which a DSC is applied.

Design-Build Contract Pricing

The predominant way that DB is procured in the public sector requires that the design-builder commit to a firm fixed price before the project's geotechnical design is complete (Mahdi and Alreshaid, 2005). Thus, the risk of cost overruns for unforeseen geotechnical site conditions is increased since the geotechnical investigations necessary for each project will likely be completed after contract award, during the design process. Some public clients have the view that using DB shifts the full risk of differing site conditions to the contractor (Christensen and Meeker, 2002). The basic flaw in this approach is that contractors cannot accurately value the risk of geotechnical uncertainty before a thorough site investigation is completed. If they are forced to price the risk, they will include contingencies that may either price themselves out of the procurement or, if they do win the contract, be insufficient for addressing actual conditions, further intensifying the bias to inflate the contingency. Many sophisticated contractors will simply refuse to compete for a contract where they have unlimited risk of differing site conditions (Centennial, 2004; Loulakis et al., 1995), and 90% of the design-builders interviewed stated that the amount of detail available in the RFP had an impact on project quality. Following the recommendations made by Hung et al. (2009) effectively limits both the contractor's and the client's risk. The client only pays for the actual costs incurred if and when these conditions are actually encountered, as opposed to the unliquidated contingency for a problem that may never emerge.

All of this creates potential risks to both parties that are not present in a DBB delivery process (Washington, 2004). In a technical sense there is a wide spectrum of potential geotechnical risks, but in DB procurement, there is only one: *actual conditions will materially differ from those upon which the project's price was predicated.*

Methodology

The researchers used qualitative comparative analysis as the overarching research method to leverage its ability "to blend the in-depth knowledge obtained from small-N studies of cases with the inferential power of statistical large-N studies…[and] determine causal relationships between 'causal conditions' (similar to independent variables) and 'outcome conditions' (similar to dependent variables)" (Jordan et al., 2011). Specifically, the researchers conducting this study needed to compare case study contractor interview output with the output from a survey and the content analysis of DB procurement documents. Since the study topic was inherently a variable mixture of technical geotechnical engineering and the legalistic construction procurement process, this relatively new approach was selected to lend rigor to the research protocol. Three research instruments formed the study's data collection plan.

First, a review of the literature on DB contracting with a focus on geotechnical risk was completed. Both US and international documents were searched. The literature was then used to develop the content of an on-line survey of US DOTs. The survey questionnaire was designed using the principles prescribed by Oppenheim (1992) for survey questionnaire design. The researcher's underlying hypothesis for the survey was:

Geotechnical risk varies inversely with the amount of site-specific geotechnical information that is provided to competing design-builders during procurement.

DOT Survey

Since the research was sponsored by the American Association of State Highway and Transportation Officials (AASHTO) and funded by the National Academies' National Cooperative Highway Research Program (NCHRP), the survey was issued to the members of the AASHTO Subcommittees on Construction and Design in each of the 50 US state DOTs. The subcommittee members were asked to forward the survey to the person best-qualified to respond from an overall departmental basis. Responses were received from 42 DOTs yielding an overall response rate of 84%. Table 5-2 shows the locations of the respondents and their positions at the time of the questionnaire. The table shows that the survey received responses from a cross-section of senior engineers with design-build experience. Design professionals made up roughly 60% of the response for those DOTs that use DB. The rest were either construction field personnel or DB project managers. Hence, the collective response from the sample covers the entirety of DB project delivery from planning through construction completion and administrative/legal close-out.

Procurement Document Content Analysis

The second instrument was a content analysis performed on DB procurement documents from 26 states in addition to DB policy documents/guidelines from 12 state DOTs and 5 federal agencies. This type of analysis can be used to develop "valid inferences from a message, written or visual, using a set of procedures" (Neuendorf 2002). The primary approach is to develop a set of standard categories into which words that appear in the text of a written document, in this case a DB procurement or policy document, can be placed, and then the method utilizes the frequency of their appearance as a means to infer the content of the document (Weber 1985).

Design-Build Industry Interviews

The final research instrument consisted of structured interviews with design-builders to validate potential conclusions and effective practices found in study. The Government Accountability Office method states that structured interviews can be used where "information must be obtained from program participants or members of a comparison group… or when essentially the same information must be obtained from numerous people for a multiple case-study evaluation" (GAO 1991). Both these conditions apply to this study; therefore, the tool is appropriate for the research.

Since geotechnical risk is often quantified in terms of cost, the above hypothesis was modified for the interviews to read:

> *The amount of the contingency for geotechnical risk varies inversely with the amount of site-specific geotechnical information that is provided to competing design-builders during procurement.*

Table 5-2 Survey Respondent Demographics

DOT Respondents with Design-build Experience				DOT Respondents without Design-build Experience	
State	**Position**	**State**	**Position**	**State**	**Position**
Alaska	Construction engineer	**Nevada**	Geotechnical engineer	**Alabama**	Construction engineer
Arkansas	Design project manager	**New Jersey**	Construction engineer	**Connecticut**	Geotechnical engineer
California	Design project manager	**New Mexico**	Geotechnical engineer	**Illinois**	Design project manager
Colorado	Design-build project manager	**New Hampshire**	Design project manager	**Iowa**	Design project manager
Florida	Construction engineer	**North Carolina**	Geotechnical engineer	**Kansas**	Construction engineer
Idaho	Construction engineer	**North Dakota**	Design project manager	**Nebraska**	Design project manager
Indiana	Geotechnical engineer	**Ohio**	Design project manager	**New York**	Design project manager
Kentucky	Construction engineer	**Oregon**	Geotechnical engineer	**Oklahoma**	Construction engineer
Louisiana	Geotechnical engineer	**South Carolina**	Design project manager	**Wyoming**	Design project manager
Maine	Geotechnical engineer	**South Dakota**	Geotechnical engineer		
Maryland	Materials engineer	**Tennessee**	Construction engineer		
Massachusetts	Design-build project manager	**Texas**	Materials engineer		
Michigan	Design-build project manager	**Utah**	Geotechnical engineer		
Minnesota	Geotechnical engineer	**Vermont**	Construction engineer		
Mississippi	Construction engineer	**Virginia**	Design-build project manager		
Missouri	Construction engineer	**Washington**	Materials engineer		
Montana	Design-build project manager				

Since it is impossible to know exactly how much contingency is being allocated to the perceived geotechnical risk, the researchers asked the entities that are at risk in a DB project for both the geotechnical design and the final project's construction to describe the impact on the proposed project contingency of the amount of geotechnical information that is available at the time a firm, fixed price must be submitted. The interviews were treated in the same manner as summary case studies because each set of interviewees had a unique perspective that was formed by the market in which it competed. Two primary criteria were established for selecting a design-builder.

First, the specific firm had to be one that had completed a DB project in at least one of the states that responded to the questionnaire. Secondly, it needed to have competed for at least one of the DB projects represented in the procurement document content analysis. Two secondary criteria were established for further filtering the pool of potential DB firms that qualified by the primary criteria, and those were to present a reasonably broad distribution of firms geographically and to have a sample that included small as well as large DB firms. Ultimately, interviews of 11 design-builders whose markets encompass over 30 states were conducted. They ranged in size from a regional bridge contractor that only worked in Utah to three national firms.

The following discussion reports the effective tools used by US state departments of transportation (DOT) to deal with the geotechnical conundrum described above and provides information on commonly used practices for managing geotechnical risks in DB projects.

Analysis and Results

Analysis of the Survey and Content Analyses

In traditional DBB construction projects, the design and construction are performed under two separate contracts. In many cases, the agency performs the design itself and then advertises for construction contractors to submit tender offers, termed "competitive bids" in US construction contracting jargon, on the construction documents. In DB, one entity takes on the responsibility for both design and construction. As a result, the agency on a DB project has less direct control over the day-to-day details of design development, as design is being done in conjunction with an awarded construction contract, which has fixed obligations to meet a schedule and a price. The analysis of the survey and content analysis strove to keep this fundamental difference in mind and seek effective practices that reconcile the design-builder's need to design to a fixed budget and contractual schedule with the agency's need to diligently oversee the geotechnical design process.

Table 5-3 contains the results of the content analysis and the survey responses regarding the amount of geotechnical information contained in typical DB RFPs. The FHWA Report to Congress on DB effectiveness (Federal, 2006) differentiated between the survey responses of those agencies that had completed 5 or more DB projects and those with less experience. Using the FHWA study's approach, this study's results were split by the number of DB projects the agency had completed to differentiate between agencies that were relatively new to DB and those with multi-project experience. The table shows that experienced agencies furnish more information than inexperienced agencies. The literature exposed one possible reason for the difference. There is a school of thought that maintains that furnishing specific geotechnical data in a DB project assigns all the risk of DSC to the client, and as such guarantees a DSC claim (Loulakis and Shean, 1996). Table 5-2 would seem to support this notion since the inexperienced agencies consistently include less information than the more experienced agencies. Taking the WSDOT (2006) policy discussed in the literature review above with Table 5-2 and remembering that the design-builder will probably complete the geotechnical investigations as part of the design process leads to the conclusion that furnishing as much geotechnical information *as is available at the time the project is advertised* effectively mitigates the risk by providing a clear definition of the site conditions at the time competitive proposals were submitted rather than hoping to avoid a claim by not furnishing any information.

Table 5-3 RFP Content Analysis and DOT Survey Results Regarding DB RFP Geotechnical Content

Geotechnical Information Included in DB RFP (in ascending level of detail)	Percentage of the Total of All RFP Observations		Percentage of the Total of All Survey Responses	
	RFP Content Analysis		**DOT Survey**	
	DOTs with less than 5 DB projects	**DOTs with 5 or more DB projects**	**DOTs with less than 5 DB rojects**	**DOTs with 5 or more DB projects**
Reconnaissance Report	3%	8%	0%	7%
Geotechnical Data Report	11%	26%	7%	22%
Geotechnical Summary Report	8%	13%	4%	11%
Preliminary Geotechnical Design Report	3%	21%	9%	11%
Geotechnical Design Report	5%	0%	4%	11%
Geotechnical Baseline Report	0%	3%	2%	11%

Managing Pre-award Geotechnical Uncertainty

Uncertainty, by definition, is a lack of information. Geotechnical uncertainty is reduced as site investigations, test reports, and geotechnical engineering is completed. In DB, geotechnical uncertainty is high during the procurement phase, and the client's primary tool to mitigate risk is through selecting a competent design-builder with the requisite experience to complete the design and construction.

Evaluation criteria are typically found in both the RFQ and the RFP, and can be expressed as standards for the qualifications of key geotechnical personnel, past experience on projects with similar geotechnical issues, and technical criteria for the proposed geotechnical design and construction approach. In the solicitation document content analysis, 37 of 46 of the project documents had some form of evaluation criteria for geotechnical factors explicitly listed in the document. Of those 37 projects, over two-thirds evaluated the qualifications of the project's geotechnical personnel. Next, 62% evaluated the design-build firm's past experience designing and building projects with similar geotechnical requirements. Slightly over one-third included geotechnical evaluation criteria in the technical and/or price evaluation plan. In the survey, 94% of experienced and 53% of inexperienced respondents evaluated the qualifications of the design-builder's project geotechnical personnel. Past geotechnical experience was rated at 65% and 33% respectively. In 53% of the experienced DOT responses local experience was also rated with only 20% of the inexperienced DOTs asking for that information. This data leads to the conclusion that evaluating the geotechnical-specific qualifications, experience, and technical approach is an effective means to manage pre-award geotechnical risk by requiring well qualified personnel, firms with a record of successfully completing DB projects with geotechnical issues, and an understanding of the design-builder's approach to solving geotechnical issues prior to DB contract award.

Allowing alternative technical concepts (ATC) to be proposed is a third method for dealing with pre-award geotechnical risk. ATCs furnish a means to "seek innovation from the private sector to reduce project costs and add technical enhancements" (Papernik and Farkas, 2011) without giving up control of the design process. Figure 1 is a compilation of the results from each research instrument. The fact that 71% of the experienced agencies included geotechnical ATCs in their DB projects testifies to the effectiveness of this particular practice. An Australian study found that the "road industry had the greatest propensity to invest in [research]… [and was] able to expertly judge the value of innovation ideas proposed by the industry" (Manley and McFallan, 2006), which validates the US observations in Figure 1. ATC procedures typically include the use of confidential "one-on-one" meetings with each competitor where, in addition to offering ATCs, the DB team can also seek clarifications of RFP content. These meeting are called "proprietary meetings" by some agencies that permit competing design-builders to clarify RFP intent and ask questions that might lead to the submission of an ATC. The overall effect of meetings is to reduce the uncertainty with regard to interpreting geotechnical evaluation criteria and to permit the design-builders to offer solutions to geotechnical design problems with which they are more confident. In theory, this process should lead to reduced contingencies in the price (Christensen and Meeker, 2002).

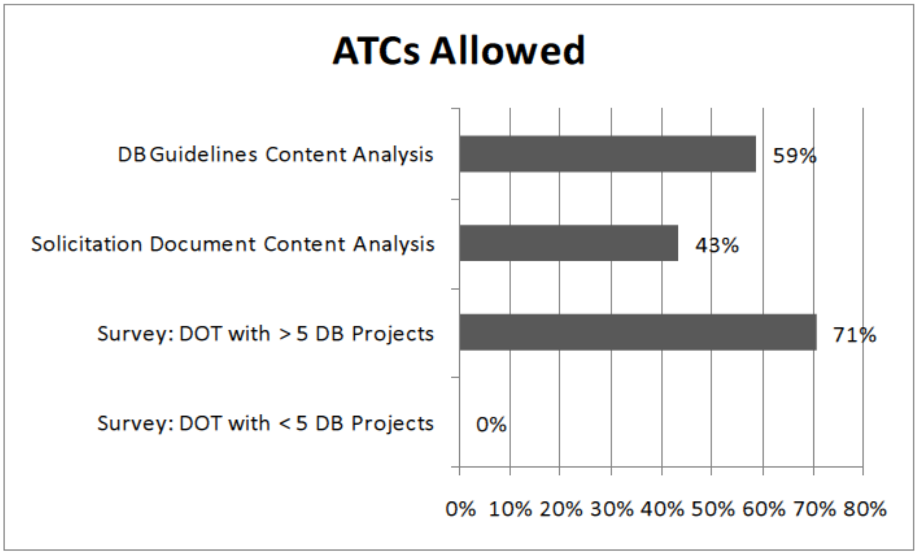


Figure 5-1 Research instrument output regarding alternative technical concept use

An example of just how valuable the confidential ATC process can be was found in a DB project in Minnesota. This particular project involved replacing a bridge over the Mississippi River atHastings. The foundation on the north side of the river rested on extremely poor alluvial soils that resulted in the need to jack the existing bridge up nearly 46 centimeters over its 30-year service life (Molenaar et al. 2012). The original engineer's estimate was about US $220.0 million. The DB RFP for the project included a "performance criterion of less than 2 inches [5 centimeters] of total settlement complete within three months of embankment construction" (Minnesota, 2010). The winning DB contractor proposed a confidential ATC to found the north approach on a "column-supported fill," a technical approach that had not yet been tried in North America (Molenaar et al. 2012). The ATC also offered to furnish and install instrumentation to monitor actual settlement over time as well as a three-year, instead of a 3-month, warranty against differential settlement. The contract was awarded at approximately US$130.0 million, and roughly US$80.0 million could be ascribed to the reduction in geotechnical risk by the approved ATC (Molenaar et al. 2012). While this is an extreme example of using confidential one-on-one meetings with competing DB contractors to mitigate pre-award geotechnical risk, it amply demonstrates that furnishing a mechanism where new ideas can be considered and clarifications to RFP requirements can be sought has the potential to accrue real cost and time savings to the public client.

Mitigating Post-Award Geotechnical Risk

The fact that US public transport agencies typically select DB to accelerate project delivery (Federal, 2006) limits the amount of pre-award geotechnical investigation an agency can do and, hence, makes post-award agency design approval a major hurdle to starting construction (Christensen and Meeker, 2002). Therefore, the geotechnical design package must be completed as expeditiously as possible (Koch et al., 2010), permitting the agency to reduce the impact of geotechnical risk as expeditiously as possible after award (Kim et al., 2009). Hence, the geotechnical design review process can act as either a barrier to releasing geotechnical design packages for construction or a conduit that facilitates the early discovery and resolution of significant geotechnical design issues.

The literature review found that the number of required design reviews by the clients varies across the US. However, NCHRP Synthesis 376 (Gransberg et al., 2008) identified three main approaches, and its DB RFP content analysis showed the percentages of use:

- No formal review prior to final (release-for-construction) design review (15%),
- One review prior to the final design review (56%),
- Multiple reviews prior to the final design review (29%).

The fact that 71% of those projects had one or no intermediate design reviews prior to the final review is noteworthy. In those cases, the agencies still provided oversight and informal comments, but made a concerted effort to not delay the design-builder's progress by imposing its design preferences for the project via multiple review and comment processes.

In many of the documents reviewed in the content analysis, the design-builder is directed to request informal reviews that allow the client to provide more frequent input to ensure that the final design will meet the contract requirements. These reviews are often called "over-the shoulder" or "oversight" reviews to indicate that the design process will not stop to wait for comments from the informal review process. The primary issue when using this process is for the agency to demonstrate that it has discharged its statutory responsibility of "due diligence."

> The Arizona DOT follows a procedure described as follows:
> "Over-the-shoulder-reviews are performed while the design is being developed. They are proactive in nature, informal, interactive, and intended to catch omissions and oversights that may lead to a major redesign of the work" (Arizona, 2001).

Arizona also uses a design review procedure that is uniquely well-suited to geotechnical design deliverables. It is called the "early construction review" and is reserved for design product that will be released for construction before the design is 100% complete. "The intent is to ensure that enough detail has been provided in the plans to allow construction to begin and that ADOT's minimum design standards are maintained" (Arizona, 2001). This process reinforces the due diligence requirements and allows the agency to obtain the necessary level of comfort with the design quality of early geotechnical features of work scheduled in support of achieving an aggressive project delivery period. Table 5-4 contains the result of the content analysis for this topic. It shows that experienced agencies are comfortable with the use of a non-traditional design process to supplement the final review.

Table 5-4 Content analysis design review output

Design Review Types	DOT < 5 DB Projects	DOT ≥ 5 DB Projects
Single or multiple design reviews before final	5	9
Over-the-shoulder	1	17
Optional early design reviews	3	12

The design-builder interviewees were asked to rate the impact of a number of components to the DB design process on the quality of the final constructed geotechnical features. The majority(67%) felt that the use of geotechnical performance criteria/specifications had a major impact and over half cited being given detailed design criteria also promoted design and construction quality. A majority cited multiple design reviews (7 of 11), sequential design reviews by different agency design personnel (8 of 11), and the agency personnel's willingness to accept over-the shoulder design reviews (9 of 11) as challenges to timely completion that could potentially negatively impact quality on all DB projects.

The required use of agency-mandated geotechnical specifications and design details on DB projects reduces the agency's need to be involved during the actual design process. This then permits the expeditious review of geotechnical engineering products and facilitates the use of design oversight practices such as the over-the-shoulder review. The literature (Higbee, 2004; Christensen and Meeker, 2002; Papernik and Farkas, 2011) and the agency DB guidelines (Washington, 2004; Department, 2010; Arkansas, 2006) promote the concept that prescriptive design requirements in the DB process limits the ability of the design-builder to innovate.

However, obtaining innovative design solutions requires the agency to spend the time necessary to satisfy its statutory due diligence requirements, which could potentially create schedule delay nullifying the benefits gained from the innovative design (Koch et al. 2010).

Impact of Geotechnical Risk Management Practices on Project Quality

All the above discussion is ultimately about controlling the quality of the constructed final product. The DOT survey and design-builder interviews asked the respondents to gauge the impact on quality of a list of project factors. Table 5-5 shows that both the clients and the industry agree that geotechnical qualifications and experience have the most impact. They also agree on the benefits of involving the contractor in the design process and the value of performance criteria. The major difference is the perception of the value of agency involvement during the proposal phase. This correlates to the use of one-on-one meetings to clarify RFP requirements and to propose ATCs. Obviously the design-builders appreciate the opportunity to ask questions and clarify ambiguities before they have to submit a lump sum proposal for a multi-million dollar project. The other disconnect regards the perceived value of geotechnical quality management plans. Again the design-builders felt the plans had a much higher impact on final quality than the agency respondents. Since these are typically submitted in part in the proposal and in full after award before work begins, the industry perception may be due to the fact that the review and approval process further clarifies and quantifies the clients' expectations with regard to final geotechnical requirements. The results of this analysis lead to the conclusion that the use of proposal phase one-on-one discussions and post-award development and agreement on quality management plans are effective geotechnical risk management practices. In essence, the analysis argues that the geotechnical quality management system be similar to that used in DBB and probably different than the quality management system that will apply to the rest of the DB project.

Table 5-5 Impact on Final Project Quality

Research Instrument	DOT Survey			Design-builder Interviews		
Factor	**Very/High Impact**	**Some/ Slight Impact**	**No Impact**	**Very/High Impact**	**Some/ Slight Impact**	**No Impact**
Qualifications of the Design-Builder's geotechnical staff	89%	11%	0%	91%	9%	0%
Design-Builder's past geotechnical project experience	85%	15%	0%	82%	18%	0%
Agency interactivity with geotechnical design team during proposal phase	26%	48%	26%	73%	27%	0%
Early contractor involvement in geotechnical design	63%	37%	0%	73%	27%	0%
Use of geotechnical performance criteria/specifications	67%	19%	15%	64%	36%	0%
Level of agency involvement in the geotechnical QA process	37%	56%	7%	55%	45%	0%
Use of agency specifications and/or design details	67%	33%	0%	45%	45%	10%
Level of detail expressed in the procurement documents	63%	33%	4%	45%	45%	10%
Quality management plans	41%	56%	4%	82%	18%	0%
Warranty provisions	22%	48%	30%	18%	55%	27%

Conclusions

The study sought to answer research questions on managing pre-award geotechnical risk and mitigating/retiring that risk expeditiously after award. The survey, the content analyses, the interviews, and the literature provided a rich source of information from which to draw conclusions and answer the research questions.

Experience has shown that the most effective approach to managing pre-award risk is to conduct a thorough geotechnical investigation before awarding the construction. However, the combination of pressure to expedite project delivery by the federal government and financial incentives to implement DB increase the potential that state DOTs will tackle major projects with thorny geotechnical issues using DB. Ultimately, the issue becomes whether or not the client is willing to pay via the design-builder’s contingencies for geotechnical risks that may go unrealized. Therefore, the primary finding is a strong recommendation that an agency use extreme caution before selecting DB project delivery on a project involving high geotechnical risk.

Given that recommendation, if an agency finds it must deliver such a project using DB, it must then aggressively manage geotechnical risk in an expeditious manner to achieve cost, schedule, and construction quality goals. To that end the following conclusions that answer the first research question are offered:

- US DOTs consider DB to be an effective tool for accelerating project delivery.
- Achieving an aggressive schedule requires that the geotechnical design be completed as soon as practical to avoid delaying the start of construction.

US state DOTs manage geotechnical risk during the pre-award phase through RFP requirements:

- for well qualified and experienced geotechnical personnel;
- by limiting potential geotechnical design solutions to ones with which the agency has previous experience.

A number of effective geotechnical risk management tools were also identified.

- Agency interactivity during the proposal phase had a high or very high impact on final project quality.
- Communication with competing design-builders is enhanced during proposal preparation phase by using confidential one-on-one meetings to clarify RFP intent, resolve ambiguities in the RFP geotechnical data, and to present potential geotechnical ATCs.
- Confidential ATCs create a mechanism for competing design-builders to clarify the magnitude of the geotechnical risk before quantifying it in the price proposal.
- The use of risk sharing clauses that quantify in dollar terms the geotechnical risk a design-builder is exposed to with the agency assuming responsibility for differing conditions cost above that threshold was found to be an effective practice.

In answer to the second research question, effective tools to retire geotechnical risk after award by expediting design review are as follows:

Minimize the number of interim design reviews before the final release for construction acceptance review of geotechnical design.

- Maximize the use of both formal and informal over-the-shoulder geotechnical design reviews to resolve issues and concerns as they arise rather than stopping the design production by requesting a complete package and a period in which the design-builder cannot move forward until comments are received and addressed.
- Permit the release of geotechnical design packages for construction before the remainder of design is complete to begin excavation to identify and resolve any differing conditions as soon as practical.

The final effective practice is crafting explicit DSC clauses that permit expeditious resolution of discrepancies between pre-award and post-award geotechnical conditions. The research found that furnishing all the geotechnical information on hand when the project is advertised and building the DSC clause in a manner that makes it specific to the available geotechnical data rather than merely using a standard boilerplate DSC found in DBB projects was an effective alternative.

The above conclusions are limited in their application to the US markets from which the data was drawn. Because there are significant differences in contracting laws and regulations internationally, the reader is cautioned against generalizing these conclusions without thoroughly reviewing the applicable legal environment in which the conclusions may seem to apply. Nevertheless, geotechnical risk is inherent to all projects regardless of location. Thus, the effective practices that relate to the expeditious review of geotechnical design product to accelerate the ability to begin excavation and determine actual site conditions as soon as practical could be generalized to all DB projects.

The conclusions do support a common theme regarding managing geotechnical risk in DB projects. The public agency can best manage these risks by creating an environment of information-rich communications with its industry partners before the procurement starts, during the procurement process itself, and after award of the DB contract. Clearly, there is no “magic” contract clause that can adequately absolve the public agency of geotechnical risk. Therefore, actively managing risk and expediting the identification, quantification, and resolution of geotechnical risk is in the best interest of the agency and the design-builder, as well as the taxpayer that ultimately must pay the bills.

CHAPTER 6. LIFE CYCLE COST EVALUATION OF ALTERNATIVES TO THE NUCLEAR DENSITY GAUGE FOR COMPACTION TESTING ON DESIGN-BUILD PROJECTS

McLain, K.W., and Gransberg D.D. "Life Cycle Cost Evaluation of Alternatives to the Nuclear Density Gauge for Compaction Testing on Design-Build Projects." *Journal of Construction Engineering and Management*, ASCE, (Submitted October 2015).

Chapter 6 investigates the life cycle costs for the nuclear density gauges and its alternatives for use in MoDOT compaction testing. Costs per tests were calculated and are also presented in this chapter.

ABSTRACT

When the Missouri Department of Transportation (MoDOT) implemented design-build (DB) contracting, it revised its quality assurance program and shifted most of the compaction testing to the design-builder. As a result, fewer compaction tests were performed by state personnel and the need for speedy quality control testing by the agency to facilitate construction production disappeared. This paper reports the results of a study conducted by the to evaluate three alternatives to the Nuclear Density Gauge (NDG) using life cycle cost analysis and cost index number theory. The study's objective was to investigate alternative soil compaction test devices and provide input to a decision regarding whether or not MoDOT should retain or replace the NDG. Despite the NDG successful track record, the ease of employment and speed with which the compaction results are delivered comes with a price in terms of life cycle costs. The NDG is regulated by the Nuclear Regulatory Commission and entails an onerous, on-going

administrative workload to permit its continued use. The NDG also incurs additional certification, storage and disposal costs, not found in non-nuclear compaction testing alternatives. This paper reports the results of a life cycle cost (LCC) analysis of NDG and three alternatives: dynamic cone penetrometer (DCP), electrical density gauge (EDG) and the sand cone (SC). The study finds that the SC and DCP are the most cost effective but are the least cost effective when measured on a basis of timely results. Thus, the NDG replacement/retention decision becomes one of how fast are compaction tests required by the agency. Since MoDOT has adopted contractor acceptance testing in its DB program, it now only conducts verification testing of contractor test results. Thus, the paper recommends that the NDG be replaced.

Background

Design-build (DB) project quality assurance (QA) programs require that an owner revise its traditional design-bid-build (DBB) QA process to account for the fact that the design-builder is providing the project's final design (Gad et al. 2015). Project delivery is often modeled as a three-legged stool where the legs are cost, schedule, and quality (Chan 2013; Goetsch and Davis 2014; Karlen et al. 1997). DBB quality is defined by the construction documents upon which construction contractors can bid (Ellis et al, 1991), the time is specified by the contract completion date, leaving cost as the only variable leg of the stool to ensure a level platform (Ellicott, 1994). Thus, DBB project delivery is a "system where the constructor tells the owner how much it will cost to deliver the quality defined in the design within the specified period of performance" (Gransberg et al 2006). DB procurement normally demands that lump sum price be offered by the design-builder with scope being established within a collection of performance criteria and a specified performance period (Ernzen and Feeney 2002). This leaves quality as defined during the design process as the variable leg in the DB stool.

As shown above, the design-builder is now in a position where the resultant level of quality is now a function of both the fixed price and the contract schedule. Therefore, a successful design-builder must produce a final design that can be built inside the cost and time constraints, and the owner must not allow its QA program to impede progress without a solid, defendable reason. The issue is exacerbated by the increased pace that usually accompanies a DB project (Stefani 2004), creating an environment where delay claims can become extremely expensive (Kandell 2014). This issue led MoDOT and other state DOTs to adopt the use of contractor acceptance testing (Smith 2001; Turochy et al. 2007) as described in a Federal Highway Administration (FHWA) Technical Advisory 6120.3 (2004). The overall impact is that the owner has transferred some of its traditional QA field testing responsibilities to the design-builder and no longer needs as large a component of in-house inspectors and testing equipment on DB projects (Ernzen and Feeney 2002).

The results of the shift to verification testing on DB projects was so promising that in 2013 MoDOT adapted its DB QA program for use in its traditional DBB projects. In doing so, it made the construction contractor responsible for the bulk of the QA/QC field testing on MoDOT construction projects (Ahlvers et al. 2013). On projects involving large structural fills, achieving the specified compaction is the key quality function that must be properly deployed for the project to perform as intended over its service life (Arditi and Lee 2004). The nuclear density gauge (NDG) has been the tool of choice for both MoDOT and its contractors because it is easy to employ and gives immediate feedback on site.

The QA policy change shifted the compaction performance risk to the contractor, reducing the number of field compaction tests conducted by MoDOT inspectors. The upshot was that MoDOT inspectors no longer needed compaction test results as quickly as it did in the 1980s when the NDGs were originally fielded. The impact of the switch from using MoDOT NDGs for front-line QC tests where timeliness of compaction test results drive contractor production to a QA verification role removed timeliness as a critical characteristic for in-house compaction testing. While this paper is not arguing that timely verification of the accuracy of contractor test results is no longer important, the shift in roles removed failing MoDOT test results as a potential barrier to progress and reduced the need for speed in compaction test results. Put another way, if the MoDOT QA verification test results do not correlate with the contractor's QC results, the discrepancy has become a contractual problem where before timely test results were a production problem.

Research has established that when given a choice testing techniques, engineers generally choose the option that involves the highest level of technology (Schein 1996). However, enhanced technology comes with a cost, and the additional life cycle cost increment must be justified by a corresponding increase in value. Therefore, the primary research questions investigated in this study are as follows:

- Do the benefits of easy employment and speedy test results provided by the nuclear density gauge (NDG) justify its life cycle cost for MoDOT projects?
- Are there alternatives to the NDG that provide a better value?

The Missouri Department of Transportation (MoDOT) has been using the NDG as its primary technology for compaction testing for nearly 35 years, and currently has nearly 56 units distributed across its 7 districts. The NDG has been found to have the following primary benefits:

- Speed for obtaining the results.
- Requisite level of precision.
- Portable and compact.
- Measure both moisture and density.

Given the role change and the need to conduct considerably fewer tests, MoDOT decided to re-evaluate its use of the NDG in light of the large number of administrative requirements for training, certification, calibration, storage, and hazardous waste disposal that form the NDG's administration and logistics tail. While its benefits are well documented, the department began to question whether they provided adequate value for money. The Virginia DOT defines value for money (VfM) as: "A project is said to have positive VfM when, relative to other procurement options, it is forecast to deliver and/or is demonstrated to have delivered the optimum combination of life cycle costs and service quality that will meet the objectives of the project" (VDOT 2011). It is important to note the dual metrics of "life cycle costs and service quality."

Life cycle cost analysis (LCCA) and cost indices are tools used to quantitatively evaluate a product or process (Riggs and West 1986). Pittenger et al. (2011) maintain that "…LCCA [can be used] to determine cost effectiveness and return on investment … [for] transportation decision-making … in transportation projects." LCCA relates the initial capital costs of investment along with the long-term usage costs of the product or process. Cost indices were first proposed by Riggs and West (1986) and provide a means to permit the engineer to measure the "bang for the buck." One study says that cost index number theory "seeks to combine cost and engineering measurements into a single index that can permit the direct comparison of two or more alternatives simultaneously and thus provide a measure of cost effectiveness on an engineering property basis… [and] compare a more expensive technology with a less expensive technology to determine if the incremental cost difference between the two alternatives is offset by enhanced engineering performance" (Gransberg and Zaman 2005). Thus, using both metrics to evaluate potential alternatives provides the analyst with two independent measures with which to compare the costs and the benefits of several alternatives over their services lives while including a measure of return on investment in engineering terms. Hence the use of LCCA and cost indices provides a similar set of evaluation criteria for the NDG and its technical alternatives.

In the past, MoDOT has used both LCCA and cost indices sparingly when evaluating technical alternatives for QA and QC testing for embankment and roadway construction. Each past investigation was limited to a single option and often the decision to not replace the NDG was a function of finding a more pressing requirement for the available funds. In a nutshell, the justification to expend the funds to replace a technology that is performing satisfactorily and is already available must be compelling if there are other unfulfilled requirements competing for the same block of funds. Therefore, MoDOT commissioned this study to make a comprehensive analysis on virtually all alternatives is using LCCA and cost indices as the evaluation tool in the investigation effort. The methodology described in the next section is designed to focus on VfM rather than merely capital costs. Therefore this paper reports the results reached in determining viable alternative testing methods for soil compaction in roadway and embankment construction.

Compaction Testing Alternatives

There are three classes of compaction measuring devices or tests. The three classes and the possible alternatives in each class are as follows:

- Density and Moisture Gauges
    - Nuclear Density Gauge (NDG)
    - Electrical Density Gauge (EDG)
    - Soil Density Gauge (SDG)
- Volume Replacement/Volume Measurement
    - Sand Cone (SC)
    - Density Drive Sampler (DDS)
- Stiffness/Modulus Measurement
    - Light Weight Deflectometer (LWD)
    - Dynamic Cone Penetrometer (DCP)
    - Electronic DCP (DCP-E)

This paper evaluates one alternative from the three different classes. It should be noted that the paper reports the results of the pilot test for a larger study that will eventually evaluate all the alternatives shown above. Table 1 describes three alternatives under investigation along with each option’s advantages and disadvantages.

Table 6-1 Summary of Comparisons of Commonly Used Alternatives and the NDG

	NDG	EDG	SC	DCP
Measurement Method	A retractable probe is lowered into the soil through a pre-drilled hole. The probe emits gamma radiation though the tested soil and then to detectors in the gauge to measure density. Moisture measurement is done through a neutron source and detector located inside the gauge.	Measures the electrical dielectric properties and moisture levels of compacted soil using high, radio frequency traveling between darts driven into the soil being tested.	Uses premeasured container of sand to fill excavated hole in soil. The volume of used sand is determined. The moisture content of the removed soil is determined by other methods.	Operates by dropping an 8 kg mass a height of 575 mm (22.6 in). Impact causes the probe to be driven in the ground. A dynamic Penetration Index (DPI) is giving in units of mm/blow and is recorded versus depth
Advantages	1. Quick measurements for both density and moisture. 2. Portable.	Portable and lightweight.	Apparatus, accessories and consumables are inexpensive.	1. Simple to use with minimal required training 2. Standard unit relatively inexpensive. 3. Electronic DCP can be operated by one person
Disadvantages	1. Must be licensed by the NRC. 2. Operators must go through initial training and annual recertification. 3. Special storage requirements. 4. Hazardous material disposal requirements.	Must be calibrated against other compaction testing device with a minimum of five testing points but for better correlation need 8 points or more.	1. Destructive test. 2. Can be time consuming. Moisture determination done in separate second step. 3. Hard to use in base material, rocky soil, and very soft plastic soils. 4. Operator technique may impact the test results.	1. Hard to use in gravelly soils. 2. DCP needs to be operated by two person team. One to stand up the device and apply loads the other to read the side scale. 3. Moisture determination done in separate second step.

The most obvious difference is the requirement for specialized training and recertification necessary when using the NDG. It was this issue that provides the motivation for the study. While all other options to perform compaction testing require initial training, NDG is the only one that is regulated by the US Nuclear Regulatory Commission (NRC). Additionally in the words of one author:

> “The nuclear density gauge is the main device used for measuring the field density of compacted layers of unbound materials. However, the use of this device entails extensive

regulations and prohibitive costs associated with its handling, storage, calibration, and maintenance and the transportation of radioactive materials.” (Nazzal 2014).

The same study reported that a survey of US DOTs and Canadian Ministries of Transportation found that “the majority were interested” in finding non-nuclear methods to measure compaction, largely because of the administrative and logistics issues associated with the NDG. MoDOT was one of those DOTs, and the remainder of this paper will detail the analysis of the NDG against the EDG, SC, and DCP on a LCC basis to determine the relative cost effectiveness of each alternative.

Life Cycle Cost Analysis Fundamentals

Barringer and Weber (1996) state that LCCA is not an exact science and researchers and statisticians will get different answers using similar sets of data. The differing answers are neither wrong or right only reasonable or unreasonable. LCCA estimates are never as accurate as their inputs, but with reasonable inputs and good judgment, LCCA allows for examining costs and comparing competing methodologies. The FHWA encourages the use of LCCA for the comparison of alternatives in the design, construction and maintenance of all types of transportation assets (Walls and Smith 1998) In essence LCCA is a mechanism whereby a public agency can justify purchasing an alternative that is not the lowest initial cost. In other words, LCCA allows the agency to quantitatively demonstrate to the taxpayer that the agency is making purchasing decision that provides good VfM. Thus, it is logical to look compare alternatives for measuring compaction using LCCA.

The underlying foundation of LCCA involves discounting all the costs and benefits during an alternative's service life to a single point in time where they can be compared (Beatty 2002). FHWA encourages the use of present value analysis (Walls and Smith 1998), which in the opinion of one author is an analog for the lowest bid, a decision criterion that permeates the public construction sector (Gransberg and Scheepbouwer 2010). However, there is an emerging opinion that since most public agencies receive funding on a fiscal year by fiscal year basis that Equivalent Uniform Annual Cost (EUAC) is a more appropriate approach since it reflects the annual impact on the agency budget (Pittenger et al. 2011). To apply LCCA to the comparison of compaction measurement alternatives, the following input parameters had to be determined:

1. The annual costs incurred by the system and/or mandated by regulations or testing standards.
2. The life of a method or system under average testing conditions.
3. The appropriate interest rate.

Water Environment Research Foundation's (WERF) Life Cycle Cost Tool specifies that the following typical costs be included in the analysis (WERF 2011):

- Acquisition Costs
- Operating Costs – cost for repairs, and spares
- Maintenance Costs – corrective, preventative, and predictive
- Disposal Costs

Inputs can be either deterministic or probabilistic, especially for costs related to operation and maintenance cost for a system. The authors placed initial and annual costs of each of the four compaction testing alternative into the following eight categories:

- Purchase Price (P)
- Annual Training Cost (T)
- Consumables (C)
- Disposal Costs (D)
- User Cost (U)
- Annual Calibration and Verification Costs (V)
- Storage Costs (S)
- Licensure Costs (L)

Fundamentals of Cost Index Number Theory

Cost index number theory is essentially a variation of classic utility theory (Riggs and West 1986). This theory permits the analyst to calculate a unit cost of quality for use in financial decision-making. In a nutshell, to be viable an alternative must furnish an increase in quality that is greater than its increase in cost. In layman's terms, to be adopted for use the alternative must give "more bang for the buck." This is particularly useful if the new technology turns out to be marginally more expensive than the traditional technologies. Thus, the analyst furnishes a justification for spending a bit more money up front to receive a commensurately better final product. This type of analysis is founded on life cycle cost fundamentals and is particularly applicable to public transportation projects (Aktaş et al. 2011).

An important aspect of cost index number theory that must be understood is its ability to establish relative relationships between alternatives. If one relies only on bottom-line dollar values to make management decisions, the decision-maker is disregarding the relative qualitative merits of each alternative (Pittenger et al. 2012). Therefore, the end-user of a construction project will always be given the minimum level of quality. This attitude is deeply ingrained in organizations like MoDOT who are required by law to award construction projects to the lowest bidder. In the low-bid paradigm, the engineer specifies the minimum acceptable level of quality in the plans and specifications. The construction contractor bids the cost of delivering the minimum level of quality and the inspector checks to make sure the minimum level of quality is received in the final product. The “minimize initial cost” without regard to quality mentality can permeate an organization’s business practices. Cost index numbers provide a means to take a longer term approach to technical decision-making while retaining an objective decision-making criterion based on quantifiable parameters. Therefore, the challenge to the engineering analyst is to accurately portray the qualities of each alternative in a quantitative fashion that allows costs to be associated with those qualities that best describe the differences in alternatives.

A cost index number portrays the cost required for acquiring, maintaining or constructing a product, as measured in money, resources or time. A cost index is usually given as a ratio of cost per unit of measure and is a useful parameter that can assist in comparing alternatives for compaction test devices with regard to the long-term cost effectiveness of each option.

Replacing the NDG promises to reduce the administrative workload to manage the NRC training and certification requirements. MoDOT spends many precious dollars each year to meet NRC requirements regarding the use, storage and disposal of NDGs, and replacing it permits those resources to be applied elsewhere. As a result, it is important to prove that a non-nuclear compaction testing alternative does indeed deliver a product whose quality is commensurate with its LCC. The product of a compaction test can be measured in terms of how long it takes to complete a test and the time between test completion and the availability of results. Thus, times associated with each alternative’s procedure create a means to quantify its value. Additionally, the cost to the state for each test is another measure. In this case, Equations 6-1 and 6-2 below were used to develop cost indexes to measure the cost effectiveness of each compaction test device.

Time/Cost = Average time to perform test in minutes /EUAC cost for each device (6-1)

Cost/Test = EUAC cost for each device/ average annual number of tests (6- 2)

Previous Study Analysis

A similar study was conducted by Cho et al. (Cho et al. 2011) used a LCC cost comparison over the assumed fifteen year useful life of an NDG and compared the annual ownership cost for the NDG with the cost of the Pavement Quality Indicator (PQI) for measuring asphalt pavement densities. The NDG measures density for both asphalt pavement and soil. Therefore, Cho et al. used the average cost of the LWD and EDG for the soil density measurements to create a comparable utility. In essence, they evaluated the possibility of replacing the NDG with two other devices. The pricing assumptions (Cho et al., 2011) made for the NDG and for the PQI, LWD and EDG are shown in Table 6-2. The results of this study are shown in Figure 6-1.

Table 6-2 LCC Singular Device Comparison from Cho et al., (Cho et al. 2011)

Ownership Cost	NDG	LWD	EDG	PQI
Initial Cost	$6950	$8675	$9000	$8200
Radiation and Cert. Class	$750	0	0	0
Safety Training	$179	0	0	0
Hazmat Cert.	$99	0	0	0
RSO Training	$399	0	0	0
TLD Badge Monitoring	$140/yr	0	0	0
Maintenance and Calibration	$500/yr	0	0	$500/yr
Leak test	$15	0	0	0
Shipping	$120	0	0	0
Radioactive Materials License	$1600	0	0	0
License Renewal	$1500/yr	0	0	0
Reciprocity	$750	0	0	0

In Figure 6-1, one can see the initial costs for the NDG is lower than those of the PQI plus average density device. However, after the fourth year, the NDG's LCC surpasses the PQI + average of non-nuclear gauges. Thus, the decision to replace the NDG appears to be warranted. The remainder of the paper will detail a similar analytical approach for the same decision in Missouri.

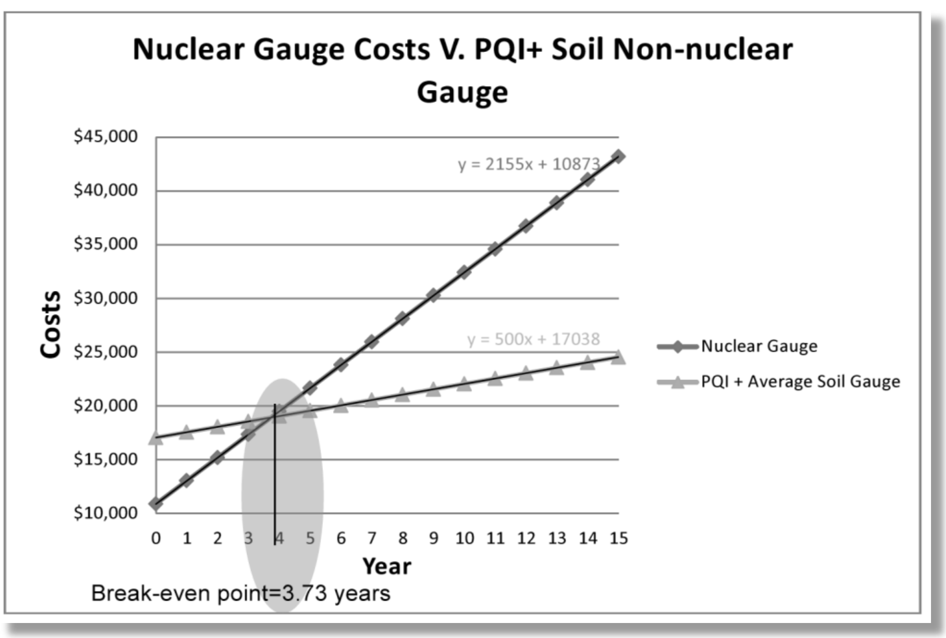


Figure 6-1 Break Even LCC Comparison from Cho et al. (Cho et al. 2011)

Methodology

The study used two primary research instruments to collect the necessary data to provide input to the analysis. It also entailed a pilot test to validate the approach to the LCCA and the cost index number analysis. The first research instrument was a survey of all MoDOT project offices where NDGs are stationed. The second instrument was a structured interview of MoDOT resident engineers, construction inspectors, laboratory technicians and most importantly, radiation safety officers (RSO) who have the responsibility to oversee the NDG training, certification, and operations. The interviews were used to collect actual cost data on the NDGs currently in deployed throughout the state and verified by checking equipment purchase invoices and other data maintained in the MoDOT central office in Jefferson City.

Data Collection Methods

Data collection and gathering of hard numbers was considerably easier for the NDG (MoDOT standard compaction testing device) because of existing records, experience, requirements and procedures. Records for NDG usage were readily available from required sign-out/check-in registers maintained at the project offices. The MoDOT RSO provided costs for NDG devices, calibration costs and frequency as well as the costs to dispose of spent nuclear material. Costs for testing alternatives and consumables came from invoices for purchased items or from the producers or distributors of rented or borrowed devices.

Simple time and motion studies were run in the field for each alternative method as it was applied during normal usage. Times were collected in the same location by the same technicians on the same portion of compacted fill. The results of the time and motion studies were validated by other agencies that use or routinely use the alternative devices for compaction testing that MoDOT is considering adopting.

NDG field usage was determined by the project office survey mentioned above. The questionnaire was developed from the literature review and assembled in accordance with the protocol established by Oppenheim (Oppenheim 1992). The questionnaire was sent to all 29 project offices to gather information on amount of usage. The questionnaire asked the respondents to determine two usage rates: during the construction season (March to November) and during the construction off season (November to March). Nineteen of the 29 project offices returned the survey, which yields a 66% response rate. The summary of the responses are shown in Table 6-3. From the project office survey, the following results were obtained:

- Average number of times NDG used per week during construction season: 1.16 times
- Average time NDG used per week during construction season: 1.26 hours
- Construction season: 8 months = 32 weeks
- Total Average NDG usage: 37 times each construction season.
- Total duration: 46.65 hours per office
- 29 Project offices
- Total annual duration for MoDOT= 1353 hours

Table 6-3 Nuclear Gauge Usage from MoDOT Project Offices for 2013

Project Office	District	Number of gauges	Usage per week during construction season	Duration of testing period construction season (hrs)	Usage per week during off season	Duration of testing period during off season
St. Joseph	NW	2	2	Unknown	1	Unknown
Chillicothe	NW	2	2	Unknown	0	0
Maryville	NW	2	7	Unknown	0	0
Troy	NE	3	0.25	Unknown	0	0
Hannibal	NE	3	0.75	Unknown	0	0
Nashua	KC	2	3	1	0	0
Marshall	KC	2	0.5	0.5	0	0
Lee Summit	KC	3	1	0.5	0	0
St. James	CD	1	0.25	0.33	0	0
Jefferson City	CD	2	5	1	0	0
Camdenton	CD	1	0.25	1	0	0
Columbia	CD	2	2	Unknown	0	0
Chesterfield	SL	1	0.1875	Unknown	0	0
Clinton	SW	1	0.5	Unknown	0	0
Branson *	SW	0	0.1	Unknown	0	0
Joplin	SW	2	0.367	Unknown	0	0
Jackson	SW	2	0.5	2	0	0
Poplar Bluff	SE	2	1	4	0.5	2
Willow Springs	SE	2	0.03	1	0	0
***Branson uses the Springfield Project Office Nuclear Gauges**						

.

LCCA Assumptions

In determining EUAC for the alternative soil compaction test devices, a number of assumptions had to be made since the systems being evaluated that are not currently in standard use with MoDOT. The major assumptions are as follows:

- Two test devices are to be assigned to the project offices in the same manner as the nuclear density gauges. This assumption is very conservative as the possibility exists that after the contractors take over the much of the compaction testing responsibility, MoDOT may not need both devices.

- Since an average of 37 compaction tests using the NDG were run by each project office in the 2013 construction season, the same number of tests for the alternative test methods was assumed.
- The costs associated with personnel time and transportation to receive the required calibration procedures was not included.
- No residual value for the equipment was assumed at the end of its useful life.
- Training costs and times were assumed to be constant for each testing alternative. While these times and costs should decrease over the lifetime of the device, they are also dependent on personnel turnover in the project offices.
- Construction inspectors and construction technician currently conducting NDG tests would be conducting compaction tests using possible alternate devices for the NDG.

Analysis and Results

Life Cycle Cost Analysis Results

The life cycle cost results expressed as the EUAC for the NDG and competing alternative testing devices are shown in Figure 6-2. The EDG was the most expensive to own and operate over its life cycle followed by the NDG. The DCP and the SC had EUACs lower than the NDG.

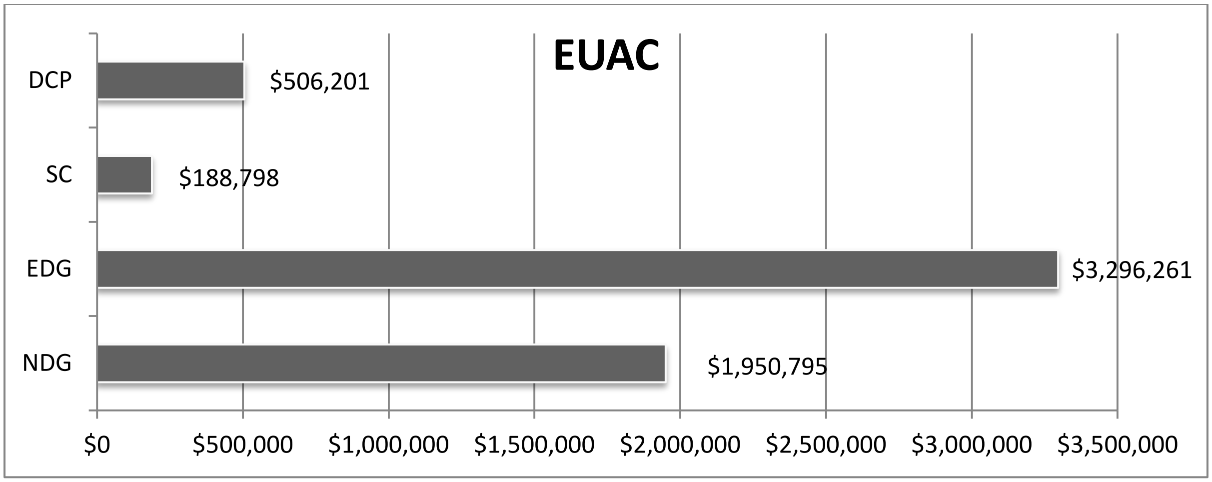


Figure 6-2 EUAC to MoDOT for Each Compaction Testing Device

The cost comparison shows the fifteen year cost for owning and operating nuclear gauges to MoDOT at approximately 30 million dollars. But cost projections predict the EDG could cost the Department an additional 20 million dollars over the cost of the NDG. Table 6-4 provides the MoDOT ownership cost totals for the period analysis. It shows the input data for the eight cost categories discussed in a previous section. The major difference between the NDG and the alternatives under analysis is that the only costs beyond the initial procurement and training of personnel are for labor and consumable supplies; whereas, the NDG has significant ownership costs throughout its useful life.

Table 6-4 EUAC for 56 units and training for 28 Project Office Locations

Ownership Cost Category	NDG	EDG	SC	DCP
Purchase (P)	$436,800	$772,800	$11,200	$84,000
Training (T)	$92,135	$60,904	$31,327	$17,332
Consumable (C)	NA	NA	$672	$1,736
Labor (M)	$29,064	$144,157	$112,000	$151,132
Disposal (D)	$44,800	NA	NA	NA
User Cost (U)	$13,407	NA	NA	NA
Verification/Calibration (V)	$6,785	NA	NA	NA
Storage (S)	$2,751	NA	NA	NA
Licensure (L)	$6,400	NA	NA	NA

Table 6-5 illustrates an analysis that compares each device's ownership costs as a percentage of total EUAC. Its purpose is to measure the effect of device specific ownership costs. Given the assumption that MoDOT will conduct the same number of compaction tests each year without regard to the device in use, the labor costs will be roughly equivalent for all four devices when taken as an annual lump sum. Therefore, the relative difference in the labor costs and purchase cost taken as a percentage of EUAC provides another measure of cost effectiveness. One can see that the NDG's and EDG's purchase costs are over 90% of its EUAC whereas the SC's largest cost is the labor in actually running the compaction test.

Table 6-5 Percent of EUAC for Compaction Test Devices

Device	P Purchase	T Training	C Consumable	M Labor	D Disposal	U User Cost	V Calibration	S Storage	L Licensure
NDG	89.6	4.7	0	1.5	2.3	0.7	0.3	0.6	0.3
EDG	93.8	1.8	0	4.4	0	0	0.00	0	0
SC	23.7	16.6	0.4	59.3	0.0	0.0	0.0	0.0	0.0
DCP	66.4	3.4	0.3	29.9	0.0	0.0	0.0	0.0	0.0

Cost Index Results

The cost index number analysis provides a "bang for the buck" evaluation of cost effectiveness. The results are shown in Table 6-6. The EUAC/Test index highlights the cost or potential cost for MoDOT every time an inspector or construction technician leaves the project office to perform a compaction test. For the EUAC/Test index, the decision makers for choice of compaction test device should be noting the lesser numbers, which for this study is the sand cone and density drive sampler. The testing time per EUAC underscores the relationship of time to perform the test to its cost. The decision makers need to consider the larger numbers because the

EUAC's magnitude is large to the testing time in minutes. This index may need to be looked at an individual device basis to lower the magnitude of the EUAC to testing time.

Table 6-6 Cost Index Summary

Device	EUAC/Test	EUAC/Testing Time ($/Minute)
NDG	1883	188.32
EDG	3182	127.22
SC	182	9.11
DCP	489	48.85

Conclusions

The LCC and cost indexes for the differing test methods and devices are tools that will factor into MoDOT's decision to eventually select a compaction testing alternative to the NDG. Costs along with accuracy, repeatability, and testing performance in differing soils, and ease of use in testing will all be used in determining the best compaction testing system or device for quality assurance and control practices on MoDOT projects. The conclusions drawn from this pilot study are as follows:

- Both the NDG and the EDG have a greater annual life cycle cost than the SC and the DCP. The EDG's EUAC is greater than the NDG's EUAC.
- The life cycle cost per test index show the SC and DCP to be the most cost effective.
- The fact that MoDOT has shifted the bulk of the annual compaction testing program to the contractor shifts those tests taken by MoDOT technicians to a QA verification testing role and reduces the advantage of having immediate feedback that is the major advantage of the NDG and the EDG.

Therefore, given all of the above discussion and analysis, the pilot test has confirmed the that replacing the NDG with an alternative testing device will accrue tangible long-term benefits to MoDOT and release scarce operations and maintenance funding for other purposes.

CHAPTER 7. MODOT QUALITY MANAGEMENT AND CORRELATION OF POTENTIAL ALTERNATIVES TO THE NUCLEAR DENSITY GAUGE

McLain, K.W. and D.D. Gransberg, "Missouri Department of Transportation's Quality Management and Correlation of Potential Alternatives to the Nuclear Density Gauge," *International Journal of Quality and Innovation,* Submitted March 31, 2015.

This chapter looks at the development of MoDOT's Quality Management compaction testing program and the effect the program has had on the frequency of compaction testing throughout the Department. In order to determine whether MoDOT contractors would be to be required to use the same compaction testing technology as MoDOT if it chose to replace the NDG, the chapter also reviews past investigations between measured soil density and corresponding modulus/Clegg Impact Values and reports correlation values from testing conducted on several Missouri construction projects.

Abstract

The Missouri Department of Transportation's (MoDOT) past and present Quality Control and Quality Assurance programs for construction are examined. MoDOT's present Quality Management program along with a small number of grading projects has lowered the number of Quality Assurance (QA) soil compaction tests completed in the past two years. The Department would like to rid itself of using the Nuclear Density Gauges because of burdensome Federal regulations, required training, security and licensing fees. Linear and multiple regression analysis was performed to see if a correlation between nuclear density gauge dry densities values and Light Weight Deflectometer modulus values/ Clegg Hammer Clegg Impact Values exist. These

relationships or lack thereof will determine the technology used by construction contractors to perform compaction quality control testing if MoDOT moves away from using nuclear density gauges for soil density verification.

Background

The Missouri Department of Transportation (MoDOT) has been using the Nuclear Density Gauge (NDG) as its primary technology for compaction testing for nearly 35 years, and currently has about 56 units distributed across its seven districts. The NDG has been found to have the following primary benefits:

- Speed for obtaining the results.
- Requisite level of precision.
- Portable and compact.
- Measure both moisture and density.

The Department changed its quality assurance (QA) program in 2013 and made the construction contractor responsible for the bulk of the quality control (QC) compaction testing. This process is termed Quality Management by MoDOT. The United States Nuclear Regulatory Commission (U.S.N.R.C.) requires training, licensing and security that have become a hindrance to both MoDOT and the contractor. This role change along, combined with a decreased number of major grading projects leads to the need to conduct considerably fewer tests, MoDOT therefore found it prudent to re-evaluate its use of the NDG in light of the large number of administrative requirements for training, certification, calibration, and storage. The initial evaluation of alternatives involved the following classes of non-nuclear testing devices (Berney and Kyzar 2012):

I. Electrical Density and Moisture Gauges
   a. Electrical Density Gauge (EDG)
   b. Soil Density Gauge (SDG)
II. Volume Replacement/Volume Measurement
   a. Balloon (RB)
   b. Sand Cone (SC)
   c. Density Drive Sampler
III. Stiffness/Modulus Measurement
   a. Light Weight Deflectometer (LWD)
   b. Dynamic Cone Penetrometer (DCP)
   c. Clegg Soil Impact Tester
   d. GeoGauge (GG)

This paper focuses on the relationship of the Light Weight Deflectometer and the Clegg Soil Impact Tester (a.k.a. Clegg Hammer) and how they correlate with the NDG dry density results in differing materials. This relation discussed in the paper can affect the contractors' QC process if MODOT elects to uses the Clegg Hammer or the LWD to measure the modulus and the contractor wishes to test compaction with the NDG.

Quality Management

MoDOT’s definition of Quality Management (QM) is: “A process that gives the contractor the primary role and responsibility for incorporating quality into the project, where quality is included in the planning and scheduling of project activities. Quality is managed by the contractor with QC testing and inspection. QA by MoDOT is conducted at specified stopping or hold points.” (Ahlvers et al 2013).

MoDOT’s present QM system was an evolutionary process that began in 2000 when QA/QC process for asphalt was initiated. Soon thereafter (2002), MoDOT implemented a QA/QC program for Portland cement concrete pavements (PCCP). Next, the release of the *Missouri Standard Specifications for Highway Construction* (MSSHC) in 2004 increased QA/QC activity. MSSHC was developed to move the department towards increased usage of performance specifications (Ahlvers et al. 2013). The performance-oriented QM system originated in 2007 as a result of MoDOTs initial large design-build (DB) projects. These include the New I-64 project in St. Louis (2007- 2009), Kcicon in Kansas City (2008-2011), the state-wide Safe and Sound Bridge Program and the New Mississippi River Bridge Project in St. Louis (2010-2014). In 2012 MODOT implemented the QM program used on its DB projects on 46 design-bid-build (DBB) pilot projects across the state. The pilot program was successful and the Department initiated full implementation of QM in 2013 on all projects.

Prior to 2000, the majority of the QC and all the QA activities were conducted by the MoDOT Construction and Materials Division on highway projects. This changed when MoDOT implemented a QA/QC program for asphaltic concrete pavement projects. The composition of the asphalt mixture was specified in Section 403 of the MSSHC, but the job mix formula was developed and submitted by the contractor for MoDOT approval. The contractor under the supervision of a MoDOT materials inspector would collect and submit representative samples of the asphalt binder and mineral aggregates to the MoDOT Central Laboratory for testing. If the tests on the samples passed, the contractor was then required to build test strips for each different mixture of a quantity of at least 2,000 tons to determine the compactive effort needed to obtain the required density. In all cases except stone mastic asphalt, MoDOT personnel performed asphalt pavement density testing using nuclear density measurements. Nominal thickness was tested by the Geotechnical Section with an auger truck equipped with water tanks and pavement core barrels. Past inspection and testing for PCCP was regulated by Section 502 of the MSSHC performed by MoDOT personnel with the contractor providing the field laboratory. Payment was based on results of profilograph measurements provided by the contractor with QA and pavement thickness measurements conducted by MoDOT.

Pavement Quality Management Program

Implementation of the QM program for pavements shifted the entire QC away from MoDOT inspectors and assigned it to the project's contractor. Under Section 403 of the MSSHC the contractor was required to maintain the necessary test equipment as well as qualified personnel to perform all QC construction and material inspection and sampling. The contractor started submitting a bituminous QC plan which named a contractor representative in charge of QC. The contractor designated testing lots and sub-lots with the number of cores cut for density. The QC plan included a proposed third party for use in dispute resolution. The plant calibration was now conducted by the contractor who was required to produce and retain calibration records.

The contractor's designated testing lab also had to retain copies of all test methods, procedures and results. Payment for the asphalt pavement was centered on pay factors based on density, asphalt content, voids in the mineral aggregate, and air voids.

QA/QC for PCCP inspection was initiated in 2002. Like asphalt pavements, the contractor was required to submit a QC plan. Again a third party was identified by the contractor and approved by MoDOT to resolve potential inspection and testing conflicts. Qualified testing personnel employed by the contractor were to be present on the paving portion of the project. The need for qualified inspectors employed by contractors prompted MoDOT to start an inspector and testing technician qualification program. The contractor defined lots and sub-lots for sample retrieval. For every 7,500 square yards at least one QC cylinder was prepared for compressive strength testing. QA cylinders were prepared every 30,000 square yards. Pavement cores were also taken every 7,500 square yards for pavement thickness determination. QC also included determination of air content, slump, gradation, and deleterious material which was previously all determined by MoDOT.

A Quality Level Analysis document was introduced where sample test results for thickness and compressive strength were investigated for average results and the variability of results, based on the mean and standard deviation. These statistical results were used to formulate pay factors for the PCCP portion of the project, with the pay factors based on each lot's thickness and compressive strength test results. The QC/QA process for pavement construction ended the need for the majority of pavement thickness cores produced by the Department. Pavement coring ceased to be a major function of the Geotechnical Section by 2006.

Design-Build Quality Management

The QA/QC process then morphed into a QM plan (QMP) for all processes on DB projects. QM was used with great success on the previously noted DB projects. The DB team was co-located with the MoDOT project team, where impromptu meetings could be held to solve problems that flared up quickly. The DB team included a full time QC manager who insured that construction means and methods complied with the specifications and that the materials installed met the submitted and MoDOT approved specifications. The QC Manager was not only in charge of the main contracting arm of the DB team but the myriad of subcontractors as well. Shown below in matrix form from the Kcicon project, the responsibility of the QC Manager and the associated risk should the Manager fail to fulfill its assigned responsibilities.

Table 7- 1 QM Checklist for Subcontractors Activities (Paseo Corridor Constructors 2008)

		Risk Level		
	Quality Management Activity	**High**	**Moderate**	**Low**
1	Establish a single point of contact for both sides.	X	X	X
2	Perform a site visit or assess subcontractor operations prior to mobilizing.	X		
3	PCC Engineers to perform reviews on critical submittals to the owner.	X	X	
4	Pre-activity meetings shall be conducted for major operations.	X		
5	Identification of Top 5 quality focus points for subcontractor’s work.	X	X	X
6	Require subcontractor to submit a weekly schedule.	X		
7	Identify and defined hold points for all inspections and tests required.	X	X	X
8	PCC engineers or superintendents to perform inspections at defined intervals.	X	X	

MoDOT inspectors worked with design-builder’s foremen and inspectors to perform quality management tasks at certain steps during construction. The DB QMPs included planned stopping points for testing. These are commonly called “hold or witness points.” Below is an example of hold points or witness points guidelines based on guidelines submitted by Paseo Corridor Constructors; this highlights nuclear density gauge testing points and quantities. Listed are the major areas in which stopping points were used.

Table 7-2 Quality Management Inspection Items (Paseo Corridor Constructors 2008)

Feature	**Testing**
Bridge Embankment	▪ Testing frequency of 1 nuclear density test per 500 yd^3 of fill in bridge embankment
Embankment	▪ Testing frequency of 1 nuclear density test per 2000 yd^3 of fill ▪ Moisture requirement for A-2-6 through A-7 soils
Pipe Placement / Structural Backfill	▪ Testing frequency of 1 nuclear density test per 200 yd^3 of fill
Subbase / Paving	▪ Testing frequency of 1 nuclear density test per 2000 tons or proof roll for reclaimed asphalt
Asphalt Placement	▪ QA testing frequency of 1 nuclear density test per 500 tons
MSE Wall	▪ Testing frequency of 1 nuclear density test per 200 feet of fill on every other lift
▪ Inspection Items -- Items that are looked at by QA and QC during construction	

New MoDOT Quality Management Program for Design-Bid-Build Projects

The DB QM procedures with a number of small enhancements became the present DBB project QM system. The central elements of the MoDOT QM program for DBB projects are as follows:

1. The contractor employs a full time Quality Manager.
2. The contractor develops and utilizes a Quality Management Plan.
3. Certified technicians and inspection staff are provided by the contractor.
4. MoDOT provides the QA personnel for the project. (Ahlvers et al. 2013)

The QMP is the strategy for instilling quality into a project. Before the start of work the contractor must submit a draft QMP before the preconstruction conference. The project's Resident Engineer and the contractor meet to negotiate and iron out the details. When an agreement is reached a "final" QMP is signed thus making it a contractual document. This document can be revised to fit the needs of the project with further negotiations between the contractor and the Resident Engineer and District/Central Office Construction personnel.

The QMP contains an Inspection and Test Plan (ITP). MoDOT has established a base ITP with minimum testing frequencies. The contractor can advocate changes from the ITP testing frequencies. The changes are reviewed by the Resident Engineer and, depending on the proposed changes; the contractors' ITP may be reviewed by District and Central Office Construction staff.

For materials sampling and testing the contractors' testing personnel must be listed in the quality management plan. If conflicts arise during inspection and testing an independent third party may be used to resolve the conflict. The contractor accepts and collects all material paperwork and tickets for materials delivered to the project site.

The MoDOT QM process addresses appropriate responses to any non-conforming work and deficient work that may occur. The definitions for these two categories are as follows:

- Non-conforming work: "Completed work that does not meet the contract requirements", (Ahlvers et al. 2013).
- Deficient work: "In-progress work that does not meet the contract requirements". (Ahlvers et al. 2013.

A non- conformance report (NCR) keeps a record of deficient or non-conforming work. Either QC inspectors or QA inspectors can issue an NCR with an expectation that the QC inspectors will discover and issue the majority of the NCRs. With the issuance of an NCR, the contractor is required to propose a resolution to the problem. The options the contractor is usually faced with are:

- Leave “as is”
- Fix/repair the problem
- Replace it

The QA inspector or Resident Engineer will approve or disapprove the proposed resolution and once the NCR is resolved MoDOT closes the issue.

Density Testing Requirements in the New QM Program

The ITP mandates a minimum QC density testing frequency of one test per lift per 500 feet per activity. Under the specification an activity is defined as predetermined item of work in a distinct location. The minimum QA density testing frequency is one test per day. These testing frequencies are for both the placement and compaction of embankment and compaction in cut. The approved tests for compaction according to Section 203 of the MSSHC are AASHT0 T 191 (Sand Cone), AASHTO T 205 (Rubber Balloon), and AASHTO T 239 (Nuclear Density Gauge) with the nuclear density gauge being both the preferred and most often used testing method. However, the new QM program requires less involvement by MoDOT personnel, which calls into question the continuing cost effectiveness of maintaining two nuclear density gauges (NDG) in each RE office. Before the new QM program was implemented and MoDOT personnel were conducting QC density tests, an argument could be made that the results of the tests needed to be available as soon as practical to facilitate the identification of nonconformance with compaction standards and their remedies in a manner that did not compromise the contractors' production. Nonetheless, the shift of all QC testing to the contractor made it the master of its own destiny and removed MoDOT from the production interruption equation. Therefore it is important to compute the change in NDG usage by MoDOT personnel both before and after the QM program change.

The frequency of nuclear density testing for the 2013 construction season running from March to November was established by sending a survey to the 29 MoDOT Resident Engineer (RE) offices in the state with 20 responses recorded and shown in Table 7-3. The same 20 offices were surveyed again for NDG usage for the 2014 construction season with 18 responses. The number of times the nuclear density gauge was used in the field per RE office during the construction season (approximately 32 weeks) has dropped from 37, approximately once a week to 22, which is roughly once every two weeks.

Table 7-3 MoDOT Resident Office NDG Usage

	2013	2014
Average time used per work week per RE office	1.16	0.68
Duration of tests(hrs)	1.26	1.20
Average NDG usage times per RE Office per construction season	37	22
Total duration of usage per RE (hrs)	46.65	26.35
Cost per test	$1881	$3144

The costs per test were generated by dividing calculated equivalent uniform annual cost (EUAC) for ownership, operation, security and maintenance of MoDOT's NDGs by the average usage times in a construction season.

The program to initiate QM on all projects has successful. As with any new initiative, there has been a learning curve for both contractors and MoDOT personnel. Now, with two construction seasons completed under the QM directive, procedures and responsibilities for both QA and QC have been learned, discussed and adjusted as required. There is a desire in the department to change compaction testing methods and do away with the NDG. To make the decision, MoDOT needed to evaluate the life cycle cost of current alternative technologies and compare that to the life cycle cost of adopting emerging technology that is compatible with Intelligent Compaction (IC) construction processes. The different alternatives will influence what MoDOT's contractors are allowed use for QC compaction testing. Their choice of testing technology will need to produce a measurement that is either the same property as the MoDOT technology or a well-defined correlation between different properties reported by different test methods. The non-nuclear Soil Density Gauge (SDG) readings are the same as the NDG (dry density and percent moisture). Thus, the contractor could still use the NDG for soil compaction QC if MoDOT adopted the SDG. The Modulus/Clegg Impact Value (CIV) based testing also has shown promise and has been implemented by several departments of transportation. However, a reliable correlation must be established between these modulus/CIV based testing procedures and density reported by the contractor's NDG or the contractor will most likely be required to utilize its own modulus/CIV based testing equipment for QC tasks. If no correlation exists then both compaction testing QA and QC will have to be conducted with the same method and/or equipment.

Earlier research studies have investigated the relationship between NDG readings and modulus/ stiffness/CIV readings. In Meehan et al. (2012), the research team used a simple linear regression approach to determine if a relationship exists between NDG dry density results and modulus readings from the light weight deflectometer (LWD), dynamic cone penetrometer (DCP), and the Humboldt GeoGauge (GG) (a.k.a. as the Soil Stiffness Gauge) and the results are shown in Table 7-4. The coefficients of determination display either a low correlation between the NDG and the DCP readings (0.22 -0.40) or essentially no correlation between NDG and the modulus readings for the LWDs and the SSG (0.068- 0.026).

Table 7-4 Coefficient of Determinations from Linear Regression Comparisons with NDG (Meehan et al. 2012)

Dependent Variable	Coefficient of Determination (R^2)
Soil Stiffness Gauge SSG/ GeoGauge GG	0.027
LWD 300	0.026
LWD 200	0.068
DCP -M	0.401
DCP-A	0.219
Notes: LWD 200 = Zorn LWD plate diameter of 200 mm; LWD 300 = Zorn LWD plate diameter of 300 mm; DCP-A = average; DCP-M = weighed mean; Method from White et al. (2007); Results from embankment constructed with sandy silt soil (SM)	

Similar results were found by Li (2013) in which linear regression was used in the comparison of the NDG Dry Density results to LWD modulus, GG stiffness, DCP California Bearing Ratio (CBR) , and modified Clegg Hammer results (MCH).The coefficients of determination varied with device compared, material tested and density of that material shown in Table 7-5.

Table 7-5 Coefficient of Determination from Comparisons to NDG from 3 Michigan Test Sites (Li 2013)

Location	**Material**	**R^2 GG**	**R^2 LWD**	**R^2 DCP**	**R^2 MCH**	**Comments/Notes**
Hancock	Gravel	0.2646	0.0192	0.2246	0.0100	Uncompacted test pads
Calumet	Gravel	0.9766	0.4119	0.6979	0.7029	Fully compacted test pads
Iron River	Sand	0.0321	0.1032	0.1358	0.1896	Fully compacted test pads with IC roller

The author noted that the poor correlation between the NDG and the other devices at the Hancock site could have resulted from the fact that test pad was uncompacted and only four measurements were taken. The author also commented that the good correlations for the tests conducted at the Calumet site might not be representative due to the limited number of tests (4) and that further assessment was needed. The author concluded that simple regression analysis did not show good correlation for the tests conducted on sand at the Iron River site due to soil heterogeneity and moisture content variation.

Meehan et al. (2012) and Li (2013) demonstrated that coefficients of determination can improve by the use of multiple regression analysis in which moisture or IC factors such as amplitude, vibration frequency and roller speed are considered. The introduction of compacted moisture content vastly improved the correlation of the nuclear gauge to the other compaction test devices as compared to linear regression seen in the table below.

Table 7-6 Multiple regression analysis that includes the effect of compaction moisture content (Meehan et al. 2012)

Dependent Variable	Coefficient of Determination R^2
Soil Stiffness Gauge SSG/ GeoGauge GG	0.266
LWD 300	0.475
LWD 200	0.439
DCP -M	0.580
DCP-A	0.571

Additionally, multivariate regression of the same data does increase the correlation between NDG readings and modulus readings, but introduces anther level of complexity into the calculation/comparison of NDG density readings to modulus readings. This process would be difficult for the average construction inspector to generate.

Methodology

If MoDOT (QA) and contractors (QC) use different compaction measurement systems, there must be a relation or correlation between modulus measurements with density measurements. To establish if two variables are related one must build an empirical model based on observed data. The following empirical model is developed from a scatter diagram of NDG data and density data from the TransTech Soil Density gauge, CIV data from the 10 KG Clegg Impact Hammer and modulus/ stiffness data from the LWD and DCP. From Montgomery, Runger and Hubele (2007), if a relationship exists between two variables then a response variable Y is related to a regressor or predictor variable x in a simple linear regression model:

$$Y = \beta_0 + \beta_1 + \varepsilon \qquad (7\text{-}1)$$

Where: β_0 and β_1 are unknown regression coefficients and

ε is a random error.

For the above linear regression model there is an expected value of Y for each value of x. β_0 is the Y axis intercept and β_1 is the slope of the line or the mean change in Y for a unit change in variable x. The linear regression consists of finding the best fit straightline through the points on the scatter plot. This best fit line is determined using by minimizing the sum of the squares of the vertical deviations. This estimation process used to determine β_0 and β is called the method of least squares. For this study, the independent variables or predictors are the NDG dry densities with the dependent variables being LWD modulus values or Clegg impact values (CIV).

MoDOT conducted these comparison tests in order to become familiar with the alternate testing devices, testing procedures, testing times, costs, and ease of use. An additional goal of the tests was to provide valid local results in which road construction contractors may have confidence.

The series of comparative tests were conducted at locations on active or recent grading projects. The tests were usually conducted in the following manner with some changes depending on the devices being tested.

1. The test location was smoothed out and the first test was conducted with the TransTech Soil Density Gauge (SDG).

2. IA pilot hole for the NDG probe was driven in the middle of the test area. t

3. The NDG was placed and the probe extended into the hole.

4. The first compaction test was run for a 4 minute count.

5. After the first test was complete the NDG was turned 180 degrees and another 4 minute test was run.

6. A Zorn LWD with a 300mm plate was placed over the outline of the first NDG test and a standard six drop test was conducted.

7. A 10 kg Clegg Hammer, 4 drop test was conducted over the outline of the second NDG test.

For compaction testing configuration refer to Figure 3-3.

Analysis and Results

To determine the strength of the relationship between x and y or how well the data fits the regression line the coefficient of determination is used. The coefficient of determination ranges for 0 to 1. An R^2 of 0 means that y cannot be determined from x. An R^2 of 1 means that y can be predicted from x without error. A coefficient determination of 0.8 means that 80 percent of the variation can be explained by the linear relationship between x and y with the other 20% being unexplained. The graph in Figure 7-2 is an example of the a linear regression model and Table 7-7 shows the calculated coefficients of determination for samples taken in the same locations using the two different compaction testing devices.

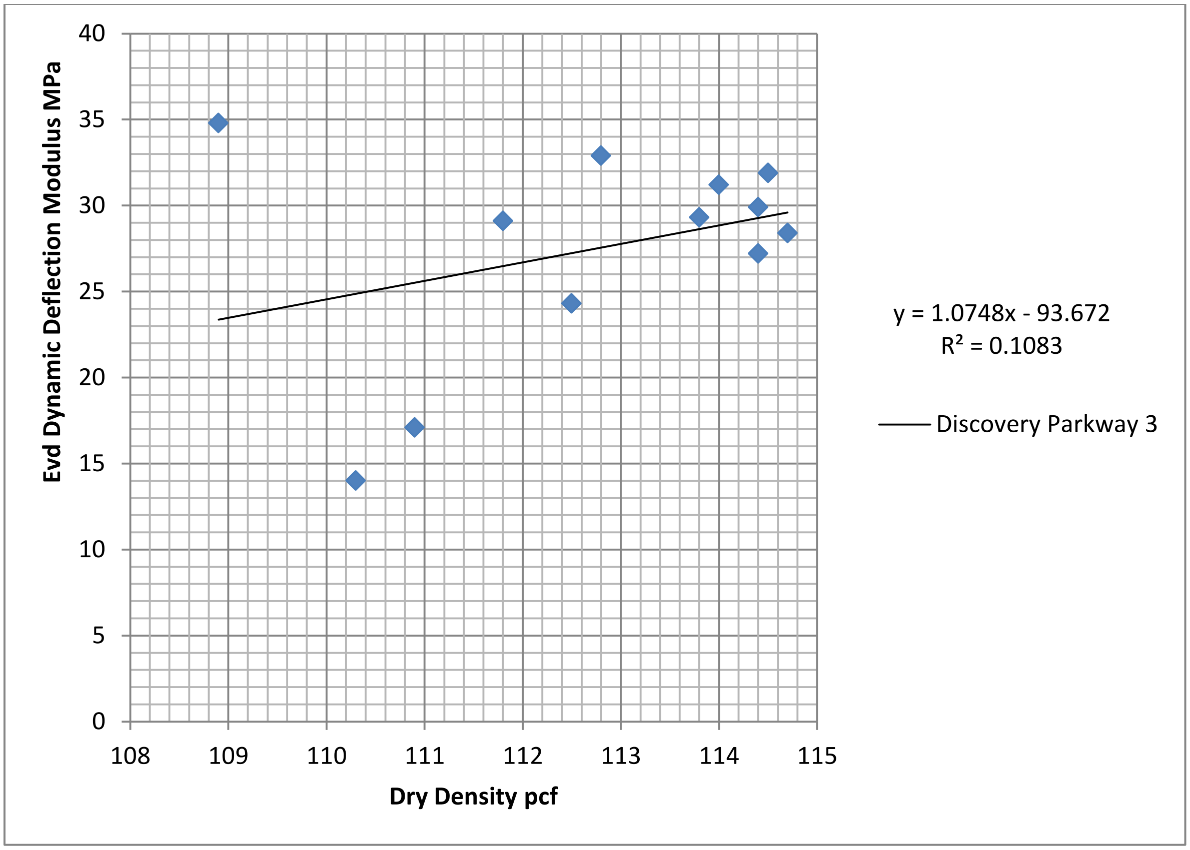


Figure 7-1 Dry Density VS LWD modulus values at the Discovery Parkway, Boone Co.

Table 7-7 Linear Regression Results

Location	Dependent Variable	Independent Variables	Linear Equation	R^2
Capital Quarries Cole Co.	LWD modulus	Dry Density (DD)	Evd= -156.57 + 2.082(DD)	0. 3616
Discovery Parkway Boone Co.	LWD modulus	Dry Density (DD)	Evd= -93.672 + 1.0748(DD	0.1083
Route 50 Osage C0. East CO. RD 401 and 604	CIV	Dry Density (DD)	CIV = 12.846 + 0.022(DD)	0.0024
Route 50 Osage C0. East Co Rd 602	CIV	Dry Density (DD)	CIV =17.56 + 0.2588(DD)	0. 676

As mentioned previously, multiple regression (a.k.a. least squares multiple linear regression) generates better predictions than simple linear regression. A multiple regression equation can take the form of:

$$\text{Dependent Variable} = C0 + V1 \times C1 + V2 \times C2 + \ldots + Vn \times Cn \quad (7\text{-}2)$$

Where: C0 =Intercept value

V1 = Value of first independent variable

C1 = First coefficient linked to first independent variable

V2 = Value of second independent variable

C2 = Second coefficient linked to second independent variable

n = number of independent variables

For multiple linear regression models for comparisons between Modulus values/Clegg Impact Values (CIV) and dry density (DD) and percent moisture (m %) the equation takes the form of

$$\text{Modulus Values/Clegg Impact Values (CIV)} = C_0 + DD \times C_1 + m\% \times C_2 \quad (7\text{-}3)$$

Multiple regression analysis was run using a commercial spreadsheet. In order to validate the mode, 70 percent of the compared values are used to build the multiple regression models, while the remaining 30 percent of the comparisons were reserved to evaluate the model's performance. When going through the validation steps using Microsoft Excel, p-values are calculated for the generated model. Low p-values $p< 0.05$ indicates that the independent variable is expected to be a significant addition to the model because changes in the independent variables value are associated to changes in the dependent variable. When maximizing the coefficient of determination, independent variables with p values greater than 0.05 should be removed from the equation. Generally the p values for percent moisture were greater than 0.05 but were not removed because the purpose of the multiple regression was to assess the effects of moisture on the prediction of modulus or CIV values.

Table 7-8 Multiple Regression Results

Location	Dependent Variable	Independent Variables	Linear Equation	R^2
Capital Quarries Cole Co.	LWD modulus	Dry Density (DD) %Moisture (%M)	Evd= -36.7169 + 1.566674(DD) - 9.64184(%M)	0.27617
Discovery Parkway Boone Co.	LWD modulus	Dry Density (DD) %Moisture (%M)	Evd= 153.1378 - 0.87329(DD)- 1.63597(%M)	0.282274
Route 50 Osage C0. East CO. RD 401 and 604	CIV	Dry Density (DD) %Moisture (%M)	CIV = 2.518564 + 0.12262(DD) - 0.40817(%M)	0.304006
Route 50 Osage C0. East Co Rd 602	CIV	Dry Density (DD) %Moisture (%M)	CIV = -32.5083 + 0.374489(DD) + (0.183126)(%M)	0.780453

A commercial neural network program was also used to in an effort to improve predictions for CIV and LWD modulus from patterns from a data set of nuclear gauge dry density and moisture readings along with corresponding modulus and CIV readings. The network is initially trained from the data points and the relationship between the points. The network can then predict a value from data fed into it. Comrie (1997) demonstrated that neural networks for predicting ozone concentration performed better than regression models. For predicting ozone concentrations for eight cities, the coefficient of determinations, R^2, increased on the average of 13.75%, using neural networks compared to multiple regression.

Table 7-9 Neural Network Results

Location	Dependent Variable	Independent Variables	Linear Equation	# of observations (n)
Capital Quarries Cole Co.	LWD modulus	Dry Density (DD) %Moisture (%M)	Evd= -218.04.7169 + 3.490(DD) – 14.23(%M)	n=9
Discovery Parkway Boone Co.	LWD modulus	Dry Density (DD) %Moisture (%M)	Evd= 141.94 - 0.2566(DD)-5.660(%M)	n=12
Route 50 Osage C0. East CO. RD 401 and 604	CIV	Dry Density (DD) %Moisture (%M)	CIV = 11.29 + 0.01805(DD) - 0.1759(%M)	n= 23
Route 50 Osage C0. East Co Rd 602	CIV	Dry Density (DD) %Moisture (%M)	CIV = -14.64 + 0.2373(DD) - 0.05549 (%M)	n=18

Conclusions

The conclusions reached in this chapter are as follows:

- The results from both linear regression and the multiple regression show that there is no definitive relationship between LWD modulus and Clegg CIV values in both lean clays to clays (Route 50, Osage Co. & Discovery Parkway, Boone County) and in sand (Capital Quarries, Cole Co.).
- The Clegg did show moderate relationship to the NDG density with coefficient of determination values of 0.676 (simple linear regression) and 0.780453 (multiple regression) at the testing site east of County Rd 602 on the Route 50 Project.

- However the Clegg and NDG density had a very poor to poor relationship at another site on Route 50 project (east of CO. Rd 401 and 604) with R^2 values of only 0.0024 from simple linear regression and 0.3046 with a multiple regression analysis.

Therefore no definitive relationship between NDG and Modulus/CIV results could be found from Missouri test sites. This confirms the findings of Meehan and Li and leads the authors to conclude that in the QM process for testing compaction of soils in embankments and cuts, both QC an QA must be verified with the same testing apparatus and method. Therefore, if MoDOT decides to cease using the NDG then its contractors will not be able to use it either.

CHAPTER 8. COMPARATIVE ANALYSIS OF REPEATABILITY AND REPRODUCIBILITY OF COMPACTION TESTING

McLain, K .W., Bumblauskas, D.P., D. J. White, D. J., and Gransberg, D.D., "Comparative Analysis of Repeatability and Reproducibility of Compaction Testing," *Transportation Research Journal Part B: Methodological*. Submitted November 13, 2015.

This chapter introduces and discusses the concept of repeatability and reproducibility for the light weight deflectometer and the dynamic cone penetrometer, two devices that show promise as possible replacement for the MoDOT's standard nuclear density gauge. This Chapter looks at 3 distinct methods of data statistical assessment to determine the repeatability and reproducibility of the two different testing devices evaluated in a construction grading project site.

Abstract

Investigating possible alternatives for soil compaction testing to potentially replace the standard nuclear density gauge entails examining all the testing methods for reproducibility and repeatability. To ensure good road base course performance, the quality control and assurance testing for soil compaction testing must be consistent with one another not only in procedure but measurement as well. The objective is to find the appropriate means for measuring repeatability and reproducibility in the field for the testing devices. This paper considers both measurement system analysis in the form of gauge repeatability and reproducibility and statistical data assessments including: (1) coefficient of variation: (2) statistical error compared to the mean; (3) one-way ANOVA; and (4) hypothesis test for paired samples. It was found that the field test data

(Dynamic Cone Penetrometer and Light Weight Deflectometer results) exceeded gauge repeatability and reproducibility acceptance guidelines. Hypothesis testing and one-way ANOVA results depend on the level of confidence employed by the tester/analyzer and the expertise of the analyzer to make a decision to reject or accept that the data means produced by the testers are essentially equal. This study found that coefficients of variation and standard error to the mean produced the best results that fit with previous studies and are generally understood by individuals without expertise in statistics.

Background

The standard method for testing construction subgrade compaction for the Missouri Department of Transportation (MoDOT) is the American Association of State Transportation Officials (AASHTO) test method T310, *In Place Density and Moisture Content of Soil and Soil-Aggregate by Nuclear Methods (Shallow Depth).* The nuclear density gauge (NDG) has been the standard instrument for compaction testing since the late 1970's. The reliability of the nuclear density gauge (NDG) and the ability to determine gravimetric moisture content have made routine compaction testing quick and straightforward for field inspectors. However, the speed and convenience of the NDG comes with a price which includes inspector licensure with the federal government and required safety training, as well as special storage, transport, and field security procedures. Annual audits of Resident Engineer Offices where the gauges are stored and dispatched to local projects are also mandated. The department allows the use of AASHTO test methods T191 (Sand Cone) as an alternatives to NDG testing, however, this test is rarely employed because it is considered comparatively slow and require samples be sent to a dedicated laboratory for moisture content analysis.

The Construction and Materials Division of MoDOT is actively investigating alternatives guided by the cost and time of the processes required when using the NDG and the AASHTOWare™ Pavement ME Design, where mechanistic-empirical (ME) design procedures no longer are based on density-moisture content requirements (Pavement design is increasing emphasizing the importance of achievement of minimum subgrade/base modulus rather the density-moisture alone.

The investigative process described in this paper looks at selected alternative testing equipment and methods. Price, portability, testing time, ease of use, calibration requirements, accuracy, repeatability and reproducibility were parameters considered in comparing the various tests. To have confidence in a method and avoid conflicts between owner and contractor, the equipment and its associated testing protocol need to be accurate, repeatable and reproducible from operator to operator (gauge repeatability and reproducibility or GRR). While GRR has been used extensively in other applications, such as production manufacturing, quality control, and process improvement, the technique has rarely been used in field soil compaction applications. Therefore the objective of this research is to apply GRR in the comparative analysis of compaction testing devices and include its output in the decision process for choosing viable alternatives to the NDG.

Tested devices for this study included the Zorn ZFG 2000 Light Weight Deflectometer (LWD), and Dynamic Cone Penetrometer (DCP). The investigation of repeatability and reproducibility was done in the field on active construction sites rather than in the laboratory with technician-prepared soil filled drums/tubs or test strips (Mazari et al. 2013)

Methodology

Field Testing Procedure

Four sites were used to assess repeatability and reproducibility of measurements for the LWD, and DCP. Tests of the three alternatives and the NDG were conducted on the following four construction project structural fills with the details on testing procedures and soil types are shown below in Table 8-1:

Table 8-1 Testing Locations, Soil types and Compaction Devices

Devices Tested	Sites	County	Soil Type	Procedure
LWD/DCP/SDG	Route 364 @K	St. Charles	CL	1 site /5 trials/ 2 testers
LWD/DCP/SDG	Route 364 @ N	St. Charles	CL	1 site /5 trials/ 2 testers
LWD/DCP/SDG	Capital Quarries	Cole	Manufactured sand	1 site /5 trials/ 2 testers
LWD/DCP	Missouri River Bridge @ Rte 54	Callaway	CL to SC	1 site /10 trials/ 2 testers
LWD/DCP	Discovery Parkway	Boone	SC	10 sites/5 trials/ 2 testers

Figure 8-1 Nuclear Density Gauge (NDG)

Figure 8-2 Dynamic Cone Penetrometer (DCP)

Figure 8-3 Light Weight Deflectometer

Prior to testing, the test locations were smoothed out using a hand shovel or a nuclear density gauge scraper plate .A nuclear density gauge (NDG) is used to produce two differing test areas and also a point of comparison. A nuclear gauge reading is taken and then the gauge is turned 180 degrees and a subsequent reading is taken. In the limits of the outline of the nuclear gauge test five DCP readings per two testers are taken approximately three inches apart (see Figure 8-4). This procedure usually limits the number of testers to two. In the second NDG test area five test trials of the LWD per tester are conducted, with the first tester performing the seating blow. This paper reports the testing results from the LWD and DCP devices conducted at the Discovery Parkway project located just south of Columbia, Missouri.

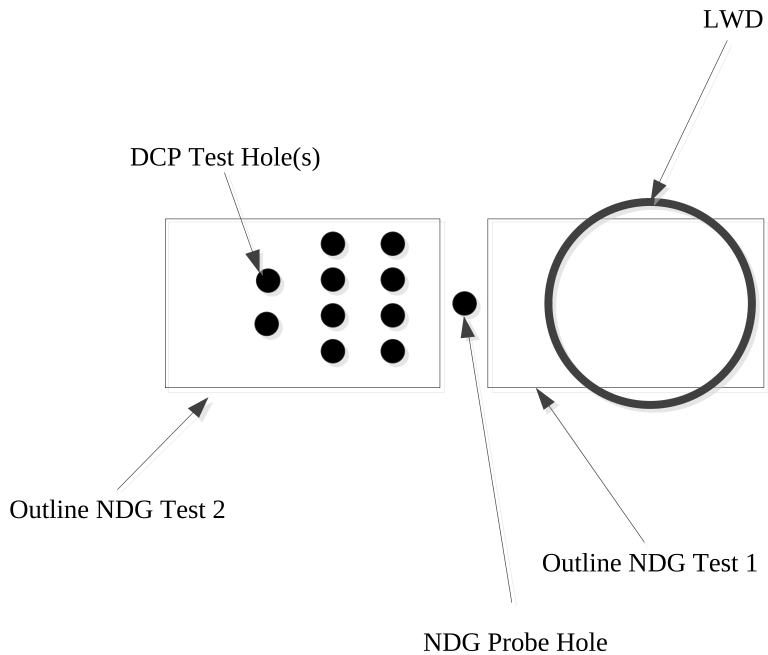


Figure 8-4 Testing Setup

Accuracy and Precision of Measurements

The terms accuracy and precision are often used synonymously, which is not technically correct. The following definitions are used by the authors for accuracy, precision and other related terms that will be used throughout the report. The following definitions are used in this study:

- *"Accuracy*: The "proximity" to the true value or an accepted reference value.

- *Precision*: The "proximity" of repeated readings of one to another. A random error that is part of the observed measurement system.

- *Bias*: The difference between a reference value and the average of measurement. An error introduced by inaccuracy and is inherent to the system.

- *Linearity*: The change in over the normal operation range. An error introduced by inaccuracy and is inherent to the system.

- *Stability*: The change in bias over time often referred as drift.

- *Repeatability*: The variation in measurement acquired by a measuring instrument or device used by one appraiser or operator measuring the characteristic of the same part several times. Also referred to as Equipment Variation.

- *Reproducibility*: The variation in the average of measurements made by differing appraisers or operators using the same gage or device while measuring a part characteristic. Also referred to as Appraiser Variation." (AIAG, 2010).

AIAG Method

The AIAG Methods are defined by the Measurement System Analysis (MSA) Manual (4th edition). The MSA manual covers three different methods of analysis

- The Range method
- The Average and Range (A&R) method
- Analysis of Variance (ANOVA) method

The authors used the A&R method to investigate compaction test devices. The A&R method can estimate both repeatability and reproducibility with differing parts' role in the precision error of measurement. The A&R method can also estimate total precision error of measurement. This method allows for differing parts to be measured by several operators with several trials. The differing soil locations are the differing parts and are being measured by the compaction test devices several times with different operators. The A&R Method, however, does not consider the operator and device interaction.

For purposes of illustration, the MSA measurement unit analysis and percent total variations are presented Equations 8-1 through 8-10 below.

Repeatability:

$$EV = \bar{\bar{R}} \text{ x } K_1 \qquad (8\text{-}1)$$

Where:

EV = Equipment Variation (Repeatability)

$\bar{\bar{R}}$ = Average range of trials

K_1 = Constant that depends on the number of trials conducted and is the inverse of d^*_2 (from Table C1 MSA Manual)

Reproducibility:

$$AV = \sqrt{(\bar{x}_{DIFF} \text{ x } K_2)^2 - \left(EV^2/(nr)\right)} \qquad (8\text{-}2)$$

Where:

AV = Appraiser Evaluation (Reproducibility)

$\bar{x}_{DIFF}$ = Maximum average readings per appraiser minus minimum aver readings per appraiser

K_2 = Constant that depends on the number of appraisers and is the inverse of d^*_2 (from Table C1 MSA Manual)

$n =$ number of parts

$r =$ number of trials

Repeatability and Reproducibility:

$$GRR = \sqrt{(EV)^2 + (AV)^2} \quad (8\text{-}3)$$

Where:

GRR = Repeatability and Reproducibility

Parts Variation:

$$PV = R_P \times K_3 \quad (8\text{-}4)$$

Where:

PV = Parts Variation

R_P = Range of part averages

K_3 = Constant that depends on the number of parts (from Figure III-B 16: Gage Repeatability and Reproducibility Report, MSA Manual)

Total Variation:

$$TV = \sqrt{GRR^2 + PV^2} \quad (8\text{-}5)$$

Where:

TV = Total Variation

GRR = Repeatability and Reproducibility

Percentage of Total Variations:

$$\%EV = 100[EV \;/\; TV] \tag{8-6}$$

$$\%AV = 100[AV \;/\; TV] \tag{8-7}$$

$$\%GRR = 100[GRR \;/\; TV] \tag{8-8}$$

$$PV = 100[PV \;/\; TV] \tag{8-9}$$

$$\%EV + \%AV + \%PV \neq 100 \tag{8-10}$$

Wheeler's HG Method

Wheeler (2009) proposed an alternate to the AIAG GRR method, which he called "an honest GRR study". It is designated as the HG Method in this report. The HG Method differs from the AIAG method in that the sum of the components of measurement (Equations 11 through 15) equals the Total Variation.

$$\%EV = EV^2 \;/\; TV^2 \tag{8-11}$$

$$\%AV = AV^2 \;/\; TV^2 \tag{8-12}$$

$$\%GRR = GRR^2 \;/\; TV^2 \tag{8-13}$$

$$\%PV = PV^2 \,/\, TV^2 \tag{8-14}$$

$$\%EV + \%AV + \%PV = 100 \tag{8-15}$$

The question that arises is: are the AIAG GRR and HG GRR accurate measurement systems for determining repeatability and reproducibility of the DCP and LWD compaction testing methods. The MSA manual furnishes general GRR criteria guidelines as shown in Table 8-2.

Table 8-2 AIAG MSA Manual Acceptance Guidelines

GRR less than 10 percent	**Acceptable or Good**
GRR from 10 to 30 percent	Marginal or Acceptable for some applications
GRR greater than 30 percent	Unacceptable

Coefficient of Variation and Standard Error to the Mean

Coefficients of Variation (COV) of the results were calculated from the trials performed by the two differing operators. The COV is defined as the ratio of the standard deviation to the mean.

$$COV = {}^{\sigma}/_{u} \tag{8-16}$$

Where:

COV = Coefficient of Variation

σ = Standard Deviation

u = Mean

The COV is useful because it is dimensionless and measurements using other units and differing means can be compared. In contrast, standard deviations themselves are in the context of the measured data and cannot be effectively compared to data with differing units.

The standard error (SE) is the standard deviation of a sampling distribution (Montgomery et al. 2007). The standard error can give an indication of the estimation of precision for a grouping of data. Equations 17 and 18 are used to calculate the SE and average SE.

$$SE = \sigma \,/\, \sqrt{n} \tag{8-17}$$

$$SE_{avg} = \bar{x} \,/\, SE \times 100 \tag{8-18}$$

Where:

$SE =$ Standard Error

$n =$ Number of Trials

$SE_{avg} =$ Standard Error in comparison to the sample mean.

$\bar{x} =$ Average value of trials (sample mean)

Testing results were also analyzed using the statistical method of one-way ANOVA. The one-way ANOVA compares the means of data from differing groups (aka two differing operators performing compaction tests). The ANOVA statistic tests the null hypothesis. For one way ANOVA the general assumptions are normality, equal variance and independence of errors (Seltman 2015).

$$H_0 = u_1^* = u_2^* = u_3^* = u_k^* \tag{8-19}$$

Where:

$H_0 =$ null hypothesis

$u^* =$ group or operator mean

$k =$ number of groups

If the one-way ANOVA test statistic gives a significant result then the null hypothesis is rejected and we accept the alternate hypothesis H_A.

$$H_A \neq u_1^* \neq u_2^* \neq u_3^* \neq u_k^* \tag{8-20}$$

To test the null hypothesis ANOVA uses the F-Statistic. The F statistic ratio is calculated form the following:

$$F = MS_{between} \,/\, MS_{within} \tag{8-21}$$

Where:

w

MS_{within} = Mean of Squares within a group or operator

$$MS_{within} = ss_{within} \,/\, df \tag{8-22}$$

$$MS_{between} = ss_{between} \,/\, df \tag{8-23}$$

Where:
$ss_{within} =$ Sum of squared deviations from the mean within a group or operator
$ss_{between} =$ Sum of squared deviations from the mean between groups or operators
$df =$ Degrees of freedom

Generally, F-Statistics are near 1.0 when the null hypothesis is true and usually larger when the alternative hypothesis is true. The F-statistic can be compared to the F-critical. If the F-statistic is less than F-critical then the null hypothesis is thought to be true. Also the p-value can be compared to the alpha value or significance level, usually 0.05. To keep the null hypothesis, the p-value must be larger than α. The authors used commercially available programs to perform the one-way ANOVA statistical tests.

Hypothesis Testing

Like one-way ANOVA, hypothesis testing statistics requires that the null hypothesis has no significant difference between the means of groups or testers as seen in equation 19. The alternative hypothesis states that the test means are significantly different as presented in equation 20. The test statistic (Equation 24) compares a mean of a test to a hypothesized value or overall test mean value as exhibited below:

$$z_{stat} = \bar{x} - u_0 \ / \ SE \quad (8\text{-}24)$$

Where:

z_{stat} = Test statistic

$\bar{x}$ = Mean of a test

SE = Standard Error. from equation 8-17.

The z_{stat} divides the area under a normal distribution curve into rejection and nonrejection regions for the null hypothesis. From the z_{stat}, a p-value is calculated. This is easily done using a statistical computer program. The p-value supplies support against or for keeping the null hypothesis. The p-value is compared against a threshold value called the level of significance or alpha (α).The authors have applied the following statistical convention (Gertsman 2015):

Table 8-3 Significances of Established P-Values

p-value compared to alpha value α	Observed Difference
p-value > 0.10	Not Significant
p-value ≤ 0.10	Marginally Significant
p-value ≤ 0.05	Significant
p-value ≤ 0.01	Highly Significant

Analysis and Results

The DCP readings from the Discovery Parkway project in blows per Inch (BPI) for 8+ inches of penetration for five separate trials for the 10 sites are presented below for testers A and B.

Table 8-4 DCP Results Discovery Parkway

Test Site #	Tester A Average of 5 Trials BPI	Tester B Average of 5 Trials BPI
Site 1	0.3338	0.3336
Site 2	0.3767	0.2407
Site 3	0.3237	0.2872
Site 4	0.3104	0.2861
Site 5	0.3261	0.2907
Site 6	0.2897	0.2724
Site 7	0.2977	0.2770
Site 8	0.3392	0.2915
Site 9	0.2819	0.2734
Site 10	0.2932	0.2852

The results are for five trials of 10 sites on the Discovery Parkway project site. The Zorn LWD results come in two forms: dynamic deflection modulus, (E_{vd}) in mega-newtons per squared meters, (MN/m^2), and settlement (s) in millimeters (mm).

Table 8-5 LWD results Discovery Parkway

Test Site #	Tester A Average of 5 Trials E_{vd} (MN/m^2)	Tester B Average of 5 Trials E_{vd} (MN/m^2)
Site 1	4.40	4.60
Site 2	3.92	4.10
Site 3	3.98	4.04
Site 4	3.98	4.30
Site 5	3.86	4.04
Site 6	3.62	3.96
Site 7	3.40	3.82
Site 8	3.82	4.02
Site 9	3.80	3.98
Site 10	3.26	3.70

The AIAG and HG reproducibility and repeatability measurement results are shown in Table 8-6 below. For both the AIAG and HG methods the % GRR (Repeatability and Reproducibility) exceeds the 30 percent failure threshold of acceptability. The considerable numbers were not unanticipated given the testing was conducted on soil using standard construction compaction techniques employing heavy equipment.

Table 8-6 Gauge R&R DCP Discovery Parkway 10 Sites

	AIAG Method	HG Method
%EV Repeatability: Equipment **Variation**	72.21	52.15
%AV Reproducibility: Appraiser Variation	55.14	30.41
%GRR Repeatability and Reproducibility	90.86	82.55
%PV Parts Variation	41.77	17.45
%PV+%EV+%AV	169.12	100

The AIAG protocol states that if the range for an individual trial exceeds the calculated Upper Control Limit (UCL_R) for range for the entirety of the trials, that that trial(s) be redone or discarded and the upper control limit be recalculated for the remaining trials. The Upper Control Limit is calculated from the following equation:

$$UCL_R = \overline{\overline{R}} \times D_4 \qquad (8\text{-}25)$$

Where:

UCL_R =Upper Control Limit

$\overline{\overline{R}}$ = Average of ranges for all trials

D_4 = factor based on the number of trials

For the 10 sites a UCLR of 0.1507 was calculated. For tester A, Sites 2, 3, and 8 ranges met or exceeded the Upper Control Limit, and using AIAG protocol, Sites 2, 3 and 8 were removed for both Testers, giving seven remaining sites for which to evaluate AIAG and HG Gauge R&R. This is shown in Figure 8-5.

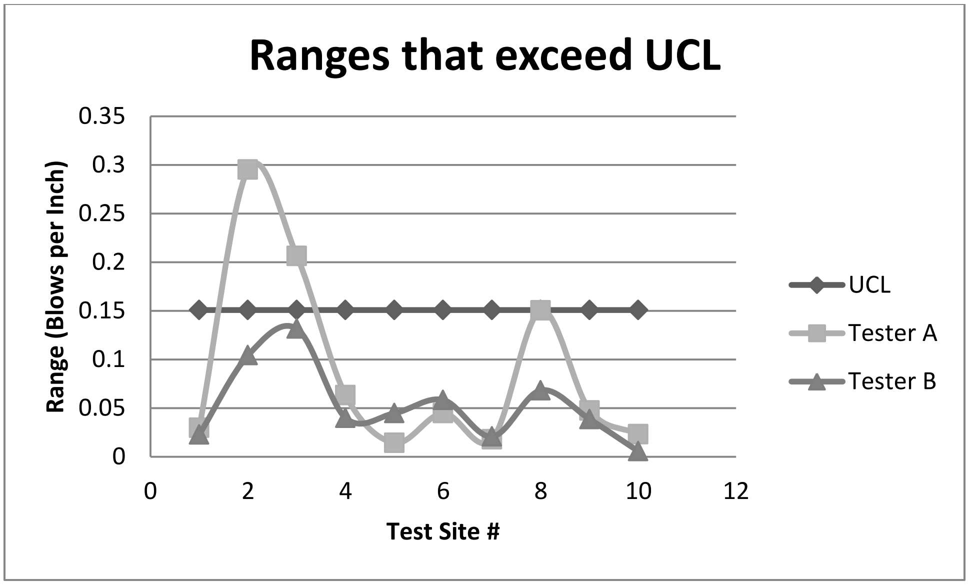


Figure 8-5 Ranges Exceeding Upper Control Limits for DCP Gauge R&R

Taking Sites 2, 3, and 8 out of the calculation lowered GRR by about 23 percent points but the figure was still over the maximum acceptance level by about 38 percentage points. The parts variation increased from 10 site set-up to the 7 site scenario because of the decrease in the number of parts (soil sites) where the K_3 constant increased.

Table 8-7 Gauge R&R DCP Discovery Parkway 7 sites

	AIAG Method	HG Method
%EV Repeatability: Equipment Variation	53.25	28.36
%AV Reproducibility: Appraiser Variation	41.97	17.62
%GRR Repeatability and Reproducibility	67.81	45.98
%PV Parts Variation	73.50	54.02
%PV+%EV+%AV	168.73	100

The LWD AIAG and HG gage R&R 10 Site tests displayed similar results for the DCP 7 Site results with the LWD GRR being 10 percent lower than the AIAG accepted DCP results. The GRR results still exceeded AIAG standards for an acceptable system.

Table 8-8 Gauge R&R LWD Discovery Parkway (Dynamic Deflection Modulus)

	AIAG Method	HG Method
%EV Repeatability: Equipment Variation	34.48	11.89
%AV Reproducibility: Appraiser Variation	45.37	20.58
%GRR Repeatability and Reproducibility	56.99	32.47
%PV Parts Variation	82.17	67.53
%PV+%EV+%AV	162.03	100

Mazari et al. (2013), in laboratory conditions found for Zorn LWD, referred as a Portable Impulse Plate Load Device (PIPLD), the following Average and Range method results:

Table 8-9 Gauge R&R Results for LWD in Laboratory Environment

EV% - Repeatability	**AV% - Reproducibility**	**R&R**	**% PV –Parts Variation**
1	0.1	1	99

The good repeatability and reproducibility results were from the research team rigidly controlling the moisture content and density of the soil being tested. The standard deviation of moisture content for all prepared lifts and specimens were 0.5 % with a range of 0.9 % . The mean moisture content was 0.1% below optimum moisture content OMC.

The optimum moisture content for the Discovery Parkway site was 15.5%. Nuclear gauge moisture measurement for the 10 LWD subsites averaged 15.66 % with a range of 3 percent and standard deviation of 0.92%. The maximum dry density for the site was reported at 111.5 pcf. The average nuclear gauge readings for dry density of the 10 subsites were 106.7 pcf. The standard deviation for the site was 2.26 pcf with a range of 7.5 pcf.

COVs were calculated for the DCP. The COV values improved by about a factor of two when the three sites 2, 3, and 8 were removed through the AIAG conventions due to the large ranges encumbered by Tester A. White et al. (2009) conveyed COVs for DCP of 20% to 32%, measuring 12 in. deep on test strips. White et al. (2009), also reported COVs ranging from 29% to 61% for Zorn LWDs tested in cohesive to granular subgrades. Prima 100 LWDs (Alshibli et al. 2005) tested in laboratory conditions had COVs that ranged from 1.2 % in clay to 55.8% percent in sands, but for eight clay samples (soils like that found on Discovery Parkway site) the average COV was 18.2%. Nazzal et al. (2007) reported Prima 100 LWD COV results that varied from 2.1% to 28.1% for various highway construction bases and subgrades. It was noted that COV value decreased as the LWD elastic moduli increased.

Table 8-10 Percent Coefficients of Variation

	Tester A	Tester B	Combined
DCP Disc. Parkway 10 sites Blows/Inch	8.59	7.68	9.92
DCP Disc. Parkway 7 sites Blows/ Inch	4.94	6.78	6.91
LWD (Modulus) Disc. Parkway 10 sites	8.00	5.82	7.64

Removing the outliers, (sites with test ranges outside AIAG specifications) decreased the COV for tester A approximately 3.5 % and the combined COC by about 3 percent. The COVs calculated for the DCP were lower than found in White et al. (2009), but more variation can be expected in DCP tests conducted in granular subgrades. The COVs for the tested Zorn LWD trended on the lower end when compared to the Prima 100 LWDs, but were in the range of reported results.

In attempting to take the soil variation from the entire site out and give an indication of reproducibility between testers , COV's for individual LWD trials were calculated and compared, as shown in Table 8-11. Percent change from average varied as little of 1.23 percent to 119 percent.

Table 8-11 LWD COV for Individual Test Sites

Trial No.	Tester A COV	TESTER B COV	Average COV	Difference in COV	Percent Change from Avg. COV
Trial 1	0.0407	0.0238	0.0323	0.0169	52.40
Trial 2	0.0338	0.0154	0.0246	0.0184	74.80
Trial 3	0.0101	0.0198	0.0150	0.0097	64.88
Trial 4	0.0582	0.0147	0.0365	0.0435	119.34
Trial 5	0.04209	0.037	0.0395	0.0051	12.87
Trial 6	0.0506	0.0342	0.0424	0.0164	38.68
Trial 7	0.0372	0.0256	0.0314	0.0116	36.94
Trial 8	0.0305	0.0186	0.0246	0.0119	48.47
Trial 9	0.0166	0.0246	0.0206	0.0080	38.83
Trial 10	0.0245	0.0242	0.0244	0.0003	1.23

The Standard Error in percent of averages of around 2 percent was calculated for the DCP and LWD. They show a good accurate point of estimate for both average blows per inch for 8 inch depth for the DCP and modulus readings with the Zorn LWD for the 10 sites on the Discovery Parkway Site.

Table 8-12 Standard Error in Percent of Average

	Tester A	Tester B	Combined
DCP Disc. Parkway 10 sites Blows/Inch	2.72	2.43	2.22
DCP Disc. Parkway 7 sites Blows/Inch	1.87	2.56	1.85
LWD (Modulus) Disc. Parkway 10 sites	2.53	1.84	1.71

One way ANOVA and Hypothesis test results for paired samples are displayed below. The statistical methods can be used to look at the reproducibility of each tester. The p- values for both methods were generated by commercially available software. The methods differ as to what significance level to reject or fail to reject the null hypothesis, that the difference between the means of the test results conducted by the two different testers are essentially equal.

Table 8-13 ANOVA and Hypothesis Test

	Significance Level	One Way ANOVA		Hypothesis Test Paired Samples	
		Null Hypothesis	**p-value**	**Null Hypothesis**	**p-value**
DCP Disc. Parkway 10 sites Blows/Inch	1%	Fail to Reject	0.0100	Fail to Reject	0.0236
	5%	Reject	0.0100	Reject	0.0236
	10%	Reject	0.0100	Reject	0.0236
DCP Disc. Parkway 7 sites Blows/Inch	1%	Fail to Reject	0.1575	Fail to Reject	0.0106
	5%	Fail to Reject	0.1575	Reject	0.0106
	10%	Fail to Reject	0.1575	Reject	0.0106
LWD (Modulus) Disc. Parkway 10 Sites	1%	Fail to Reject	0.0653	Reject	0.0001
	5%	Fail to Reject	0.0653	Reject	0.0001
	10%	Reject	0.0653	Reject	0.0001

Conclusions and Limitations

For the MoDOT personnel and partnering contractors looking at the systems, the most understandable and useful statistics are the Coefficient of Variation and the Standard Error in Percent of Average. The COV is also a useful comparative element since it is unit-less; this allows for comparison among the differing testing devices that produce dissimilar test results. The key in understanding the concept of COV is the test data with the smaller COV is less dispersed than the variable with the larger COV (IDRI 2015) (Note using the same units or engineering parameters). In a field test comparing the ten differing sites, the COV displays the amount of variation in the sites. Individual site COV’s show the variability between testers.

The LWD Tests for each of the 10 sites were conducted with Tester A performing the initial three seating blows then conducting 5 sets of three drops. They recorded the average dynamic modulus and settlement after each three drops. Then Tester B repeated the process excluding the initial seating blows. Figure 8- 2 shows that Tester B had higher average modulus readings than Tester A. This would indicate that after the initial three seating blows that the soils of the Discovery Parkway project were still being compacted from Tester A drops.

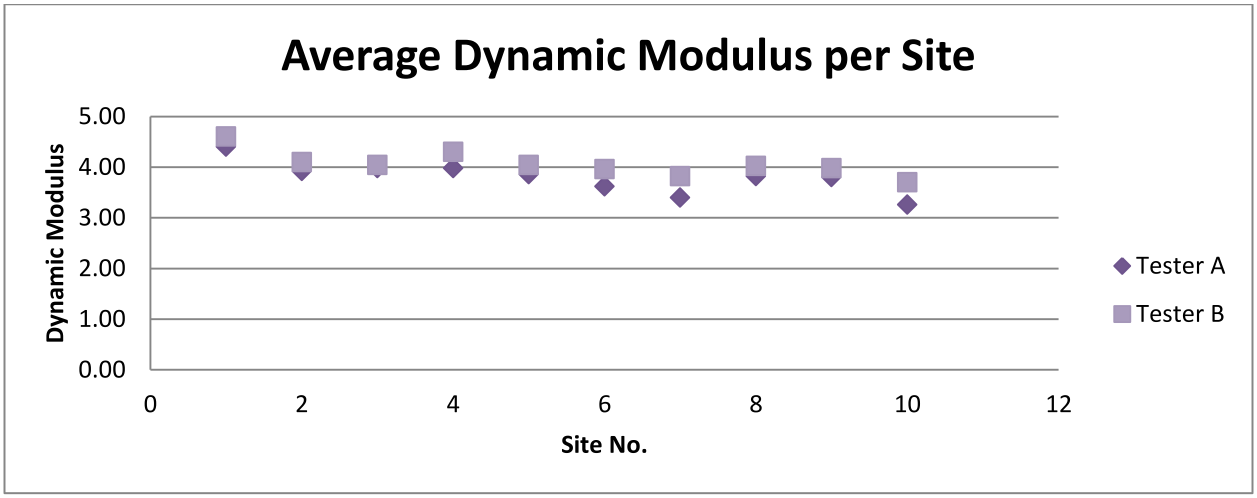


Figure 8-6 LWD Average Dynamic Modulus per Tester per Site

The difference in average modulus measurement ranged from 0.06 MPa to 0.44 MPa with the average difference being about 0.254MPa. A stiffer soil site would have displayed less variation and given a better indication of repeatability and reproducibility especially within individual trials.

The DCP Tests were not true repeatable tests because the test is a destructive test and the distinct soil columns were obliterated. The test had to be averaged over the outline of the initial nuclear density test. This procedure introduced further variability into the measurements. The soil and degree of compaction on the overall test sites varied under concentrated testing terms but was fairly uniform in standard construction procedures.

The AIAG GRR method's stated thresholds or limits are subjective and there is no support for the limits in the MSA manual (Wheeler 2009). When Equipment Variability (repeatability) is found to be greater than Appraiser Variation (Reproducibility) as seen in the DCP tests, the probable causes are that the gage needs repair or replacement or there is excessive within part variation (Pandiripalli 2010). The excessive within part variation is likely for the DCP tests in which every trial for both testers was an individual test in a varying medium (soil on a project). For Reproducibility greater than Repeatability, appraisers or operators need better training or the testing equipment needs to be recalibrated. In the case of LWD testing the part (soil) was changed in the testing process by becoming more dense, producing a higher modulus.

Limitations

The AIAG and HG methods are designed more for manufactured parts or laboratory prepared specimens. The gage R&R tests for the LWD would provide more consistent results if conducted on manufactured plates or on known varying stiffness rubber pads (White et al. 2009). The AIAG and HG method require the removal or replacement of data if range of measurement between trials exceeds a calculated Upper Control Limit. If the tests were conducted correctly, that data has value and has significance. It can mean variation in the soil or a malfunction in the instrument and should be investigated as real data or an anomaly before removal from a data set. Other challenges include the One Way ANOVA and Hypothesis Test for Paired Samples which are often not thoroughly understood by construction personnel without previous research, work experience or subject matter expertise. Secondly, the two tests are not definitive tests (Nuzzo

2014). When Fisher introduced the concept of the P value in the 1920's, it was envisioned to be an informal method to determine if the data produced results that warranted further examination. Fisher intended the P value to part of a process that used both data and background knowledge to point to a scientific conclusion (Nuzzo 2014). The level of confidence is an additional query for field testing. The P value condenses data from a null hypothesis; it cannot indicate the basis for the data. The decision maker needs to have sufficient background on the data. The alpha value at 0.05 has become the standard and has been accepted by researchers as statistically significant or noteworthy. There are no guidelines as to what alpha value/ level of confidence to use when investigating field data vs. lab data. This is a decision for the tester or other informed decision maker.

Conclusions

Coefficient of Variation and Standard Percent of Error are the preferable methods to measure repeatability and reproducibility in the field LWD, where the ground is sufficiently compacted to where after the seating drops are performed the soil modulus remains fairly uniform. The One way ANOVA and Hypothesis Test for Paired Samples take experience in statistics and also in compaction testing to choose the appropriate P or alpha value. To take soil variation out of the analysis and measure a "more true "repeatability and reproducibility of the testing devices to where the AIAG and Wheeler Gauge R&R methods can be used the testing medium must have strict compaction and moisture control (Mazari et al.2013) or use a manufactured medium such as rubber bearing pads that are manufactured under a quality control system. Rubber bearing pads are used by MnDOT to calibrate LWD devices (MnDOT presentation August 26, 2014).

CHAPTER 9. CONSOLIDATED CONCLUSIONS AND LIMITATIONS

The dissertation is an assembly of five journal articles, each of which contains the information and analysis that supports its own set of conclusions. As a result, this section will merely recapitulate the detailed conclusions contained in the previous chapters and put them in context with each other. It also will discuss the most important conclusions and provide recommendations on how these findings may be used by MoDOT to make the NDG replacement decision. These findings can also be used by other organizations as a procedural template for making similar decisions regarding alternative quality management testing technologies.

The literature review on alternative compaction testing methods and devices revealed that no in-depth life cycle cost analysis had been completed for a state DOT for both density and modulus/stiffness test methods. Additionally no cost indices had been advanced for the total cost per test for state DOTs or testing consultants. The analysis demonstrated that the research conducted by the author of this dissertation had not be done before and that the development of a rational methodology for quantitatively developing a set of metrics that can be used to compare compaction testing alternatives is indeed needed and should be viewed as a contribution to the body of knowledge in this field. Additionally, little field testing had been reported in the literature and as a result, it was concluded that most of the knowledge that is available comes from the controlled conditions of the geotechnical materials laboratory. This reaffirms the need to develop a methodology that is based on field test results made simultaneously in the same locations using multiple technology alternatives to the NDG.

DB project delivery transfers many of the quality management tasks from the DOT to the design-builder. This occurs with a contract that is typically awarded before the DB project's geotechnical studies and investigations have been completed. Therefore, on DB projects, state

DOTs must manage geotechnical risk in a prompt, concise and efficient method to achieve the aims of DB procurement; that being a faster schedule, lower cost and positive construction quality. The best practice for managing geotechnical risks is to have open lines of communication with prospective design-builders as well as the eventual winning DB team. Clear communication can point out gaps in information that can permit the DOT gather more data to fill in gaps, thus reducing potential risk for the bidders and reducing the cost of the project. Treating the design-builder as a project partner opens up paths of communication which should lead to a fast, cost effective and quality design and construction in a project that is beneficial and expedient not only to the state DOT but the driving public as well.

The life cycle cost and cost index analysis for the different test methods and devices provides a tool to allow MoDOT to make a rational, informed decision when selecting a compaction testing alternative to the NDG. Costs along with accuracy, repeatability, and field testing performance in various soils, and ease of use in testing will all be used in determining the best compaction testing system or device for quality assurance and control practices on MoDOT projects. MoDOT through its Quality Management Initiative has shifted the bulk of the annual compaction testing program to its contractors, relegating those tests taken by MoDOT technicians to a QA verification testing role and eliminating the need immediate feedback which is the primary advantage of the NDG.

No definitive relationship between NDG and modulus/CIV results could be found from Missouri test sites. This confirms the findings of Meehan et al. (2012) and Li (2103) and leads the authors to conclude that the new QM compaction testing process both the contractor’s QC/QA testing and MoDOT verification testing must be made using the same testing apparatus and method. Therefore, if MoDOT decides to replace the NDG with a modulus/stiffness device then its contractors will have to shift to the same modulus/stiffness device.

Measuring repeatability and reproducibility of the prospective Light Weight Deflectometer (LWD) and Dynamic Cone Penetrometer (DCP) devices cannot be satisfactorily completed in the field. The study found that these metrics are best completed in small controlled moisture and compaction test areas. Field testing on project areas contain too much variation in compaction and moisture notwithstanding the notion that the areas are "uniform" areas defined by general construction processes and methodologies. LWD reproducibility and repeatability along with device defects can be measured and checked using manufactured rubber bearing pads constructed concrete pads.

As some of the alternative devices come into conventional use and purchase prices come down this can modify the life cycle cost calculations presented in this report. Also, as technology such as time domain reflectometry improves, the tested devices this could change the way that that some of the devices are perceived and used.

CHAPTER 10. CONTRIBUTIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

The primary contribution made by this research was to field test and comparatively evaluate nine different types of compaction devices on the same soils, at the same time, and under the same field conditions. Past research has largely been restricted to evaluating a limited number of options in the laboratory. Hence, the findings did not relate to the field performance of each alternative, nor were all possible alternatives tested in a single set of experiments. Based on this research, MoDOT can now confidently make the decision regarding whether or not to replace its inventory of NDGs on a basis of comparative performance in the conditions in which it must operate.

The second major contribution was to extend the comparative analysis of compaction testing alternatives outside the technical realm of laboratory testing and integrate the financial and production aspects of each alternative to its ability to reliably produce satisfactory results at an acceptable life cycle cost. The immediate contribution is to give MoDOT a means for measuring each alternative's "bang for the buck." However, at a higher level, the methodology developed to complete this analysis could be used as a framework to complete a similar analysis of almost any quality control testing alternatives. This is significant because the literature clearly shows that engineers will almost always default to the highest technology alternative without respect to cost. Thus, the methodology developed to complete the evaluation of compaction test alternatives showed that the NDG's speed and technical precision comes a much higher cost to the agency than the slower, low-tech sand cone method.

The research also validated MoDOT practitioners' sense that with the shift of compaction testing responsibilities from the agency to the contractor that the NDG's high life cycle cost was no longer justified. The mere fact that the contractor no longer had to control the production of structural embankments to the speed at which MoDOT inspectors could determine if compaction specifications had indeed been met removed the “need for speed” embodied in the NDG technology.

Other contributions are as follows:

- An algorithm for calculating the cost of ownership of the present NDGs and the potential cost of testing alternatives was developed.
- The research has also disclosed an estimated number of density tests performed by the Department in 2013 and 2014 and the associated costs with performing those tests. An algorithm for computing a cost index was developed and was shown to act as a valuable metric for comparing the financial performance of differing technologies.
- The research verified the literature and found that there is no correlation between density readings with accompanying recorded modulus/stiffness readings and/ or Clegg Impact Values when employed on MoDOT projects. This finding argues that both MoDOT and its contractor will have to use the same devices if a change is made from the NDG to a modulus/stiffness technology for compaction testing.

The next step for future research is to integrate both MoDOT and contractor compaction test data into a common database such as SiteManager® and or SharePoint® so that it can be tracked and checked for accuracy. The analysis of this data over time can also be used further refine compaction test practices and procedures. Inspectors will learn through experience and past data; acceptable target compaction test values for commonly used soil/aggregate materials. This was demonstrated during a scanning tour of MnDOT construction office and projects in the

Detroit Lakes Area (MnDOT District 4), where acceptance for 3 soil/aggregate materials are based on acknowledged LWD target values. Once the central compaction test database is assembled, cost indices should be formulated to provide MoDOT practitioners with a performance metric to evaluate the trends in the life cycle cost per test for each district as well as the entire department. In doing so, MoDOT not only will be complying with the performance measurement mandate contained in MAP-21 but also will be able to identify those testing devices which are not being fully utilized, triggering an investigation and a decision as to whether they should be retained

BIBLIOGRAPHY

Abu-Farsakh, M.Y., Alshibli, K., Nazzal M., and Seyman E. (2004). *Assessment of in situ test technology for construction control of base courses and embankments*. Louisiana Transportation Resource Center, Baton Rouge, LA.

Ahlvers, D, Brucks, D., Kincaid, J. (2013). Quality management review, *Engineering policy guide*, http://epg.modot.mo.gov/files/b/b2/150_QM_Training_for_Contractors.pdf. (July 25, 2014).

Alshibli, K.A., Abu-Farsakh, M. and Seyman, E. (2005). "Laboratory evaluation of the geogauge and light falling weight deflectometer as construction tools*." J. of Mat. in Civil Eng.*, 17(5),560-569.

AASHTO (2014). "Density of Soil In-Place by the Sand Cone Method." T191, AASHTO, Washington D.C.

AASHTO (2013). "In Place Density and Moisture Content of Soil and Soil-Aggregate by Nuclear Methods (Shallow Depth)." T310, AASHTO, Washington D.C.

AASHTOWare version 2.2 (2015), (computer software), AASHTO, Washington D.C.

ASTM (2007). "Standard Test Method for Density and Unit Weight of Soil in Place by the Sand-Cone Method." D1556-07, ASTM International, West Conshohocken, PA.

ASTM (2007). "Standard Test Method for Determination of the Impact Value (IV) of a Soil." D5874-02, ASTM International, West Conshohocken, PA.

ASTM (2008). "Standard Test Method for Density and Unit Weight in Place by the Rubber Balloon Method." D2167-08, ASTM International, West Conshohocken, PA.

ASTM (2008). "Standard Test Method for Measuring Stiffness and Apparent Modulus of Soils and Soil-Aggregate In-Place by Electro-Mechanical Method." D6758-08, ASTM International, West Conshohocken, PA.

ASTM (2009). "Standard Test Method for Density of Use of the Dynamic Cone Penetrometer in Shallow Pavement Applications." D6951/D6951M-10, ASTM International, West Conshohocken, PA.

ASTM (2010). "Standard Test Method for Density of Soil in Place by the Drive-Cylinder Method." D2937-10, ASTM International, West Conshohocken, PA..

ASTM (2010). "Standard Test Method for In-Place Density and Water Content of Soil and Soil-Aggregate by Nuclear Methods (Shallow Depth)." D6938-10, ASTM International, West Conshohocken, PA.

ASTM (2011) “Standard Test Method for Measuring Deflections using a Portable Impulse Plate Load Test Device.” E2835-11, ASTM International, West Conshohocken, PA.

ASTM (2013). “Form and Style for ASTM Standards.” ASTM International, West Conshohocken, PA.

Arditi, D. and D. Lee, (2004). “Service quality performance of design/build contractors using quality function deployment,” *Construction Management and Economics*. London: Vol. 22(2), pp. 123-157.

Arizona Department of Transportation (AzDOT) (2001). *Design-Build Procurement and Administration Guide, 2nd Edition*. Arizona DOT Construction Group, Phoenix, AZ.

Arkansas State Highway and Transportation Department (AHTD) (2006). *Design-Build Guidelines and Procedures,* AHTD, Little Rock, AR.

Aktaş, B., D.D. Gransberg, C. Riemer, and D. Pittenger, (2011)“Comparative Analysis of Macrotexture Measurement Tests for Pavement Preservation Treatments, *Transportation Research Record No.2209, Journal of the Transportation Research Board*, National Academies, Washington, D.C., pp. 34-40.

Automotive Industry Action Group (AIAG) (2010). Measurement System Analysis (4th ed). AIAG, Detroit, MI.

Baecher, G.B. (1987). *Geotechnical Risk Analysis User’s Guide*, *Federal Highway*

Administration Report RD-87-011.US Department of Transportation, Federal Highway Administration, Washington, D.C., 50.

Barringer, H. P. and Weber, D. P. (1996). “Life cycle cost tutorial.” *Fifth International Conference on Process Plant Reliability*, Gulf Publishing Company, Houston, TX, 3-1 – 3-58.

Baynes, F.J. (2010). “Sources of geotechnical risk”, *Quarterly Journal of Engineering Geology*

and Hydrogeology, 43**,** 321–331.

Beatty, T.L. (2002). *Life-Cycle Cost Analysis Primer*. US Department of Transportation: Federal Highway Administration, Washington, D.C.

Berney, E.S., and Kyzar, J.D. (2012) “Evaluation of nonnuclear soil moisture and density devices for field quality control.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 2310, 18-26.

Berney, E.S., Kyzar, J.D., and Oyelami, O.L. (2012) *Device Comparison for determining field soil moisture content*, U.S. Army Corps of Engineers, Washington, D.C.

Blanchard, B. (2007). “Design-build lessons learned Florida DOT”, *Proceedings Louisiana Transportation Engineering Conference,* 12-14 March, Baton Rouge, LA, 6-14.

California Department of Transportation (CalTrans) (2010). Design-Build Demonstration Program,Sacramento, CA. <http://www.dot.ca.gov/hq/oppd/designbuild/db.htm >(October 30, 2012).

Cardell-Oliver, R., Smettem, K., Kranz, M., & Mayer, K. (2004). "Field testing a wireless sensor network for reactive environmental monitoring (soil moisture measurement)." *Intelligent Sensors, Sensor Networks and Information Processing Conference*, Melbourne, Australia.

Centennial Contractors, Inc. (2004). *Design-Build Project Manual*, Centennial Contractors, Inc.,Vienna, VA, 137.

Chan, E. (2013). When Australian Construction SME Implementing TQM. *World Review of Business Research*, *3*(3), 167-178.

Christensen M.R. and Meeker, L.E. (2002). "Design-build projects – lessons learned from the contractor's perspective" *Proceedings American Railway Engineering and Maintenance-of-Way Association*, 3-9 December, Lanham, MD, 14.

Cho, Y.,Yong-Rak, K., Kabassi, K., Zhuang, Z., Im, H., Wang,C., Bode, T. (2011) *Non-nuclear method for density measurement*, Report No. SPR1(10) P335, Nebraska Transportation Center, Lincoln, Nebraska.

Clark, G.T. and Borst A. (2002). "Addressing risk in Seattle's underground." *PB Network*,January, 34-38.

Comrie, A.C. (1997). "Comparing neural networks and regression models for ozone forecasting." Journal of the Air & Waste Management Association, 47(6),653-663.

Cowles, M., and Davis, C. (1982). " On the origins of the .05 level of statistical significance." *American Psychologist*., 37(5), 553-558.

Department of Defense (DOD) (2010.) Unified Facilities Criteria (UFC), UFC 1-300-09N 25 May 2005 Including Change 7, Washington, DC.

Design-Build Institute of America (DBIA) (2013). 2013 Design-build state laws for transportation procurement, < http://www.dbia.org/NR/rdonlyres/231CFB85-2483-4D8A87DF-BF72193C61D7/0/tran130103.pdf> (February 2, 2013).

Dhawale, R. and Raut, D.N. (2013). 'Evaluating measurement capabilities by gauge R&R using ANOVA for reliability." *International Journal of Engineering Research and Application* 3(3), 726-730.

Diekmann, J.E., Songer, A.D. and Pecsok, R.S. (1987). "Risk analysis for revenue dependent infrastructure projects." *Journal of Construction Management and Economics*, 15 (4), 377–382.

Dwyre, E.M., Batchko, Z. and Castelli, R.J. (2010). “Geotechnical baseline reports for foundation projects.” *Proceedings GeoFlorida 2010: Advances in Analysis, Modeling & Design (GSP 199)*,21-25 August, Orlando, FL, 1-10.

Ellis, R. D., Herbsman, Z. and Kumar, A., (1991). “Evaluation of the FDOT design/build program,” *Final Report*, Submitted to Florida Dept. of Transportation, State Project No. 99700-7543-010, Department of Civil Engineering, University of Florida, Gainesville, Florida.

Erchul, R.A. and Meade, R.B. (1990). Using the modified Clegg impact hammer to evaluate to evaluate adequacy of compaction. Department of Civil Engineering, Virginia Military Institute, Lexington, Virginia.

Erchul, R.A. (1999). “Compaction comparison testing using a modified impact soil tester and nuclear density gauge.” *STP 1358, Field Instrumentation for Soil and Rock, American Society for Testing and Materials*, West Conshohocken, PA, 382-391.

Ernzen, J. and Feeney, T. (2002). “Contractor-led quality control and quality assurance plus design-build: who is watching the quality?”, *Transportation Research Record*, 1813, Transportation Research Board of the National Academies, 253-259.

Farrar, K., Vetter, D., Hill, B., and Esposito (2005). *Evaluation of soil Compaction Devices*, Gas Technology Institute, Des Plaines, IL.

Fay, D.S., and Gerow K. (2013). “A biologist’s guide to statistical thinking and analysis.” WormBook, Available: http://www.wormbook.org/chapters/www_statisticalanalysi/statisticalanalysis.html (August 5, 2015).

Federal Highway Administration (FHWA), (2004). Use of Contractor Test Results in the Acceptance Decision, Recommended Quality Measures, and the Identification of Contractor/Department Risks, Technical Advisory 6120.3, August 2004.

Federal Highway Administration (FHWA) (2006). Design-Build Effectiveness Study – as required by TEA21Section 1307(f): final report, US Department of Transportation, Federal Highway Administration, Washington, DC.

FHWA (2013). MAP-21: Moving Ahead for Progress in the 21st Century, <http://www.fhwa.dot.gov/map21/> (February 2, 2013).

Federal Register (2007) Design-build contracting, *Final Rule, Rules and Regulations,* 72 (156), August 14, 2007, 45329.

Florida Department of Transportation (FDOT) (2011). “Alternative technical concepts reviews,” *FDOT Design-build Guidelines*, 2**,** 21-27.

Gad, G.M. S.A. Adamtey, and D.D. Gransberg, (2015). "Quality Management Approach Trends in State Design-build Transportation Projects," *Transportation Research Record, Journal of the Transportation Research Board,* National Academies, Washington, D.C., (in press).

Gertsman, B.B. (2006). “Introduction to Hypothesis Testing.”, StatPrimer, www.sjsu.edu/faculty/gertsman/StatPrimer/hyp-test.pdf (June 16, 2015).

Goetsch, D. L., & Davis, S. B. (2014). *Quality management for organizational excellence*. Pearson.

Government Accountability Office (GAO) (1991). *Using Structured Interviewing Techniques*, 7

GAO/PEMD-10.1.5, Government Accountability Office, Washington, D.C., 191.

Gransberg, D.D., and Molenaar, K.R (2004). “Analysis of owner’s design and construction quality management approaches in design-build projects.” *Journal of Management in Engineering,* ASCE, 20 (4), 162-169.

Gransberg, D.D., Molenaar, K.R., and Datin J.N. (2008*). Quality Assurance in Design-Build Projects, NCHRP Synthesis 376.* Transportation Research Board National Academies, Washington, DC, 39-42.

Gransberg D.D. and Loulakis, M.C. (2012). *Geotechnical Information Practices in Design-Build Projects, NCHRP Synthesis 429*.Transportation Research Board, National Academies, Washington, DC.

Gransberg, D.D., and E. Scheepbouwer, (2010). “Infrastructure Asset Life Cycle Cost Analysis Issues,” *2010 Transactions*, AACE, Int‘l, Atlanta, Georgia, pp. CSC.03.01- CSC.03.8.

Gransberg, D.D**.** and M. Zaman, (2005). **“**Analysis of Emulsion and Hot Asphalt Cement Chip Seal Performance**,”** *Journal of Transportation Engineering,* 131(3), 229-238.

Hatem, D.J. (2011). “Risk allocation for subsurface conditions and defective design.”

Brierley, G.S. Corkum D.H. and Hatem, D. J.(eds), *Design-Build Subsurface Projects,* Society for Mining, Metallurgy and Exploration, Englewood, CO, 22-28.

Higbee, J.B. (2004). ”Geotechnical issues with large design–build highway projects.” *Transportation Research Record, 1868*, TRB, National Research Council, Washington, DC, 147-153.

Ho, K. K. S., Leroi, E., & Roberds, W. J. (2000). “Quantitative risk assessment: application, myths and future direction.”, *Proceedings International Conference on Geotechnical & Geological Engineering*, GeoEng, London, 14-19 July, 269-312.

Hung, C.J., Monsees, J., Munfah, N. and Wisniewski, J. (2009). *Technical Manual for Design and Construction of Road Tunnels – Civil Elements, Technical Report FHWA-NHI-10-034*.National Highway Institute, New York, 35-87.

Institute for Digital Research and Education (IDRE), (2015). “Faq: what is the coefficient of variation.” University of California at Los Angeles, (UCLA), http://www.ats.ucla.edu/stat/mult_pkg/faq/general/coefficient_of_variation.htm (July1, 2015)

Jaksa, M. B. (2000). “Geotechnical risk and inadequate site investigations: A case study.”, *Australian Geomechanics*, 35(2), 39-46.

Jordan, E., Gross, M. E., Javernick-Will, A. N., & Garvin, M. J. (2011).”Use and misuse of qualitative comparative analysis.” *Construction Management and Economics*, 29(11), 1159-1173.

Joubert, J.W., and Meintjes, S. (2015). “Repeatability and reproducibility: implications of using GPS data for freight activity chains.” Transportation Research Part B.

Kandell, J. (2014) “Contract Dispute Delays Panama Canal’s Expansion,” *Institutional Investor*, http://www.institutionalinvestor.com/Article/3318308/Banking-and-Capital-Markets-Emerging-Markets/Contract-Dispute-Delays-the-Panama-Canals-Expansion.html?ArticleId=3318308#.VPIO6oc5C70

Karlen, D. L., Mausbach, M. J., Doran, J. W., Cline, R. G., Harris, R. F., and Schuman, G. E. (1997). Soil quality: a concept, definition, and framework for evaluation (a guest editorial). *Soil Science Society of America Journal*, *61*(1), 4-10.

Kim, H., Prezzi, M., and Salgado, R. (2010). *Use of dynamic cone penetration and Clegg hammer test for quality control of roadway compaction and construction*. Publication FHWA/IN/JTRP-2010/27.Joint transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, IN.

Kim, K.J., Kreider, C.A., and Valiquette, M.D. (2009). “North Carolina Department of Transportation’s practice and experience with design–build contracts geotechnical perspective.” *Transportation Research Record*, *2116*, Transportation Research Board of the National Academies, Washington, D.C., 47–52.

Koch, J.E., Gransberg D.D, and Molenaar, K.R.(2010). *Project Administration for Design-Build: A Primer for Owners, Engineers, and Contractors*, ASCE Press, Reston, VA, 118-145.

Kraft E., and Molenaar, K.R. (2013). “Project quality assurance organization selection for highway design and construction projects.” *Construction 2013*, *Transportation Research Record: Journal of the Transportation Research Board*, 2347, 29-37.

Lenke, L.R., McKeen, R.G. and Grush, M.P. (2003). “Laboratory evaluation of GeoGauge for compaction control.” *Transportation Research Record: Journal of the Transportation Research Board*, 1849, 20-30.

Li, J.(2013). "Evaluation Of Intelligent Compaction Control In The M-189 Reconstruction Project At Iron River, Michigan", Master's Thesis, Michigan Technological University, Hougton MI, < http://digitalcommons.mtu.edu/etds/736>

Loulakis, M.C. and Shean, O. J. (1996). “Risk transference in design-build contracting.” *Construction Briefings, 2nd series*, Federal Publications, Washington, DC, 12.

Loulakis, M.C., Wagner, B.P. and Splan, H.C. (1995). “Differing site conditions.”, R.S. Brams and C. Lerner, *Construction Claims Deskbook*, Aspen Law and Business, New York, 1995, 131-166.

Mahdi, I.M., Alreshaid, K. (2005). “Decision support system for selecting the proper project delivery method using analytical hierarchy process.” *International Journal of Project Management,* 23**,** 564-572.

Maher, A., Bennert, T., and Gucunski, N. (2002). *Evaluation of the Humboldt Stiffness Gauge*. Publication No FHWA-NJ-2002-002, New Jersey Department of Transportation, Trenton, NJ.

Manley, K. and McFallan, S. (2006). “Exploring the drivers of firm-level innovation in the construction industry.” *Construction Management and Economics,* 2(9), 911-920.

Mann, P.S. (2010). *Introductory Statistics, 7th Edition*. John Wiley and Sons, Inc, Hoboken, NJ.

Mazari M., Garcia, G., Garibay, J., Abdallah, I. and Nazarian. S. (2013). “Impact of modulus based device variability on quality control of compacted geomaterials using measurement system analysis.” *Proceedings of the Transportation Research Board 92nd Annual Meeting*, Washington D.C. 13p.

Meehan, C.L., and Hertz, J.S. (2011) *Using electrical density gauges for field compaction control*, Delaware Center for Transportation, University of Delaware, Newark, DE.

Meehan, C.L., and Hertz, J.S. (2013). “Using a complex-impedance measuring instrument to determine in situ soil unit weight and moisture content.” *Geotechnical Testing Journal*, 36(1), 119-137.

Meehan, C.L., Tehrani, F.S., and Vahedifard F. (2012). “A comparison of density-based and modulus-based in situ test measurement for compaction control.” *Geotechnical Testing Journal*, 35(3), 387-39.

Mendez, V., (2010). *Every Day Counts: Innovation Initiative*, Federal Highway Administration, Washington, DC, 1-2.

Minnesota Department of Transportation (MNDOT) (2003). “Approach to alternative technical concepts.” *MnDOT Design-Build Program White Paper No. MN-11*, St Paul, MN. 1.

MnDOT (2010). “Request for Proposals, Book 1 – TH 61.” *Hastings Bridge Design-Build Project S.P. 1913-64*, Minnesota DOT, Metro District, St. Paul, MN, 7.

Mississippi Department of Transportation (MDOT) (2005) Request for Proposals, Addendum 1, A Design-Build Project Bridge Replacement on US 90 Biloxi to Ocean Springs Bridge

Jackson and Harrison Counties, Mississippi, Contract No. ER/BR-0003-01(099) 104556/301000 –Biloxi/Ocean Springs Bridge, Jackson, MS.

Missouri Department of Transportation (MoDOT) (2013). "Quality Management: Section 200 Checklists."<http://www.modot.org/business/contractor_resources/Quality_Management/Checklists/Section_200.htm >(Nov.13, 2014).

MoDOT (2014). "MoDOT Standard Inspection and Testing Plan (ITP)."< http://www.modot.org/business/contractor_resources/Quality_Management/> (Nov.13, 2014).

Missouri State Highway Department, (MSHD). (1975). "Field evaluation of a direct transmission type nuclear moisture-density gauge." *Missouri Cooperative Highway Research Program Final Report No. 74-2*, Missouri State Highway Department, Jefferson City, Missouri.

Missouri Highways and Transportation Commission (2011). *Missouri Standard Specifications for Highway Construction*. Jefferson City, MO.

Missouri Highways and Transportation Commission (2014). *Missouri Standard Specifications for Highway Construction*. Jefferson City, MO.

Molenaar, K.R., Gransberg, D.D., West, N.J.N. and Kraft, E. (2012). *Alternative Quality Management Systems for Highway Construction, National Cooperative Highway Research Program Report 10-83*. Transportation research Board of the National Academies, Washington, D.C., 121-124.

Montgomery, D.C., Runger, G.C., and Hubele, N.F. (2007). *Engineering Statistics*. Wiley and Sons, Hoboken, NJ.

Mooney, M.A., Nocks, C.S., Selden, K.L., Bee, G.T., and Seseney C.T (2008). "Improving Quality Assurance of MSE wall and bridge approach earthwork compaction." *Report No. CDOT-2008-11*, Colorado Department of Transportation, DTD Applied Research and Innovation Branch, Denver, CO.

Mosier, R.C.,. Pittenger, D.M., and. Gransberg, D.D. (2014). "Carbon Footprint Cost Index: Measuring the Cost of Airport Pavement Sustainability," *Compendium*, 2014 Transportation Research Board Annual Meeting, Paper 14-3214, National Academies, 24.

Nazzal, M. (2014) *Non-Nuclear Methods for Compaction Control of Unbound Materials, NCHRP Synthesis 456.* Transportation Research Board National Academies, Washington, DC.

Nazzal, M. D., Abu-Farsakh, M.Y, Alshibli, K. and L. Mohammad L. (2007) "Evaluating the light falling weightdeflectometer device for in situ measurement of elastic modulus of pavement layers." *Transportation Research Record: Journal of the Transportation Research Board, No. 2016*, Transportation Research Board of the National Academies, Washington, D.C., 13–22.

Neuendorf, K.A. (2002). *The Content Analysis Guidebook,* Sage Publications, Thousand Oaks, CA, 216.

Nuzzo, R. (2014) Scientific method: Statistical errors, *Nature.com*, http://www.nature.com/news/scientific-method-statistical-errors-1.14700 (July 23, 2015)

Oberguggenberger, M., and Fellin W. (2002). "From probability to fuzzy sets: the struggle for

meaning in geotechnical risk assessment." *Probabilistics in GeoTechnics: Technical and Economic Risk Estimation,* United Engineering Foundation, Essen, Germany 29 - 38.

Ooi, P.S.K, and Pu, J. (2003) "Use of stiffness for evaluating compactness of cohesive pavement geomaterials." Transportation Research Record: Journal of the Transportation Research Board, 1849, 11-19.

Oppenheim, A. N. (1992). *Questionnaire Design, Interviewing and Attitude Measurement*, 24 Continuum, London.

Pandiripalli, B. (2010). "Repeatibility and reproducibility studies: a comparison of techniques". M.S. Thesis,Wichita State University, Wichita, KS.

Papernik, B.G. and Farkas, D.J. (2011). "Using alternative technical concepts to improve design build and PPP procurements." *Nossaman E-Alerts*, http://www.nossaman.com/using-alternative-technical-concepts-improve-designbuild-ppp (May 26, 2011).

Perkins, R.A. (2009). "Sources of changes in design–build contracts for a governmental owner." *Construction Engineering and Management,* 135(7), 32-41.

Pittenger, D.M., Gransberg, D.D.M. Zaman, M., and Riemer, C. (2011). "Life cycle cost-based pavement preservation treatment design." *Transportation Research Record No. 2235, Journal of the Transportation Research Board*, 28-35.

Rathje, E.M, Wright, S.G., Stokoe II, K.H., Adams A., Tobin R., and Salem, M. (2006). *Evaluation of non-nuclear methods for compaction control*. Texas Department of Transportation, Research and Technology Implementation Office, Austin, TX.

Riggs, J. L. and T. H. West, (1986) *Essentials of Engineering Economics,* McGraw-Hill, Inc., New York.

Salter, A. and Torbett, R. (2003). "Innovation and performance in engineering design." *Journal of Construction Management and Economics*, 21, 573-580.

Schaefer, V.R., Stevens, L.J, White, D.J., and Ceylan H. (2008). *Design guide for improved quality of roadway subgrades and subbases, TR-525*. Iowa Highway Research Board, Ames, IA.

Schein, E.H. (1996) "Three cultures of management: The key to organizational learning." *Sloan management review*, 38(1), 9-20.

Seltman, H.J. (2015) " Chapter 7: one-way ANOVA." *Experimental Design and Analysis*, Carnegie Mellon University, Pittsburgh, PA. 171-190.

Siddiki, N., Walker, R., Sommer, K., Coffman, J., Kim, D. (2015)., "Implementing dynamic cone penetrometer and lightweight deflectometer in Indiana." *TR Circular E-C199: Geotechnical Research Deployment*, Transportation Research Board, 32-40.

Siekmeier, J., Pinta, C., Merth, S., Jensen, J., Davich, P., Camargo, F., and Beyer, M. (2009). *Using dynamic cone penetrometer and light weight deflectometer for construction quality assurance*. Minnesota Department of Transportation, Office of Materials and Road Research, Mapelwood, MN.

Smith, N.C. (2001). "Getting What You Paid For: the Quality Assurance and Acceptance Process for Transportation Projects," *Proceedings*, American Bar Association Forum on the Construction Industry, Denver, Colorado, 17-18.

Smith, R. (2008). "An evolving view of geotechnical engineering – a focus on geo-risk Management." *Geotechnical Special Publication No. 178*, GeoCongress 2008: New Orleans, LA, 231-239.

Songer, A.D. and Molenaar, K.R. (1996). "Selecting design-build: private and public sector owner attitudes." *Management in Engineering*,12(6), 47-53.

Stamm, S. (2013). " A comparison of gauge repeatability and reproducibility methods". PhD Dissertation, Indiana State University, Terre Haute IN.

StatTools version 6.1.1, (2013), (computer software), Palisades Corporation, Ithaca, New York.

Tehrani . F.S. and Meehan C.L. (2010). "The effect of water content of light weight deflectometer measurements. "*GeoFlorida 2010: Advances in Analysis, Modeling and Design* . 930-939.

Texas Turnpike Authority (2001). "Request for proposals to construct, maintain and repair the SH130 turnpike through an exclusive development agreement." Texas Department of Transportation, Austin, TX.

Tufenkjian, M. (2007). Review of geotechnical services, California Department of Transportation, Final Report, USDOT Federal Highway Administration, Washington, D.C., 76.

Turochy, R.E., Willis, J.R., and Parker, F. (2007). "Comparison of Contractor Quality Control and Georgia Department of Transportation Data for Quality Assurance of Hot-Mix Asphalt," *TRB 85th Annual Meeting*, CD-ROM, TRB, National Research Council, Washington, D.C., 16.

Vanden Berge, D.R. (2003) "An experimental and theoretical investigation of the clegg hammer". Thesis for Masters of Science in Civil Engineering, Michigan Technological University, Houghton, MI.

Van Staveren, M.T. (2007). ‘Extending to geotechnical risk management’, ISGSR2007, 1st International Symposium on Geotechnical Safety and Risk, 22-26 June, Shanghai, 1-12.

Virginia Department of Transportation (VDOT) (2011), *Value for Money Guidance*, Office of Transportation Public Private Partnerships, Richmond, Virginia, 39.

Walls, J., and Smith, M.R. (1998). *Life cycle cost analysis in pavement design- interim technical bulletin FHWA-SA-98-079.* Federal Highway Administration, Washington D.C.

Washington State Department of Transportation (2004*). Guidebook for Design-build Highway Project Development*, Washington State Department of Transportation, Olympia, WA.

Water Environment Research Foundation (WERF) (2001). *Life cycle cost tool.* WERF, Alexandria Virginia.

Wheeler, D. J. (2009). “An honest gauge R&R study.” Paper presented at the 2006 American Society for Quality/American Statistical Association Fall Technical Conference, Columbus, OH, http://www.spcpress.com/pdf/DJW189.pdf (April 5, 2015).

Weber, R.P. (1985). *Basic Content Analysis*. Sage Publications, Beverly Hills, CA.

White, D. J., Becker, P., Vennapusa, P.K.R., Dunn, M.J., and White, C.I. (2013). “Assessing soil stiffness of stabilized pavement foundations.” *Transportation Research Record: Journal of the Transportation Research Board*, No. 2335, 99-109.

White, D.J., Vennapusa, P.K., Zhang, J. Geieslman, H., and Morris M. (2009). “Implementation of intelligent compaction performance based specifications in Minnesota.” Minnesota Department of Transportation Research Services Section, St. Paul MN.

Yale Statistics (1997). “Linear regression.” *Introduction to statistics*, <http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm (March 8, 2014).