AN ABSTRACT OF THE THESIS OF

Suraj Darra for the degree of
Master of Science in
Civil Engineering presented
on November 20, 2009.
Title: Development of an Improved System for Oregon to Accurately Quantify Dense-Graded Hot Mix Asphalt Pavement Density.

Abstract approved:

____

Dr. Todd V. Scholz

The Oregon Department of Transportation (ODOT) standard specifications require a minimum density for the construction of dense-graded hot mix asphalt (HMA) pavements. Currently, the standard specifications call for density measurements for quality control (QC) and quality assurance (QA) testing to be made using nuclear density gauges that are calibrated using reference blocks and correlated (adjusted) to densities from pavement cores. QC measurements (performed by the contractor) are verified by QA measurements (performed by ODOT); then ODOT utilizes the contractor's QC results for determining acceptance of the pavement (in conjunction with other specification criteria). Hence, appropriately accepting an HMA pavement, based on the in-place density criterion, relies on the accuracy of the density measurements. However, density measurement results using nuclear gauges have been observed by ODOT to be questionable on a number of projects, and repeatability and reproducibility with the same gauge and between gauges has been unattainable.

The overall objective of the project described herein was to develop a system that accurately quantifies density of dense-graded HMA pavements. More specifically, the objectives were to: 1) investigate the efficacy of the various methods used by ODOT and other agencies for determining in-place HMA density; 2) assess current practices used by

ODOT and other agencies for determining in-place HMA density using nuclear gauges; 3) conduct field and laboratory testing and analyses to determine the most accurate and reliable state-of-the-practice means for determining in-place HMA density; 4) provide recommendations for changes to current practices to improve accuracy and reproducibility of in-place HMA density measurements using nuclear gauges; and 5) provide recommendations for alternate means for determining in-place HMA density.

To satisfy the first objective, a literature review was conducted and summarized herein. For the second objective, practices employed by ODOT were observed and assessed. For the third objective, state-of-the-practice methods used to measure HMA density (i.e., tests on core samples using the saturated surface-dry and automatic vacuum sealing methods, nuclear density gauge measurements and electromagnetic density gauge measurements) were investigated. Numerous statistical comparisons of the results were made to determine the best combination of measurement methods to ensure accurate assessment of HMA density for a variety of construction scenarios. Finally, the findings from the first three objectives were used to formulate the recommendations identified in the fourth and fifth objectives.

Some of the significant findings that are based on the supporting evidence from the research are as follows: 1) Nuclear gauge densities should be adjusted to core densities for future in-place density testing of HMA pavements; 2) Cores should be tested in accordance with CoreLok testing procedure and not by using the SSD methods; 3) Cores should be extracted from the overlapping portion of the footprints of the nuclear gauge measurements; 4) It is sufficient to take two nuclear gauge readings (perpendicular and parallel to the direction of paving) rather than four readings; 5) Under certain conditions, the core adjustment factors obtained from the bottom lift can be used to adjust the nuclear gauge densities on at least the next two overlying lifts of the same pavement; and 6) The electromagnetic gauge adjustment factors could be used on more lifts than nuclear gauge adjustment factors to adjust measurements; and, electromagnetic gauge densities were not significantly different from core densities.

Development of an Improved System for Oregon to Accurately Quantify Dense-Graded Hot Mix Asphalt Pavement Density

by

Suraj Darra

A THESIS

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Master of Science

Presented November 20, 2009

Commencement June 2010

Master of Science thesis of Suraj Darra presented on November 20, 2009.

APPROVED:

____

Major Professor, representing Civil Engineering

____

Head of the School of Civil and Construction Engineering

____

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

____

Suraj Darra, Author

ACKNOWLEDGEMENTS

Firstly, I would like to thank my parents, Arthi Darira and Madhusudan Darira, my brother, Yogesh Darra, and, my maternal and paternal grandparents. Without their love and support, I would not have reached so far in life. I am also grateful to the rest of my family and relatives.

Next, I would like to thank my major professor, Dr. Todd V. Scholz for his unstinted support, guidance, and patience. With his help, I have come a long way in the last two years. From him, I understood the true meaning of professionalism and technical accuracy. It is with great gratitude that I would like to thank him for always guiding me in the right direction and for his invaluable words of encouragement and motivation.

I would also like to thank my other professors at Oregon State University, specially Dr. John A. Gambatese and Dr. David N. Sillars. Their depth of understanding and the interest they took in my education knows no bounds. The knowledge they have imparted to me are the tools that will take me ahead in my career and life. I thank them for their insightfulness and intuitiveness.

I thank the Oregon Department of Transportation and the Federal Highway Administration and their technical advisory board for funding this project.

I thank Dr. Christoph K. Thomas for being the Graduate Council Representative of my thesis defense committee.

I thank all my friends at Oregon State University for being there in my life and making my life wonderful. I specially thank Vikas, Rama, Shweta, Ravi, Gaurav, Divya, Aruna, Vineeth, Madan, and Pratheeksha.

A special thanks to Mr. Deepak K. Lakhi, Mr. and Mrs. Omprakash Darira, and Mr. Ashwini Kumar for being ‘catalysts’ to help me achieve my goals faster.

Finally, I would like to thank God for everything in life, including this research opportunity and international educational experience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDIX TABLES

Table Page

1.0 INTRODUCTION

The density of hot mix asphalt (HMA) pavements is such an important factor in ensuring that the pavement has a long service life, it is used by several states as a partial basis for paying the contractor. Based on the proximity of the measured pavement density value to the target value specified by the state, the state department of transportation pays the contractor accordingly; typically a bonus for exceeding the target value, or a penalty for not achieving the target value (or removal and replacement if sufficiently below the target value). The target value set by the Oregon Department of Transportation (ODOT) is at least 92% of the maximum theoretical density (MTD) for dense-graded paving mixtures exposed to very heavy traffic volumes or heavy truck traffic (54). To ensure that the density is within acceptable limits ODOT requires that the contractor develop and execute a quality control plan to measure pavement density at a frequency based on tonnage of material placed. ODOT currently tests 10% of the locations tested by the contractor to assess the quality of the contractor's measurements (54).

The service life of a pavement can be improved if it is compacted to at least the target density, but not greatly exceeding the target. For this, it is always important to verify the density of a pavement while it is within the compactable temperature range. If the pavement density value falls below the target value, it must be further compacted until the target density is obtained; or, it must be removed and replaced if the density far exceeds the target value (unless accepted as is by the state). This is because the service life of the pavement is reduced if the density of the compacted pavement is too low or too high, thereby resulting in frequent repairs or replacement of the pavement, which can be quite expensive. Such expenses can be avoided if there is a system which can measure density accurately and quickly. If the density of the pavement can be accurately measured while the pavement mat is still within compaction temperature range, then the pavement can be suitably constructed to last longer and, thus, prevent expenses for early repair or replacement.

Since density of the compacted pavement plays an important role in determining and ensuring the longevity of an HMA pavement, this research initiative was undertaken by the Oregon Department of Transportation to develop a system that accurately quantifies the density of dense-graded HMA pavements. This involved critically evaluating how ODOT currently measures HMA density, an investigation and evaluation of what other agencies do to measure HMA density, and conducting testing and analysis of alternate ways of measuring HMA density.

1.1 Background

Traditionally, the density of hot mix asphalt pavements has been measured using cores cut out from the pavements. The density tests on cores are performed in a laboratory and the results are known to be most accurate (1 and 3). However, the process of obtaining density results from cores takes so long that the pavement has already cooled down below the suitable temperature range for effective compaction. Thus, density results obtained by cores can be suitably used for quality assurance but not for quality control. For quality control, nuclear gauges are used by the state of Oregon. Nuclear gauges can be used to measure density quickly and quite accurately; thus, they are used to indicate whether or not further compaction is necessary while the pavement is still within a suitable temperature range. However, nuclear gauges are not as accurate as cores in density measurement (33). The number of nuclear gauge density readings and the method of obtaining them, and the number of cores and the type and method of density tests to be performed on them vary from state to state in the United States. While some states prefer to use only cores or only nuclear gauges, other states prefer to use a combination of both for QC and QA purposes.

A more recent method of measuring density of compacted pavement involves use of electromagnetic gauges, which are said to be the least accurate and unreliable means of measuring density (5, 7, 16, 22, 32, 29, and 39).

1.2 Objectives

The primary objective of this research was to develop a system that accurately quantifies the density of dense-graded hot mix asphalt pavement to be used by the state of Oregon for QC and QA purposes. More specifically, the objectives of this research effort were to:

1. Investigate the efficacy of the various methods used by ODOT and other agencies/entities for determining in-place HMA density.
2. Assess current practices used by ODOT and other agencies/entities for determining in-place HMA density using nuclear gauges.
3. Conduct field and laboratory testing and analysis to determine the most accurate and reliable state-of-the-practice means for determining in-place HMA density.
4. Provide recommendations for changes to current practices to improve accuracy and reproducibility of in-place HMA density measurements using nuclear gauges.
5. Provide recommendations for alternate means for determining in-place HMA density.

1.3 Scope

Chapter 2 of this report provides a summary of a literature review that was conducted to investigate research work performed by others regarding the efficacy of methods for measuring in-place HMA density including nuclear gauges, electromagnetic gauges, and measurements on cores. Also, methods of using the gauges to obtain the most accurate density results and factors that affect different types of gauges were investigated. In addition, other states' practices in utilizing density as part of HMA acceptance were reviewed.

Based on the findings from the literature review, an experiment design was developed for the field and laboratory investigations as detailed in Chapter 3, whereas Chapter 4 summarizes the findings from the field study, and Chapter 5 summarizes the findings from the laboratory study. Chapter 6 provides details of the statistical analyses performed on the various data obtained, and Chapter 7 provides a discussion of the results. Finally, Chapter 8 provides conclusions and recommendations based on all the findings from this research effort.

2.0 LITERATURE REVIEW

One of the tasks undertaken as part of this effort was to conduct a literature review to determine the state-of-the-practice regarding hot mix asphalt pavement density measurement in the United States. This literature review covers methods of density measurement that utilize cores, nuclear gauges, and electromagnetic gauges, either solely or as a part of a differences system. In addition, states' practices in utilizing density as part of hot mix asphalt acceptance are summarized. The sources for literature included the Transportation Research Information Services (TRIS) online database, Oregon State University Library (for journal publications), states' department of transportation websites (for density testing methodology and Quality Control/Quality Assurance specifications), and gauge manufacturers' websites (for updated information and research studies conducted in evaluating gauge performance). The findings are synthesized below.

2.1 Methods for measuring in-place density of hot mix asphalt pavements

In-place density of hot mix asphalt pavements is commonly determined by extracting cores and measuring the bulk specific gravity of the cores in a laboratory and converting the results to density, or by directly measuring the pavement density through use of nuclear density gauges, or electromagnetic gauges. The former method is often referred to as a "destructive method," whereas the methods employing a nuclear or electromagnetic gauge are "non-destructive." The following sections provide a brief overview of each method.

2.1.1 Density measurement on cores

Cores are extracted from a pavement after the pavement is laid, thus making it a destructive process. Bulk specific gravity measurements are then made on the cores in a laboratory. The methods of measuring bulk specific gravity of cores include: a) saturated surface-dry method, b) paraffin-coated method, c) automatic vacuum sealing method, and d) dimensional analysis. The bulk specific gravity of the core specimen determined by any of these methods is then multiplied by the unit density of water to obtain the density of the hot mix asphalt core sample. These methods are described briefly in the following paragraphs.

a) Saturated surface-dry specimen method

The saturated surface-dry method is described in AASHTO T 166 (41) and ASTM D 2726 (42). According to AASHTO T 166, this method should not be used for specimens that contain open or interconnected voids and absorb more than 2.0% water by volume.

Briefly, test procedure is as follows: 1) determine the dry mass of the specimen (A); 2) submerge the specimen in water for a specified period and determine its submerged mass (C); and 3) remove the specimen from the water, wipe off the surface moisture and determine its saturated surface-dry mass (B). The bulk specific gravity is then determined by dividing the dry mass by the difference between the saturated surface-dry mass and submerged mass as indicated in Equation 2.1.

$$\text{Bulk Specific Gravity} = \frac{A}{B-C} \tag{2.1}$$

b) Paraffin-coated specimen method

The paraffin-coated specimen method is described in AASHTO T 275 (47) and ASTM D 1188 (48). AASHTO T 275 indicates that this method can be used for specimens that contain open or interconnected voids and/or absorb more than 2.0% water by volume because the specimen is coated with paraffin to prevent water from penetrating the surface connected voids (12).

Briefly, the test procedure is as follows: 1) determine the dry mass of the specimen (A); 2) coat the entire specimen with melted paraffin and allow it to cool before recording the cooled specimen mass (D); 3) submerge the specimen in water for a specified period and determine its submerged mass (E); and 4) determine the specific gravity of paraffin (F). Equation 2.2 is then used to calculate the bulk specific gravity of the specimen.

$$\text{Bulk Specific Gravity} = \frac{A}{D - E - \left(\dfrac{D-A}{F}\right)} \tag{2.2}$$

The *parafilm* method is the same as the paraffin method except that, instead of coating the specimen with paraffin, it is wrapped in parafilm and the mass of the wrapped specimen is recorded as D.

c) Automatic vacuum sealing method

The automatic vacuum sealing method is described in AASHTO T 331 (49), ASTM D 6752 (50), and CoreLok operator’s manual (53). It can be used to determine the specific gravity of both fine- and coarse-graded mixtures.

Briefly, the test procedure is as follows: 1) determine the dry mass of the specimen at room temperature (A); 2) place it in a plastic bag and evacuate the air from the bag using a vacuum chamber, and then determine the mass of the sealed specimen in air (B); and 3) determine the mass of the sealed specimen submerged in water (E). Using the apparent specific gravity (F) of plastic sealing material at 25°C (77°F) provided by manufacturer, Equation 2.3 is then used to calculate the bulk specific gravity of the specimen.

$$\text{Bulk Specific Gravity} = \frac{A}{B - E - \left(\dfrac{B-A}{F}\right)} \tag{2.3}$$

d) Dimensional analysis

In this method, the density of the pavement is directly obtained by calculating the mass and the volume of the pavement section. The mass is obtained by weighing the pavement section; and, the volume is calculated based on the dimensions of the pavement section. The density is then calculated by dividing the mass of the pavement section by its dimensional volume.

2.1.2 Density measurements using nuclear devices

An alternative method of measuring density (instead of using cores) is to use a nuclear density gauge. The density values obtained using nuclear density gauges can be compared with those obtained using cores to determine an adjustment factor for the gauge measurements.

Nuclear density gauges can measure density of hot mix asphalt pavements in a span of a few minutes. The methods of density measurements using nuclear density gauges include: 1) backscatter mode, and 2) direct transmission mode. However, for measuring density of bituminous pavements, the backscatter mode is generally used (3). The direct transmission mode is used for measuring the density of materials such as soils and aggregate. The backscatter method of density measurement is described briefly in the following paragraphs.

The procedure for calibrating and operating each nuclear density gauge is provided by the manufacturer of the gauge. The methods for measuring density using nuclear gauges, calibrating gauge, correlating gauge readings with core densities are described in AASHTO T 310 (43), ASTM D 2950 (44), WAQTC TM 8 (55), ODOT TM 304 (56), ODOT TM 327 (57). A state may either follow one of the above methods, or it may develop its own procedure for determining density using a nuclear density gauge.

The basic principle employed in measuring density using the backscatter mode is as follows: 1) the gauge contains a radioactive source that emits gamma photons from the surface of the pavement down into the pavement; 2) the materials in the pavement (hot mix asphalt) absorb some of the photons, while the remaining photons are scattered back into a detector located in the gauge (13, 27); and 3) the number of gamma photons reaching the detector is used to calculate the density of the pavement (27). It should be noted that the nuclear gauge is calibrated using blocks of material of known density (typically three blocks of differing density) prior to use in the field.

Nuclear density gauges are set to display a density value after taking measurements which last thirty seconds, one minute, or four minutes. A majority of the states that use nuclear gauges use a time span of one minute to take density measurements (as observed in Appendix A). Some of the factors that vary between states that use nuclear gauges are: 1) the number and duration of readings that must be averaged to get the density value at a particular location, and 2) the direction and angle of rotation of gauge between readings (Appendix A).

2.1.3 Density measurements using electromagnetic devices

The procedure (calibration and operation) for measuring density using electromagnetic devices is provided by the manufacturer of the gauge. ASTM D 7113-05 (58) and AASHTO TP 68-04 (59) also describe procedures for determining density using electromagnetic devices.

During operation the electromagnetic device sends out an electrical sensing field into the pavement (27). This field is produced by applying an electrical charge to a conductor (5 and 36). The pavement is a dielectric or non-conductor, which impedes the flow of this electrical current, and thus decreases its strength. This decrease in strength can be measured using dielectric constants. The dielectric constant of the pavement mix can be measured by an electromagnetic sensing device. The dielectric constants of individual components of the mix are already known since they are constant values (27, 36). When the pavement is compacted, the quantity of each pavement component (e.g., aggregate, binder, etc.) remains the same, but the air void content decreases, thus changing the ratio of the volume of air to that of other components, which causes a change in the dielectric constant of the system and, thus, a change in the electric signal (5). The gauge measures this dielectric constant and change in electric signal. Then it applies a correction for sensor impedance, surface moisture and temperature variation based on data input (which has to be done every time there is a change in site conditions).

Finally, it relates this to a proportional density value since the amount and type of each pavement material remains constant except for air (5). The electromagnetic device can measure density in a few seconds. Significantly, it is not required to have a license to operate such devices (27, 36, and 39).

2.2 Summary of the Quality Control and Quality Assurance procedures currently used in Oregon for measuring in-place density.

To determine in-place density, firstly, five locations are randomly identified on each sublot (1 sublot = 1000 tons) and nuclear gauge measurements are made on them. The nuclear gauges are calibrated in accordance with ODOT TM 304. Two nuclear gauge density readings in the backscatter mode are taken and averaged at every location such that the gauge is rotated 90º after each reading as determined in figure 3.1. The average of these two readings provides the density at that location. The procedure followed to handle and operate nuclear gauges is described in WAQTC TM 8. In this manner, the density is obtained at five random locations and their average is determined. The target density on most types of hot mix asphalt concrete pavements is 92% of the moving average maximum density (MAMD) for that job mix formula (JMF) on the given lift as mentioned in section 00745.49 (b) of the Oregon Standard Specifications for Construction, 2008. The maximum density is determined in accordance with AASHTO T 209 and the MAMD is calculated in accordance with ODOT TM 305. This is the quality control testing and is the responsibility of the contractor.

Then, for acceptance or verification testing, ODOT identifies 10% of the QC tests and measures the density with their nuclear gauge. If the densities are within acceptable limits, it is used in the calculation of the pay factor. If the densities are not within acceptable limits, then an independent assurance verification testing is performed by a third party to determine the in-place density and verify contractor results. At any time, the contractor or ODOT can request for core correlations (in accordance with section 00745.49 (b) of Oregon Standard Specifications for Construction 2008) if they are skeptical about the density results. Core correlations are performed in accordance with ODOT TM 327.

2.3 Studies on the various methods of density measurement

This section provides a comparison of the different methods of measurement of density (i.e., by using cores, nuclear gauges, and electromagnetic gauges). It also provides findings regarding calibration of nuclear and electromagnetic gauges, differences of density results obtained through the different methods, and performance and accuracy of the different methods and different gauge models and manufacturers in comparison with each other.

2.3.1 Studies on the measurement of cores

This section provides findings that compare the different test standards for measuring the bulk specific gravity of cores. This includes the accuracy and variability of the various methods, as well as the factors that affect the measurements.

The densities obtained from cores are considered to be more accurate and precise (1, 3); they have lower variability than nuclear gauge and electromagnetic gauge measurements (1); thus, they are used as reference values (2). However, the coring process and laboratory procedure is time consuming (3), and transportation and testing of cores increase the quantity, duration, and cost of work (1). Core samples cannot be extracted until the pavement or mat has cooled (1, 2, and 3). Researchers have indicated that core density values are typically made available a day after the pavement has cooled (1, 2). Extraction of cores from the pavement is destructive (2, 3) and can possibly result in premature pavement failure (2) due to localized points of weakness along the pavement (1). Additionally, the pavement requires patching after the cores have been extracted (3).

2.3.1.1 Issues with the SSD method

Although the SSD method is a very reliable method, there are certain issues with this method (as observed from the literature studies). These issues are explained in detail below.

Problems associated with the SSD method at high air void content

Cooley Jr. et al. indicated that core samples with high air void contents are likely to have large voids interconnected to the sample surface; thus allowing water to quickly enter the sample during submersion in water, and quickly exit the sample after removal from water (12). This indicates that density of core specimen is overestimated when the air void content is more (interconnected voids) and when the air voids are larger (5). Cooley et al. also found that coarse-graded specimens have larger, interconnected internal air voids as compared with fine-graded specimens (12).

At high air void content, the saturated surface-dry method overestimates density and bulk specific gravity (and underestimates air void content). When air voids approach 8 to 10 percent, the mix becomes permeable and results in an overestimated bulk specific gravity and density and thus an underestimated air void content (6). This occurs because, water that infiltrates the sample when it is submerged, flows out of the core sample through the large interconnected air voids while it is being removed from the water bath (6, 12). This results in a false or less-than-actual saturated surface-dry mass indicating a decreased or underestimated sample volume, which, in turn, overestimates bulk specific gravity and density and underestimates air void content (12). From the above findings, it can be inferred that the density and bulk specific gravity is overestimated and the air void content is underestimated when one uses the saturated surface-dry method; this is especially true for coarse-graded mixes, which tend to have larger air void content than fine-graded mixes (6, 12). Thus, the saturated surface-dry method is not appropriate for coarse mixes due to interconnected air voids and significant surface irregularities (40).

Willoughby and Mahoney recommended to use the CoreLok device (which employs the Automatic Vacuum Sealing method described in AASHTO TP 69 and ASTM D 6752) when the air void content exceeds 8 percent (25).

Variability of SSD results

According to Spellerberg and Savage, “For 9.5, 12.5, and 19.0-mm mixtures containing non-absorptive aggregates, AASHTO T 166 is an excellent method for determining bulk specific gravity. The AASHTO T 166 measurement of bulk specific gravity is a very precise process (for specimens with 3% air voids and 0.5% absorption), indicated by minimal error. The error is so small that it may be impracticable to attempt to find ways to improve AASHTO T 166” (11). For the same mixes, they also found that the within laboratory and between laboratory coefficients of variation for AASHTO T 166 (SSD method) were much better than those obtained using the ASTM D 6752 (automatic vacuum sealing method) (11). Cooley Jr. et al. also found that the results from the CoreLok device (automatic vacuum sealing method) were slightly more variable than those obtained from the AASHTO T 166 method (12). They concluded that this variability might have been due to the overall lesser experience of the personnel in handling the CoreLok device (12).

Factors responsible for variability in SSD results

Spellerberg and Savage found that “For 12.5 and 19.0 mm mixtures containing non-absorptive aggregates, approximately 90 percent of the variation in AASHTO T 166 bulk density test results for 150 mm diameter superpave gyratory test specimens can be attributed to the mixing and compaction process” (11). For these mixtures, they concluded “most of the variation in AASHTO T 166 test results can be attributed to an individual laboratory’s inability to produce uniform specimens” (11). They also found that “The added variability resulting from use of different equipment, and mixing and compaction procedures in multiple laboratories does not appear to be very significant” (11). However, the results obtained using AASHTO T 166 consist of additional and unspecified reproducibility errors, which are caused by system level influences such as operators, the coring process, and the specific method utilized (38).

Alternative methods to alleviate the problems associated with the saturated surface-dry method

Gedney et al. recommended that cores with rough surfaces must be paraffin coated (15). Cooley Jr. et al. indicated that some of the alternatives to alleviate the problems associated with the saturated surface-dry method include using parafilm, paraffin wax, rubber membranes, and masking tape. Other alternatives include using compounds like zinc stearate that are hydrophobic which would prevent water from penetrating into the sample. However, these alternatives were not adopted by Cooley Jr. et al. because they caused increased variability in bulk specific gravity measurements or, damaged the sample such that additional testing could not be performed on the sample (12).

2.3.1.2 Advantages of automatic vacuum sealing method (utilizing CoreLok device) over SSD method

The CoreLok device utilizes an Automatic Vacuum Sealing procedure (similar to ones described in AASHTO TP 69 and the ASTM D 6752) to measure the bulk specific gravity of core samples. The bulk specific gravity (and density) of core samples do not differ significantly when they are measured in different laboratories and by different personnel in accordance with AASHTO TP 69 (26). Padlo et al. confirmed this fact in their studies; they found that when cores were tested (in accordance with AASHTO TP 69) by different personnel in three different laboratories, the mean difference observed was 0.06 to 0.29 as a percent of MTD, which is well within AASHTO TP 69 standards (26).

When the Automatic Vacuum Sealing method (wherein the CoreLok device utilizes the procedure described by AASHTO TP 69 or ASTM D 6752) was compared to the saturated surface-dry method (wherein cores are mostly tested in accordance with the procedure described in AASHTO T 166 or ASTM D 2726), literature studies indicate that the former method provides more accurate measurements than the latter method.

Problems associated with the SSD method (and overcome by automatic vacuum-sealing method utilizing CoreLok device)

Spellerberg and Savage found that "For 9.5, 12.5, and 19.0-mm mixtures containing non-absorptive aggregates, the bulk specific gravity values – from specimens with 3 percent air voids and 0.5 percent absorption – obtained using AASHTO T166 (2.469, 2.495, and 2.497) are significantly higher than those obtained using ASTM D6752 (2.456, 2.475, and 2.472)" (11). Even Cooley Jr. et al. found that bulk specific gravity results of coarse graded mixes (like coarse-graded superpave and stone matrix asphalt) measured using the AASHTO T 166 were significantly higher than those of the CoreLok – when the water absorption exceeds 0.4 percent (12). But, at low levels of water absorption (high density values), the bulk specific gravity measured by the two test methods are similar (i.e., they do not differ significantly) (12). This indicates that the AASHTO T 166 method overestimates bulk specific gravity at high water absorption levels (unlike the CoreLok device) (12). Hence, Cooley Jr. et al. recommended that at water absorption above 0.4%, the CoreLok must be used to measure density (12).

Relation between coarse and fine graded specimen, air void size and content, and water absorption content

The AASHTO T 166 method also overestimates bulk specific gravity at high air void levels (unlike the CoreLok); and high air void content with larger air voids is more prevalent in coarse-graded specimen (and mixes) than in fine-graded specimen (and mixes) (12). Even Williams and Hall indicated that fine graded mixes (12.5mm) have smaller air voids when compared to coarse graded mixes (37.5mm) (24).

From the above findings, it can be inferred that the air void content and the water absorption are directly related to each other such that, high air void content leads to greater water absorption (and vice versa); and low air void content indicates to low water absorption (and vice versa).

Cooley Jr. et al. also found that, for coarse gradations, as density (or compaction) of the sample increases, the bulk specific gravity measurements obtained from the AASHTO T 166 and the CoreLok method become closer (12). Generally, higher density and higher compaction mean lesser air void content and lesser water absorption (and vice versa) (5, 12). In other words, based on all the above studies, it can be inferred that when a sample (or mix) has higher density or is highly compacted (low air voids and less water absorption), the AASHTO T 166 and the CoreLok measurements are precise; but at low densities or low compaction (high air voids and more water absorption), AASHTO T 166 overestimates the bulk specific gravity.

Influence of gyration level on SSD and CoreLok BSG results of coarse-graded specimen

For coarse specimens, as the gyration level decreases, there is an increase in the difference between AASHTO T 166 and CoreLok measurements of bulk specific gravity. In other words, AASHTO T 166 overestimates bulk specific gravity to a greater extent at lower gyration levels than at higher gyration levels. This was indicated by Cooley Jr. et al. in their studies; they noticed that, among the coarse specimens (coarse-graded superpave and stone matrix asphalt specimens at gyrations of 15, 50, and 100 respectively), the largest difference between AASHTO T 166 and CoreLok measurements of bulk specific gravity was observed in the 15 gyrations (0.041 and 0.045 for Coarse and SMA mixes); slightly lesser difference was observed in the 50 gyrations (0.012); and least difference was observed in the 100 gyrations (0.008) (12).

Summary of Factors influencing BSG results obtained using SSD method

Thus, it can be concluded that, unlike the CoreLok method, the AASHTO T 166 overestimates the bulk specific gravity of a sample (or mix) if the mix is a coarse mix (i.e., utilizing coarse-graded aggregates), if the density or compaction is low, if the gyration level is low, if the air void content is high, or if the water absorption is high (12). However, no significant difference exists between the CoreLok and the AASHTO T 166 measurement of bulk specific gravity for mixes that utilize fine graded aggregates, have high densities (or compaction), have high gyration levels, have low air void content, or have low water absorption levels (12).

2.3.1.3 Advantages of automatic vacuum sealing method over other core testing methods

Overall, the Automatic Vacuum Sealing method (which is described in AASHTO TP 69 and ASTM D 6752 and which is utilized by the CoreLok device) seems like a more viable option for measuring bulk specific gravity, when compared to the saturated surface-dry method (which is described by AASHTO T 166 and ASTM D 2726). Cooley Jr. et al. suggested that “the CoreLok vacuum-sealing device could be used to determine the bulk specific gravity of compacted hot mix asphalt samples with greater accuracy than conventional methods such as water displacement, parafilm, and dimensional analysis.” Cooley Jr. et al. also cited another source to indicate that “the CoreLok vacuum-sealing device provides a better measure of internal air void content of coarse graded mixes than other conventional methods” (12, page 6). Willoughby and Mahoney indicated that the Automatic Vacuum Sealing method provides a better estimate of density than the saturated surface-dry method and the Paraffin (and Parafilm) method; this is especially significant for samples that have air void content greater than 8% (25). Even Padlo et al. recommended that all cores must be tested for density in accordance with AASHTO TP 69 (i.e., the Automatic Vacuum Sealing method) (26).

2.3.1.4 Disadvantages of CoreLok device (automatic vacuum sealing method)

Even though the CoreLok device seems like the most viable option for measuring bulk specific gravity and air void content, it still has some shortcomings. Spellerberg and Savage found that “For 9.5, 12.5, and 19.0-mm mixtures containing non-absorptive aggregates, when ASTM D 6752 testing errors occur, they result in densities that are considerably lower than the expected density” (11). Cooley Jr. et al. suggested that even the density values obtained from the CoreLok device can be slightly overestimated due to bridging of the plastic bag over a sample's surface voids. This over estimation can be greater for laboratory prepared samples, which have texture on all sides as opposed to field cores (12). The plastic bags specified in ASTM D6752 are susceptible to leakage due to pin holes (11). Some labs (of the total of 18 labs) reported that the CoreLok bags used for samples were sometimes punctured; this could lead to water infiltration or losing of vacuum over time (12). Specific concerns with the ASTM D 6752 include the durability of the plastic bags (susceptibility to pin holes), possibility of the bag contacting the side of the bath, and air being trapped under the bag while immersed (11). It may not be possible to dry back specimens that take on water during immersion (11).

Spellerberg and Savage recommended that if AASHTO T 166 (saturated surface-dry method) is used to determine the bulk specific gravity of specimens during the mix design process it should also be used in quality control testing (11). They also recommended that if ASTM D 6752 (Automatic Vacuum Sealing method) is used to determine the bulk specific gravity of specimens during the mix design process it should also be used in quality control testing (11).

2.3.1.5 Disadvantages of dimensional analysis method

The Dimensional Analysis method overestimates the internal air void content and the volume of the core sample, and thus it results in an underestimated density value (12, 13). This was also indicated by Choubane et al. in their studies; they found that when the densities obtained from three core testing methods (Florida test method FM 1-T 166 similar to AASHTO T 166 except that immersion of sample was for 2 $\pm$ 1 minutes instead of 4 $\pm$ 1 minutes, ASTM D 1188 method using parafilm, and dimensional analysis method i.e., weight and dimensions of cores were measured) were compared, based on R^2 value of .93 to .97, they seemed to be quite similar. But, from ANOVA F-value based on 95% confidence interval, there was indication of significant differences between the three values per location. The density values obtained from the dimensional method were always lower. There was better differences between the FM 1-T 166 (SSD) and the ASTM D 1188 (Parafilm) method (maximum difference was 17 kg/m^3 and minimum difference was 2 kg/m^3 and mean difference was 11.1 kg/m^3) (13).

2.3.2. Studies on nuclear gauges

This section provides information that compares different standards for measuring density using nuclear gauges. This includes information on variability of gauge density readings, accuracy of various methods to measure density of hot mix asphalt using nuclear gauges, comparison of different nuclear gauges by manufacturer and model, methods for accurate calibration of gauges, and factors that affect nuclear gauge density readings. This section also provides information related to accuracy of gauge density readings when compared to density readings obtained by cores. Information is provided on calibration of the nuclear gauge and differences of the nuclear gauge readings with corresponding core densities to obtain more accurate results. Other included information relates to factors affecting differences of gauge and core readings, and variability of gauge readings.

General advantages of nuclear gauges

Nuclear gauges use a test procedure that is fast and non destructive (1, 2, and 3). They can be used to provide results in less than ten minutes (1). The principal advantage of using nuclear gauges is that the pavement density can be measured immediately (while the hot mix asphalt is in a workable state) allowing real-time adjustments to rolling patterns (during compaction) so as to achieve desired density (2). In other words, density can be measured while the mat is still hot, so additional compaction can be applied if necessary (3). Thus, a greater number of measurements are possible, which allows for a more accurate estimate of pavement density (2). One requires licensing to be able to operate a nuclear gauge (1). Certified Technical Training is required for a technician to implement the Quality Control/Quality Assurance process (30).

2.3.2.1 Factors that influence nuclear gauge density measurements

There are a number of significant factors that have considerable influence on the nuclear gauge measurements. Various details that describe these factors and the level to which they influence the nuclear gauge measurements are indicated below.

Influence of lift thickness on nuclear gauge readings

Nuclear gauges are affected by pavement thickness. As the thickness of a pavement decreases, the density results obtained from a nuclear gauge becomes more variable. This is because, if the thickness is less, the nuclear gauge reading is influenced by underlying layers. Esch found that a minor change in compaction (compaction affects thickness) can cause a huge change in density reading (21). Gedney et al. found that nuclear gauges are sensitive to thin overlays and underlying material, such that the readings obtained are an average of the overlay and the underlying material (15). They also found that nuclear gauges are sensitive to the top 2 inches of pavement material, and it gives greatest weight to materials that are closest to the bottom of the gauge (15). The sub base material influences density readings more than the base material. While Huot et al. observed that the base course material has a negligible effect on density variability (19), Parikh indicated that the sub base material significantly affects the density reading (33). Parikh also found that, if thickness is 1.5 to 2 inches, then accuracy of nuclear gauge density measurements is affected by one percent; and if top layer thickness is 1 inch or lesser, then accuracy is affected by one to two percent (33). Manufacturers recommend that nuclear gauges must be used only if the pavement thickness is greater than 2.54cm (or 1 inch) (1). While Sanders et al. found that nuclear gauge density readings are influenced by the mix type (1), Parker and Hossain observed that the density measurement variability is lesser in thicker binder mix layers than in thinner ones (3). They suggested that this can be due to the effect of mat thickness on density determination and due to slower cooling of thicker mats that permits better compaction control (3). Even Schmitt et al. found that mat thickness influences variability of nuclear gauge density reading (31). Choubane et al. found that the thin lift gauge results were highly variable in spite of being corrected for the effect of underlying layers using the previous density data (13). Padlo et al. observed that as the mat thickness increases, the error or variability between gauge and core readings decreases. They found that mat thickness and aggregate source caused significant difference between gauge and core results in the range of 0.3 to 1.2% of MTD (26). They recommended that the thickness of some pavements must be entered into certain gauge so that the influence of underlying layers, while determining density, is minimized. However, they also pointed out that entering the exact depth doesn't minimize this effect significantly because different gauges measure density up to different depths; thus it is necessary to enter a lesser-than-actual depth in order to eliminate the effect of an underlying layer in determining the density of the current layer (26).

Influence of aggregates on nuclear gauge readings

Huot et al. observed that the aggregate size influenced the nuclear gauge readings such that it caused density variation to the range of $\pm 2\%$ (19). Esch found that minor changes in gradation, asphalt content and compaction can cause huge changes in density readings (21). Mix composition and mix aggregate type also significantly affects nuclear gauge density reading (33, 1, and 3). Parker and Hossain suggested that different acceptance limits should be set for different mix types. The acceptance limits should be set using standard deviation of measurements for specific mix types (3).

Influence of surface texture on nuclear gauge readings

Huot et al. found that the surface texture has a big influence on nuclear gauge readings (19). Esch found that the correlations (between gauge and core readings) are affected by surface roughness and insufficient depth of penetration (or influence) of nuclear gauges as compared to cores (21). Gedney et al. found that the density values obtained from nuclear gauges were significantly lower than those of cores for pavements with rough surface texture; this is due to the inclusion of surface voids in gauge's sensitive volume. The correlations improve when the pavement is smoothened (15). Rough surface texture causes the gauge to show a reduced density reading than the actual density reading (26). A rough surface texture has an uneven surface and thus more air voids will be present between gauge and pavement surface; thus a less-than-actual density reading will be displayed by the gauge (26).

Influence of gauge type, testing location, and site condition on nuclear gauge readings

Type of gauge (or gauge model), project conditions, and location are factors that significantly affect gauge readings. Burati and Elzoghbi found (while measuring the joint density) that the between-gauge differences is affected by project condition since it significantly varies from project to project (35). Also, Stroup-Gardiner and Newcomb found that little differences existed between core density readings and nuclear density readings; and a two-way ANOVA test indicated that the factors responsible for significantly different results between nuclear and core readings were gauge type, testing location, and site condition. (14). Even Choubane et al. indicated that the gauge model and the testing location are factors that significantly influence the gauge results. They maintained that variability differs from gauge to gauge and, within each gauge, from location to location (13). Padlo et al. also found in their studies that the type of gauge used to measure pavement density is a cause for significant density variation. The differences between gauge results were statistically significant. The mean difference in reading between 2 gauges at the same location was 0.3 to 1.36 as a percent of maximum theoretical density. No two gauges produced statistically similar results (out of the six nuclear gauges used) (26).

Other factors that influence nuclear gauge readings

Some of the other factors that affected nuclear gauge density reading were differences in temperature, using latex as a modifier, milling of old pavement surfaces, air gap, and nuclear gauge operator. Huot et al. found that a difference in temperature of almost 100^0F can cause a difference in density by 1 pcf (19). Parker and Hossain found that using latex as a modifier in surface mixes significantly increases variability of mat density (3). They also found that milling and surface treatment of surfaces of old pavements significantly decreases density variability in overlay mats; milling has a greater effect by increasing uniformity of overlay mat compaction (3). Even Padlo et al. recommended milling of pavements because it decreased the variability of nuclear gauge density readings (26). Troxler Jr. and Dep found that air gap has a significant effect on the gauge reading. When gauge is placed right on the joint and there is 1mm elevation difference between adjacent lanes, the density is under estimated by about 2pcf (18). The operator of the nuclear gauge is also responsible for causing significant variations in density reading (1, 13). When Choubane et al. correlated the manufacturer representative-operated nuclear gauge readings (obtained from Troxler model 3440 and 3401) with corresponding core results (obtained using the ASTM 1188 Paraffin & Parafilm method), they found that correlation was very good (R^2 was greater than 0.89). They also found that among various gauge operators (State DOT, Contractor, and Manufacturer Representative), the manufacturer's results indicated best correlations. Thus, one requires licensing to be able to operate a nuclear gauge (1). Certified Technical Training is required for a technician to implement the Quality Control/Quality Assurance process (30).

2.3.2.2 Factors that do not influence nuclear gauge density measurements

Literature studies indicate that factors such as gauge direction, rubber pads, sand, coal tar, standard count, and date of calibration do not have any significant influence on nuclear gauge readings and/or do not cause significant variability in gauge readings. The position of the gauge (transverse or longitudinal), while taking measurements, is an insignificant factor which does not affect gauge readings. The mean difference between density obtained in transverse and longitudinal direction was 0.05% of MTD. Thus, it was concluded that this factor does not influence density reading (26). Burati and Elzoghbi found that orienting the gauge in a perpendicular direction to the joint (rather than parallel) gave density values that were closer to core densities obtained from the joints; however, they also concluded that it was not yet appropriate to replace gauges with cores for Quality Assurance purpose (35). Placing rubber pads or material of known density between gauge and irregular surface does not improve results from a nuclear gauge; it is an insignificant factor that does not affect gauge reading (26). Sanding of the pavement and sealing of coal tar did not seem to cause variability in gauge density measurements (14). Even the standard count and date of calibration did not appear to be a cause for varied density readings (14).

2.3.2.3 Number and duration of nuclear gauge measurements

Esch suggested a formula (Number of readings = (3.92*Standard Deviation/Range) ^{2}) to calculate the number of readings required to achieve a 95% confidence level (21). Stroup-Gardiner and Newcomb observed in a paired t-test (with a 99% confidence) that there was no significant difference between the 15 second, 1 minute and 4 minute readings (14). Even Parikh found that the count duration did not have significant effect on nuclear density values. He recommended that to be assured of accuracy within $\pm$1pcf of true density, four readings (using special calibration mode) are required per sub lot (200 tons) (33). He found that laboratory analysis indicated no significant difference in average of four readings taken at one, two, or four minute counts (33). Schmitt et al. indicated that a total of at least 2 replicate density readings (one-minute readings) are required per site (or testing location) to eliminate erroneous results and to control variability (31). Relative densities (i.e., percentage of theoretical maximum density) determined from measurements on adjacent cores obtained at each test site (i.e., where the nuclear gauge measurements were made) varied from 0.0% to 0.5%, whereas the relative densities obtained from nuclear gauge measurements varied from about 0.8% to 1.2% based on 1-minute counts. It should be noted that the researchers determined that a 15-second count resulted in a variation in relative density of about 2% to 3% and that a 4-minute count resulted in a variation in relative density of about 0.8 to 1.0% (very similar to the 1-minute count) (2). Sebesta et al. recommended taking 2 one-minute readings (in backscatter mode), with the nuclear gauge rotated 180^0 after the first reading. The average of the 2 readings must be recorded as the nuclear density measurement at that location (10). Padlo et al. also found that one-minute readings provide more accurate results than thirty-second readings. They found that using one-minute readings (instead of thirty-second readings) improved accuracy of gauge results by 0.45% of maximum theoretical density (26). They recommended taking 4 one-minute readings at each location to measure pavement density using a nuclear gauge (26). Troxler Jr. and Dep also found that counting time (or reading duration) significantly affected gauge results; and they suggested that for best precision, 2 one-minute readings are recommended per location or spot (18). Schmitt et al. recommended increasing the nuclear gauge density tests from 7 to 15 (per 750-ton lot) to increase the confidence in the estimated pavement density (37).

2.3.2.4 Comparison of methods of operating different nuclear gauges in various modes

Nuclear gauges can be operated in different modes to measure the density of asphalt pavements. Some of the modes are standard backscatter mode, backscatter with air gap mode, and thin lift mode. The mode of operation of the gauge can significantly influence the density reading. Esch found that air gap testing with one surface and one air gap count generally produced standard deviation's approximately 20% higher than normal backscatter testing using the average of three one minute counts. Therefore they recommended that the air-gap method should not be used (21). Later, Gedney et al. found that the nuclear gauge was more sensitive to chemical composition and surface roughness in the backscatter mode, and hence the better option is the backscatter with air gap mode. They also found that the backscatter mode was effective in measuring density of pavements up to a minimum thickness of 1.8 inches (15). Campbell Pacific Nuclear (CPN) recommends using deeper backscatter for layers greater than 2 inches thick and shallower AC position for layers 1.4 inches thick (15). Parikh found that the mode of operation of the gauge influences the density reading, and it can possibly cause variability in results (33). When comparing normal, special calibration and surface void modes the analysis indicates that mean nuclear densities with special calibration mode was same as mean core densities. For the special calibration mode, a 95 % confidence interval for difference showed that difference between core and nuclear gauges was less than 2pcf. Thus special calibration mode seems favorable for best correlation (33). He suggested that the normal mode of operation can be used to measure density in normal aggregates like limestone and granite. The surface void mode of operation can be used for coarse mixes having less than 40% fines passing the number 8 sieves. The special calibration mode is used when the gauge readings are significantly different from core results due to asphalt mix composition and mix aggregate type. Special calibration mode is most suitable for thin layer density. Special calibration mode is recommended when operating the nuclear gauge because it showed least difference between gauge and core densities (1 to 2pcf) (33). Padlo et al. also found that the mode of operation of the gauge (Backscatter or Thin Lift) significantly affects nuclear gauge reading. They observed that both the modes showed significant differences in density results obtained (probability of difference in results between 2 modes being greater than 0.1% of MTD was 0.97 i.e., a very high probability). The thin lift results were on an average 0.33% of MTD higher than that of backscatter mode. This occurred in spite of receiving equal measure of radiation, so the difference might be due to the difference in algorithm (26). Even though statistically significant difference between the results obtained from thin lift mode and backscatter mode was observed, there still wasn't enough evidence to suggest that one of the methods is the better of the two (26).

2.3.2.5 Recommendations to improve nuclear gauge calibration and differences procedures

Huot et al. found, while comparing gauge and core readings, that a tolerance of ±2% could result in a 95% confidence interval on a calibration curve. Also, accuracy could be improved if more than one calibration was taken (19). Esch suggested that at least 7 to 10 correlation sets are required to find the average correction factor (21). Gedney et al. found that using a control strip ensures that density measurements are precise (15). They suggested taking the average of 10 gauge density measurements on the control strip to get the control strip density; when field measurements are made on the pavement (with the same mix), the average of 5 gauge readings must exceed 98% of control strip density (15). They also suggested that, if necessary, the control strip density itself could be a percentage of laboratory density (Marshall or Theoretical) (15). Sanders et al. found that calibration procedure and measurement procedure are factors that significantly affect nuclear gauge density reading (1). Parker and Hossain found that nuclear gauges need to be calibrated every time for specific conditions like mix composition, layer thickness, and underlying conditions or else the readings obtained will not be accurate and will vary widely (3). Accuracy of nuclear gauge can be occasionally determined by initially making a simultaneous comparison and differences with core density measurements. Then they might be used for Quality Assurance purpose instead of cores (6). Padlo et al recommended selecting one location on a paving project so that 4 one-minute readings can be taken every day to monitor the gauge and ensure its accuracy (26). Troxler Jr. and Dep found that the problem with factory calibration is that the gauge is calibrated based on the average asphalt mix. In order to make the actual field calibration on a job site more accurate, the gauge must be calibrated for that particular mix by correlating the readings to core density readings. They suggested that these readings must be obtained from random locations in the middle of the pavement to estimate a constant bias between the two methods (18).

2.3.3 Studies on electromagnetic gauges

Literature studies indicate that a lot of research has been conducted on the following electromagnetic gauges: the Pavement Quality Indicator (Pavement Quality Indicator Model 300 and Pavement Quality Indicator Model 301) and the PaveTracker (PaveTracker 1B, PaveTracker Model 2701, and PaveTracker Model 2701-B). These gauges might possibly be the most efficient electromagnetic gauges available. A lot of literature studies exist to identify the various pros and cons of these electromagnetic gauges. This includes an identification of factors that affect (and factors that do not affect) specific electromagnetic gauge models; it also includes other related conclusions and recommendations.

General advantages of PQI

Rogge and Jackson preferred to use a electromagnetic gauge (Pavement Quality Indicator) for field use. This was because it weighed less than 4.5kg and used a capacitance (complex impedance technique). The device makes use of a technology that employs constant current, low frequency, and complex impedance. The device can be used to create an electrical sensing field that is established in the material by a flat sensing plate, and measurements are thus taken. This approach allows the depth of measurement to be controlled precisely (16). Kvasnak et al. found that the Pavement Quality Indicator can be set to measure density up to a controlled minimum depth of 2 inches without the influence of underlying layers (36, 40). However, if the measuring depth of the gauge is more than the pavement thickness, then the density value obtained is the combined density of the composite material (i.e., combined density of the pavement and its underlying material) (5).

2.3.3.1 Factors that influence electromagnetic gauge measurements

From the literature studies, a lot of information was obtained on factors that influence the electromagnetic gauge measurements. The details on individual factors and the degree or extent to which they influence the measurements are provided below.

Influence of temperature on electromagnetic gauge readings

Earlier models of the Pavement Quality Indicator (namely the Pavement Quality Indicator 100 and the Pavement Quality Indicator Model 300) and the PaveTracker are affected by the surface mix temperature (34, 17, and 5). Gauge reading decreased with a decrease in surface mix temperature (17, 5). Liao et al., found that, on an average, the gauge reading decreased by 16kg/m^3 with a temperature drop of 50^{0}C (5). Likewise, they was also found that the gauge reading (measured using a PaveTracker) increased with the increase in surface temperature; and this is more profound for coarse mixes (19mm) than for fine mixes (9.5mm) (5). However, it was found that the latest electromagnetic gauges (Pavement Quality Indicator Model 301 and PaveTracker Model 2701-B) are hardly affected by pavement or mat surface temperature (27, 24). Williams and Hall verified this by measuring the density using these two electromagnetic gauges at two temperatures (high temperature of 180^{0}F i.e., just behind the finish roller; and low temperature was after mat had cooled considerably i.e., approximately 100^{0}F). They found through their analysis (t-tests) that the electromagnetic devices' measurements were unaffected by temperature (i.e., t-tests' results indicated no significant difference, p-value>0.05). Thus, it can be assumed that manufacturer's have corrected these gauges to account for temperature variability while providing density readings (24).

Influence of moisture on electromagnetic gauge readings

Surface moisture and internal moisture also significantly affect gauge readings obtained from earlier gauge models (namely the Pavement Quality Indicator Model 300 and the PaveTracker) (34, 17, and 5). Increase in the surface moisture (without internal moisture) causes the gauge reading to display a lower or decreased value, while increase in internal moisture (with or without surface moisture) causes the gauge reading to display a higher or increased value (17, 5). When compared to cores, it was observed that gauge (PaveTracker) and core readings increased linearly with increase in surface moisture (5). However, when subjected to surface and internal moisture (for 24 hours), it was noticed that the gauge and core measurements were non linear, and the gauge readings decrease as the core readings increase (5). Kvasnak et al. also found that gauge density measurement results of wet pavements were more variable than those of dry pavements (36, 40). Surface and internal moisture also affect the newer electromagnetic gauge models (namely the Pavement Quality Indicator Model 301 and the PaveTracker Model 2701-B) (24, 27). This was indicated by Williams and Hall in their studies. They measured density on two conditions: a 'dry' condition, where there was no visible sign of water; and a 'wet' condition, (resembling a SSD condition) where water was sprayed on the surface to fill the voids and then a towel was used to blot the surface. It was found that moisture caused a significant change in density. It affected the 37.5 mm mix (which has larger voids) more than the 12.5 mm mix. The 'wet' condition caused an increased density reading (by 1pcf). Thus it was recommended that all visible signs of water must be removed from the surface. The PaveTracker Plus™ Model 2701-B was a little more sensitive than the Pavement Quality Indicator Model 301 (24). Smith and Diefender found that the Pavement Quality Indicator Model 301 could provide valid results when the moisture index was below 10; however, if the moisture index was above 10, the readings were significantly affected. A higher moisture index indicated higher density reading and more deviation from core reading (27).

The Pavement Quality Indicator Model 300 seemed to be affected by environmental conditions; it also seemed to be affected by the density gradient created by the influence of rolling wheel compactor (32).

Influence of type of mix and aggregate size on electromagnetic gauge readings

Electromagnetic gauges are affected by the type of mix (i.e., coarse or fine) (17, 9, 37, 36, and 24). Sargand et al. found that the PaveTracker displayed more accurate results for measurements that were made on fine mixes than on coarse mixes (17). Larsen and Henault observed that in spite of changing the algorithm of the Pavement Quality Indicator and the PaveTracker (to provide improved results), the measured density values were affected by the mix design (i.e., aggregate size and proportion, and particle segregation) (9). Even Schmitt et al. found that electromagnetic gauges are affected by mixture- and project-specific factors like: aggregate source, design loads, laboratory air voids, asphalt content, aggregate specific gravity, and pavement layer thickness (37). The newer gauge models like the PaveTracker Model 2701 were also affected by various factors like: contractor, traffic level, aggregate type, NMAS, roller passes, and moisture condition (i.e., wet and dry conditions). Pavement Quality Indicator Model 301 was affected in the single mode by factors like: station, roller passes, temperature, and distance across pavement width. In multi mode (method of data collection) it was affected by factors like: site, station, pavement width, contractor, aggregate type, binder content, roller pass, and distance across pavement width (36). Williams and Hall indicated that the gauges were significantly affected by large aggregate size and void size (37.5mm mix). They measured density at the same spot (for all locations) using both the gauges (the PaveTracker Plus^{TM} Model 2701-B and the Pavement Quality Indicator Model 301). It was found that the gauge type did not show significant differences in the 12.5mm mix, which has smaller voids. However, it showed significant differences in the 37.5mm mix, which has larger voids (this was indicated in the ANOVA table with p-value <0.0001) (24). Liao et al. found that gauge measurements on fine mixes (9.5mm) have better correlation with cores than differences of coarse mixes (19mm). This is indicated by the coefficient of determination (R^2) for density results from fine mix (0.94), which is larger than that of coarse mix (0.81). This might be due to the smooth texture of fine mix leading to better surface contact between gauge and pavement (5).

Influence of air voids and sand on electromagnetic gauge readings

Air voids seem to significantly affect the electromagnetic gauge reading (18, 37, 40, 36, and 24). Troxler Jr. and Dep found that when a factory calibrated gauge is placed right on the joint – with a 1mm elevation difference between adjacent lanes, the density is under estimated by 0.2pcf at 95pcf, 2pcf at 123pcf and 6pcf at 160pcf (18). Williams and Hall found that sprinkling of sand (to fill voids and level the pavement surface) causes a decrease in density, which is more significant in mixes with smaller voids; and the PaveTracker Plus^{TM} Model 2701-B is more sensitive to it. They measured density under two site conditions: the first condition was to brush off all the sand from the surface and then measure the density leaving it susceptible to air voids, the second condition was to sprinkle sand on the surface in order to level it such that it fills only the voids and the remaining sand it dusted off. Generally, more voids decreases density. However, the results indicated that it is better not to try to level the surface by filling the voids with sand since it tended to prevent solid contact between gauges and surface. Sprinkling of sand (to level the surface) caused a decrease in density (by approximately 3pcf) in 9 of 10 locations. Also, sand affected the mixes with smaller voids more significantly and the PaveTracker was more sensitive to it (24).

Influence of area of surface on which electromagnetic measurements are taken

Area of laboratory prepared specimen also significantly affects gauge reading (17). Sargand et al. indicated that when a mat of side 12 inches was reduced to side 6 inches, the average gauge density decreased by 2.4pcf. Similarly, when mat of size 12 inches was increased to 22.5 inches, density reading increased by 0.4pcf (17). Liao et al. found that the surface area of a specimen can significantly affect the gauge density readings if the surface area of the specimen being measured is less than two times the area of the contact plate of the PaveTracker gauge. This change in gauge readings when surface area changes was observed when sample size was changed from side 30cm to side 15cm, causing average decrease in density value by 38kg/m^3, and so sample of side 15cm is not recommended. Similarly, when sample size was increased from side 30cm to side 57.2cm, the average increase in density value was 6.4kg/m^3 (5).

Influence of gauge position on gauge readings

The direction of the gauge (i.e., parallel or perpendicular) also shows considerable variance in gauge results. Williams and Hall found that the measurements taken parallel to the direction of paving (parallel orientation means that a technician facing the direction of paving could read the gauge display screen directly) were approximately 2.4pcf higher than when placed perpendicular to the paving direction (24). They also found that rotating the gauge increased the contact and caused changes in reading; also, “picking up and placing the gauge” caused further compaction which led to increased density by approximately 3 pcf (24). Thus, they did not recommend rotation or lifting of the gauge in between readings (24).

2.3.3.2 Factors that do not influence gauge readings

It was confirmed that battery voltage and mineral filler do not influence (or affect the accuracy and variability) the electromagnetic gauge reading. Battery voltage has no effect on the gauge (PaveTracker) reading (hence the gauge can be used at the safe, low battery voltage of 8.2V before it can be recharged) (10). Also, the use of mineral filler does not affect accuracy or variability of density reading with Pavement Quality Indicator Model 300 (32).

2.3.3.3 Number of electromagnetic gauge readings at a given location

Number of readings significantly influences density values. Sebesta, S. et al. recommended taking 5 three-second readings (with the Pavement Quality Indicator) by beginning in the center and then moving clockwise around the center, and moving the instrument at least 50 mm (2 inches) between readings. The average of the 5 readings is to be recorded as the Pavement Quality Indicator density measurement at that location (10). Troxler Jr. and Dep recommended that, for best precision, one must take 4 to 8 readings (on places slightly shifted about the measurement locations by about 2 in.) when they use the PaveTracker Model 2701 gauge. In this way, it can provide 0.5 to 1.2pcf precision at 1-sigma level (18). Williams and Hall took one, four, and eight readings at the same location to assess the repeatability and variability of readings. It was found that it was best to average four reading to get the reading at a particular location. One reading generally showed an overestimated density, and eight readings didn't show significant difference when compared with four readings (24).

2.3.3.4 Methods to calibrate electromagnetic gauges to get more accurate density results

The electromagnetic gauge can be made to produce more accurate readings by correcting it for errors. Schmitt et al. recommended that electromagnetic gauges must be calibrated daily at each project site (to offset the effect of various factors) (37). They also found that nuclear gauge readings were the most appropriate means to calibrate an electromagnetic gauge (37). They also recommended that technicians using electromagnetic gauges take at least 30 readings within each lot to ensure acceptable confidence levels (37). Rao suggested that an optimal adjustment function must be calculated by correcting the raw electromagnetic readings to match core readings. This should include an additive shift and a slope factor (20). They suggested that the gauge readings can be corrected (using the *slope* factor) by multiplying a constant correction factor (or slope term) to the raw gauge readings (20). Also, the gauge readings can be corrected using the slope and intercept method, wherein, the electromagnetic readings are adjusted by multiplying a slope term (C3) and adding an intercept term (C4) (20). They recommended that a calibration factor must be determined using the slope function and based on 10 calibration points; and this must be implemented for agency use. Daily calibration for each mix design is recommended when the project involves multiple days of paving. The use of 10 test points and a correlation using the *slope* method is optimal to develop accurate calibration factors. Finally, the study recommends that independent calibrations be established for each day of paving (20). Also, Smith and Diefender suggest that for each project, if the gauge is calibrated on a test strip before actual field testing and measurement is performed, then it can be used for quality control (Quality Control) and quality assurance (Quality Assurance) (27).

However, certain studies exist to suggest that calibration of gauges doesn't improve differences between results obtained from gauges and cores (29, 32, 7, 36, and 40). Prowell and Dudley found that the Pavement Quality Indicator Model 300 was insensitive to pavement changes and also displayed poor differences to AASHTO and CoreLok densities (29). Hausman and Buttlar indicated that poor differences between Pavement Quality Indicator Model 300 and AASHTO T166 method is due to influence of density gradient created by rolling wheel compactor (32). Later, Hurley et al. found that calibrating gauges (Pavement Quality Indicator Model 301, 300, and PaveTracker) for offset and slope didn't improve differences between densities obtained from gauges and cores (7). Also, statistical evaluations by Kvasnak et al. indicate that the majority of density readings obtained via electromagnetic gauges and cores differed (36). They found that in most cases, when readings obtained from PaveTracker, Pavement Quality Indicator and cores were compared, it was observed that PaveTracker showed the lowest readings, core reading was in between that of PaveTracker and Pavement Quality Indicator, and Pavement Quality Indicator showed highest readings (36). Variability of the Pavement Quality Indicator readings tended to be more than that of PaveTracker and core readings. When quality indices were compared for the unadjusted data, Pavement Quality Indicator and core results differed on just 1 out of 14 sites but after correction they differed on 7 sites. When quality indices were compared for the unadjusted data, PaveTracker and core results differed on 11 out of 14 sites but after correction they differed on just 1 site. Pavement Quality Indicator showed most variability, followed by PaveTracker. Cores showed least variability. When a new algorithm was developed for the Pavement Quality Indicator and results obtained from this new algorithm were compared with the results obtained from the old algorithm, it was observed that the new algorithm yielded better results (36).

2.3.3.5 Comparison of electromagnetic gauges and recommendations for QC/QA purpose

Earlier studies by Romero indicate that the Pavement Quality Indicator Model 300, the Pavement Quality Indicator Model 300+ (i.e., modified algorithm which gives improved results), and the PaveTracker cannot be used for Quality Assurance purpose, but they can be used for Quality Control purpose (39). Allen et al. and Andrewski also found that the Pavement Quality Indicator Model 300 can be used for quality control (8, 34). Hurley, G.C. et al. found that the Pavement Quality Indicator Model 301 gives more accurate results than the Pavement Quality Indicator Model 300 and the PaveTracker (7). Smith and Diefender found that the Pavement Quality Indicator Model 301 can be used for Quality Control and Quality Assurance; provided it is calibrated daily on a test strip for each project before actual testing and measurement is performed (27).

Romero recommended that the electromagnetic gauges (Pavement Quality Indicator Model 300 and PaveTracker) should be operated by experienced professionals, and that a reference standard should be developed for the Pavement Quality Indicator just like the PaveTracker (39). Kvasnak et al. suggested that to ensure the appropriate implementation of electromagnetic gauges, there is a need for additional work that considers the following elements: 1. Utilization of test strips 2. Increase the electromagnetic gauge testing frequency 3. Evaluate new electromagnetic gauges that have entered the construction industry (e.g., Pavement Quality Indicator models 302 and 303) (36).

2.3.4 Studies comparing cores, nuclear gauges, and electromagnetic gauges

This section provides findings that compare nuclear and electromagnetic gauges across various parameters. This includes information on variability of gauge density readings, accuracy of density results obtained using nuclear gauges vs. electromagnetic gauges, comparison of different gauges by manufacturer and model, methods for accurate calibration of gauges, comparison of density results and their proximity to results from corresponding cores, and various factors that affect gauge density readings.

Irrespective of the method i.e., gauge or core, random sampling is required for an unbiased estimate of density or a true representation of in-place density (31, 7). The random selection of test site is dependent on random station and offset (31).

2.3.4.1 Advantages of electromagnetic gauges over nuclear gauges (cost and use)

Electromagnetic gauges have a few advantages over nuclear gauges in terms of cost and time. Hurley et al. observed that obtaining density measurements using electromagnetic gauges are cheaper compared to using nuclear gauges; also, electromagnetic gauges measure density faster than nuclear gauges (7). Electromagnetic gauges provide density readings in a few seconds, while nuclear gauges generally display acceptable density readings in a minute, or in some cases, more than a minute (27). Also, electromagnetic gauges are easier to use and weigh approximately half as much as nuclear gauges (27). Electromagnetic gauges do not require licensing unlike nuclear gauges (27).

2.3.4.2 Advantages of nuclear gauges over electromagnetic gauges (accuracy)

Nuclear Gauges, however, are found to be more accurate than electromagnetic gauges. A majority of the literature indicates that densities obtained from nuclear gauges correlate better with cores than densities from electromagnetic gauges. Rogge and Jackson found that the density values obtained from a nuclear gauge correlated better with those of cores than corresponding density differences of electromagnetic gauge (Pavement Quality Indicator) and cores (16). Electromagnetic gauges were never able to match the standards set by the nuclear gauge when they were tested for correlations with cores (22). When nuclear and electromagnetic gauges were compared to check which gauge provides density values that correlate better with those of cores, it was observed that densities obtained from nuclear gauges correlated better with core densities (32, 29). In spite of changing the algorithm of the Pavement Quality Indicator Model 300 to provide improved density readings, it was found that the nuclear gauge readings always correlated better with cores densities than the corresponding density differences of the Pavement Quality Indicator (and the PaveTracker) with cores’ densities. Neither the improved Pavement Quality Indicator Model 300 nor the PaveTracker provided better results than the nuclear gauges when being compared with cores for differences purpose (39). It was suggested that the gauge accuracy and gauge and core density readings differences might improve if a slope function was calibrated to match gauge and core data (10). However, Hurley et al. found that even after applying correction for offset and slope, the electromagnetic gauges (Pavement Quality Indicator Model 301, Pavement Quality Indicator Model 300, and PaveTracker) do not correlate with cores as closely as a nuclear gauge can correlate with cores (7). Neither the Pavement Quality Indicator Model 300 nor the PaveTracker Model 1-B correlated well with cores or with nuclear gauge readings. Thus, they were recommended for Quality Control purpose but not Quality Assurance purpose (17). Liao et al. observed that the difference in density measurements between PaveTracker and cores is greater and more significant than corresponding difference between nuclear and core density measurements; thus, they indicated that nuclear gauges perform better than electromagnetic gauges. When they compared the differences of density values obtained from a nuclear gauge and a PaveTracker with values obtained from cores, they found that nuclear gauge correlates better with cores. A contractor would receive a pay factor of 1.02 based on core results, 1.00 based on nuclear gauge results, and 0.99 based on PaveTracker gauge results (5).

2.3.4.3 Comparison of factors influencing measurements from various nuclear gauges, electromagnetic gauges, and core tests

The various factors and their influence on the measurements obtained using the various gauges and cores can be described as follows:

Influence of mix temperature and moisture on gauges

Factors like mix temperature and moisture seem to have a direct affect on nuclear and electromagnetic gauge readings. Sebesta, S. et al. found that nuclear and electromagnetic gauges (Pavement Quality Indicator, PaveTracker, and Troxler 3450 in thin lift mode) are influenced by the mix temperature and by moisture. They found that gauge reading dropped with a decrease in mix temperature (10). They also found that the presence of moisture increases gauge readings such that the Pavement Quality Indicator is most affected by moisture and the nuclear gauge is least affected (10).

Influence of air gap on nuclear and electromagnetic gauge reading

Another factor that affects both nuclear and electromagnetic gauge readings is the air gap between the base of the gauge and the surface of the pavement. Air gap leads to significant variability or under representation in nuclear and electromagnetic gauge density readings. Thus, to avoid this error, gauges must be used only on a flat location and only when two lanes near joint are of same elevation (18).

Influence of nominal maximum aggregate size (NMAS) and binder type on density values

NMAS and binder type do not seem to affect gauge readings and core values. Smith and Diefender found that no differences existed to justify that NMAS and binder type had a significant effect in causing variation in density reading obtained from nuclear gauge, electromagnetic gauge, and cores (27).

Comparison of mean densities between nuclear and electromagnetic gauges

Rogge and Jackson found that five of six projects had mean core densities greater than mean nuclear gauge and Pavement Quality Indicator readings. They found that the Pavement Quality Indicator and nuclear gauge readings were generally within 3 % of each other. The mean field density measurements were generally within about 10% of mean core densities (16). Rao et al. found that the nuclear gauge measurements were higher than the electromagnetic gauge measurements. They also found that the PaveTracker readings were mostly greater than the Pavement Quality Indicator measurements (20). Even Schmitt et al. found that electromagnetic gauge density readings were consistently lower than nuclear gauge density readings. The Pavement Quality Indicator Model 301 read 11.2 to 27.2 pcf lower than the nuclear gauge; the

Pavement Quality Indicator Model 300 ranged from 4.2 to 26.6 pcf lower; the PaveTracker varied from 1.8 to 17.7 pcf lower (37).

Comparison of variability amongst gauges' readings

Sebesta et al. found that the laboratory studies on samples indicated that the repeatability standard deviation was below 0.5 pcf for the electromagnetic gauge and below 1 pcf for the nuclear gauge (10). The mean standard deviation values of the electromagnetic gauge data were lesser than those for the nuclear gauge measurements (20). When comparing the Pavement Quality Indicator and the nuclear gauge with cores in terms of a percentage difference (standard deviation, average difference, and absolute value), Smith and Diefender found that the nuclear gauge displayed lesser variability than the Pavement Quality Indicator on 7 of 8 projects. They also found that the histogram from the residuals of the regression analysis showed a normal distribution for the nuclear gauge (indicating that a linear regression is a good model) whereas the residual plot for the Pavement Quality Indicator indicated that it might still be influenced by other factors (27).

Selection of gauge based on lift thickness

According to Troxler Jr. and Dep, selection of a gauge is dependent on the top layer thickness. For a thickness equal to or less than 1 in., thin layer nuclear gauge or electromagnetic gauge is recommended. For 1 to 2 in. thick layer, thin-layer nuclear gauge is recommended. For 2 to 3 in. thick layer, any nuclear gauge in back scatter mode is recommended (18).

2.3.4.4 Comparison of core density and nuclear gauge density to improve core correlation

When comparing the density results (obtained from cores and nuclear gauges) for variability or differences, it was found that the density results obtained from cores show lesser variability than densities obtained from nuclear gauges (1, 2). Sanders et al. indicated that density values obtained from nuclear gauges are not as accurate as those obtained from cores (1). When they averaged 3524 measurements, they found that a majority of the density values measured from cores were higher (by an average value of 14.42 kg/m^3) than those measured from nuclear gauges (i.e., in 69.5% of the cases) (1). This can be due to the possible influence of the type of mix on the nuclear density gauge (1). Schmitt et al. found that the variability of densities determined from cores was less than the variability of densities determined from nuclear gauge measurements in 9 of 14 hot mix asphalt overlay projects (i.e., in 64% of cases) (2). Schmitt et all, in a study of 14 hot mix asphalt overlay projects in Wisconsin comparing densities determined from cores to those determined by nuclear gauge measurements, found that nearly 74% of the variation between the two measurements was explained by simple regression of core density measurement versus nuclear gauge density measurement. By including core thickness as a step function (thick or thin) and including theoretical maximum specific gravity as a continuous function, they were able to improve the regression model to explain nearly 75% of the variability (2).

Esch found that an accuracy level of ±1% could be achieved if 7 samples were extracted for core testing and 9 measurements were made using a nuclear gauge (21). Prowell and Dudley used the Troxler 4640-B thin lift nuclear gauge to take 2 readings in the parallel direction and 1 reading in the perpendicular direction (to paving) and they found that the nuclear gauge results correlated better with core densities obtained from AASHTO T 166 than with corresponding core densities obtained from the CoreLok device (29). For coarse mixes, core method is much more accurate than nuclear gauge (33).

Romero and Kuhnow suggested that using more cores could result in slightly more accurate correlation (22). Correlation value of 0.85 or greater signified good correlation and 0.60 or lesser signified bad correlation, and the in between values were used to account for confidence interval. They found that if the number of cores used for correlation was increased then the correlation improved. But still not more than 46% was in the good correlation range for electromagnetic gauges. When cores were increased from ‘5 or lesser’ to the range of ‘6 to 12’ cores, then there was 2 to 3% overall increase in the percentage of cores where there was ‘good correlation’ (22).

Comparison between densities from gauges and cores after accounting for the influence of moisture

Smith and Diefender conducted a paired t-test between gauge densities and core densities (for nuclear and electromagnetic gauges vs. cores); the result was a significant difference between density values of gauges and cores. However, they found that more sites have an equal gauge versus mean core density for the Pavement Quality Indicator than for the nuclear gauge. When a multiple regression equation was developed to obtain the average difference between the electromagnetic gauge density reading and the moisture index values with the core density measurements, it was found that the differences improved between the electromagnetic gauge and the cores. The R^2 range was 0.36 to 0.72 (i.e., if it is above 0.72 then there is good correlation). Three of eight projects showed a good differences (earlier it was two) and only three of eight projects showed a bad differences (earlier it was four). Also, when the corrected Pavement Quality Indicator readings and the nuclear gauge readings were compared with the core readings, it was found that the corrected Pavement Quality Indicator readings was better than the nuclear gauge readings for 7 of 8 projects (i.e., it better correlated with core readings). Thus after being corrected for moisture index, the Pavement Quality Indicator was found to be more suitable than the nuclear gauge (4640-B) for measuring density on dense-graded hot mix asphalt (27). When the algorithm of the Pavement Quality Indicator Model 300 was changed, it was referred to as the Pavement Quality Indicator+. The correlation between the density readings obtained using the Pavement Quality Indicator+ and the cores indicated a 14% improvement over the cores-Pavement Quality Indicator differences (27). It was also found that when pooled data was used to compare the Pavement Quality Indicator and the nuclear gauge density readings with those of corresponding cores, a poor correlation existed. So, a unique regression was developed for each project and it was observed that there was a higher R^2 value for this than for a pooled site differences. Thus, it was realized that when projects were independently treated, mix specific calibration could improve the correlation between gauge and core density measurements.

2.4 Summary

To fulfill the first objective described in section 1.2, the efficacy of various methods of density measurements were investigated and summarized. This included a description of the principle behind each method of measurement, factors affecting the methods of measurement, factors not affecting the methods of measurement, correlation between the various methods of measurements, and suggestions and recommendations to improve certain methods of measurements or to improve the correlation between methods so that a more accurate density closer to the true density can be obtained in the future. The entire findings from the literature review studies can be summarized as follows:

The densities obtained from cores were most accurate. However, the entire coring method is most time consuming and can weaken the pavement. Cores exhibit a density gradient with lesser density along the perimeter. Many agencies including the Oregon Department of Transportation use the AASHTO T 166 to calculate the density. However, density, when obtained using the saturated surface-dry method (described in AASHTO T 166), tends to be overestimated at high air void content (greater than 8%) or when a greater part of the surface has interconnected air voids. This is more profound in coarse mixes. For fine and coarse mixes (with water absorption greater than 0.4%), the density results from the Automatic Vacuum Sealing method are more accurate than those from the saturated surface-dry method. At high water absorption levels (greater than 0.4%), low densities, low levels of compaction, and lower level of gyration, the AASHTO T 166 overestimates the density. This answers the fifth research question indicated in section 3.1 and provides evidence to suggest that the SSD method must not be used at water absorption levels of 2% and instead, the CoreLok method should be used as recommended.

Nuclear gauges are fast and non-destructive. They correlate better with cores when compared to electromagnetic gauges. Nuclear gauges are affected by pavement thickness; a decrease in thickness causes readings to become more variable. The base coarse material has a negligible effect and the sub base material has a significant effect on the gauge readings. Gauge reading variability is lesser for a thicker binder mix layer than a thinner binder mix layer. Mix composition, mix aggregate type and size, and surface texture or roughness influence gauge readings. Rough surface texture causes the gauge to show an underestimated density. A temperature difference of 100^0F causes a density difference of 1pcf. Literature studies also indicate that using latex as a modifier increases variability, while milling and surface treatment decreases variability of gauge readings and improves core differences. Air gap can lead to an underestimated density reading. Factors like gauge direction, rubber pads, sand, coal tar, standard count, and date of calibration do not cause any significant variability in gauge readings. Factors like type of gauge (or gauge model), project conditions, and location significantly affect gauge readings. It is recommended to take one minute readings. At least 2 readings are required per location.

Electromagnetic gauges are lighter in weight, measure density fastest and are most cost-effective. Earlier models like the PQI 300 were affected by surface mix temperature; the gauge reading is directly proportional to the surface mix temperature (density decreases by 16kg/m^3 with a temperature drop of 50^0C). Later models like the PQI 301 and the PT 2701-B were unaffected by surface temperature. However, all electromagnetic gauges were affected by the surface and internal moisture. More moisture makes the readings more variable; coarse mixes were more affected by moisture than fine mixes. There is better correlation between gauge densities and core densities when the moisture index is below 10. Even the mix type or mix design has a significant effect on gauge reading such that there is better correlation between gauge and core densities for a fine mix than for a coarse mix. Air voids and sprinkling of sand affect the gauge and cause it to display an underestimated density. Battery voltage and mineral filler do not affect gauge accuracy or variability. It is best to calibrate the gauge daily, and then take four to five readings at each location for a more accurate estimate of density. Majority of the studies indicate that electromagnetic gauges can be used for Quality Control and one study also indicates that the PQI 301 can be used for Quality Assurance.

A comparison between the densities from the tests on cores, the nuclear gauges, and the electromagnetic gauges, indicates that the densities are significantly different from each other at 95% confidence level. The correlation between the nuclear gauge and core densities is better than the correlation between the electromagnetic gauge and core densities.

3.0 EXPERIMENT DESIGN

A major part of this research effort was to conduct field and laboratory studies. Prior to conducting the studies, several research questions were developed to address the specific needs of ODOT. This chapter presents the research questions, a summary of the experiment design, and a description of the projects investigated.

3.1 Research questions

Several questions were formulated to establish a framework for conducting the field and laboratory studies. These were developed at the request of ODOT personnel to address specific issues concerning nuclear gauge measurements on pavements with multiple lifts, on pavements with different mix designs, on new construction versus mill and inlay projects, and whether or not differences between core densities and nuclear gauge densities are necessary to obtain accurate results using nuclear gauges. The questions are listed below while section 3.2 summarizes the experiment designs for the field and laboratory studies.

1. Is it necessary to determine nuclear gauge calibration factors for each lift of hot mix asphalt (HMA) placed in multiple lifts with the same mix design for each lift?

2. For projects incorporating different mix designs for individual lifts (e.g., use of a rich binder base), is there a significant difference between nuclear gauge calibration factors for the different mixtures?

3. On mill & fill (inlay) plus overlay projects, is there a significant difference between the nuclear gauge calibration factors for the fill (inlay) material and the overlay material?

4. Is there a significant difference between nuclear gauge calibration factors derived from cores obtained from the corner of the overlapping portion of the nuclear gauge measurement footprint (Figure 3.1) and cores obtained from the center of the overlapping portion of the footprint (Figure 3.2)?

5. ODOT currently allows use of the saturated surface-dry (SSD) method for determining the bulk specific gravity of a core even if the water absorption of the core is greater than 2%. Is it appropriate to continue to allow this?

6. Is there a significant difference between densities of HMA blocks cut from pavements using a masonry blade and cores cut using conventional diamond-tipped barrels? (not included as a part of thesis)

7. Can a dedicated location marked on a particular pavement lift be used to spot-check nuclear gauge readings to determine if recalibration of a particular nuclear gauge is necessary? (not included as a part of thesis)

8. A nuclear gauge calibrated using calibration blocks from one Region Office may not have the same calibration factors when calibrated using the calibration blocks in a different Region Office. Is the current calibration procedure valid? (not included as a part of thesis)

3.2 Experiment design

1. Is it necessary to determine nuclear gauge calibration factors for each lift of hot mix asphalt (HMA) placed in multiple lifts with the same mix design for each lift?

   To gather sufficient information to convincingly answer this question, a project was investigated where the HMA was placed in multiple lifts. This was done to determine the influence of as-built HMA lift thickness and lift density on nuclear gauge measurements of overlying lifts. Also, the project was selected such that the HMA was placed in three lifts to provide information of HMA density of a successively thicker substrate (i.e., density of first lift of HMA over aggregate base versus density of second lift of HMA over first lift of HMA over aggregate base, and so on). To accomplish this, 30 test locations were predetermined along a stretch of the project that was bounded by the first day of paving of the first lift. The test locations were randomly selected according to the procedure described in AASHTO T 52-05 and ASTM D 3665-07 (52). Following placement of the first lift of HMA, nuclear gauge measurements were made and cores were taken at 10 of the 30 predetermined locations. Similarly, nuclear gauge measurements were made and cores were taken at 10 different locations of the 30 predetermined locations for the second and third lifts of HMA paving. Details on the number of cores, number of nuclear and electromagnetic gauges' readings, and various methods of testing are provided in the table 3.1. As indicated in the table 3.1, 10 cores were extracted from the bottom lift after it was paved. Then, 10 cores were extracted after the middle lift was paved; these cores were composite (i.e., consisting of bottom lift and middle lift). So they were tested in the composite form and then the cores were cut and separated into different lifts and re-tested. Finally, 10 cores were extracted after the top lift was paved; these cores were also composite (i.e., consisting of the bottom lift, middle lift, and top lift). So these cores were tested in the composite form before being cut, categorized, and re-tested as cores from individual lifts. Also, it is important to mention that the top lift (or third lift) is the top lift of the base course (which was later overlaid with a wearing course). This work was carried out on the Fort Hill – Wallace Creek Bridge project on OR 18.

Table 3.1 – Experiment design for investigating HMA density on new pavement construction projects placed in multiple lifts having the same mix design

Pavement Lift	Number of Sample Locations	Number of Nuclear Gauges	Number of Nuclear Gauge Measurements per Location (for each gauge)	Number of Electromagnetic Gauges	Number of Electromagnetic Gauge Measurements per Location (for each gauge)	Number of Cores	Number of BSG Measurements via Water Immersion Method	Number of BSG Measurements via Vacuum Sealing Method
First Lift (Bottom Lift)	30	3	4	2	5	30	30	30
Second Lift (Middle Lift)	20	3	4	2	5	20	20	20
Third Lift (Top Lift)	10	3	4	2	5	10	10	10
First and Second Lift	10	3	4	2	5	10	10	10
First, Second, and Third Lift	10	3	4	2	5	10	10	10

2. For projects incorporating different mix designs for individual lifts (e.g., use of a rich binder base), is there a significant difference between nuclear gauge calibration factors for the different mixtures?

   This is similar to the above question, except it includes investigation of two lifts with different mix designs. To gather sufficient information to convincingly answer this question, a stretch of a project was selected wherein 20 test locations were predetermined. The test locations were randomly selected according to the procedure described in AASHTO T 52-05 (ref) and ASTM D 3665-07 (52). Following placement of the first lift of HMA, nuclear gauge measurements were made and cores will be taken at 10 of the 20 predetermined locations. In a similar fashion, nuclear gauge measurements were made and cores were extracted at 10 different locations of the 20 predetermined locations for the second lift of HMA paving, with the second lift having a different mix design from the first lift. Details on the number of cores, number of nuclear gauges' readings, and various methods of testing are provided in the table 3.2.

   As seen in the table 3.2, ten cores were extracted from the first or bottom lift when it was paved. Then, ten cores were extracted after the second or top lift is paved; these cores are composite; that is, consisting of first (bottom) lift and second (top) lift. So they were tested in the composite form and then the cores were cut and separated into first (bottom) and second (top) lifts and re-tested. This work was carried out on the South Medford Interchange project on Interstate 5.

Table 3.2 – Experiment design for investigating HMA density on new pavement construction projects placed in two lifts having different mix designs

Pavement Lift	Number of Sample Locations	Number of Nuclear Gauges	Number of Nuclear Gauge Measurements per Location (for each gauge)	Number of Cores	Number of BSG Measurements via Water Immersion Method	Number of BSG Measurements via Vacuum Sealing Method
First Lift (Bottom Lift)	10	3	4	10	10	10
Second Lift (Top Lift)	10	4	4	10	10	10

3. On mill & fill (inlay) plus overlay projects, is there a significant difference between the nuclear gauge calibration factors for the fill (inlay) material and the overlay material?

   To gather sufficient information to convincingly answer this question, 20 test locations were predetermined along a stretch of a mill & fill project that was bounded by the first day of paving of the inlay lift. The test locations were randomly selected according to the procedure described in AASHTO T 52-05 and ASTM D 3665-07 (52). Following placement of the inlay lift of HMA, nuclear gauge measurements were made and cores were taken at 10 of the 20 predetermined locations. Similarly, nuclear gauge measurements were made and cores were taken at 10 different locations of the 20 predetermined locations for the overlay lift. Details on the number of cores, number of nuclear and electromagnetic gauges' readings, and various methods of testing are provided in the table 3.3. In the table 3.3, ten cores were extracted from the fill (inlay) lift after it is paved. Then, ten cores were extracted after the fill (overlay) lift was paved; these cores were composite (i.e., consisting of inlay lift and overlay lift). So they were tested in the composite form and then the cores were cut and separated into inlay and overlay lifts and re-tested. This work was carried out on the North Fork Little Butte Creek – Great Meadow Snow Park project on OR 140.

Table 3.3 – Experiment design for investigating HMA density on mill and inlay projects placed in the lifts having the same mix design

Pavement Lift	Number of Sample Locations	Number of Nuclear Gauges	Number of Nuclear Gauge Measurements per Location (for each gauge)	Number of Electromagnetic Gauges	Number of Electromagnetic Gauge Measurements per Location (for each gauge)	Number of Cores	Number of BSG Measurements via Water Immersion Method	Number of BSG Measurements via Vacuum Sealing Method
First Lift (Inlay Lift)	20	3	4	2	5	20	20	20
Second Lift (Overlay Lift)	10	3	4	2	5	10	10	10
First and Second Lift	10	3	4	2	5	10	10	10

4. Is there a significant difference between nuclear gauge calibration factors derived from cores obtained from the corner of the overlapping portion of the nuclear gauge measurement footprint (Figure 3.1) and cores obtained from the center of the overlapping portion of the footprint (Figure 3.2)?

    10 locations were predetermined along a project by random selection according to the procedure described in AASHTO PP 52-05 (ref) and ASTM D 3665-07 (52). Nuclear gauge measurements were made in accordance with WAQTC TM 8 at all 10 locations. Five of the locations were then cored as indicated in figure 3.1, and the other five locations cored as indicated in Figure 3.2.

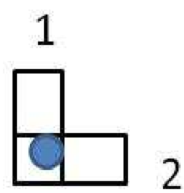


Figure 3.1 – Core derived from the corner of the overlapping portion of the 'L-shaped' nuclear gauge measurement footprint.

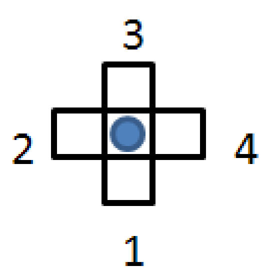


Figure 3.2 – Core derived from the center of the overlapping portion of the 'cross-shaped' nuclear gauge measurement footprint.

5. ODOT currently allows use of the saturated surface-dry (SSD) method for determining the bulk specific gravity of a core even if the water absorption of the core is greater than 2%. Is it appropriate to continue to allow this?

To best answer this research question, additional review of the literature with emphasis on studies that have investigated this was performed. Hence, no additional coring was performed for this part of the study.

3.3 Projects evaluated

To fulfill the objectives of the research project, three paving projects were selected. The following sections briefly describe each of the projects.

3.3.1 Fort Hill - Wallace Bridge section on Salmon River Highway (OR 18)

The mile point co-ordinates from start point to end point of this project were M.P. 23.82 (Station ‘A,’ 39+50) to M.P. 26.28 (Station ‘B,’ 170+00). The base course for this new work (new construction) project was constructed in three lifts, all with the same mix design. The base course was capped with a wearing course, but all work described herein was conducted on the base course lifts. The project was selected to answer the first and fourth research questions. A total of six nuclear gauges and two electromagnetic gauges were used in this project. The pavement thickness varied nominally from 2 to 3 inches throughout the project.

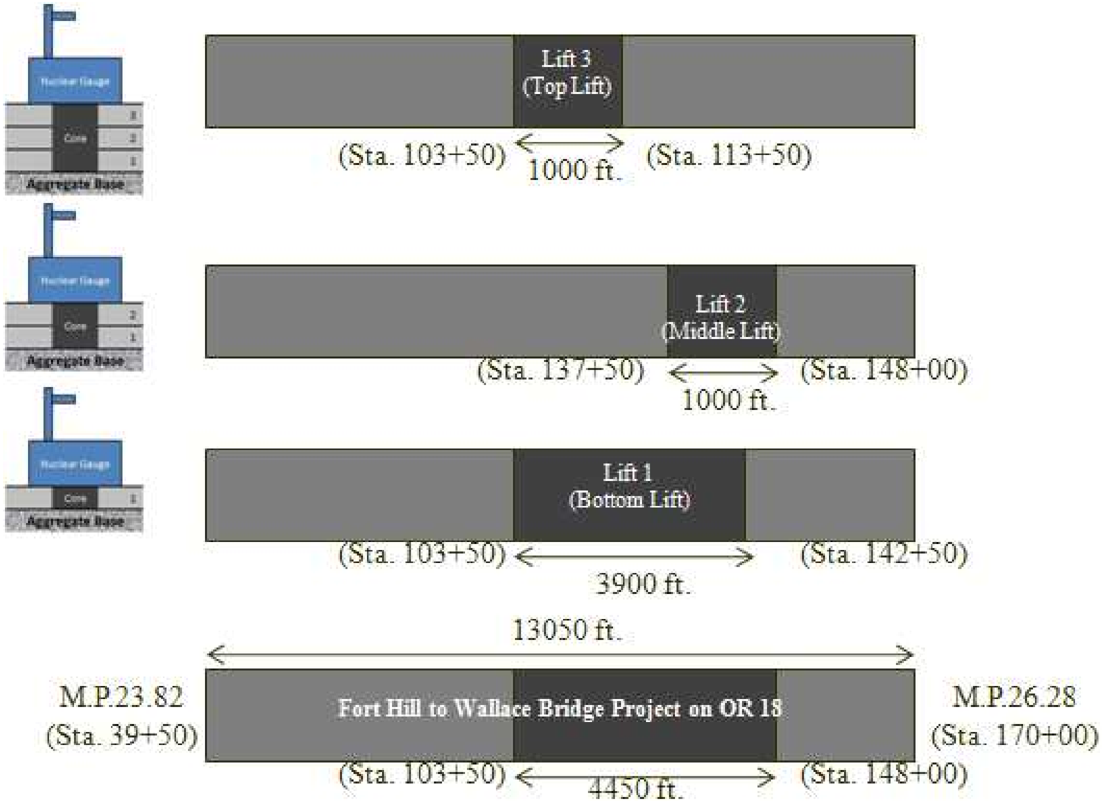


Figure 3.3 – Diagrammatic representation of sampling limits for each lift on project on OR 18

3.3.2 South Medford Project Interchange on Interstate 5 (I-5)

This was a new pavement construction (new work) project wherein the first lift of the base course had a different mix design than the remaining lifts. The field work was done on the first two lifts of the base course. It was selected to answer the second research question. Four nuclear gauges (and no electromagnetic gauges) were used on this project. The pavement thickness was a nominal 2 inches throughout the project.

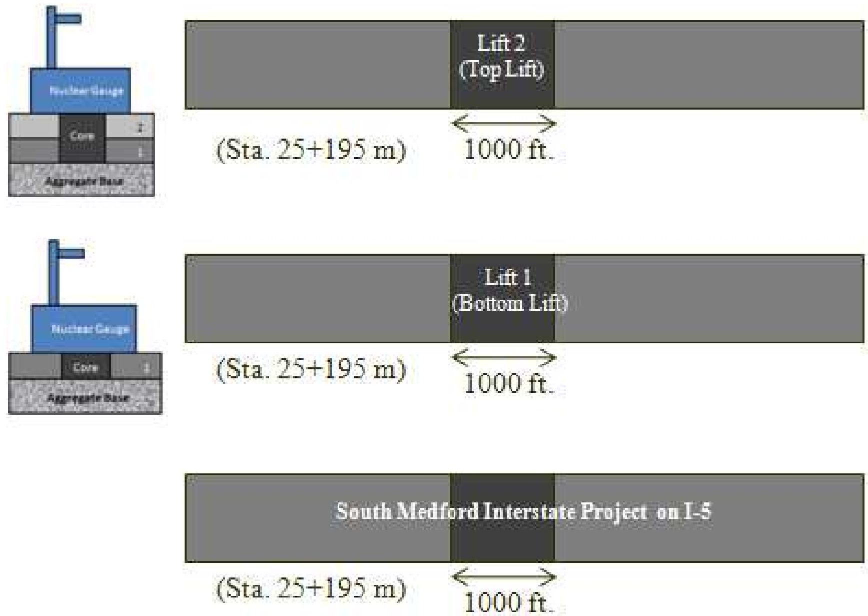


Figure 3.4 – Diagrammatic representation of sampling limits for each lift on project on I-5

3.3.3 North Fork Little Butte Creek to Great Meadow Snow Park on Oregon Highway 140 (OR 140)

The mile point co-ordinates from start point to end point of this project are M.P. 16.04 (Station 'A,' 11+50) to M.P. 30.92 (Station 'B,' 800+00). This project was a mill and inlay with overlay having the same mix design for both lifts. It was selected to answer research question 3. Three nuclear gauges and two electromagnetic gauges were used on this project. The pavement thickness was nominally 2 inches throughout the project.

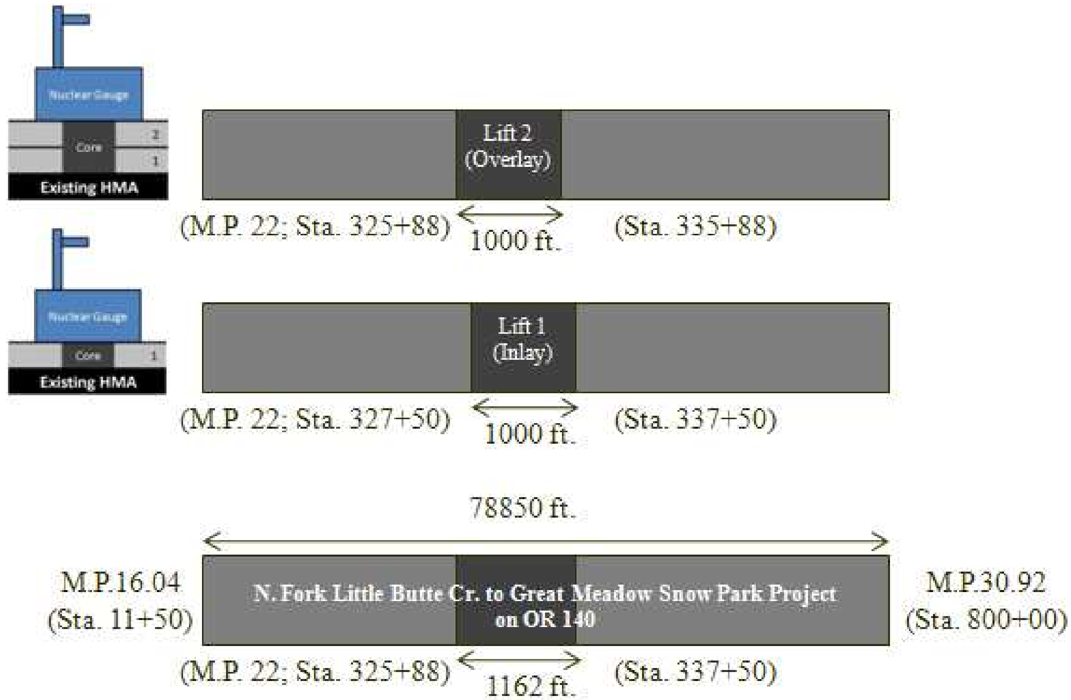


Figure 3.5 – Diagrammatic representation of sampling limits for each lift on project on OR 140

4.0 FIELD STUDY

The field study was undertaken to obtain density measurements using nuclear and electromagnetic gauges at predetermined locations along each of the projects investigated. Following the density measurements, cores were obtained from the locations where gauge measurements were made and subsequently tested as detailed in chapter five of this document. The following sections provide details and results of the field study.

4.1 Procedure for Performing Field Studies

On the OR 18 project, three lifts of HMA with the same mix design were placed for the base course and data were mostly collected along the same stretch on three days (on each day a different lift was paved). On each of the three days of paving, ten locations were selected randomly in accordance with ASTM D 3665-07 (52) along the fixed stretch of the pavement being constructed. At each of these locations, first sand was spread out on that spot or location, and then it was leveled with a wooden bar. This was done to minimize the effect of surface irregularities of the pavement on the gauge density readings. Then, the nuclear gauges were used to obtain the density values at the location. Next, the electromagnetic gauge was used to obtain the density values at the same spot or location. Finally, a core was drilled out from that location with a conventional diamond tipped core barrel. The core location is indicated in figure 4.1. This was repeated for all the locations on each of the three days of paving. The same procedure used on the OR 18 project was also used on the OR 140 project which was a mill and inlay with overlay project (i.e., two lifts total) and on the I-5 project which had two lifts of different mix designs. To answer the fourth research question (i.e., whether there is a significant difference between the core average differences for cores drilled from the corner of the 'L-shaped' footprint and cores drilled from center of cross-shaped footprint as indicated in the figures 3.1 and 3.2 in the previous chapter), the wearing course on the OR 18 project was selected for field testing. Ten locations were selected randomly in accordance with ASTM D 3665-07 (52) along the fixed stretch of the pavement being constructed. From these, five alternate locations were tested by taking measurements and extracting cores as indicated from the 'L-shaped' footprint in figure 3.1, and the remaining five alternate locations were tested by taking measurements and extracting cores as indicated from the 'cross-shaped' footprint in figure 3.2.

Loose HMA was also obtained from the pavements immediately following the paving operation (i.e., immediately before compaction) to measure the theoretical maximum specific (Rice) gravity for that particular lift. The loose HMA was obtained from each lift by randomly sampling the mat (using a cookie cutter) behind the paver at different times during the paving operation to account for potential variability in the theoretical maximum specific (Rice) gravity. This was done in accordance with the method described in AASHTO T 168 (51). Since this data was obtained after testing the loose HMA in the laboratory in accordance with AASHTO T 209 (45), it is summarized in the next chapter i.e., Laboratory Studies.

4.2 Method of Taking Readings using Nuclear and Electromagnetic Gauges

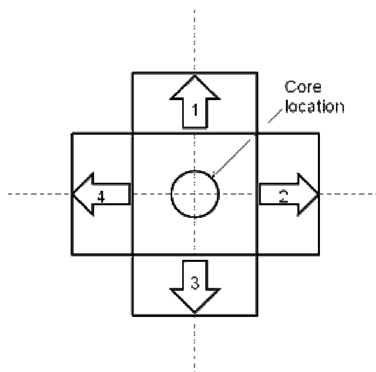


Figure 4.1 – Primary Nuclear Gauge Measurement Orientations and Core Location

The process by which nuclear and electromagnetic readings were taken on the project sites (OR 18 and OR 140) is described here as follows:

To gather the nuclear gauge measurement data for questions 1, 2, and 3, the nuclear gauge readings were made in the sequence shown in figure 4.1. The nuclear gauge measurements on the various lifts of the hot mix asphalt pavements were made in accordance with the WAQTC TM 8, except that the gauge was rotated about its longitudinal centerline as indicated by orientations 1 and 2 in the figure 4.1, plus an additional two measurements were made at each location as indicated by orientations 3 and 4 in figure 4.1. This sequence allowed determination of the average density at a given location to be based on two versus four orientations.

The electromagnetic gauges were used to measure density such that there were five readings per sample location – one in the center (i.e., directly over the core location) and four surrounding it such that they were positioned at 90° from each other.

4.3 Test Results

Tables 4.1 to 4.9 contain summary data of nuclear gauge measurements and, where appropriate, electromagnetic gauge measurements from each paving lift of each project evaluated during the field study. Each table provides information regarding the gauge operator, gauge identification, date of last calibration, and average density obtained using the average of 4 readings and 2 readings at each location. Further details (e.g., type of mix, date of testing, lift thickness, distance and offset of each test location, etc.) are provided in Appendix B.

Tables 4.1 to 4.3 summarize the gauge data collected to investigate density measurements on new construction projects where the base course of the pavement was placed in multiple lifts with each lift having the same mix design. Table 4.1 includes the data for the first (bottom) lift over the aggregate base, table 4.2 includes the data for the second (middle) lift, and table 4.3 includes the data for the third (top) lift of the base course. For each nuclear gauge used, note the similarity of average density based on two readings versus four readings. Note also the exceptionally low coefficient of variation (less than about 2% in most cases) indicating very low variance about the average density of the 10 sample locations. Also, note that the coefficient of variance of the nuclear gauges was higher than that of the electromagnetic gauges in most of the instances. In addition, there was at most a 3 lb/ft^3 difference in average density between nuclear gauge measurements. However, it must be mentioned that this was only observed on the top lift; while, the bottom and middle lifts had differences that were at most 1.6 lb/ft^3 and 1.7

lb/ft^3 respectively). Finally, it can be seen from this data that the average densities obtained from the electromagnetic gauges were consistently lower than those obtained from the nuclear gauges.

Table 4.1 - Summarized data obtained from lift 1 of OR 18

Gauge Make, Model, and Serial Number	Nuclear Gauges					
	Troxler 3430 # 38804		Troxler 3430 # 38806		Troxler 3440 # 25714	
Last date of calibration	N/A		4/28/2008		3/7/2008	
Gauge Operator	Ray		Dean Chess		Paul Sloan	
Sample Location	Avg Density, lb/ft^3		Avg Density, lb/ft^3		Avg Density, lb/ft^3	
	Four readings	Two readings	Four readings	Two readings	Four readings	Two readings
1	155.5	155.6	153.7	153.7	155.0	154.5
2	149.2	148.7	149.8	149.5	147.5	147.5
3	147.4	148.8	150.5	150.2	148.0	148.1
4	151.0	150.9	152.7	152.6	150.9	151.1
5	152.6	152.6	152.7	153.4	150.4	150.7
6	156.6	157.1	156.1	156.0	154.3	154.3
7	158.1	157.9	157.7	158.0	156.7	157.0
8	151.7	151.9	152.7	153.1	150.7	151.2
9	156.4	156.5	156.8	156.8	155.2	154.7
10	154.4	154.5	155.7	155.2	153.9	153.5
Average, lb/ft^3	153.3	153.4	153.8	153.8	152.3	152.2
Standard Deviation, lb/ft^3	3.5	3.4	2.7	2.7	3.2	3.1
Coefficient of Variation, percent	2.3%	2.2%	1.7%	1.8%	2.1%	2.0%

Table 4.2 – Summarized data obtained from lift 2 of OR 18

Gauge Make, Model, and Serial Number	Nuclear Gauges						Electromagnetic Gauges			
	Troxler 3440 # 30860		Troxler 3430 # 38806		Troxler 3440 # 25714		PQI 301		PQI 300	
Last date of calibration	2/5/2008		4/28/2008		3/7/2008		N/A		N/A	
Gauge Operator	Josh Huber		Dean Chess		Paul Sloan		Larry Ilg		Larry Ilg	
Sample Location	Avg Density, lb/ft^3		Avg Density, lb/ft^3		Avg Density, lb/ft^3		Average Density, lb/ft^3	Average Moisture Content, percent	Average Density, lb/ft^3	Average Moisture Content, percent
	Four readings	Two readings	Four readings	Two readings	Four readings	Two readings				
1	152.3	152.4	152.0	152.0	151.1	150.9	149.54	9.82	149.44	7.32
2	151.9	151.3	151.1	151.2	149.8	150.1	150.04	10.2	149.48	7.52
3	149.9	150.5	148.8	148.2	147.1	146.9	147.74	9.44	148.06	7
4	152.1	151.6	150.8	150.8	149.9	150.0	147.02	8.72	146.24	6.32
5	154.4	154.1	153.2	153.1	152.2	152.3	148.48	9.22	147.68	6.62
6	151.9	151.8	151.9	151.6	150.7	150.9	148.6	9.52	148.42	7.1
7	152.1	152.2	152.1	152.3	151.1	151.3	148.44	9.36	147.76	7.12
8	152.2	151.5	150.8	150.1	151.0	151.7	148.46	9.28	148.2	6.88
9	153.0	152.8	152.2	151.8	151.1	151.4	148.02	10.12	149.58	7.44
10	152.7	152.6	152.4	152.8	152.0	152.2	149.5	9.72	148.34	7.14
Average	152.2	152.0	151.5	151.4	150.6	150.7	148.6	9.5	148.3	7.0
Standard Deviation	1.107	0.988	1.227	1.418	1.439	1.573	0.907	0.443	1.018	0.367
Coefficient of Variation	0.73%	0.65%	0.81%	0.94%	0.96%	1.04%	0.61%	4.65%	0.69%	5.20%

Table 4.3 – Summarized data obtained from lift 3 of OR 18

Gauge Make, Model, and Serial Number	Nuclear Gauges						Electromagnetic Gauges							
	Troxler 3440 # 30860		Troxler 3430 # 38806		Troxler 3440 Plus #		PQI 301 (sand)		PQI 301 (no sand)		PQI 300 (sand)		PQI 300 (no sand)	
Last date of calibration	2/5/2008		4/28/2008		1/24/2008		N/A		N/A		N/A		N/A	
Gauge Operator	Josh Huber		Dean Chess		Terry Vann		Larry Ilg		Larry Ilg		Larry Ilg		Larry Ilg	
Sample Location	Avg Density, lb/ft^3		Avg Density, lb/ft^3		Avg Density, lb/ft^3		Average Density, lb/ft^3	Average Moisture Content, percent	Average Density, lb/ft^3	Average Moisture Content, percent	Average Density, lb/ft^3	Average Moisture Content, percent	Average Density, lb/ft^3	Average Moisture Content, percent
	Four readings	Two readings	Four readings	Two readings	Four readings	Two readings								
1	151.5	151.5	151.6	152.0	155.0	155.0	148.1	8.1	N/A	N/A	148.4	6.0	N/A	N/A
2	154.3	153.5	154.9	155.3	155.4	155.1	149.9	8.8	N/A	N/A	150.5	6.6	N/A	N/A
3	152.5	153.3	151.9	152.2	154.7	154.9	147.9	8.2	148.1	8.6	148.1	6.1	147.0	6.2
4	150.6	150.3	149.8	150.1	152.9	153.1	147.4	8.1	147.7	8.3	147.8	5.9	147.2	6.1
5	150.0	149.8	148.6	147.6	152.8	152.5	147.7	8.6	146.2	8.3	147.8	6.3	146.0	6.4
6	151.2	151.6	151.0	151.1	153.1	152.9	149.3	9.2	149.1	9.3	149.7	6.7	147.4	6.7
7	148.9	148.6	147.8	149.3	152.3	152.3	148.2	8.5	148.0	8.9	148.2	6.2	147.3	6.5
8	150.3	150.1	150.2	150.1	152.8	152.7	148.3	8.7	148.4	9.0	148.3	6.4	147.2	6.5
9	145.1	144.9	144.5	145.1	148.4	148.5	145.2	7.8	144.6	7.9	145.7	5.8	143.9	5.8
10	149.4	149.6	148.9	149.3	152.2	152.2	148.3	8.9	147.9	9.1	148.7	6.5	146.9	6.7
Average	150.4	150.3	149.9	150.2	152.9	152.9	148.0	8.5	147.5	8.7	148.3	6.2	146.6	6.4
Standard Deviation	2.4	2.5	2.8	2.8	2.0	1.9	1.2	0.4	1.4	0.5	1.3	0.3	1.2	0.3
Coefficient of Variation	1.61%	1.65%	1.84%	1.84%	1.29%	1.26%	0.83%	5.16%	0.98%	5.63%	0.85%	4.68%	0.81%	4.96%

Tables 4.4 and 4.5 summarize the gauge data collected to investigate density measurements on new construction projects where the base course of the pavement was placed in multiple lifts with each lift having a different mix design. Table 4.4 includes the data for the first (bottom) lift over the aggregate base and table 4.5 includes the data for the second (top) lift of the base course. For each nuclear gauge used, note the similarity of average density based on two readings versus four readings. Also, note the exceptionally low coefficient of variation (less than about 2% in most cases) indicating very low variance about the average density of the 10 sample locations. In addition, there was at most a 1.8 lb/ft^3 difference in average density between nuclear gauge measurements.

Table 4.4 – Summarized data obtained from lift 1 of I-5 Project

Gauge Make, Model, and Serial Number	Nuclear Gauges					
	Troxler 3430 # 38996		Troxler 3430 # 38806		CPN MC3 Portaprobe # 7443	
Last date of calibration	2/26/2009		2/12/2009		7/23/2008	
Gauge Operator	John Groth		Dean Chess		Mick King	
Sample Location	Avg Density, lb/ft^3		Avg Density, lb/ft^3		Avg Density, lb/ft^3	
	Four readings	Two readings	Four readings	Two readings	Four readings	Two readings
1	146.7	146.3	146.4	145.8	146.1	145.8
2	141.6	141.4	141.7	141.8	142.7	143.2
3	144.8	145.0	144.0	144.5	145.4	145.5
4	143.9	143.9	142.5	142.8	144.9	145.0
5	145.9	146.0	144.9	145.1	145.4	145.4
6	144.7	144.9	143.7	143.3	144.4	144.2
7	146.6	146.5	145.3	144.9	146.9	147.0
8	146.3	146.0	144.8	145.1	146.5	146.7
9	146.3	146.5	145.1	145.5	146.6	147.0
10	144.1	143.7	143.8	143.8	144.9	144.8
Average	145.1	145.0	144.2	144.2	145.4	145.4
Standard Deviation	1.6	1.6	1.4	1.3	1.3	1.2
Coefficient of Variation	1.11%	1.13%	0.97%	0.90%	0.87%	0.85%

Table 4.5– Summarized data obtained from lift 2 of I-5 Project

Gauge Make, Model, and Serial Number	Nuclear Gauges							
	Troxler 3430 # 38996		Troxler 3430 # 38806		CPN MC3 Portaprobe # 7443		Humboldt 5001C # 1344	
Last date of calibration	2/26/2009		2/12/2009		7/23/2008		4/8/2009	
Gauge Operator	John Groth		Dean Chess		Mick King		Rod Savage	
Sample Location	Avg Density, lb/ft^3		Avg Density, lb/ft^3		Avg Density, lb/ft^3		Avg Density, lb/ft^3	
	Four readings	Two readings	Four readings	Two readings	Four readings	Two readings	Four readings	Two readings
1	144.5	144.4	142.6	142.8	144.1	143.8	N/A	N/A
2	146.0	146.0	144.3	144.2	145.5	144.1	145.7	145.9
3	145.8	146.1	143.8	143.7	145.1	145.1	144.7	145.5
4	145.0	144.8	143.2	143.6	144.1	143.7	144.7	144.4
5	142.9	142.7	140.8	140.8	143.0	143.2	142.3	141.8
6	144.0	143.6	141.8	142.1	143.7	143.8	144.3	145.0
7	140.8	141.1	139.7	140.0	141.1	140.4	139.7	139.0
8	143.1	143.2	141.1	141.0	143.6	143.4	142.7	142.3
9	142.8	143.2	140.8	140.7	142.4	141.9	142.1	141.9
10	144.2	144.1	142.8	143.1	143.3	143.8	144.2	144.2
Average	143.9	143.9	142.1	142.2	143.6	143.3	143.4	143.3
Standard Deviation	1.6	1.5	1.5	1.5	1.3	1.3	1.8	2.2
Coefficient of Variation	1.09%	1.06%	1.04%	1.05%	0.89%	0.91%	1.28%	1.57%

Tables 4.6 and 4.7 summarize the gauge data collected to investigate density measurements on mill and fill plus overlay projects with both lifts having the same mix design. Table 4.6 includes the data for the first (bottom or inlay) lift over the aggregate base and table 4.7 includes the data for the second (top or overlay) lift of the pavement. For each nuclear gauge used, once again note the similarity of average density based on two readings versus four readings. Note also the exceptionally low coefficient of variation (less than about 2% in most cases) indicating very low variance about the average density of the 10 sample locations. Also, note that the coefficient of variance of the nuclear gauges was always higher than that of the electromagnetic gauges. In addition, there was at most a 1.9 lb/ft^3 difference in average density between nuclear gauge measurements. Finally, it can be seen from this data that the average densities obtained from the electromagnetic gauges were consistently lower than those obtained from the nuclear gauges.

Table 4.6 – Summarized data obtained from lift 1 of OR 140

Gauge Make, Model, and Serial Number	Nuclear Gauges						Electromagnetic gauges	
	Troxler 3430 # 38996		Troxler 3430 # 38806		Troxler 3430 # 35601		PQI 301	PQI 300
Last date of calibration	7/31/2008		4/28/2008		8/18/2008		N/A	N/A
Gauge Operator	John Groth		Dean Chess		Dean Chess		Larry Ilg	Larry Ilg
Sample Location	Avg Density, lb/ft^3		Avg Density, lb/ft^3		Avg Density, lb/ft^3		Average Density, lb/ft^3	Average Density, lb/ft^3
	Four readings	Two readings	Four readings	Two readings	Four readings	Two readings		
1	138.9	139.3	141.0	141.1	140.3	140.5	140.5	138.5
2	141.0	141.2	143.7	143.9	142.7	142.7	141.1	138.6
3	137.4	137.6	139.2	139.2	139.0	138.7	140.3	138.2
4	138.1	138.0	140.7	140.0	140.0	139.9	140.3	139.0
5	138.4	138.8	139.9	140.3	139.4	139.4	140.5	136.4
6	137.4	137.2	139.3	138.7	138.8	139.5	140.1	138.9
7	137.8	137.8	139.1	137.8	139.7	139.5	140.1	138.6
8	134.2	134.0	136.8	136.3	136.0	136.2	139.7	137.3
9	136.5	136.3	138.7	138.5	137.9	138.0	140.1	138.2
10	133.4	133.1	134.1	134.5	134.6	134.5	138.8	134.8
Average	137.3	137.3	139.2	139.0	138.8	138.9	140.1	137.9
Standard Deviation	2.2	2.4	2.5	2.6	2.3	2.3	0.6	1.3
Coefficient of Variation	1.61%	1.75%	1.82%	1.87%	1.64%	1.64%	0.42%	0.96%

Table 4.7 – Summarized data obtained from lift 2 of OR 140

Gauge Make, Model, and Serial Number	Nuclear gauges						Electromagnetic gauges	
	Troxler 3430 # 38996		Troxler 3430 # 38806		Troxler 3430 # 35601		PQI 301	PQI 300
Last date of calibration	7/31/2008		4/28/2008		8/18/2008		N/A	N/A
Gauge Operator	Ray Cunningham		Dean Chess		Dean Chess?!? (2 gauges same operator?)		Todd Scholz	Todd Scholz
Sample Location	Avg Density, lb/ft^3		Avg Density, lb/ft^3		Avg Density, lb/ft^3		Average Density, lb/ft^3	Average Density, lb/ft^3
	Four readings	Two readings	Four readings	Two readings	Four readings	Two readings		
1	140.6	140.9	140.9	140.8	141.2	140.8	139.3	138.8
2	136.6	136.6	137.6	137.6	137.2	136.8	138.2	136.9
3	137.0	137.0	138.1	137.9	135.6	134.3	138.2	136.5
4	138.0	137.9	138.3	137.7	138.3	138.4	138.4	136.8
5	135.5	135.7	135.8	136.5	135.8	136.6	137.3	135.7
6	138.0	137.9	139.3	139.0	139.6	140.0	138.2	136.9
7	142.8	142.9	143.1	142.9	143.5	143.3	139.8	139.4
8	143.0	142.8	144.0	144.7	144.1	143.9	139.8	139.8
9	142.3	142.8	142.7	142.7	143.0	143.5	139.5	138.8
10	144.7	144.1	145.1	144.9	145.7	146.3	140.2	139.7
Average	139.8	139.8	140.5	140.5	140.4	140.4	138.9	137.9
Standard Deviation	3.2	3.2	3.1	3.2	3.6	3.9	0.9	1.5
Coefficient of Variation	2.30%	2.26%	2.23%	2.25%	2.57%	2.75%	0.68%	1.09%

Tables 4.8 and 4.9 summarize the gauge data collected to investigate if core average differences derived from cores obtained from the ‘L-shaped’ footprint (indicated in figure 3.1) are significantly different from those derived from cores obtained from the ‘cross-shaped’ footprint (indicated in figure 3.2). The data was collected by performing testing on the third (top) lift of the base course on the project with multiple lifts having same mix design (OR 18 project). Ten locations were randomly marked and alternate locations were used to perform testing either having the ‘cross-shaped’ combined gauge footprint or the ‘L-shaped’ combined gauge footprint. Table 4.8 comprises of data obtained by positioning nuclear gauge such that the footprint is ‘cross shaped’. Table 4.9 comprises of data obtained by positioning nuclear gauge such that the footprint is ‘L-shaped’. For each nuclear gauge used on the ‘cross-shaped’ combined footprint, again observe the similarity of average density based on two readings versus four readings. Note also the exceptionally low coefficient of variation (less than about 2% in most cases) indicating very low variance about the average density of the 5 sample locations each. In addition, there was at most a 0.9 lb/ft^3 difference in average density between nuclear gauge measurements.

Table 4.8 – Summarized data obtained from ‘cross-shaped’ footprint of wearing course (lift 3) on OR 18 Project

Gauge Make, Model, and Serial Number	Nuclear Gauges			
	Troxler 3430 # 38806		Troxler 3430 # 38994	
Last date of calibration	2/12/2009		1/14/2009	
Gauge Operator	Dean chess		Josh Huber	
Sample Location	Avg Density, lb/ft^3		Avg Density, lb/ft^3	
	Four readings	Two readings	Four readings	Two readings
11	150.8	150.0	151.9	151.4
13	152.2	151.9	153.4	152.7
15	149.2	149.2	150.4	150.3
17	153.2	153.3	152.5	152.6
19	154.3	154.3	154.7	154.6
Average	151.9	151.7	152.6	152.3
Standard Deviation	2.0	2.2	1.6	1.6
Coefficient of Variation	1.31%	1.43%	1.06%	1.06%

Table 4.9 – Summarized data obtained from 'L-shaped' footprint of wearing course (lift 3) on OR 18 Project

Gauge Make, Model, and Serial Number	Nuclear Gauges	
	Troxler 3430 # 38806	Troxler 3430 # 38994
Last date of calibration	2/12/2009	1/14/2009
Gauge Operator	Dean chess	Josh Huber
Sample Location	Avg Density, lb/ft^3	Avg Density, lb/ft^3
	Two readings	Two readings
12	149.7	151.0
14	147.8	148.6
16	153.7	152.5
18	151.1	152.7
20	153.8	153.7
Average	151.2	151.7
Standard Deviation	2.6	2.0
Coefficient of Variation	1.70%	1.30%

4.4 Summary

The tables above contain the essential field data to help answer the various research questions and fulfill the objectives. From the tables above, the common observations can be summarized as follows:

1. The average densities obtained using the nuclear gauges are almost always higher than those obtained using the electromagnetic gauges.

2. The coefficient of variance is exceptionally low (less than about 2% in most cases) indicating very low variance about the average density of the 10 sample locations each (and in some instances, 5 sample locations as indicated in tables 4.8 and 4.9).

3. The difference in average density between nuclear gauge measurements was at most 1.9 lb/ft^3 (except on the top lift of OR 18 project as indicated in table 4.3 when it was 3 lb/ft^3, however, this was not the case when more readings were taken on the same lift for other purposes as indicated in tables 4.8 and 4.9 and the difference was at most 0.9 lb/ft^3).

4. The similarity in results between the average of four readings per location and the average of two readings per location is consistent throughout all the tables (tables 4.1 to 4.9).

5.0 LABORATORY STUDY

The laboratory study was undertaken to measure the bulk specific gravity of the cores obtained from the projects evaluated during the field study (see Chapter 4), as well as to measure the theoretical maximum specific (Rice) gravity of loose HMA obtained from the pavements immediately following the paving operation (i.e., immediately before compaction). Bulk specific gravity measurements were performed using the saturated surface-dry and automatic vacuum sealing methods. The following sections provide details and results of the laboratory study.

5.1 Labeling of Cores

The cores extracted from the projects during the field study were labeled, placed in plastic bags, and transported to the asphalt laboratory at Oregon State University (OSU) where they were stored temporarily on shelves. The cores were labeled based on the project highway number, lift number, and location number. So, a core extracted from the sixth location of the second lift of paving on the OR 18 project was labeled “18-2-6.” Further, since it was extracted from the second lift, it was a composite core that comprised of the first and second lifts; thus, the core was cut with a diamond tipped saw blade to separate the lifts. The portion comprising the first lift was then labeled “18-2-6 Lift 1”. Similarly, the portion comprising the second lift was labeled “18-2-6 Lift 2.”

5.2 Details of Laboratory Testing Methods

Two test methods were used to measure the bulk specific gravity of the cores and a single test method was used to measure the Rice gravity of the loose HMA. The following sections describe the test methods in further detail.

5.2.1 Density Measurement of Cores

The cores were tested for bulk specific gravity in accordance with SSD method as described in AASHTO T 166 Method A and C (41), as well as in accordance with the CoreLok manufacturer’s procedure (53) as outlined in Figure 5.1. Cores with more than one lift were tested for bulk specific gravity using the saturated surface-dry and vacuum sealing methods, and then they were cut at the lift interfaces and retested for bulk specific gravity using the same methods.

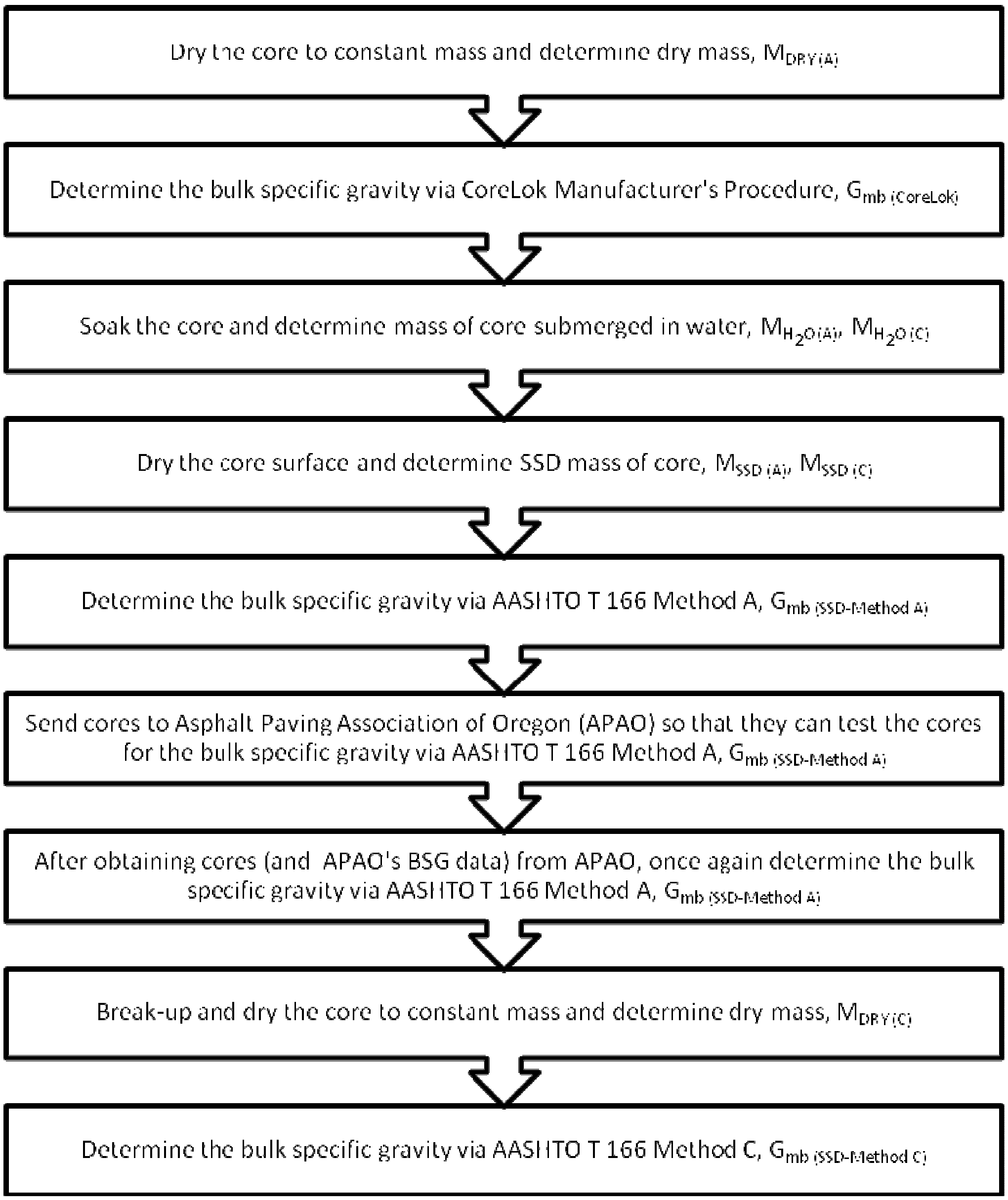


Figure 5.1 – Flowchart to describe the order in which laboratory tests were performed

After bulk specific gravity testing via AASHTO T 166 Method A and CoreLok manufacturer's method was completed at OSU, the cores were sent to the Asphalt Paving Association of Oregon (APAO) so that they could conduct bulk specific gravity tests

(AASHTO T 166, Method A) on the same cores. Thus, inter-laboratory comparative studies could be performed. Once testing at APAO was completed, the cores were sent back to OSU and retested via AASHTO T 166 Method A because there existed slight differences in the dry weight of the core specimens before and after the tests were conducted at APAO. This was because the saturated surface-dry masses and the submerged masses were required for obtaining bulk specific gravity in accordance with AASHTO T 166 Method C, which had not yet been performed.

It should be noted that, since the water used in the bulk specific gravity tests was maintained at a temperature of 25±1˚C, the density of the cores were calculated by multiplying the bulk specific gravity with the density of water at 25˚C (i.e., precisely 62.24 pcf).

5.2.2 Density Measurement on Loose Mix

Loose mix was sampled in accordance with AASHTO T 168 (51) and tested for theoretical maximum specific gravity in accordance with AASHTO T 209 (45). Quality control (QC) test results for theoretical maximum specific gravity were obtained from the contractor so as to make comparisons of theoretical maximum specific gravity obtained from the plant and from the mat.

5.3 Test Results

Various data were recorded as a part of the laboratory studies. Detailed summary tables are included in Appendix C, which includes details like project name, date of extraction, date of testing, name of person performing test, type of mix, lift thickness, core identification, sample location (i.e., station and offset), nominal field and laboratory thickness, bulk specific gravity and density (and related test procedure values needed to calculate the core density in accordance with the various procedures). This section includes condensed summaries of the data contained in Appendix C.

5.3.1 Densities Obtained from Measurements on Cores

Tables 5.1 to 5.8 summarize the density of cores from lifts on new construction pavements where each lift has the same mix design. Table 5.1 includes the data for the first (bottom) lift over the aggregate base, table 5.2 includes the data for the composite cores extracted from the second (middle) lift, table 5.3 includes the data for the core section belonging to first (bottom) lift that was cut out from that composite core, table 5.4 includes the data for the core section belonging to second (middle) lift that was cut out from that composite core, table 5.5 includes the data for the composite cores extracted from the third (top) lift, table 5.6 includes the data for the core section belonging to first (bottom) lift that was cut out from that composite core, table 5.7 includes the data for the core section belonging to second (middle) lift that was cut out from that composite core, and table 5.8 includes the data for the core section belonging to the third (top) lift that was cut out from that composite core. Note that each core testing method, in comparison to the remaining core testing methods, is consistent throughout all the lifts (for e.g., if the CoreLok method average reads lower than the SSD method A average and higher than the SSD method C average on a particular lift, then this same comparative result is observed to remain consistent throughout all the remaining lifts as well). Note also the exceptionally low coefficient of variation (less than about 2% in most cases) indicating very low variance about the average density of the 10 sample locations. In addition, there was at most a 2.1 lb/ft^3 difference in average density between various core tests' measurements.

The Figures 5.2 to 5.9 comprise of bar charts that compare the nuclear gauge measurements, electromagnetic gauge measurements (wherever used), and core results (on the new construction pavement project with multiple lifts and same mix design) for the lifts indicated in the corresponding tables 5.1 to 5.8. Here, note that the averages of the nuclear gauges' density values were always higher than those of the electromagnetic gauges and almost always higher than those of the core testing methods' density values in most of the instances. Similarly, the density results from the electromagnetic gauges and the various core tests were about the same in most of the instances.

Table 5.1 – Summarized data obtained from lift 1 of OR 18

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)		
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method C) BSG
18-1-1	106+50	8.2	3.35	2.52	147.2	155.1	155.0	154.1
18-1-2	116+91	1.6	3.74	3.43	147.7	148.0	147.2	146.6
18-1-3	120+53	2.9	2.36	2.05	148.3	148.0	147.4	146.7
18-1-4	130+80	6.6	3.54	2.95	146.0	150.0	149.6	148.5
18-1-5	142+09	8.2	3.43	2.95	150.2	150.3	149.6	148.4
18-1-6	133+38	5.7	3.39	2.95	152.7	153.0	152.2	150.9
18-1-7	124+96	4.9	3.62	2.99	154.0	155.6	155.5	155.1
18-1-8	123+30	3	3.43	2.80	148.9	150.0	149.5	148.1
18-1-9	116+56	5.4	3.82	3.39	155.1	155.0	154.7	153.5
18-1-10	115+53	4.8	4.25	3.54	151.1	154.4	154.2	152.6
Average			3.49	2.96	150.1	151.9	151.5	150.4
Standard Deviation			0.48	0.45	3.1	3.0	3.2	3.2
Coefficient of Variation			13.78%	15.10%	2.03%	1.97%	2.12%	2.12%

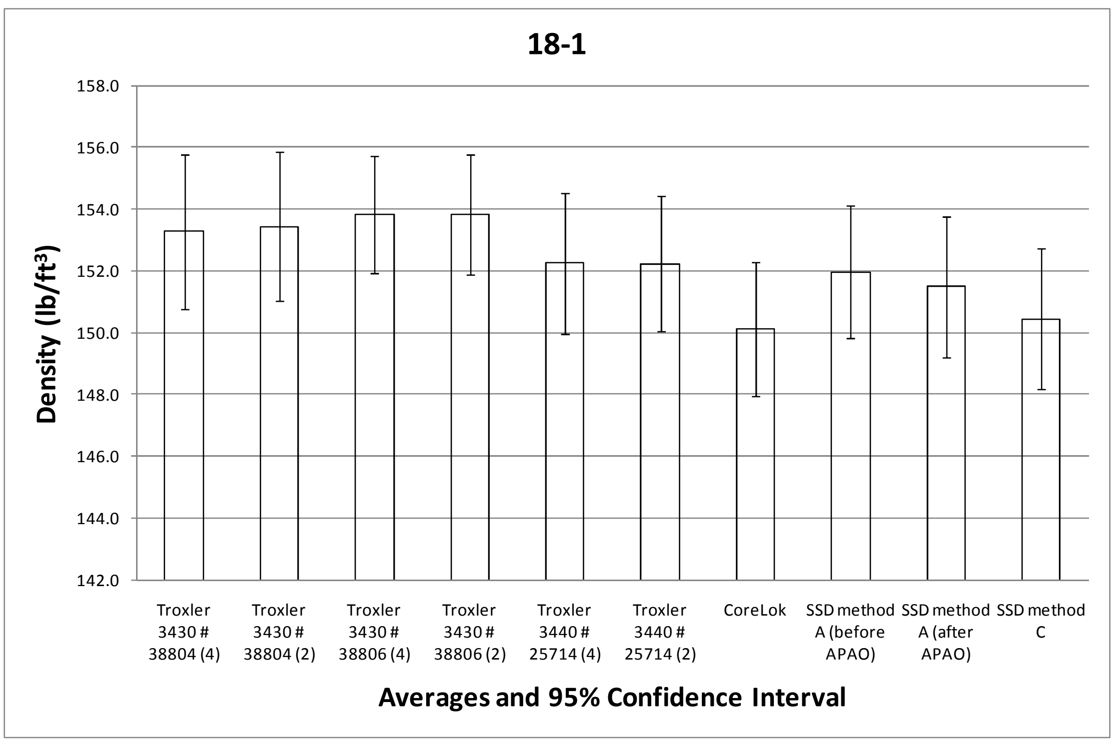


Figure 5.2 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 18-1

Table 5.2 - Summarized data obtained from composite cores (lift 1 and 2) of OR 18

Core ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (SSD Method A) BSG-before APAO
	Station	Offset				
18-2-1	140+08	5.2	5.12	4.45	149.4	149.2
18-2-2	140+53	6.4	5.28	4.29	148.2	148.7
18-2-3	141+42	10.8	5.71	4.92	148.5	148.3
18-2-4	142+06	15.1	5.71	4.96	148.6	148.7
18-2-5	144+06	8.4	5.91	5.35	148.9	149.0
18-2-6	145+08	13.6	6.06	5.31	149.8	149.8
18-2-7	145+25	3.8	5.75	4.92	148.7	148.4
18-2-8	145+94	1.9	5.79	5.12	147.0	147.5
18-2-9	146+67	7.4	5.63	4.72	148.2	150.7
18-2-10	146+97	4.1	5.39	4.80	149.7	149.6
Average			5.63	4.89	148.7	149.0
Standard Deviation			0.29	0.34	0.8	0.9
Coefficient of Variation			5.15%	6.97%	0.56%	0.60%

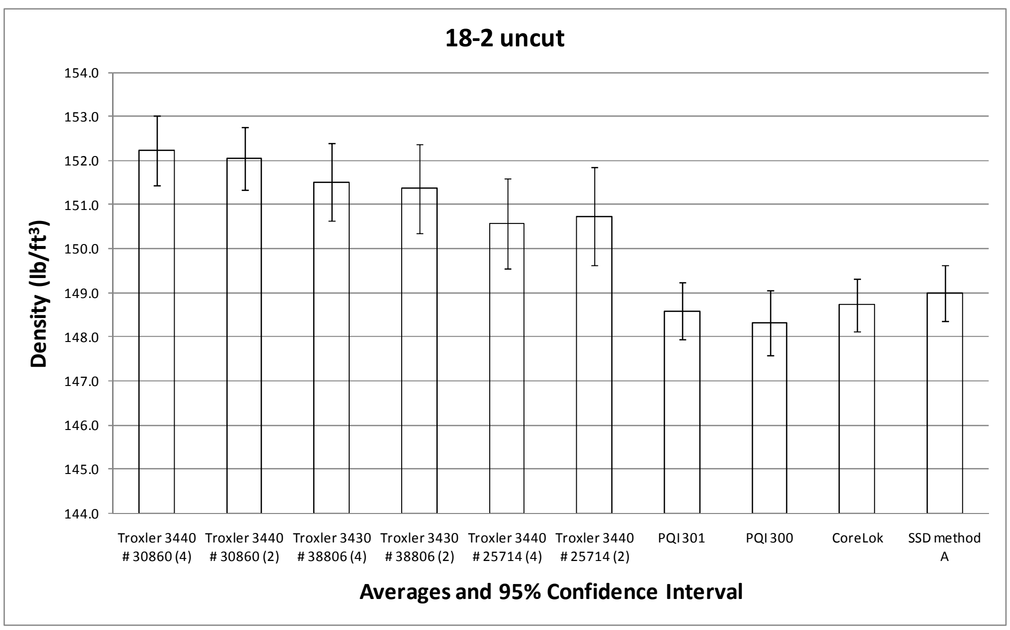


Figure 5.3 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 18-2

Table 5.3 - Summarized data obtained from lift 1 section cores of composite cores (lifts 1 and 2) of OR 18

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)			
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method A) BSG-APAO results	(Method C) BSG
18-2-1 Lift 1	140+08	5.2	0.00	2.17	149.8	149.8	149.3	149.6	148.0
18-2-2 Lift 1	140+53	6.4	0.00	2.05	149.0	149.5	148.6	149.1	147.1
18-2-3 Lift 1	141+42	10.8	0.00	2.36	151.3	151.3	150.6	151.0	149.9
18-2-4 Lift 1	142+06	15.1	0.00	2.52	149.1	149.8	147.3	147.1	146.3
18-2-5 Lift 1	144+06	8.4	0.00	2.52	148.4	148.7	147.2	147.4	146.3
18-2-6 Lift 1	145+08	13.6	0.00	2.56	150.0	149.9	149.6	149.8	149.3
18-2-7 Lift 1	145+25	3.8	0.00	2.24	148.5	148.0	146.2	146.3	145.2
18-2-8 Lift 1	145+94	1.9	0.00	2.52	145.1	145.8	144.6	145.0	143.3
18-2-9 Lift 1	146+67	7.4	0.00	2.24	152.5	152.6	151.8	152.1	151.2
18-2-10 Lift 1	146+97	4.1	0.00	2.28	145.9	150.3	148.4	148.4	147.5
Average				2.35	149.0	149.6	148.4	148.6	147.4
Standard Deviation				0.18	2.2	1.8	2.1	2.2	2.3
Coefficient of Variation				7.55%	1.50%	1.23%	1.44%	1.47%	1.58%

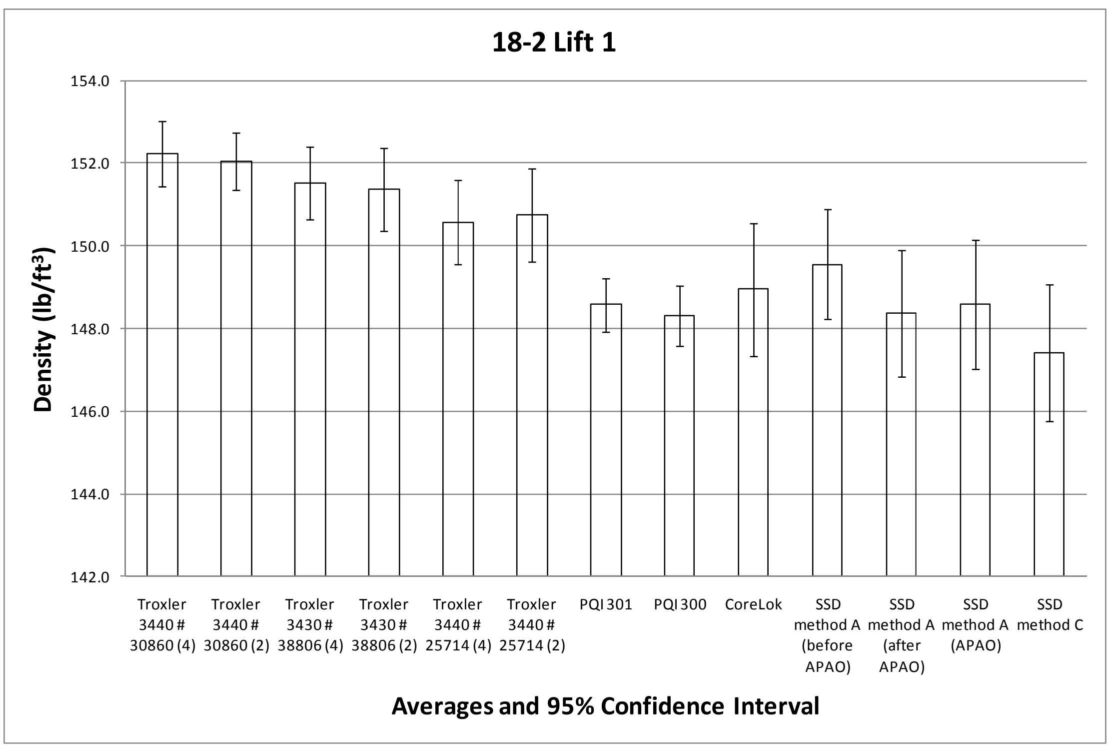


Figure 5.4 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 18-2 Lift 1

Table 5.4 - Summarized data obtained from lift 2 section cores of composite cores (lifts 1 and 2) of OR 18

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)			
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method A) BSG-APAO results	(Method C) BSG
18-2-1 Lift 2	140+08	5.2	N/A	2.24	148.5	148.5	148.8	149.3	148.0
18-2-2 Lift 2	140+53	6.4	N/A	2.24	147.1	148.2	147.1	146.6	146.6
18-2-3 Lift 2	141+42	10.8	N/A	2.56	147.1	148.3			
18-2-4 Lift 2	142+06	15.1	N/A	2.44	148.6	145.5	147.7	147.6	147.7
18-2-5 Lift 2	144+06	8.4	N/A	2.64	148.5	149.9	149.0	148.8	148.0
18-2-6 Lift 2	145+08	13.6	N/A	2.76	148.6	150.4	149.7	150.2	149.0
18-2-7 Lift2	145+25	3.8	N/A	2.60	149.0	149.6	148.5	148.6	147.8
18-2-8 Lift 2	145+94	1.9	N/A	2.60	149.4	149.9	149.4	149.7	148.7
18-2-9 Lift 2	146+67	7.4	N/A	2.36	149.0	149.8	148.7	148.8	147.7
18-2-10 Lift 2	146+97	4.1	N/A	2.44	148.0	149.0	147.9	147.9	147.5
Average				2.49	148.4	148.9	148.5	148.6	147.9
Standard Deviation				0.17	0.8	1.4	0.8	1.1	0.7
Coefficient of Variation				6.87%	0.51%	0.96%	0.56%	0.75%	0.47%

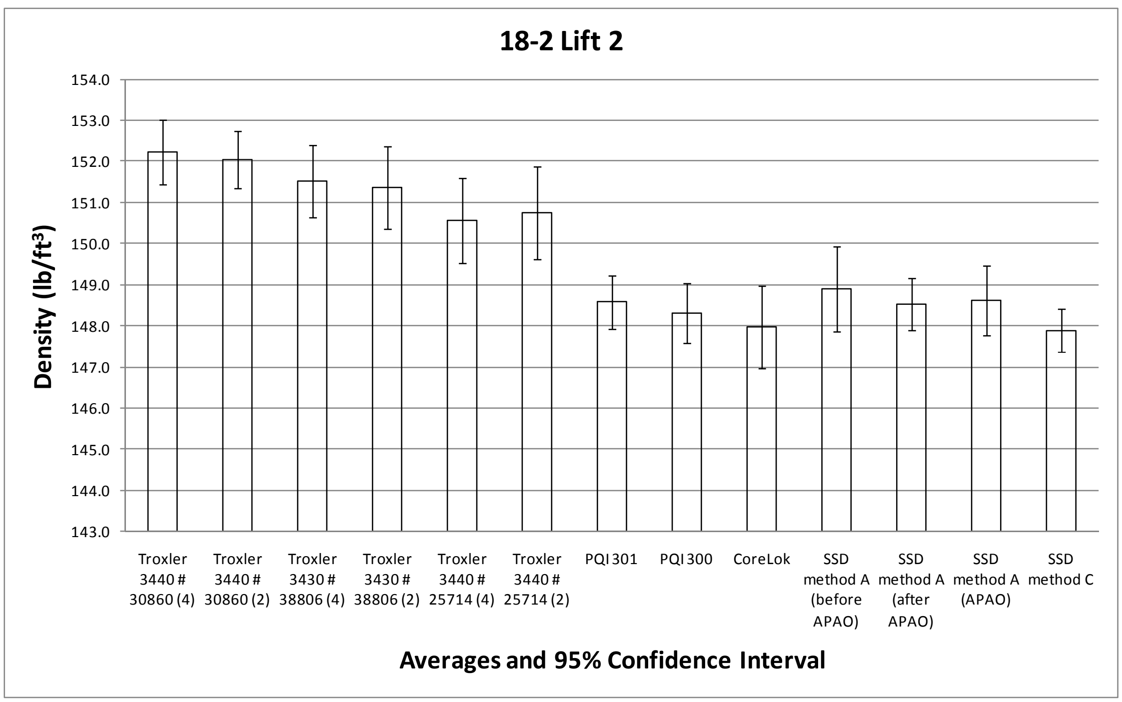


Figure 5.5 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 18-2 Lift 2

Table 5.5 - Summarized data obtained from composite cores (lifts 1, 2, and 3) of OR 18

Core ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (SSD Method A)
	Station	Offset				
18-3-1	139+57	3.8	6.30	5.51	151.1	150.4
18-3-2	139+84	9.4	6.81	5.91	155.1	152.6
18-3-3	140+86	2.2	7.36	6.38	152.2	151.5
18-3-4	141+38	10	6.81	6.18	150.1	149.7
18-3-5	142+01	5.3	6.81	6.06	149.5	149.2
18-3-6	142+51	2.9	7.28	5.51	150.5	145.7
18-3-7	143+92	3.4	7.44	6.85	148.2	147.8
18-3-8	144+98	11.5	7.24	5.63	N/A	N/A
18-3-9	146+82	2.5	7.48	6.65	145.8	145.7
18-3-10	147+36	1.5	7.20	6.30	144.5	145.8
Average			7.07	6.10	149.7	148.7
Standard Deviation			0.38	0.46	3.2	2.6
Coefficient of Variation			5.33%	7.62%	2.16%	1.75%

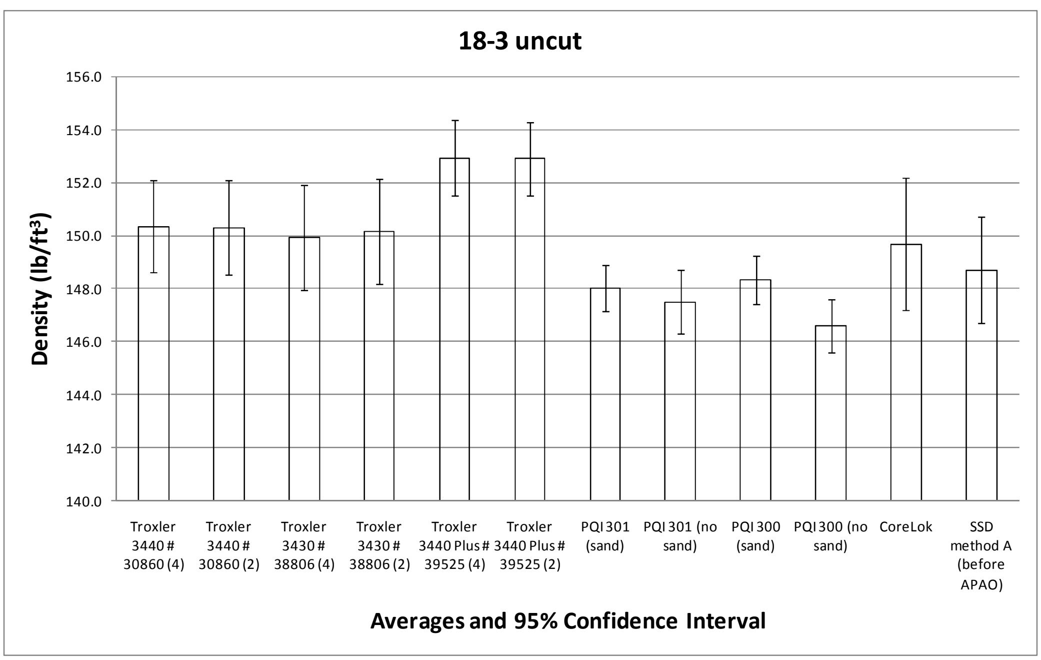


Figure 5.6 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 18-3

Table 5.6 - Summarized data obtained from lift 1 section cores of composite cores (lifts 1, 2, and 3) of OR 18

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)			
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method A) BSG-APAO results	(Method C) BSG
18-3-1 Lift 1	139+57	3.8	0.00	1.81	154.3	154.0	153.9	154.2	153.9
18-3-2 Lift 1	139+84	9.4	0.00	2.56	154.9	155.0	154.1	154.0	153.2
18-3-3 Lift 1	140+86	2.2	0.00	2.99	153.3	153.1	150.8	150.7	150.0
18-3-4 Lift 1	141+38	10	0.00	2.87	152.7	152.9	152.5	152.7	151.6
18-3-5 Lift 1	142+01	5.3	0.00	2.95	151.9	151.7	150.7	150.9	150.5
18-3-6 Lift 1	142+51	2.9	0.00	1.97	152.0	152.2	152.0	152.5	151.0
18-3-7 Lift 1	143+92	3.4	0.00	3.43	149.4	151.0	150.1	150.8	149.3
18-3-8 Lift 1	144+98	11.5	0.00	1.89	152.2	152.3	151.7	152.2	150.6
18-3-9 Lift 1	146+82	2.5	0.00	2.76	146.6	147.1	144.0	143.6	142.5
18-3-10 Lift 1	147+36	1.5	0.00	2.56	141.5	143.1			
Average				2.58	150.9	151.2	151.1	151.3	150.3
Standard Deviation				0.54	4.1	3.6	3.0	3.2	3.3
Coefficient of Variation				20.77%	2.70%	2.36%	1.98%	2.10%	2.18%

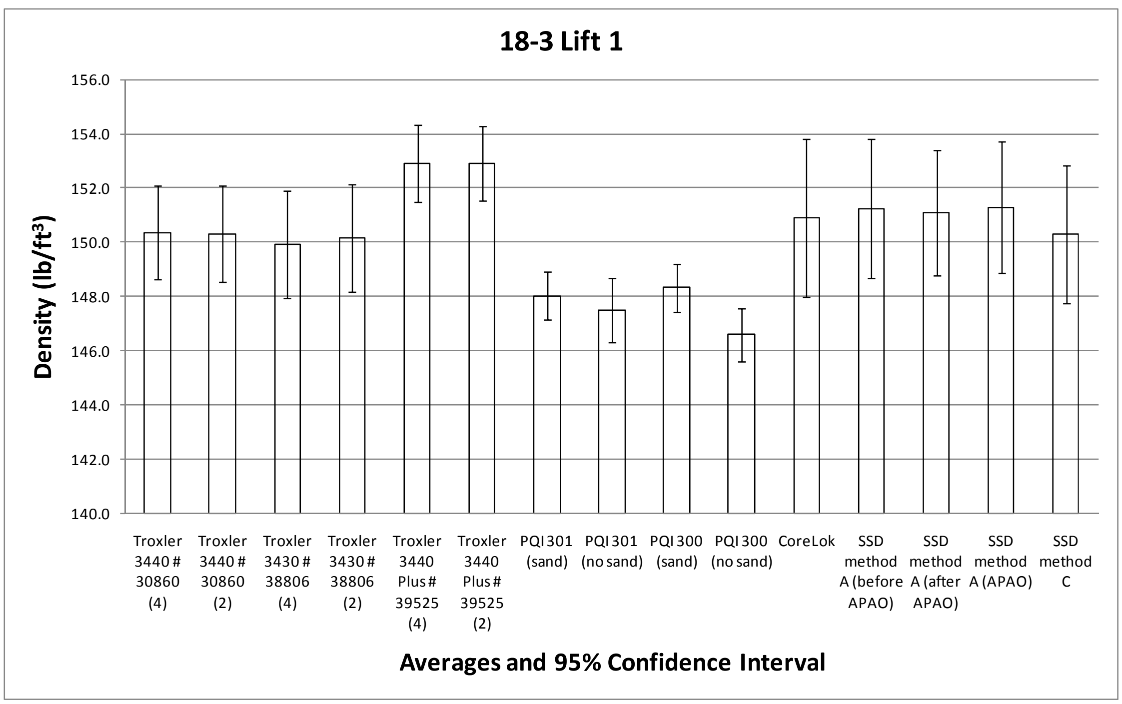


Figure 5.7 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 18-3 Lift 1

Table 5.7 - Summarized data obtained from lift 2 section cores of composite cores (lifts 1, 2, and 3) of OR 18

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)			
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method A) BSG-APAO results	(Method C) BSG
18-3-1 Lift 2	139+57	3.8	N/A	1.57	147.9	149.5	149.1	149.5	148.2
18-3-2 Lift 2	139+84	9.4	N/A	1.14	147.8	148.8	147.5	147.5	146.3
18-3-3 Lift 2	140+86	2.2	N/A	1.54	151.2	152.2	151.8	151.6	151.2
18-3-4 Lift 2	141+38	10	N/A	0.98	147.3	148.3	147.9	148.4	146.3
18-3-5 Lift 2	142+01	5.3	N/A	1.02	146.6	148.2	147.3	147.6	146.2
18-3-6 Lift 2	142+51	2.9	N/A	1.18	149.0	149.8	149.3	149.9	148.3
18-3-7 Lift 2	143+92	3.4	N/A	0.98	142.0	144.3	142.6	143.7	142.0
18-3-8 Lift 2	144+98	11.5	N/A	1.06	146.6	148.4	147.9	147.9	146.7
18-3-9 Lift 2	146+82	2.5	N/A	1.57	149.3	150.3	149.8	149.8	148.9
18-3-10 Lift 2	147+36	1.5	N/A	1.57	146.0	147.3	145.9	145.8	144.5
Average				1.26	147.4	148.7	147.9	148.2	146.9
Standard Deviation				0.27	2.4	2.1	2.5	2.3	2.5
Coefficient of Variation				21.10%	1.65%	1.39%	1.67%	1.52%	1.72%

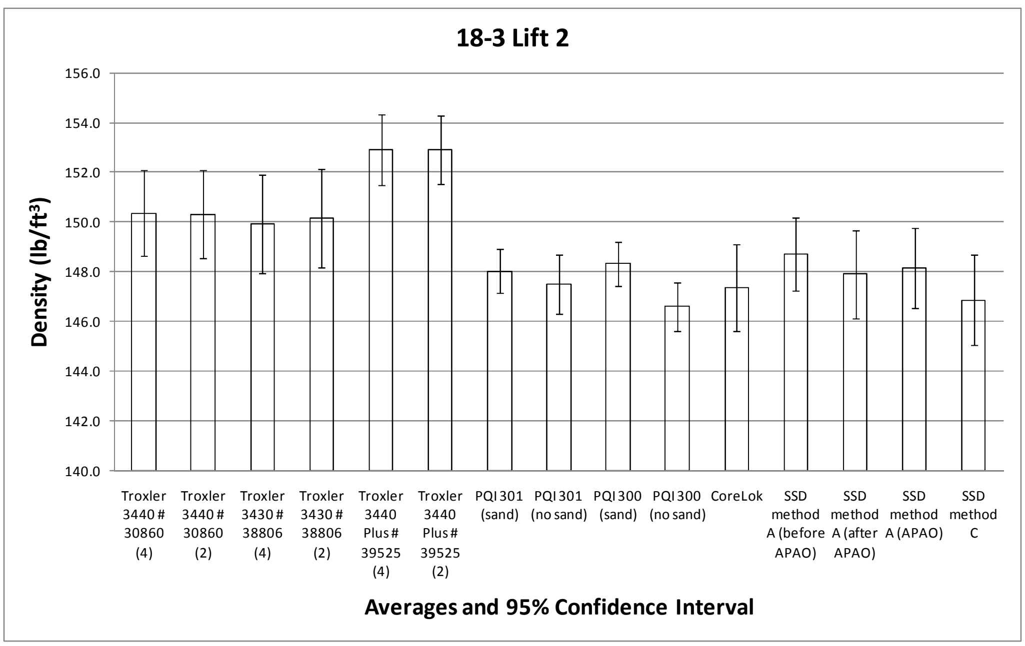


Figure 5.8 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 18-3 Lift 2

Table 5.8 - Summarized data obtained from lift 3 section cores of composite cores (lifts 1, 2, and 3) of OR 18

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)			
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method A) BSG-APAO results	(Method C) BSG
18-3-1 Lift 3	139+57	3.8	N/A	2.17	149.9	150.1	149.4	149.0	149.0
18-3-2 Lift 3	139+84	9.4	N/A	2.20	152.8	153.2	153.0	153.0	152.5
18-3-3 Lift 3	140+86	2.2	N/A	2.05	148.2	149.5	148.4	148.5	147.9
18-3-4 Lift 3	141+38	10	N/A	2.36	146.2	148.1	147.5	148.0	147.0
18-3-5 Lift 3	142+01	5.3	N/A	2.05	146.9	147.7	146.2	146.4	145.7
18-3-6 Lift 3	142+51	2.9	N/A	2.32	148.5	149.2	148.9	149.3	148.3
18-3-7 Lift 3	143+92	3.4	N/A	2.36	147.6	148.1	147.7	148.0	147.1
18-3-8 Lift 3	144+98	11.5	N/A	2.44	147.3	148.2	147.9	148.3	147.4
18-3-9 Lift 3	146+82	2.5	N/A	2.48	142.9	144.2	142.7	141.8	142.1
18-3-10 Lift 3	147+36	1.5	N/A	2.13	147.3	147.9	147.1	146.9	146.8
Average				2.26	147.7	148.6	147.9	147.9	147.4
Standard Deviation				0.16	2.5	2.3	2.6	2.8	2.6
Coefficient of Variation				7.03%	1.73%	1.52%	1.75%	1.88%	1.77%

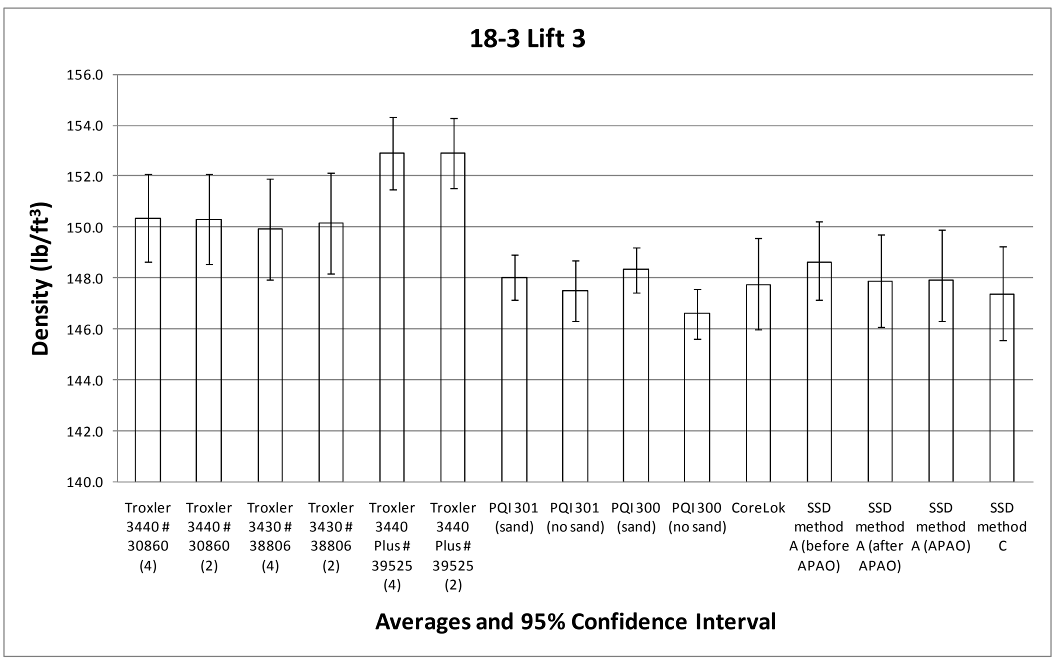


Figure 5.9 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 18-3 Lift 3

Tables 5.9 and 5.10 summarize the density of cores from lifts on the new pavement construction pavement project with each lift having different mix design. Table 5.9 includes the data for the first (bottom) lift over the aggregate base and table 5.10 includes the data for the second (top) lift. Note that each core testing method, in comparison to the remaining core testing methods, is consistent throughout all the lifts (for e.g., if the CoreLok method average reads lower than the SSD method A average and higher than the SSD method C average on a particular lift, then this same comparative result is observed to remain consistent throughout all the remaining lifts as well). Note also the exceptionally low coefficient of variation (less than about 2% in most cases) indicating very low variance about the average density of the 10 sample locations. In addition, there was at most a 1.8 lb/ft^3 difference in average density between various core tests' measurements.

The Figures 5.10 and 5.11 comprise of bar charts that compare the nuclear gauge measurements and core results (on the new pavement construction pavement project with each lift having different mix design) for the lifts indicated in the corresponding tables 5.9 and 5.10. Here, note that the averages of the nuclear gauges' density values were about the same as those of the core testing methods' density values in most of the instances. However, the density measurements of the electromagnetic gauges' measurements were almost always lower than those of the nuclear gauges' measurements and core tests' measurements.

Table 5.9 - Summarized data obtained from lift 1 of I-5 project

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)	
	Station	Offset (ft)			(SSD Method A) BSG	(SSD Method C) BSG
5.1.1	252+22	4	2.76	144.3	146.2	145.6
5.1.2	254+07	6.5	2.95	140.1	142.4	141.9
5.1.3	255+07	9.5	2.56	143.6	144.7	144.1
5.1.4	255+67	5.5	2.68	142.2	143.4	143.3
5.1.5	255+81	8.4	2.24	143.6	145.3	145.5
5.1.6	256+33	9	2.76	141.2	143.8	143.6
5.1.7	257+55	9	2.83	144.6	145.5	145.0
5.1.8	259+35	2.7	2.95	143.8	145.2	144.3
5.1.9	261+68	5.3	2.56	143.3	145.2	144.8
5.1.10	263+88	7.2	2.48	143.5	144.3	144.0
Average			2.7	143.0	144.6	144.2
Standard Deviation			0.2	1.4	1.1	1.1
Coefficient of Variation			8.29%	1.00%	0.79%	0.77%

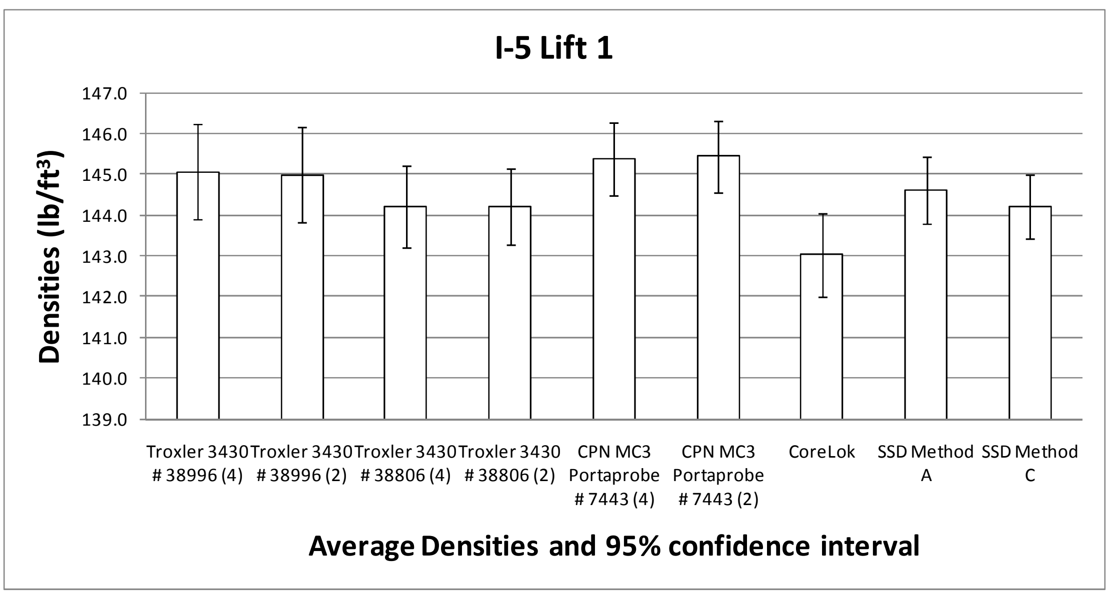


Figure 5.10 – Comparisons of densities and their 95% confidence level obtained from the various tests performed on Lift 1 of the project on I-5

Table 5.10 - Summarized data obtained from lift 2 of I-5 project

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)	
	Station	Offset (ft)			(SSD Method A) BSG	(SSD Method C) BSG
5.2.1	252+02	9.1	2.48	142.4	143.8	143.1
5.2.2	252+59	3.4	3.15	143.3	144.9	144.7
5.2.3	253+34	2	2.95	144.6	145.4	145.1
5.2.4	254+23	4.1	3.15	141.6	143.3	142.9
5.2.5	254+98	3.3	3.23	141.7	142.5	142.1
5.2.6	256+62	2.3	3.23	143.2	144.3	143.8
5.2.7	258+07	11.7	3.31	141.1	142.2	141.9
5.2.8	258+83	6	3.23	140.7	142.5	142.0
5.2.9	260+92	10.1	2.91	142.4	143.8	143.3
5.2.10	263+60	9	2.95	138.5	144.6	144.4
Average			3.1	142.0	143.7	143.3
Standard Deviation			0.2	1.7	1.1	1.2
Coefficient of Variation			8.03%	1.19%	0.77%	0.81%

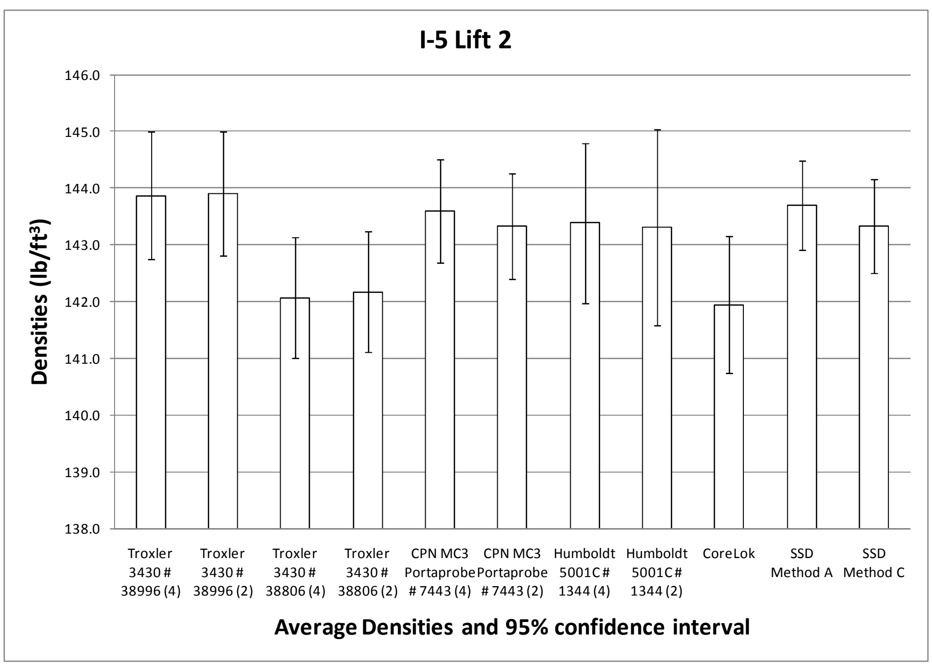


Figure 5.11 – Comparisons of densities and their 95% confidence level obtained from the various tests performed on Lift 2 of the project on I-5

Tables 5.11 to 5.14 summarize the density of cores from lifts on the mill and fill with overlay construction pavement project with each lift having same mix design. Table 5.11 includes the data for the first (bottom or inlay) lift over the aggregate base, table 5.2 includes the data for the composite cores extracted from the second (top or overlay) lift, table 5.3 includes the data for the core section belonging to first (bottom or inlay) lift that was cut out from that composite core, and table 5.4 includes the data for the core section belonging to second (top or overlay) lift that was cut out from that composite core. Note that each core testing method, in comparison to the remaining core testing methods, is consistent throughout all the lifts (for e.g., if the CoreLok method average reads lower than the SSD method A average and higher than the SSD method C average on a particular lift, then this same comparative result is observed to remain consistent throughout all the remaining lifts as well). Note also the exceptionally low coefficient of variation (less than about 2% in most cases) indicating very low variance about the average density of the 10 sample locations. In addition, there was at most a 2.9 lb/ft^3 difference in average density between various core tests' measurements (however, the remaining lifts had a difference of at most 1.7 lb/ft^3).

The Figures 5.12 to 5.15 comprise of bar charts that compare the nuclear gauge measurements, electromagnetic gauge measurements, and core results (on the mill and fill with overlay construction pavement project with each lift having same mix design) for the lifts indicated in the corresponding tables 5.11 to 5.14. Here, note that the averages of the nuclear gauges' density values were about the same as those of the core testing methods' density values in most of the instances. However, the density measurements of the electromagnetic gauges' measurements were almost always lower than those of the nuclear gauges' measurements and core tests' measurements.

Table 5.11 - Summarized data obtained from lift 1 of OR 140

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)		
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method C) BSG
140-1-1	328+55	1.3	2.09	1.26	139.9	141.7	141.3	140.2
140-1-2	329+05	11.2	2.13	1.65	140.8	142.2	141.5	141.0
140-1-3	329+21	10.2	2.09	1.46	138.9	140.4	139.9	138.7
140-1-4	329+37	6.6	2.13	1.38	139.8	141.3	140.8	140.4
140-1-5	330+50	3.5	2.36	1.65	139.9	141.3	N/A	N/A
140-1-6	330+80	5.1	2.44	1.77	139.5	140.6	140.0	138.9
140-1-7	331+40	3.3	2.36	1.77	139.5	141.1	140.3	139.3
140-1-8	332+62	12.2	2.28	1.46	137.2	139.5	138.0	137.1
140-1-9	334+03	7.8	2.52	1.85	138.2	139.8	139.5	138.4
140-1-10	334+86	1.6	2.56	1.97	134.5	136.9	136.7	135.7
Average			2.30	1.62	138.8	140.5	139.8	138.9
Standard Deviation			0.18	0.23	1.8	1.5	1.5	1.7
Coefficient of Variation			7.88%	14.00%	1.31%	1.07%	1.10%	1.19%

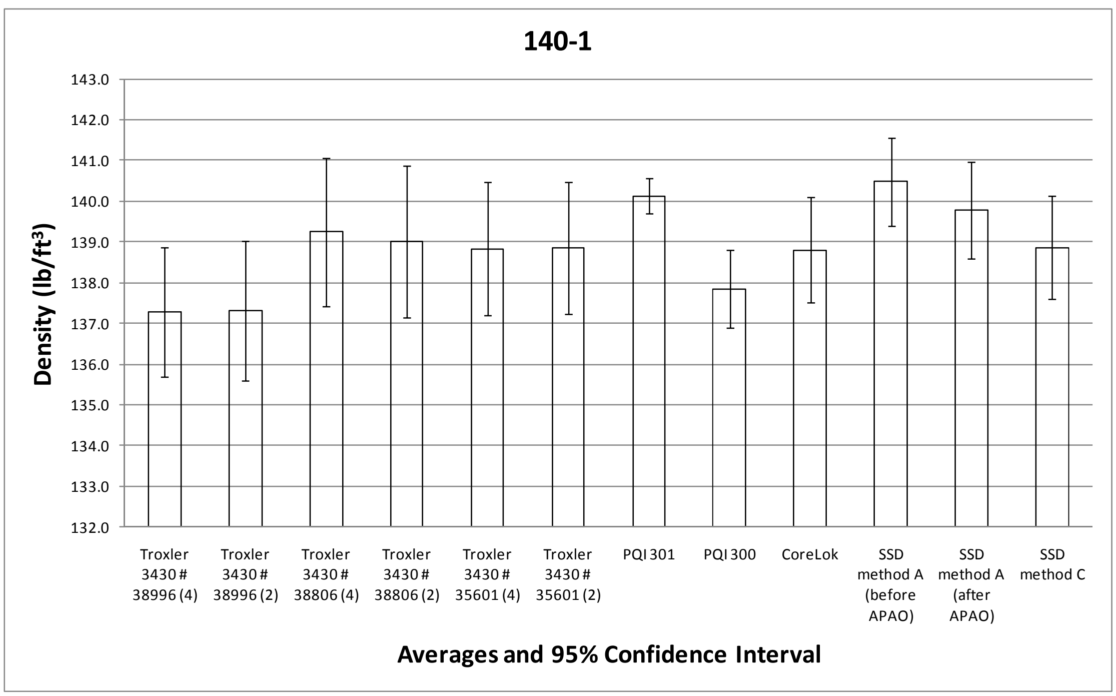


Figure 5.12 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 140-1 Lift 1

Table 5.12 - Summarized data obtained from composite cores (lifts 1 and 2) of OR 140

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (SSD Method A) BSG-before APAO
	Station	Offset				
140-2-1	326+35	11.6	4.92	3.94	142.0	142.4
140-2-2	329+40	6.2	5.04	3.94	138.3	139.5
140-2-3	330+65	12.3	4.65	3.74	N/A	N/A
140-2-4	331+02	3.6	4.61	3.74	139.5	139.8
140-2-5	333+46	2.2	5.12	4.33	131.2	138.2
140-2-6	334+02	4.1	5.31	4.41	138.6	139.5
140-2-7	334+48	11.2	4.80	3.74	142.0	142.6
140-2-8	334+78	10.5	5.00	4.49	N/A	N/A
140-2-9	334+83	11.4	5.08	4.17	N/A	N/A
140-2-10	334+94	10.1	5.24	4.57	143.2	143.6
Average			4.98	4.11	139.2	140.8
Standard Deviation			0.23	0.33	4.0	2.0
Coefficient of Variation			4.72%	7.97%	2.88%	1.43%

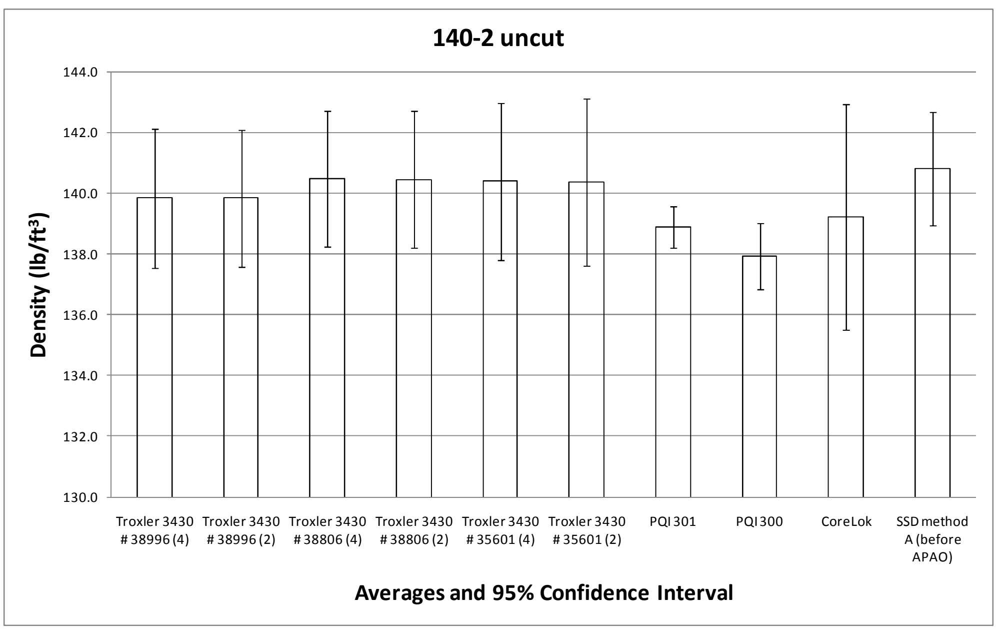


Figure 5.13 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 140-2

Table 5.13 - Summarized data obtained from lift 1 section cores of composite cores (lifts 1 and 2) of OR 140

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)			
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method A) BSG-APAO results	(Method C) BSG
140-2-1	326+35	11.6	N/A	1.77	142.5	143.1	142.8	143.1	142.1
140-2-2	329+40	6.2	N/A	1.69	140.6	141.9	141.0	141.3	139.6
140-2-3	330+65	12.3	N/A	1.46	139.0	141.1	140.4	140.8	139.1
140-2-4	331+02	3.6	N/A	1.65	140.8	141.8	141.4	141.5	140.3
140-2-5	333+46	2.2	N/A	1.97	136.9	139.0	138.5	138.4	N/A
140-2-6	334+02	4.1	N/A	1.89	138.5	140.1	139.2	139.9	138.4
140-2-7	334+48	11.2	N/A	1.50	138.4	140.3	139.2	139.0	137.6
140-2-8	334+78	10.5	N/A	1.97	135.9	139.6	139.1	139.2	137.9
140-2-9	334+83	11.4	N/A	1.65	137.3	139.2	138.4	138.0	137.4
140-2-10	334+94	10.1	N/A	2.05	140.0	141.2	140.5	140.9	139.6
Average				1.76	139.0	140.7	140.0	140.2	139.1
Standard Deviation				0.20	2.0	1.3	1.4	1.6	1.5
Coefficient of Variation				11.60%	1.46%	0.94%	1.02%	1.13%	1.09%

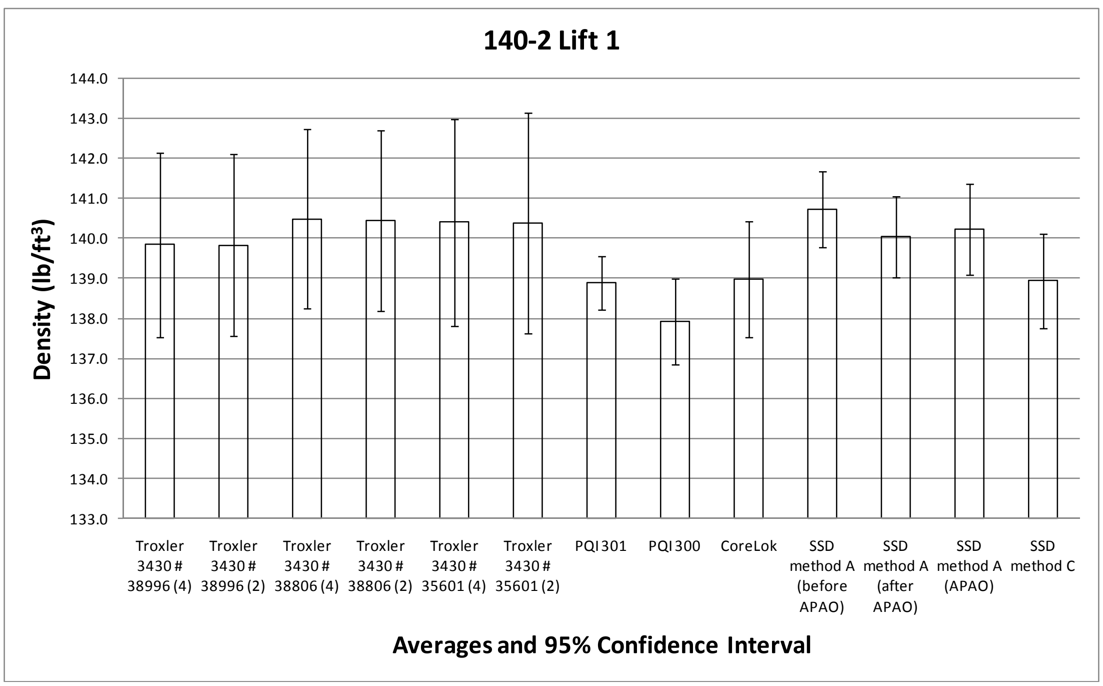


Figure 5.14 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 140-2 Lift 1

Table 5.14 - Summarized data obtained from lift 2 section cores of composite cores (lifts 1 and 2) of OR 140

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Thickness of Separated Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)			
	Station	Offset				(Method A) BSG-before APAO testing	(Method A) BSG -after APAO testing	(Method A) BSG-APAO results	(Method C) BSG
140-2-1	326+35	11.6	N/A	2.28	139.9	142.3	143.4	141.9	142.2
140-1-2	329+40	6.2	N/A	2.28	140.8	139.3	138.0	137.7	137.4
140-1-3	330+65	12.3	N/A	2.36	138.9	139.0	138.0	138.0	136.9
140-1-4	331+02	3.6	N/A	2.05	139.8	139.7	139.1	139.2	137.9
140-1-5	333+46	2.2	N/A	2.56	139.9	138.3	137.3	137.2	137.4
140-1-6	334+02	4.1	N/A	2.56	139.5	140.4	139.8	140.3	139.2
140-1-7	334+48	11.2	N/A	2.28	139.5	143.8	143.1	143.4	142.6
140-1-8	334+78	10.5	N/A	2.56	137.2	144.8	143.5	144.0	142.2
140-1-9	334+83	11.4	N/A	2.56	138.2	143.6	142.6	142.7	141.3
140-1-10	334+94	10.1	N/A	2.56	134.5	145.7	145.4	145.3	144.9
Average				2.41	138.8	141.7	141.0	141.0	140.2
Standard Deviation				0.18	1.8	2.7	2.9	2.9	2.8
Coefficient of Variation				7.50%	1.31%	1.87%	2.04%	2.05%	1.99%

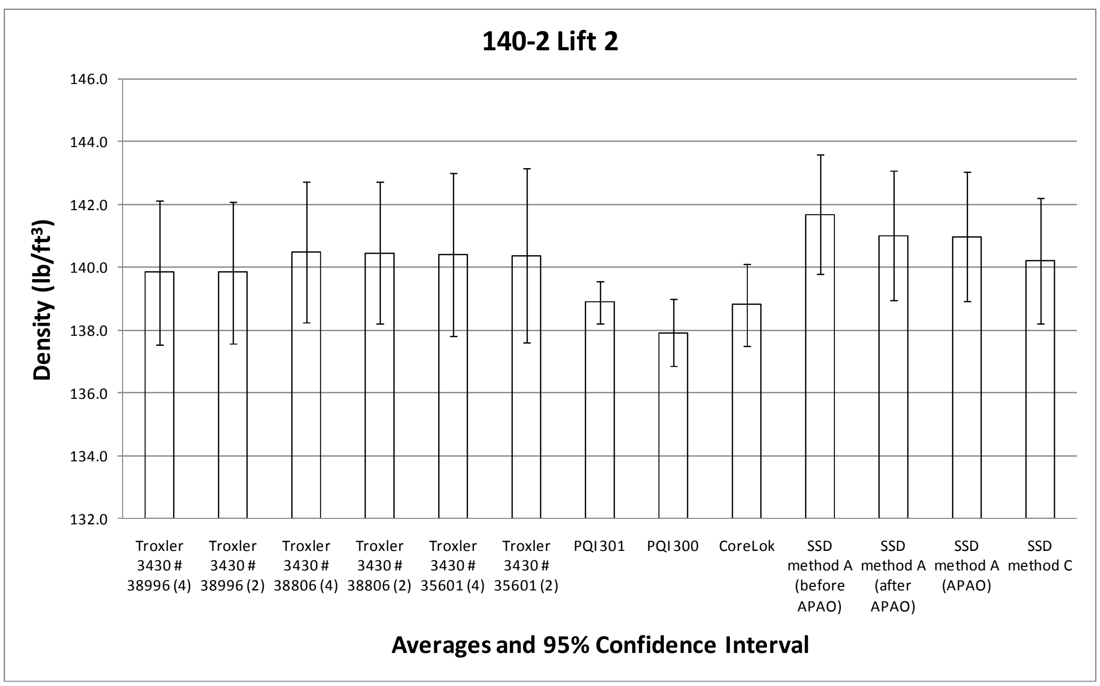


Figure 5.15 – Comparison of densities and their confidence limits (95%) obtained from the various tests performed on section 140-2 Lift 2

Tables 5.15 and 5.16 summarize the density of cores that are extracted from the inner corner of the 'L-shaped' footprint and cores that are extracted from the center of the 'cross-shaped' footprint. This is performed on the top lift of the new pavement construction project with the lifts below having same mix design. Table 5.15 includes the data for the cores extracted from the 'cross-shaped' footprint and table 5.16 includes the data for the cores extracted from the 'L-shaped' footprint. Note that each core testing method's result, in comparison to the remaining core testing methods' results, is consistent regardless of whether the core was extracted from the inner corner of the 'L-shaped footprint or whether it was extracted from the center of the 'cross-shaped' footprint (for e.g., the CoreLok method average always reads lower than the SSD method A average and higher than the SSD method C average). Note also the exceptionally low coefficient of variation (less than about 2% in most cases) indicating very low variance about the average density of the 5 sample locations. In addition, there was at most a 1.3 lb/ft^3 difference in average density between various core tests' measurements.

The Figures 5.16 and 5.17 comprise of bar charts that compare the nuclear gauge measurements and core results (based on the 'cross-shaped' footprint and the 'L-shaped' footprint on the new pavement construction project having same mix design on all lifts) based on the footprints as indicated in the corresponding tables 5.15 and 5.16. Here, note that the averages of the nuclear gauges' density values were about the same as those of the core testing methods' density values in most of the instances. However, the density measurements of the electromagnetic gauges' measurements were almost always lower than those of the nuclear gauges' measurements and core tests' measurements.

Table 5.15 - Summarized data obtained from cores extracted from the ‘cross-shaped’ footprint on lift 3 of project on OR 18

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)	
	Station	Offset (ft)			(SSD Method A) BSG	(SSD Method C) BSG
18.3.11	90+89	9	2	148.3	148.7	147.3
18.3.13	98+85	8.6	2	149.8	150.2	148.9
18.3.15	103+62	1.7	2	148.4	149.3	148.8
18.3.17	109+44	3.8	2	151.4	151.8	150.5
18.3.19	113+73	5.1	2	154.9	155.2	154.2
Average				150.5	151.1	149.9
Standard Deviation				2.7	2.6	2.6
Coefficient of Variation				1.82%	1.72%	1.75%

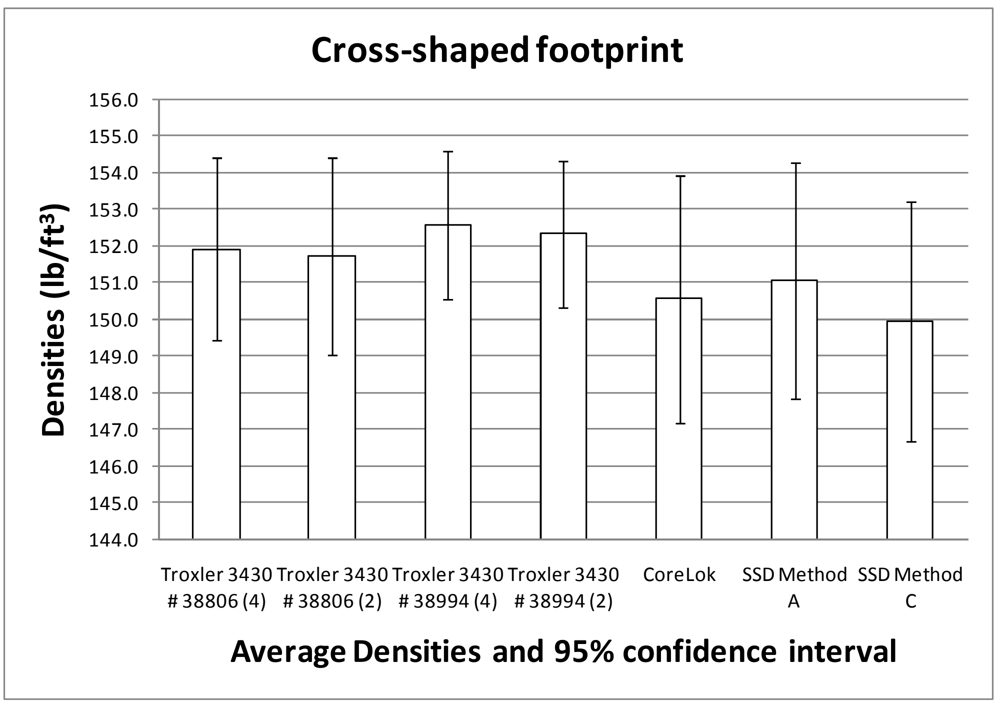


Figure 5.16 – Comparisons of densities and their 95% confidence level obtained from the various tests performed on the core extracted from the ‘cross-shaped’ footprint on Lift 3 of the project on OR 18

Table 5.16 - Summarized data obtained from cores extracted from the 'L-shaped' footprint on lift 3 of project on OR 18

Location/ Specimen ID	Sample Location		Nominal Thickness of Intact Specimen (inches)	Density obtained from CoreLok BSG (lb/ft^3)	Density obtained from AASHTO T 166 (lb/ft^3)	
	Station	Offset (ft)			(SSD Method A) BSG	(SSD Method C) BSG
18.3.12	91+88	7.4	2	148.1	149.0	147.5
18.3.14	99+71	2	2	145.1	146.4	144.8
18.3.16	108+70	4.5	2	151.0	151.4	150.2
18.3.18	111+53	3.3	2	151.3	151.9	151.0
18.3.20	113+94	4.9	2	154.2	154.9	153.5
Average				149.9	150.7	149.4
Standard Deviation				3.4	3.2	3.4
Coefficient of Variation				2.30%	2.13%	2.25%

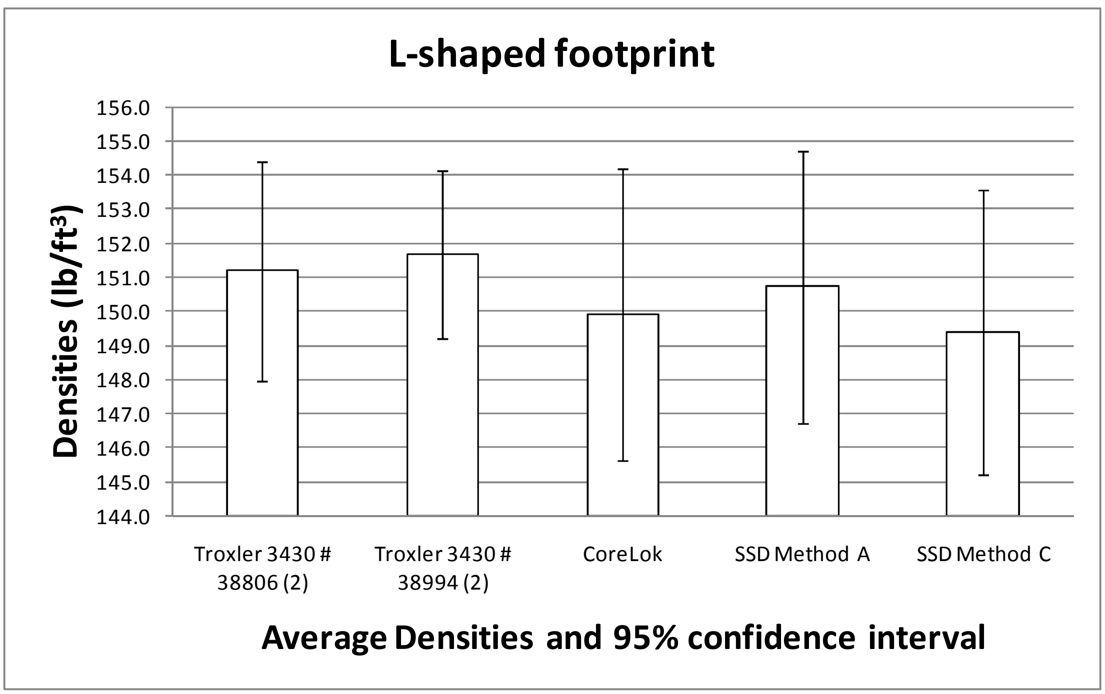


Figure 5.17 – Comparisons of densities and their 95% confidence level obtained from the various tests performed on the core extracted from the 'L-shaped' footprint on Lift 3 of the project on OR 18

5.3.2 Theoretical Maximum Specific Gravity of Loose Mix Samples

On the project with multiple lifts of same mix design, three samples were taken in accordance with AASHTO T 168, and testing was performed in accordance with AASHTO T 209. These theoretical maximum specific gravity values are contained in the table 5.17 (this includes contractor's values too). Similarly, for the project with multiple lifts of different mix design, the theoretical maximum specific gravity values are contained in table 5.18 (this does not include contractor's values). Finally, for the project with mill and fill (inlay) with overlay, the theoretical maximum specific gravity values are contained in table 5.19 (this includes contractor's values too).

Table 5.17 – Rice Gravity Data on OR 18 as obtained from contractor and from OSU

Project	Date	Sample	OSU G_{mm} Results	Contractor G_{mm} Results
OR 18	18-Aug 2008	a	2.609	2.582
OR 18	18-Aug 2008	b	2.600	
OR 18	22-Aug 2008	a	2.575	2.576
OR 18	22-Aug 2008	b	2.588	
OR 18	25-Aug 2008	c	2.592	
OR 18	25-Aug 2008	a	2.577	2.5650
OR 18	25-Aug 2008	b	2.569	

Table 5.18 – Rice Gravity Data on I-5 as obtained from contractor and from OSU

Project	Date	Sample	OSU G_{mm} Results	Contractor G_{mm} Results
I-5	14-Aug 2009	a	2.540	N/A
I-5	14-Aug 2009	b	2.499	N/A
I-5	14-Aug 2009	c	2.503	N/A
I-5	14-Aug 2009	d	2.491	N/A
I-5	14-Aug 2009	a	2.486	N/A
I-5	14-Aug 2009	b	2.485	N/A
I-5	14-Aug 2009	c	2486	N/A
I-5	14-Aug 2009	d	2.480	N/A
I-5	14-Aug 2009	a	2.446	N/A
I-5	14-Aug 2009	b	2.432	N/A

Table 5.19 – Rice Gravity Data on OR 140 as obtained from contractor and from OSU

Project	Date	Sample	OSU G_{mm} Results	Contractor G_{mm} Results
OR 140	20-Aug 2008	a	2.459	2.445
OR 140	20-Aug 2008	b	2.475	
OR 140	20-Aug 2008	c	2.465	
OR 140	11-Sep 2008	a	2.452	2.452
OR 140	11-Sep 2008	b	2.457	

5.4 Summary

The common findings from the tables above can be summarized as follows:

1. For any given testing method, the results are consistent (i.e., if the resultant values of a particular are lesser that those of another method, that remains the same throughout all the lifts for the given project).

2. Also the exceptionally low coefficient of variation (less than about 2% in most cases) indicates very low variance about the average density of the 5 or 10 sample locations.

3. In addition, there was at most a 3 lb/ft^3 difference (in most cases it was below 2 lb/ft^3) in average density between various core tests' measurements considering all the three projects tested.

4. It must be noted that the averages of the nuclear gauges' density values were about the same as those of the core testing methods' density values in most of the instances (multiple lifts with different mix design i.e., I-5, mill and fill with overlay i.e., OR 140, cores extracted from 'L-shaped' footprint vs. 'cross-shaped' footprint). However, on the project with multiple lifts having same mix design i.e., the OR 18 project, nuclear gauges' readings are mostly greater than corresponding results from the various core tests.

5. The density measurements of the electromagnetic gauges' measurements were almost always lower than those of the nuclear gauges' measurements and core tests' measurements.

6.0 DATA ANALYSES

Based on the experiment design to answer the research questions (Section 3.1), the densities were obtained on the selected projects using nuclear and electromagnetic gauges, and bulk specific gravities were obtained from cores extracted from the same locations as the gauge measurements as detailed in sections 3.2 and 3.3. The theoretical maximum bulk specific gravities (AASHTO T 209) were also obtained for each lift on the above mentioned projects. The nuclear gauge and bulk specific gravity measurement procedures, other experiment details, and supporting data are provided in Chapters 3 through 5.

Once the above mentioned data was recorded, statistical tests were performed on the data to derive various conclusions. Paired and pooled t-tests were used to evaluate the research questions. The statistical analysis procedures are provided in this chapter as are the analyses details and results for each research question listed in Section 3.1.

6.1 Scope of Statistical Analysis

Firstly, paired t-tests were performed to compare the densities obtained from the various methods and devices on every given lift to understand how significantly different they are from each other. It should be noted that the bulk specific gravities obtained from the cores were converted to densities prior to making these comparisons. These initial comparisons were made to determine if it is necessary to obtain the ‘average difference’ between the densities from the various tests on cores and the densities from the various gauges in the first place. In other words, we wanted to determine if it is necessary to adjust the average of measurements from nuclear gauges to the average of measurements obtained from cores to obtain a more accurate nuclear gauge density. This is referred to as the nuclear gauge calibration factor by ODOT. Here, it was termed as the ‘average difference’ or ‘adjustment factor’.

Secondly, pooled t-tests were performed to compare the average differences in densities (from cores and nuclear gauges) between successive lifts of HMA pavement and between two different core extraction locations within the nuclear gauge footprint. More specifically, the first, second, and third research questions were postulated to determine if the ‘average differences’ should be determined between core readings and nuclear gauge readings on each lift of hot mix asphalt for various construction scenarios (e.g., same mix design for multiple lifts, different mix designs for individual lifts, and mill and inlay with overlay). While the first research question addresses new pavement construction with same mix design, the second research question addresses new pavement construction with different mix designs, and the third research question addresses mill and inlay projects with an overlay having the same mix design as the inlay. The fourth research question was proposed to determine if there is a significant difference (at 95% confidence level) in average differences between densities obtained from cores and densities obtained from nuclear gauges when comparing a core extracted from the corner of an ‘L-shaped’ footprint to a core extracted from the center of the ‘cross-shaped’ footprint on a particular lift (see Section 3.2).

Additional research was performed by taking density measurements with electromagnetic gauges. The same research questions were also applied to electromagnetic gauges such that electromagnetic gauges were used instead of nuclear gauges to verify the necessity of obtaining average differences (between densities from tests on cores and densities from electromagnetic gauges) on individual lifts.

6.2 Paired t-tests to compare the methods of measurements on a given lift

This section presents the methodology and results of statistical analyses conducted to determine if it is necessary to correlate densities obtained from nuclear gauges with those obtained from cores.

6.2.1 Methodology

Paired t-tests were conducted to determine if the differences in average densities between the two methods of measurements (i.e., tests on cores versus nuclear gauge measurements) were equal to zero for a given lift. Any non-zero differences found would provide evidence to suggest the need for obtaining an 'average difference' between core and nuclear gauge densities on the given lift. The 'average difference' value can be used to adjust density values obtained from a particular nuclear gauge on a particular pavement lift to coincide with the average density obtained from cores on the same lift. Further, the next part of this research (section 6.3) involves answering research questions which were aimed at understanding if this 'average difference' value (between the densities from a particular testing method on cores and a particular gauge) obtained on the given lift can be used on other lifts.

More specifically, for a given set of measurements obtained from cores and nuclear gauges for a particular lift of HMA pavement, the differences between measurements were determined for each specific location within the lift and the average and standard deviation of the differences were calculated following the procedure as specified in WAQTC TM8 (55). The average and standard deviation of the differences were then utilized in the paired t-test to determine if the differences were equal to zero at a specified confidence level. Thus, the hypotheses and test statistic were as follows:

Hypotheses:

- μ_d = difference between the true mean density of the pavement lift when determined by measurements on cores and by nuclear gauge measurements
- Null hypothesis, H_0: $\mu_d = 0$ (i.e., no difference between determinations of the true mean density)
- Alternate hypothesis, H_a: $\mu_d \neq 0$ (i.e., difference between determinations of the true mean density)

Test statistic:

$$t = \frac{\overline{X}_d - 0}{s/\sqrt{n}} = \frac{\overline{X}_d}{s/\sqrt{n}}$$

Where:

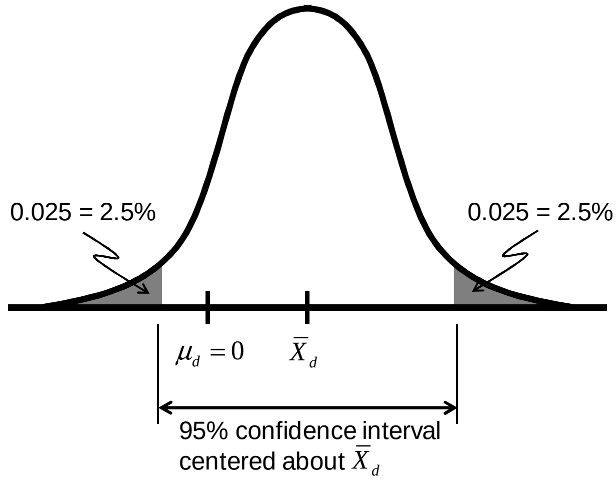
t = Paired t-statistic

$\bar{X}_d$ = Estimate of the difference between the true mean density of the pavement lift when determined by measurements on cores and by nuclear gauge measurements (i.e., mean of differences between core densities and nuclear gauge densities)

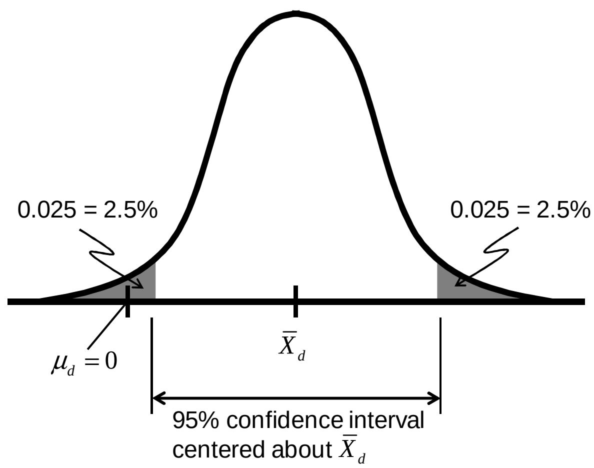
s = Standard deviation of the differences

n = Number of paired observations

The nature of the alternate hypothesis (H_a) implies that a two-tailed rejection region should be used, meaning that the difference could be positive or negative. That is, the two-tailed t-test correctly accounts for the possibility that the mean density determined from nuclear gauge measurements could be less than or greater than the mean density determined from cores. The t-test was conducted using a confidence level of 95% (significance level α = 0.05, or 5%). Figure 6.1a illustrates a result that would not allow rejection of the null hypothesis at a 95% confidence level. That is, the 95% confidence interval centered about the mean difference obtained from test results captures the value of zero indicating there is insufficient evidence to suggest a difference in the true mean pavement lift density, as determined by the two methods being compared, at this confidence level. Conversely, Figure 6.1b illustrates a result that would allow rejection of the null hypothesis (i.e., the value of zero falls outside of the 95% confidence interval).



a) Fail to reject null hypothesis



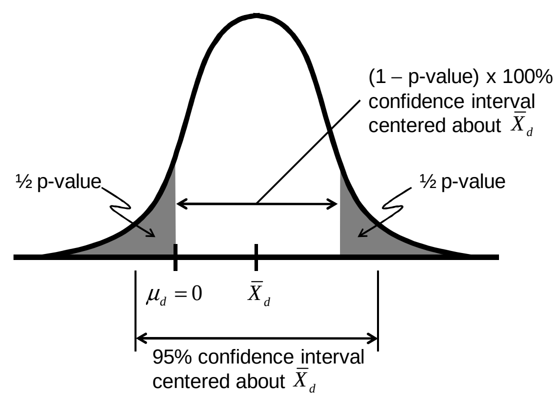
b) Reject the null hypothesis

Figure 6.1 – Illustration of interpretation of paired t-test results.

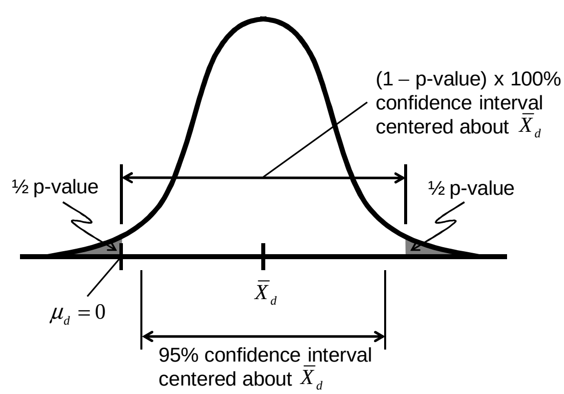
Another result that can be derived from the t-test is referred to as the p-value, which indicates the smallest level of significance (highest level of confidence) at which the null hypothesis (H_0) can be rejected. Figure 6.2 illustrates that the p-value is the sum of the shaded areas under the curve for a two-tailed test. Figure 6.2a illustrates the p-value corresponding to the scenario shown in Figure 6.1a (where the result indicated that the null hypothesis could not be rejected at the 95% confidence level). In comparing Figure 6.2a to Figure 6.1a, it can be seen that the p-value is greater in Figure 6.2a (i.e., the p-value in Figure 6.2a is greater than the significance level in Figure 6.1a).

Conversely, Figure 6.2b illustrates the p-value corresponding to the scenario shown in Figure 6.1b. In this case the p-value in Figure 6.2b is less than the significance level in Figure 6.1b, indicating that the null hypothesis is rejected at the 95% confidence level (5% significance level). In either case, the p-value can be used to determine whether or not the null hypothesis can be rejected as follows:

- p-value $\leq \alpha \rightarrow$ reject H_0 at level α
- p-value $> \alpha \rightarrow$ do not reject H_0 at level α



a) H_0 not rejected



b) H_0 rejected

Figure 6.2 – Interpretation of the p-value obtained from the paired t-test.

In addition, when the null hypothesis is rejected, the p-value indicates the strength of the evidence against H_0, with lower values indicating greater strengths of evidence. For example, a p-value of 0.01, which is much lower than a 0.05 (i.e., 5%) significance level (95% confidence level), indicates that the evidence is very strong against the null hypothesis. A p-value of 0.001 indicates the evidence is extremely strong against H_0.

6.2.2 Results

The tables contained in this section summarize the results of paired t-tests utilized to determine whether or not the null hypothesis of each comparison (i.e., mean of differences between core densities and nuclear gauge densities equal to zero) can be rejected at the 95% confidence level (5% significance level). Cells highlighted in yellow indicate that the null hypothesis can be rejected (i.e., sufficient evidence at the 95% confidence level to indicate that the difference between core densities and nuclear gauge densities is not zero), whereas cells highlighted in green indicate that the null hypothesis cannot be rejected.

The values listed in the tables represent the p-values obtained from the paired t-tests, with values less than 0.05 indicating that the null hypothesis can be rejected at a 5% significance level (95% confidence level). Values much less than 0.05 indicate very strong evidence in support of rejecting H_0. P-values greater than 0.05 indicate that the null hypothesis cannot be rejected at the 5% significance level, and can be used to determine the highest confidence level at which H_0 can be rejected by subtracting the p-value from unity and multiplying by 100% as illustrated in Figure 6.2.

6.2.2.1 OR 18 Project

Recall from Section 3.3.1 that the base course along this project was constructed in three lifts, all having the same mix design. The results presented in this section were derived from density measurements within each of the three lifts, not between lifts. Again, the purpose of the comparisons was to determine if there were significant differences between the various measurement methods and devices.

Table 6.1 presents the results of comparisons of density measurement methods and devices used to determine the density of the first (bottom) lift placed along the OR 18 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift from which the cores were obtained. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the CoreLok method for determining the density of cores and both of the saturated, surface-dry (SSD) methods; but there was significant difference between the densities from the saturated surface-dry method A and those from saturated surface-dry method C. The results also provide strong evidence to suggest significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges. Similarly, there is strong evidence to suggest significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges in all but two of the 18 comparisons.

The results in Table 6.1 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between the densities from various nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings (e.g., the average density derived from 2 readings from one nuclear gauge versus the average derived from 2 readings from another nuclear gauge), the results indicate strong evidence to suggest significant difference between gauges in 4 of the 6 possible comparisons.

Table 6.1 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the bottom lift (18-1) of the project having multiple lifts with the same mix design (OR 18)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD method C	Troxler 3430 - 38804 (4 readings)	Troxler 3430 - 38804 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3440 - 25714 (4 readings)	Troxler 3440 - 25714 (2 readings)
CoreLok	0.0532	0.1547	0.7168	0.0030	0.0014	0.0001	0.0001	0.0285	0.0275
SSD Method A (before APAO)		0.0005	0.0000	0.0089	0.0038	0.0011	0.0021	0.1863	0.3599
SSD Method A (after APAO)			0.0000	0.0026	0.0013	0.0006	0.0011	0.0080	0.0391
SSD method C				0.0002	0.0001	0.0001	0.0001	0.0001	0.0004
Troxler 3430 - 38804 (4 reading)					0.3648	0.2206	0.1965	0.0070	0.0056
Troxler 3430 - 38804 (2 reading)						0.3151	0.3001	0.0014	0.0011
Troxler 3430 - 38806 (4 reading)							0.9345	0.0012	0.0006
Troxler 3430 - 38806 (2 reading)								0.0017	0.0006
Troxler 3440 - 25714 (4 reading)									0.7825

Table 6.2 presents the results of comparisons of density measurement methods and devices used to determine the density of the first (bottom) lift of the composite pavement (comprised of first and second lifts) placed along the OR 18 project. For this set of comparisons, the nuclear gauge was not in direct contact with the lift being evaluated; it was in contact with the second (middle) lift. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the CoreLok method for determining the density of cores and the saturated surface-dry (SSD) method A but not method C. Also, there was significant difference between the densities from saturated, surface-dry method A and those from saturated, surface-dry method C. The results also provide strong evidence to suggest significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges in 4 of 6 comparisons. Similarly, there is strong evidence to suggest significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges in 20 of 24 comparisons. However, the results provide strong evidence to suggest no significant differences between the densities derived from any of the tests on cores and the densities measured from either of the PQI electromagnetic gauges.

The results in Table 6.2 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate strong evidence to suggest significant difference between gauges in all six possible comparisons. The results also indicate that there is strong evidence to suggest significant differences between the densities measured from the nuclear gauges and the densities measured from the electromagnetic gauges.

Table 6.2 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the bottom lift section (18-2 Lift 1) of the composite pavement (comprised of bottom and middle lifts) of the project having multiple lifts with the same mix design (OR 18)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD Method A (APAO)	SSD method C	Troxler 3440 - 30860 (4 readings)	Troxler 3440 - 30860 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3440 - 25714 (4 readings)	Troxler 3440 - 25714 (2 readings)	PQI 301 - density	PQI 300 - density	PQI 300 - moisture
CoreLok	0.1976	0.1803	0.4053	0.0056	0.0041	0.0038	0.0159	0.0270	0.1381	0.1219	0.6727	0.3845	
SSD Method A (before APAO)		0.0006	0.0067	0.0000	0.0051	0.0045	0.0246	0.0405	0.2585	0.2185	0.1878	0.0609	
SSD Method A (after APAO)			0.0063	0.0000	0.0012	0.0009	0.0039	0.0068	0.0429	0.0393	0.7771	0.9345	
SSD Method A (APAO)				0.0000	0.0020	0.0017	0.0067	0.0117	0.0679	0.0606	0.9918	0.6491	
SSD method C					0.0004	0.0003	0.0010	0.0018	0.0105	0.0102	0.1794	0.2002	
Troxler 3440 - 30860 (4 reading)						0.1542	0.0023	0.0057	0.0000	0.0003	0.0000	0.0000	
Troxler 3440 - 30860 (2 reading)							0.0096	0.0183	0.0003	0.0026	0.0000	0.0000	
Troxler 3430 - 38806 (4 reading)								0.1943	0.0003	0.0083	0.0000	0.0001	
Troxler 3430 - 38806 (2 reading)									0.0030	0.0390	0.0001	0.0002	
Troxler 3440 - 25714 (4 reading)										0.0791	0.0016	0.0022	
Troxler 3440 - 25714 (2 reading)											0.0016	0.0020	
PQI 301 - density												0.3057	
PQI 301 - moiture													0.0000

Table 6.3 presents the results of comparisons of density measurement methods and devices used to determine the density of the second (middle) lift of the composite pavement (comprised of first and second lifts) placed along the OR 18 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift being evaluated. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the CoreLok method for determining the density of cores and the saturated, surface-dry (SSD) method A but not method C. Also, there was significant difference between the densities from saturated, surface-dry method A and those from saturated, surface-dry method C. The results also provide strong evidence to suggest significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges. Similarly, there is strong evidence to suggest no significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges in all of the comparisons. However, the results provide strong evidence to suggest no significant differences between the densities derived from any of the tests on cores and the densities measured from either of the PQI electromagnetic gauges.

The results in Table 6.3 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that strong evidence to suggest significant difference between gauges in all of the possible comparisons. The results also indicate that there is strong evidence to suggest significant differences between the densities measured from the nuclear gauges and the densities measured from the electromagnetic gauges.

Table 6.3 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the middle lift section (18-2 Lift 2) of the composite pavement (comprised of bottom and middle lifts) of the project having multiple lifts with the same mix design (OR 18)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD Method A (APAO)	SSD method C	Troxler 3440 - 30860 (4 readings)	Troxler 3440 - 30860 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3440 - 25714 (4 readings)	Troxler 3440 - 25714 (2 readings)	PQI 301 - density	PQI 300 - density	PQI 300 - moisture
CoreLok	0.0611	0.3305	0.2630	0.0598	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2286	0.5012	
SSD Method A (before APAO)		0.2585	0.4077	0.0430	0.0001	0.0001	0.0003	0.0013	0.0081	0.0053	0.5002	0.1908	
SSD Method A (after APAO)			0.4370	0.0004	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.7442	0.7005	
SSD Method A (APAO)				0.0027	0.0000	0.0000	0.0001	0.0002	0.0002	0.0001	0.9181	0.6011	
SSD method C					0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1127	0.3548	
Troxler 3440 - 30860 (4 reading)						0.1542	0.0023	0.0057	0.0000	0.0003	0.0000	0.0000	
Troxler 3440 - 30860 (2 reading)							0.0096	0.0183	0.0003	0.0026	0.0000	0.0000	
Troxler 3430 - 38806 (4 reading)								0.1943	0.0003	0.0083	0.0000	0.0001	
Troxler 3430 - 38806 (2 reading)									0.0030	0.0390	0.0001	0.0002	
Troxler 3440 - 25714 (4 reading)										0.0791	0.0016	0.0022	
Troxler 3440 - 25714 (2 reading)											0.0016	0.0020	
PQI 301 - density												0.3057	
PQI 301 - moiture													0.0000

Table 6.4 presents the results of comparisons of density measurement methods and devices used to determine the density of the composite pavement (comprised of first and second lifts) placed along the OR 18 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift being evaluated. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the CoreLok method for determining the density of cores and the saturated, surface-dry (SSD) method A. The results also provide strong evidence to suggest significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges. Similarly, there is strong evidence to suggest significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges. However, the results provide strong evidence to suggest no significant differences between the densities derived from any of the tests on cores and the densities measured from either of the PQI electromagnetic gauges.

The results in Table 6.4 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that there is strong evidence to suggest significant difference between gauges in 4 of the 6 possible comparisons. The results also indicate that there is strong evidence to suggest significant differences between the densities measured from the nuclear gauges and the densities measured from the electromagnetic gauges. Finally, the densities from the PQI 300 and PQI 301 electromagnetic gauges are significantly different from each other at 95% confidence level.

Table 6.4 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the composite pavement (18-2 uncut; i.e., comprised of bottom and middle lifts) of the project having multiple lifts with the same mix design (OR 18)

	SSD Method A (before APAO)	Troxler 3440 - 30860 (4 readings)	Troxler 3440 - 30860 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3440 - 25714 (4 readings)	Troxler 3440 - 25714 (2 readings)	PQI 301 - density	PQI 300 - density	PQI 300 - moisture
CoreLok	0.3333	0.0000	0.0000	0.0000	0.0001	0.0036	0.0045	0.6834	0.3500	
SSD Method A (before APAO)		0.0000	0.0000	0.0000	0.0002	0.0068	0.0068	0.3205	0.0723	
Troxler 3440 - 30860 (4 reading)			0.1542	0.0023	0.0057	0.0000	0.0003	0.0000	0.0000	
Troxler 3440 - 30860 (2 reading)				0.0096	0.0183	0.0003	0.0026	0.0000	0.0000	
Troxler 3430 - 38806 (4 reading)					0.1943	0.0003	0.0083	0.0000	0.0001	
Troxler 3430 - 38806 (2 reading)						0.0030	0.0390	0.0001	0.0002	
Troxler 3440 - 25714 (4 reading)							0.0791	0.0016	0.0022	
Troxler 3440 - 25714 (2 reading)								0.0016	0.0020	
PQI 301 - density									0.3057	
PQI 301 - moiture										0.0000

Table 6.5 presents the results of comparisons of density measurement methods and devices used to determine the density of the first (bottom) lift of the composite pavement (comprised of first, second, and third lifts) placed along the OR 18 project. For this set of comparisons, the nuclear gauge was not in direct contact with the paving lift being evaluated; it was in contact with the third (top) lift. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the CoreLok method for determining the density of cores and the saturated surface-dry (SSD) method A and not method C. Also, the densities from the saturated, surface-dry method A was significantly different from those of method C. The results also provide strong evidence to suggest no significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges. Similarly, there is strong evidence to suggest no significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges in all but 6 of the 24 comparisons. However, the results provide strong evidence to suggest no significant differences between the densities derived from any of the tests on cores and the densities measured from the PQI 301 electromagnetic gauge (in un-sanded condition) and the PQI 300 electromagnetic gauge (in sanded condition).

The results in Table 6.5 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that strong evidence to suggest significant difference between gauges in 5 of the 6 possible comparisons. The results also indicate that there is strong evidence to suggest significant differences between the densities measured from the nuclear gauges and the densities measured from the electromagnetic gauges.

Table 6.5 also provides information regarding the differences between densities measured using PQI 300 and the PQI 301 electromagnetic gauge, the influence of sand on the measurements from the two electromagnetic gauges, and the difference between the moisture measurements between the two gauges. The density and moisture measurements from the PQI 300 electromagnetic gauge were significantly different from those of the PQI 301 electromagnetic gauge at 95% confidence level. Another observation is that the density and moisture measurements made using the PQI 301 in the sanded condition were not significantly different from the density and moisture measurements made using the PQI 301 in the un-sanded condition. However, for the PQI 300, the density and moisture measurements differed significantly between sanded condition and un-sanded condition.

Table 6.5 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the bottom lift (18-3 Lift 1) of the composite pavement (comprised of bottom , middle, and top lifts) of the project having multiple lifts with the same mix design (OR 18)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD Method A (APAO)	SSD method C	Troxler 3440 - 30860 (4 readings)	Troxler 3440 - 30860 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3440 Plus - 39525 (4 readings)	Troxler 3440 Plus - 39525 (2 readings)	PQI 301 (sand) - density	PQI 301 (no sand) - density	PQI 301 (no sand) - moisture	PQI 300 (sand) - density	PQI 300 (sand) - moiture	PQI 300 (no sand) - density	PQI 300 (no sand) - moisture
CoreLok	0.1403	0.0455	0.2139	0.0054	0.5827	0.5662	0.3407	0.4967	0.0697	0.0766	0.0388	0.1129		0.0607		0.0408	
SSD Method A (before APAO)		0.0109	0.0803	0.0024	0.3119	0.3147	0.1584	0.2612	0.0756	0.0842	0.0119	0.0413		0.0204		0.0127	
SSD Method A (after APAO)			0.1221	0.0010	0.2062	0.2177	0.0617	0.1831	0.0026	0.0043	0.0018	0.0055		0.0032		0.0014	
SSD Method A (APAO)				0.0002	0.1671	0.1719	0.0627	0.1492	0.0127	0.0181	0.0019	0.0067		0.0034		0.0020	
SSD method C					0.7659	0.8902	0.6559	0.9763	0.0008	0.0013	0.0184	0.0496		0.0332		0.0102	
Troxler 3440 - 30860 (4 reading)						0.7270	0.0386	0.5429	0.0000	0.0000	0.0010	0.0022		0.0023		0.0003	
Troxler 3440 - 30860 (2 reading)							0.1514	0.7075	0.0000	0.0000	0.0018	0.0048		0.0041		0.0011	
Troxler 3430 - 38806 (4 reading)								0.2169	0.0000	0.0000	0.0102	0.0099		0.0203		0.0018	
Troxler 3430 - 38806 (2 reading)									0.0001	0.0001	0.0054	0.0018		0.0101		0.0006	
Troxler 3440 Plus - 39525 (4 reading)										0.6339	0.0000	0.0000		0.0000		0.0000	
Troxler 3440 Plus - 39525 (2 reading)											0.0000	0.0000		0.0000		0.0000	
PQI 301 (sand) - density												0.2163		0.0018		0.0003	
PQI 301 (sand) - moiture													0.0558		0.0000		0.0000
PQI 301 (no sand) - density														0.0365		0.0010	
PQI 301 (no sand) - moisture															0.0000		0.0000
PQI 300 (sand) - density																0.0001	
PQI 300 (sand) - moiture																	0.0101

Table 6.6 presents the results of comparisons of density measurement methods and devices used to determine the density of the second (middle) lift of the composite pavement (comprised of first, second, and third lifts) placed along the OR 18 project. For this set of comparisons, the nuclear gauge was not in direct contact with the paving lift from which the core slices were obtained; it was in contact with the third (top) lift. The results indicate that there was strong evidence to suggest significant differences (at a 95% confidence level) between the CoreLok method for determining the density of cores and both of the saturated, surface-dry (SSD) methods. The results also provide strong evidence to suggest significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges. Similarly, there is strong evidence to suggest significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges in 16 of the 24 comparisons. However, the results provide strong evidence to suggest no significant differences between the densities derived from any of the tests on cores and the densities measured using the electromagnetic gauges.

The results in Table 6.6 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that there was strong evidence to suggest significant difference between gauges in 5 of the 6 possible comparisons. The results also indicate that there is strong evidence to suggest significant differences between the densities measured from the nuclear gauges and the densities measured from the electromagnetic gauges.

Table 6.6 also provides information regarding the differences between densities measured using PQI 300 and the PQI 301 electromagnetic gauge, the influence of sand on the measurements from the two electromagnetic gauges, and the difference between the moisture measurements between the two gauges. The densities and the moisture measurements from the PQI 300 electromagnetic gauge were significantly different from those of the PQI 301 electromagnetic gauge at 95% confidence level. Another observation is that the density and moisture measurements made using the PQI 301 in the sanded condition were not significantly different from the density and moisture measurements made using the same gauge in the un-sanded condition. However, for the PQI 300, the density and moisture measurements in sanded condition differed significantly from those in un-sanded condition.

Table 6.6 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the middle lift (18-3 Lift 2) of the composite pavement (comprised of bottom , middle, and top lifts) of the project having multiple lifts with the same mix design (OR 18)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD Method A (APAO)	SSD method C	Troxler 3440 - 30860 (4 readings)	Troxler 3440 - 30860 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3440 Plus - 39525 (4 readings)	Troxler 3440 Plus - 39525 (2 readings)	PQI 301 (sand) - density	PQI 301 (no sand) - density	PQI 301 (no sand) - moisture	PQI 300 (sand) - density	PQI 300 (sand) - moiture	PQI 300 (no sand) - density	PQI 300 (no sand) - moisture
CoreLok	0.0000	0.0062	0.0053	0.0427	0.0124	0.0109	0.0316	0.0260	0.0002	0.0002	0.4851	0.8302		0.3120		0.5910	
SSD Method A (before APAO)		0.0008	0.0118	0.0000	0.0988	0.0956	0.2348	0.1755	0.0007	0.0006	0.4191	0.3161		0.6365		0.0902	
SSD Method A (after APAO)			0.0818	0.0000	0.0365	0.0333	0.0868	0.0693	0.0005	0.0004	0.9112	0.7914		0.6771		0.3284	
SSD Method A (APAO)				0.0000	0.0469	0.0437	0.1161	0.0892	0.0005	0.0004	0.8801	0.6042		0.8608		0.2076	
SSD method C					0.0075	0.0063	0.0187	0.0156	0.0001	0.0001	0.2595	0.5515		0.1639		0.8871	
Troxler 3440 - 30860 (4 reading)						0.7270	0.0386	0.5429	0.0000	0.0000	0.0010	0.0022		0.0023		0.0003	
Troxler 3440 - 30860 (2 reading)							0.1514	0.7075	0.0000	0.0000	0.0018	0.0048		0.0041		0.0011	
Troxler 3430 - 38806 (4 reading)								0.2169	0.0000	0.0000	0.0102	0.0099		0.0203		0.0018	
Troxler 3430 - 38806 (2 reading)									0.0001	0.0001	0.0054	0.0018		0.0101		0.0006	
Troxler 3440 Plus - 39525 (4 reading)										0.6339	0.0000	0.0000		0.0000		0.0000	
Troxler 3440 Plus - 39525 (2 reading)											0.0000	0.0000		0.0000		0.0000	
PQI 301 (sand) - density												0.2163		0.0018		0.0003	
PQI 301 (sand) - moiture													0.0558		0.0000		0.0000
PQI 301 (no sand) - density														0.0365		0.0010	
PQI 301 (no sand) - moisture															0.0000		0.0000
PQI 300 (sand) - density																0.0001	
PQI 300 (sand) - moiture																	0.0101

Table 6.7 presents the results of comparisons of density measurement methods and devices used to determine the density of the third (top) lift of the composite pavement (comprised of first, second, and third lifts) placed along the OR 18 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift from which the core slices were obtained. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the CoreLok method for determining the density of cores and both of the saturated, surface-dry (SSD) methods; however, there was significant difference between the densities from the saturated, surface-dry method A and those from method C. The results also provide strong evidence to suggest significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges. Similarly, there is strong evidence to suggest no significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges. However, the results provide strong evidence to suggest no significant differences between the densities derived from any of the tests on cores and the densities measured using the electromagnetic gauges.

The results in Table 6.7 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between nuclear gauges. For the former case, the results provide strong evidence to suggest there existed significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that strong evidence to suggest no significant difference between gauges in 5 of the 6 possible comparisons. The results also indicate that there is strong evidence to suggest significant differences between the densities measured from the nuclear gauges and the densities measured from the electromagnetic gauges.

Table 6.7 also provides information regarding the differences between densities measured using PQI 300 and the PQI 301 electromagnetic gauge, the influence of sand on the measurements from the two electromagnetic gauges, and the difference between the moisture measurements between the two gauges. The densities and the moisture measurements from the PQI 300 electromagnetic gauge were significantly different from those of the PQI 301 electromagnetic gauge at 95% confidence level. Another observation is that the density and moisture measurements made using the PQI 301 in the sanded condition were not significantly different from the density and moisture measurements made using the same gauge in the un-sanded condition. However, for the PQI 300 electromagnetic gauge, the density and moisture measurements in sanded condition differed significantly from those in un-sanded condition.

Table 6.7 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the top lift (18-3 Lift 3) of the composite pavement (comprised of bottom , middle, and top lifts) of the project having multiple lifts with the same mix design (OR 18)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD Method A (APAO)	SSD method C	Troxler 3440 - 30860 (4 readings)	Troxler 3440 - 30860 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3440 Plus - 39525 (4 readings)	Troxler 3440 Plus - 39525 (2 readings)	PQI 301 (sand) - density	PQI 301 (no sand) - density	PQI 301 (no sand) - moisture	PQI 300 (sand) - density	PQI 300 (sand) - moiture	PQI 300 (no sand) - density	PQI 300 (no sand) - moisture
CoreLok	0.0005	0.5046	0.5586	0.0631	0.0000	0.0002	0.0001	0.0000	0.0000	0.0000	0.5777	0.0501		0.2583		0.4411	
SSD Method A (before APAO)		0.0009	0.0213	0.0000	0.0000	0.0007	0.0007	0.0002	0.0000	0.0000	0.1906	0.1890		0.4805		0.0012	
SSD Method A (after APAO)			0.7445	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.7583	0.0898		0.3693		0.1740	
SSD Method A (APAO)				0.0036	0.0000	0.0002	0.0001	0.0000	0.0000	0.0000	0.8390	0.3606		0.4625		0.2365	
SSD method C					0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2193	0.0039		0.0800		0.8107	
Troxler 3440 - 30860 (4 reading)						0.7270	0.0386	0.5429	0.0000	0.0000	0.0010	0.0022		0.0023		0.0003	
Troxler 3440 - 30860 (2 reading)							0.1514	0.7075	0.0000	0.0000	0.0018	0.0048		0.0041		0.0011	
Troxler 3430 - 38806 (4 reading)								0.2169	0.0000	0.0000	0.0102	0.0099		0.0203		0.0018	
Troxler 3430 - 38806 (2 reading)									0.0001	0.0001	0.0054	0.0018		0.0101		0.0006	
Troxler 3440 Plus - 39525 (4 reading)										0.6339	0.0000	0.0000		0.0000		0.0000	
Troxler 3440 Plus - 39525 (2 reading)											0.0000	0.0000		0.0000		0.0000	
PQI 301 (sand) - density												0.2163		0.0018		0.0003	
PQI 301 (sand) - moiture													0.0558		0.0000		0.0000
PQI 301 (no sand) - density														0.0365		0.0010	
PQI 301 (no sand) - moisture															0.0000		0.0000
PQI 300 (sand) - density																0.0001	
PQI 300 (sand) - moiture																	0.0101

Table 6.8 presents the results of comparisons of density measurement methods and devices used to determine the density of the composite pavement (comprised of first, second, and third lifts) placed along the OR 18 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift from which the core slices were obtained. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the CoreLok method for determining the density of cores and the saturated surface-dry (SSD) method A. The results also provide strong evidence to suggest no significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges in 4 of 6 possible comparisons. However, there is strong evidence to suggest significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges. Also, the results provide strong evidence to suggest no significant differences between the densities derived from any of the tests on cores and the densities measured using the electromagnetic gauges.

The results in Table 6.8 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between nuclear gauges. For the former case, the results provide strong evidence to suggest there existed significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that strong evidence to suggest no significant difference between gauges in 5 of the 6 possible comparisons. The results also indicate that there is strong evidence to suggest significant differences between the densities measured from the nuclear gauges and the densities measured from the electromagnetic gauges.

Table 6.8 also provides information regarding the differences between densities measured using PQI 300 and the PQI 301 electromagnetic gauge, the influence of sand on the measurements from the two electromagnetic gauges, and the difference between the moisture measurements between the two gauges. The densities and the moisture measurements from the PQI 300 electromagnetic gauge were significantly different from those of the PQI 301 electromagnetic gauge at 95% confidence level. Another observation is that the density and moisture measurements made using the PQI 301 in the sanded condition were not significantly different from the density and moisture measurements made using the same gauge in the un-sanded condition. However, for the PQI 300 electromagnetic gauge, the density and moisture measurements in sanded condition differed significantly from those in un-sanded condition.

Table 6.8 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the composite pavement (18-3 uncut; i.e., comprised of bottom, middle, and top lifts) of the project having multiple lifts with the same mix design (OR 18)

	CoreLok	SSD Method A (before APAO)	Troxler 3440 - 30860 (4 readings)	Troxler 3440 - 30860 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3440 Plus - 39525 (4 readings)	Troxler 3440 Plus - 39525 (2 readings)	PQI 301 (sand) - density	PQI 301 (no sand) - density	PQI 301 (no sand) - moisture	PQI 300 (sand) - density	PQI 300 (sand) - moiture	PQI 300 (no sand) - density	PQI 300 (no sand) - moisture
CoreLok		0.1310	0.2538	0.3337	0.7139	0.4342	0.0011	0.0015	0.0938	0.2040		0.1626		0.0548	
SSD Method A (before APAO)			0.0229	0.0411	0.1175	0.0703	0.0001	0.0001	0.4066	0.6011		0.6516		0.1522	
Troxler 3440 - 30860 (4 reading)				0.7270	0.0386	0.5429	0.0000	0.0000	0.0010	0.0022		0.0023		0.0003	
Troxler 3440 - 30860 (2 reading)					0.1514	0.7075	0.0000	0.0000	0.0018	0.0048		0.0041		0.0011	
Troxler 3430 - 38806 (4 reading)						0.2169	0.0000	0.0000	0.0102	0.0099		0.0203		0.0018	
Troxler 3430 - 38806 (2 reading)							0.0001	0.0001	0.0054	0.0018		0.0101		0.0006	
Troxler 3440 Plus - 39525 (4 reading)								0.6339	0.0000	0.0000		0.0000		0.0000	
Troxler 3440 Plus - 39525 (2 reading)									0.0000	0.0000		0.0000		0.0000	
PQI 301 (sand) - density										0.2163		0.0018		0.0003	
PQI 301 (sand) - moiture											0.0558		0.0000		0.0000
PQI 301 (no sand) - density												0.0365		0.0010	
PQI 301 (no sand) - moisture													0.0000		0.0000
PQI 300 (sand) - density														0.0001	
PQI 300 (sand) - moiture															0.0101

Based on the results summarized in tables 6.1 to 6.8, a number of observations are made as follows:

- Tables 6.1, 6.3, and 6.7 provide results wherein the cores (or slices of cores) were obtained from the same lift on which the nuclear gauge readings were taken. The methodology used to generate the results in these particular tables follows the methodology provided in WAQTC TM8 for determining nuclear gauge calibration factors. Hence, in determining whether or not development of nuclear gauge calibration factors are recommended as a mandatory requirement for dense-graded HMA paving projects, inferences drawn from the above-mentioned tables carry greater weight than those drawn from tables 6.2, 6.4, 6.5, and 6.8.

- Overall, a majority of the comparisons between the densities from the various tests on cores on 18-2 Lift 1 and 18-3 Lifts 1 to 3 (tables 6.2, 6.5, 6.6, and 6.7) indicate that they were significantly different from each other at 95% confidence level. However, a closer observation suggests that, on all the lifts (except for 18-3 Lift 2; i.e., table 6.6), a majority of the densities from the CoreLok were not significantly different from the densities obtained from the saturated, surface-dry method A at a 95% confidence level.

- Except for 18-3 Lift 1 (table 6.5), a majority of the comparisons indicate that the densities from the various tests on cores were significantly different from those of nuclear gauge measurements at 95% confidence level.

- However, a majority of the comparisons (all tables except tables 6.1 and 6.5) indicate that the densities from the various tests on cores were not significantly different from those of electromagnetic gauge measurements at a 95% confidence level.

- A majority of the comparisons (in all 8 tables) indicate that the densities from the various measurements using nuclear gauges were significantly different from each other at a 95% confidence level.

- A majority of the comparisons (in all tables) also indicate that densities from the various nuclear gauges were significantly different from those of electromagnetic gauges at a 95% confidence level.

- Also, all of the comparisons on all of the lifts indicate that the average density of the lift obtained using the average of 4 nuclear gauge measurements at every location were not significantly different from the average density of the lift obtained using the average of 2 nuclear gauge measurements at every location (as seen in all the tables).

- The densities obtained from the PQI 301 electromagnetic gauge and the PQI 300 electromagnetic gauge were significantly different from each other at 95% confidence level on the third (or top) lift (as seen in tables 6.5 to 6.8) but not on the second or middle lift (as seen in tables 6.2 to 6.4). The electromagnetic gauges were not used when the first (bottom) lift alone was paved.

- The moisture measurements obtained from the PQI 301 and the PQI 300 electromagnetic gauges were significantly different from each other at 95% confidence level on both lifts where these gauges were used (as seen in tables 6.2 to 6.8).

- On the third lift, where measurements with the electromagnetic gauges were taken in the sanded and un-sanded condition, the PQI 301 density and moisture measurements taken in the sanded condition were not significantly different (at 95% confidence level) from those taken in the un-sanded condition (tables 6.5 to 6.8). However, for the PQI 300, the density and moisture measurements taken in the sanded condition were significantly different (at 95% confidence level) from those taken in the un-sanded condition (tables 6.5 to 6.8).

6.2.2.2 I-5 Project

As indicated in section 3.3.2, the first lift of the base course along this project had a different mix design from the remaining lifts. For research purposes, measurements and samples were taken only from the first and second lift to study the influence of different mix designs on the density measurements. The results presented in this section were derived from density measurements within each of the lifts, not between lifts. Again, the purpose of the comparisons was to determine if there were significant differences between the various measurement methods and devices.

Table 6.9 presents the results of comparisons of density measurement methods and devices used to determine the density of the first (bottom) lift placed along the I-5 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift from which the cores were obtained. The results indicate that there was strong evidence to suggest significant differences (at a 95% confidence level) between the various methods for determining the density of cores. The results also provide strong evidence to suggest significant differences between densities derived from cores using the CoreLok device and densities measured using nuclear gauges. Similarly, there is strong evidence to suggest significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges in all but three of the 12 possible comparisons.

The results in Table 6.9 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between the densities from various nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from 2 readings versus 4 readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that there is strong evidence to suggest significant difference between gauges in 4 of the 6 possible comparisons.

Table 6.9 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the bottom lift (5-1) of the project having two lifts with different mix designs (I-5)

	SSD (A)	CoreLok	SSD (C)	Troxler 3430-38996 (4)	Troxler 3430-38996 (2)	Troxler 3430-38806 (4)	Troxler 3430-38806 (2)	CPN MC3 Portaprobe-7443 (4)	CPN MC3 Portaprobe-7443 (2)
SSD (A)		0.0000	0.0025	0.0429	0.1380	0.0062	0.0025	0.0025	0.0060
CoreLok			0.0005	0.0000	0.0001	0.0010	0.0004	0.0000	0.0000
SSD (C)				0.0041	0.0118	0.9522	0.9015	0.0005	0.0012
Troxler 3430-38996 (4)					0.3563	0.0007	0.0011	0.1285	0.1686
Troxler 3430-38996 (2)						0.0067	0.0048	0.0953	0.1193
Troxler 3430-38806 (4)							0.9544	0.0008	0.0022
Troxler 3430-38806 (2)								0.0002	0.0004
CPN MC3 Portaprobe-7443 (4)									0.5246

Table 6.10 presents the results of comparisons of density measurement methods and devices used to determine the density of the second (top) lift placed along the I-5 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift from which the core slices were obtained. The results indicate that there was strong evidence to suggest significant differences (at a 95% confidence level) between the various methods for determining the density of cores. The results also provide strong evidence to suggest significant differences (at 95% confidence level) between densities derived from cores using the CoreLok device and densities measured using 2 of the 4 nuclear gauges. However, there was strong evidence to suggest no significant differences between densities obtained from cores using the SSD methods and densities measured using nuclear gauges in all but 4 of the 16 possible comparisons.

The results in Table 6.10 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between the densities from various nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from 2 readings versus 4 readings for all four nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that there is strong evidence to suggest significant difference between gauges in four of the 12 possible comparisons.

Table 6.10 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the top lift (5-2) of the project having two lifts with different mix designs (I-5)

	SSD (A)	CoreLok	SSD (C)	Troxler 3430-38996 (4)	Troxler 3430-38996 (2)	Troxler 3430-38806 (4)	Troxler 3430-38806 (2)	CPN MC3 Portaprobe-7443 (4)	CPN MC3 Portaprobe-7443 (2)	Humboldt 5001C-1344 (4)	Humboldt 5001C-1344 (2)
SSD (A)		0.0062	0.0000	0.5995	0.4824	0.0002	0.0005	0.7225	0.2506	0.4458	0.4277
CoreLok			0.0266	0.0067	0.0047	0.8461	0.7257	0.0073	0.0358	0.0694	0.0981
SSD (C)				0.1094	0.0678	0.0014	0.0032	0.3869	0.9855	0.9396	0.9176
Troxler 3430-38996 (4)					0.6628	0.0000	0.0000	0.1084	0.0358	0.0249	0.0883
Troxler 3430-38996 (2)						0.0000	0.0000	0.0872	0.0444	0.0777	0.1259
Troxler 3430-38806 (4)							0.1175	0.0000	0.0011	0.0003	0.0049
Troxler 3430-38806 (2)								0.0001	0.0024	0.0005	0.0078
CPN MC3 Portaprobe-7443 (4)									0.1632	0.5604	0.5455
CPN MC3 Portaprobe-7443 (2)										0.7030	0.9376
Humboldt 5001C-1344 (4)											0.6475

Based on the results summarized in tables 6.9 and 6.10, a number of observations can be made.

- Tables 6.9 and 6.10 provide results wherein the cores (or slices of cores) were obtained from the same lift on which the nuclear gauge readings were taken. The methodology used to generate the results in these particular tables follows the methodology provided in WAQTC TM8 for determining nuclear gauge correlation (adjustment) factors.

- A majority of the comparisons in tables 6.9 and 6.10 indicate that the densities from the various tests on cores were significantly different from each other (at 95% confidence level).

- Also, a majority of the comparisons on the bottom lift (as seen in table 6.9) indicate that the densities from the various tests on cores were significantly different from those of nuclear gauge measurements (at 95% confidence level). However, a majority of the comparisons on the top lift (as seen in table 6.10) indicate that there was no significant difference (at 95% confidence level) between densities from tests on cores and densities from various nuclear gauges.

- A majority of the comparisons in tables 6.9 and 6.10 indicate that the densities from the various nuclear gauges were significantly different from each other (at 95% confidence level).

- In all of the comparisons on both the lifts (as seen in tables 6.9 and 6.10), the average density of the lift obtained using the average of four nuclear gauge measurements at every location was not significantly different from the average density of the lift obtained using the average of two nuclear gauge measurements.

6.2.2.3 OR 140 Project

As indicated in section 3.3.3, this project was a repaving project wherein, the pavement was milled and replaced with an inlay followed by placement of an overlay (both having same mix design). The density measurements and samples were taken from both lifts. The results presented in this section were derived from density measurements within both lifts, not between lifts. Again, the purpose of the comparisons was to determine if there were significant differences between the various measurement methods and devices.

Table 6.11 presents the results of comparisons of density measurement methods and devices used to determine the density of the first (bottom or inlay) lift placed along the OR 140 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift from which the cores were obtained. The results indicate that there was strong evidence to suggest significant differences (at a 95% confidence level) between the various methods for determining the density of cores. The results also provide strong evidence to suggest no significant difference between densities derived from cores using the CoreLok device and densities measured using nuclear gauges. Similarly, there was strong evidence to suggest no significant differences between densities obtained from cores using the SSD method C and densities measured using nuclear gauges in a majority of the comparisons. However, there was strong evidence to suggest significant differences (at 95% confidence level) between densities obtained from cores using the SSD method A and densities measured using nuclear gauges in a majority of the comparisons. Also, the results provide strong evidence to suggest significant differences between the densities derived from the various tests on cores and the densities measured using the electromagnetic gauges.

The results in Table 6.11 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between the densities from various nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from two readings versus four readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that there is strong evidence to suggest significant differences between gauges in five of the six possible comparisons. The results also indicate that there is strong evidence to suggest no significant differences between the densities measured from the nuclear gauges and the densities measured from the PQI 300 electromagnetic gauge. Conversely there was significant difference between the densities from the nuclear gauge and the densities from the PQI 301 electromagnetic gauge in a majority of the comparisons.

The average density from the PQI 300 electromagnetic gauge was significantly different from that obtained from the PQI 301 electromagnetic gauge.

Table 6.11 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the inlay lift (140-1) of the mill and fill (inlay) plus overlay project having same mix design (OR 140)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD method C	Troxler 3430 - 38996 (4 readings)	Troxler 3430 - 38996 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3430 - 35601 (4 readings)	Troxler 3430 - 35601 (2 readings)	PQI 301 - density	PQI 300 - density
CoreLok	0.0000	0.0002	0.3501	0.0002	0.0009	0.2057	0.6208	0.9034	0.8226	0.0088	0.0286
SSD Method A (before APAO)		0.0013	0.0000	0.0000	0.0000	0.0114	0.0113	0.0006	0.0006	0.2841	0.0000
SSD Method A (after APAO)			0.0000	0.0000	0.0001	0.1938	0.0992	0.0161	0.0186	0.3585	0.0001
SSD method C				0.0003	0.0011	0.4425	0.9839	0.7792	0.8533	0.0096	0.0165
Troxler 3430 - 38996 (4 reading)					0.7640	0.0000	0.0001	0.0000	0.0000	0.0004	0.3231
Troxler 3430 - 38996 (2 reading)						0.0000	0.0000	0.0000	0.0000	0.0010	0.3989
Troxler 3430 - 38806 (4 reading)							0.2140	0.0408	0.0387	0.1871	0.0395
Troxler 3430 - 38806 (2 reading)								0.5311	0.5796	0.1171	0.1154
Troxler 3430 - 35601 (4 reading)									0.7900	0.0408	0.0871
Troxler 3430 - 35601 (2 reading)										0.0443	0.0762
PQI 301 - density											0.0001

Table 6.12 presents the results of comparisons of density measurement methods and devices used to determine the density of the first (bottom or inlay) lift of the composite pavement comprised of the first and second lift (inlay and overlay lift) placed along the OR 140 project. For this set of comparisons, the nuclear gauge was not in direct contact with the paving lift from which the core slices were obtained, it was in contact with the second or overlay lift. The results indicate that there was strong evidence to suggest significant differences (at a 95% confidence level) between the various methods for determining the density of cores. The results also provide strong evidence to suggest no significant differences (at 95% confidence level) between densities derived from the various tests on cores (i.e., the CoreLok and SSD methods) and the densities measured using nuclear gauges. Also, the results provide strong evidence to suggest no significant differences between the densities derived from the CoreLok and SSD method C tests on cores and the densities measured using the electromagnetic gauges. However, there were significant differences between the densities derived from the SSD method A and the densities measured using the electromagnetic gauges.

The results in Table 6.12 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between the densities from various nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from two readings versus four readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that there is strong evidence to suggest no significant difference between gauges in four of the six possible comparisons. The results also indicate that there was strong evidence to suggest no significant differences between the densities measured from the nuclear gauges and the densities measured from the PQI 301 electromagnetic gauge; whereas, there were significant differences between the densities from the nuclear gauge and the densities from the PQI 300 electromagnetic gauge in a majority of the comparisons.

Table 6.12 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the inlay lift (140-2 Lift 1) of the composite pavement (comprised of inlay and overlay) of the mill and fill (inlay) plus overlay project having same mix design (OR 140)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD Method A (APAO)	SSD method C	Troxler 3430 - 38996 (4 readings)	Troxler 3430 - 38996 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3430 - 35601 (4 readings)	Troxler 3430 - 35601 (2 readings)	PQI 301 - density	PQI 300 - density
CoreLok	0.0001	0.0037	0.0013	0.8780	0.5120	0.5105	0.2612	0.2883	0.3294	0.3658	0.8973	0.2317
SSD Method A (before APAO)		0.0000	0.0041	0.0000	0.4538	0.4466	0.8327	0.8180	0.8028	0.8048	0.0047	0.0016
SSD Method A (after APAO)			0.1298	0.0000	0.8653	0.8604	0.7059	0.7325	0.7859	0.8134	0.0564	0.0105
SSD Method A (APAO)				0.0000	0.7598	0.7552	0.8242	0.8483	0.8918	0.9118	0.0476	0.0101
SSD method C					0.4572	0.4545	0.2102	0.2310	0.2848	0.3243	0.9290	0.1718
Troxler 3430 - 38996 (4 reading)						0.9598	0.0007	0.0112	0.0502	0.2213	0.2215	0.0078
Troxler 3430 - 38996 (2 reading)							0.0057	0.0215	0.0810	0.2474	0.2136	0.0065
Troxler 3430 - 38806 (4 reading)								0.8259	0.7742	0.8225	0.0488	0.0010
Troxler 3430 - 38806 (2 reading)									0.8661	0.8736	0.0572	0.0012
Troxler 3430 - 35601 (4 reading)										0.9143	0.1140	0.0063
Troxler 3430 - 35601 (2 reading)											0.1562	0.0145
PQI 301 - density												0.0007

Table 6.13 presents the results of comparisons of density measurement methods and devices used to determine the density of the second (top or overlay) lift section of the composite pavement comprised of the first and second lift (inlay and overlay lift) placed along the OR 140 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift from which the core slices were obtained. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the densities derived from the tests using the CoreLok device and the densities derived from the SSD methods. However, the results also indicate that there was strong evidence to suggest significant differences between the densities obtained from the two SSD methods. The results also provide strong evidence to suggest no significant differences (at 95% confidence level) between densities derived from the CoreLok device and the SSD method C and the densities measured using nuclear gauges. However, there was strong evidence to suggest significant differences between the densities derived from the SSD method A and the densities measured using the nuclear gauges. Also, the results provide strong evidence to suggest no significant differences between the densities derived from the CoreLok and the densities measured using the electromagnetic gauges. However, there were significant differences between the densities derived from the SSD methods and the densities measured using the electromagnetic gauges.

The results in Table 6.13 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between the densities from various nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from two readings versus four readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that there is strong evidence to suggest no significant difference between gauges in four of the six possible comparisons. The results also indicate that there was strong evidence to suggest no significant differences between the densities measured from the nuclear gauges and the densities measured from the PQI 301 electromagnetic gauge; whereas, there were significant differences between the densities from the nuclear gauge and the densities from the PQI 300 electromagnetic gauge in a majority of the comparisons.

The average density from the PQI 300 electromagnetic gauge was significantly different from that obtained from the PQI 301 electromagnetic gauge.

Table 6.13 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the overlay lift (140-2 Lift 2) of the composite pavement (comprised of inlay and overlay) of the mill and fill (inlay) plus overlay project having same mix design (OR 140)

	SSD Method A (before APAO)	SSD Method A (after APAO)	SSD Method A (APAO)	SSD method C	Troxler 3430 - 38996 (4 readings)	Troxler 3430 - 38996 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3430 - 35601 (4 readings)	Troxler 3430 - 35601 (2 readings)	PQI 301 - density	PQI 300 - density
CoreLok	0.0582	0.1433	0.1538	0.3276	0.5051	0.4973	0.2838	0.2959	0.3449	0.3730	0.9184	0.3827
SSD Method A (before APAO)		0.0132	0.0007	0.0002	0.0000	0.0000	0.0001	0.0001	0.0044	0.0227	0.0007	0.0000
SSD Method A (after APAO)			0.8571	0.0004	0.0012	0.0011	0.0906	0.1111	0.1124	0.2225	0.0081	0.0001
SSD Method A (APAO)				0.0057	0.0001	0.0004	0.0176	0.0360	0.0613	0.1801	0.0091	0.0001
SSD method C					0.2689	0.2752	0.4388	0.5063	0.6141	0.7241	0.0604	0.0008
Troxler 3430 - 38996 (4 reading)						0.9598	0.0007	0.0112	0.0502	0.2213	0.2215	0.0078
Troxler 3430 - 38996 (2 reading)							0.0057	0.0215	0.0810	0.2474	0.2136	0.0065
Troxler 3430 - 38806 (4 reading)								0.8259	0.7742	0.8225	0.0488	0.0010
Troxler 3430 - 38806 (2 reading)									0.8661	0.8736	0.0572	0.0012
Troxler 3430 - 35601 (4 reading)										0.9143	0.1140	0.0063
Troxler 3430 - 35601 (2 reading)											0.1562	0.0145
PQI 301 - density												0.0007

Table 6.14 presents the results of comparisons of density measurement methods and devices used to determine the density of the composite pavement comprised of the first and second lift (inlay and overlay) placed along the OR 140 project. For this set of comparisons, the nuclear gauge was in direct contact with the paving lift from which the cores were obtained. The results indicate that there was strong evidence to suggest no significant differences (at a 95% confidence level) between the densities derived from the tests using the CoreLok device and the densities derived from the SSD method A. The results also provide strong evidence to suggest no significant differences (at 95% confidence level) between densities derived from the various tests on cores and the densities measured using nuclear gauges in all but 1 of 12 comparisons. Also, the results provide strong evidence to suggest no significant differences between the densities derived from the CoreLok and the densities measured using the electromagnetic gauges. However, there were significant differences between the densities derived from the SSD method A and the densities measured using the electromagnetic gauges.

The results in Table 6.14 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between the densities from various nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from two readings versus four readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate that there is strong evidence to suggest no significant difference between gauges in 4 of the 6 possible comparisons. The results also indicate that there was strong evidence to suggest no significant differences between the densities measured from the nuclear gauges and the densities measured from the PQI 301 electromagnetic gauge; whereas, there were significant differences between the densities from the nuclear gauge and the densities from the PQI 300 electromagnetic gauge in a majority of the comparisons.

Table 6.14 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the composite pavement (140-2 uncut i.e., comprised of inlay and overlay) of the mill and fill (inlay) plus overlay project having same mix design (OR 140)

	SSD Method A (before APAO)	Troxler 3430 - 38996 (4 readings)	Troxler 3430 - 38996 (2 readings)	Troxler 3430 - 38806 (4 readings)	Troxler 3430 - 38806 (2 readings)	Troxler 3430 - 35601 (4 readings)	Troxler 3430 - 35601 (2 readings)	PQI 301 - density	PQI 300 - density
CoreLok	0.1385	0.8024	0.8506	0.3609	0.4974	0.2854	0.3117	0.7114	0.1990
SSD Method A (before APAO)		0.0520	0.0302	0.1838	0.1107	0.3718	0.4990	0.0017	0.0000
Troxler 3430 - 38996 (4 reading)			0.9598	0.0007	0.0112	0.0502	0.2213	0.2215	0.0078
Troxler 3430 - 38996 (2 reading)				0.0057	0.0215	0.0810	0.2474	0.2136	0.0065
Troxler 3430 - 38806 (4 reading)					0.8259	0.7742	0.8225	0.0488	0.0010
Troxler 3430 - 38806 (2 reading)						0.8661	0.8736	0.0572	0.0012
Troxler 3430 - 35601 (4 reading)							0.9143	0.1140	0.0063
Troxler 3430 - 35601 (2 reading)								0.1562	0.0145
PQI 301 - density									0.0007

Based on the results summarized in tables 6.11 to 6.14, a number of observations can be made.

- Tables 6.11 and 6.13 provide results wherein the cores (or slices of cores) were obtained from the same lift on which the nuclear gauge readings were taken. The methodology used to generate the results in these particular tables follows the methodology provided in WAQTC TM8 for determining nuclear gauge calibration factors. Hence, the above-mentioned tables carry greater weight than tables 6.12 and 6.14 in determining whether or not development of nuclear gauge calibration factors is recommended as a mandatory requirement for dense-graded HMA paving projects.

- The comparisons on both the lifts indicate that the densities derived from the CoreLok device were not significantly different from the densities derived from the SSD method C. Also, on the top or overlay lift and the composite pavement (tables 6.13 and 6.14), the densities derived from the CoreLok device were not significantly different from the densities derived from the SSD method-A tests on cores. However, on the bottom or inlay lift (tables 6.11 and 6.12) the densities derived from the CoreLok device were significantly different from the densities derived from SSD method A.

- A majority of the comparisons made on both the lifts indicate that the densities from the tests on cores using the CoreLok device and the SSD method C were not significantly different from the densities measured using nuclear gauges at 95% confidence level. Also, on the bottom or inlay lift section of the composite pavement and on the composite pavement itself (tables 6.12 and 6.14), the results indicate that there was no significant difference even between the densities derived from the SSD method A and the densities measured using the nuclear gauges. However, on the inlay lift and the overlay lift (tables 6.11 and 6.13) wherein the gauges were in direct contact with the lift from which the core slices were extracted as per the WAQTC TM 8 procedure, a majority of the comparisons indicate that the densities derived from the SSD method A tests on cores were significantly different from the density measurements from nuclear gauges at a 95% confidence level.

- Also, a majority of the comparisons made after the overlay lift was paved (tables 6.12 to 6.14) indicate that the densities derived from the tests on cores using the CoreLok device were not significantly different from those of electromagnetic gauge measurements at 95% confidence level. However, a majority of the comparisons made when only the inlay lift was paved (table 6.11) indicate that the densities derived from the CoreLok were significantly different from those of electromagnetic gauge measurements at 95% confidence level. Also, the results were inconclusive in determining whether the densities derived from the SSD methods were significantly different from the densities measured from the electromagnetic gauges at 95% confidence level.

- A majority of the comparisons made after the overlay lift was paved (tables 6.12 to 6.14) indicate that the densities from the various nuclear gauges were not significantly different from each other at a 95% confidence level. However, a majority of the comparisons made when only the inlay lift was paved (table 6.11) indicate that the densities from the various nuclear gauges were significantly different from each other at 95% confidence level.

- A majority of the comparisons made after the overlay lift was paved (tables 6.12 to 6.14) indicate that the densities from the various nuclear gauges were not significantly different from those of the PQI 301, but were significantly different from the PQI 300, electromagnetic gauge at 95% confidence level. However, a majority of the comparisons made when only the inlay lift was paved (table 6.11) indicate that the densities from the various nuclear gauges were not significantly different from those of the PQI 300, but significantly different from the PQI 301, electromagnetic gauge at 95% confidence level.

- For a given nuclear gauge used, the average density of the lift obtained using the average of 4 nuclear gauge measurements at every location was not significantly different from the average density of the lift obtained using the average of 2 nuclear gauge measurements (as seen in all the tables).
- The average densities obtained from the PQI 301 and the PQI 300 were significantly different from each other on all the lifts at a 95% confidence level.

6.2.2.4 OR 18 Project (Core Extraction Location)

Recall from Section 3.2 that a part of the research was to investigate differences between densities obtained from the corner of the 'L-shaped' and 'cross-shaped' footprints. The results presented in this section were derived from density measurements within the 'L-shaped' and 'cross-shaped' footprints, not between the two types of footprints, to determine if there were significant differences between the various measurement methods and devices. Note that section 6.3.2.4 provides results of comparisons between the footprint core locations.

Table 6.15 presents the results of comparisons of density measurement methods and devices used to determine the densities on the 'cross-shaped' footprints along the OR 18 project. The results indicate that there was strong evidence to suggest no significant difference (at a 95% confidence level) between the densities derived from using the CoreLok and the densities derived from the SSD method C. However, there was significant differences between the densities derived using the SSD method A and the densities derived using SSD method C. The results also provide strong evidence to suggest that the densities derived from CoreLok and SSD methods on cores were not significantly different (at a 95% confidence level) from density measurements using nuclear gauges. However, there was significant difference (at a 95% confidence level) between the densities derived from SSD method C tests and the densities measured using one of the two nuclear gauges.

The results in Table 6.15 also provide information regarding differences between the numbers of readings (2 versus 4) used to compute the average density from a set of measurements at a particular sample location for a particular nuclear gauge as well as differences between the densities from various nuclear gauges. For the former case, the results provide strong evidence to suggest no significant differences between the average densities derived from two readings versus four readings for all three nuclear gauges compared. When comparing results between nuclear gauges for a given number of readings, the results indicate strong evidence to suggest no significant difference between gauges in all of the possible comparisons.

Table 6.15 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the 'cross-shaped' footprint of the project having multiple lifts with same mix design (OR 18)

	SSD (Method A)	SSD (Method C)	Troxler 3430 - 38806 (4)	Troxler 3430 - 38806 (2)	Troxler 3430 - 38994 (4)	Troxler 3430 - 38994 (2)
CoreLok	0.0116	0.0907	0.0817	0.0805	0.0526	0.0462
SSD (Method A)		0.0026	0.2360	0.2801	0.1086	0.1063
SSD (Method C)			0.0560	0.0505	0.0337	0.0276
Troxler 3430 - 38806 (4)				0.2826	0.1464	0.2085
Troxler 3430 - 38806 (2)					0.1491	0.1691
Troxler 3430 - 38994 (4)						0.1866

Table 6.16 presents the results of comparisons of density measurement methods and devices used to determine the densities on the ‘L-shaped’ footprints along the OR 18 project. The results indicate that there was strong evidence to suggest significant difference (at a 95% confidence level) between the densities derived from the various tests on cores. The results also provide strong evidence to suggest that the densities derived from CoreLok and SSD methods on cores were not significantly different (at a 95% confidence level) from density measurements using nuclear gauges. However, there was significant difference (at a 95% confidence level) between the densities derived from SSD method C tests and the densities measured using one of the two nuclear gauges.

The results in Table 6.16 also provide information regarding differences between the densities from various nuclear gauges. When comparing results between nuclear gauges, the results indicate strong evidence to suggest no significant differences between the densities measured from the two gauges.

Table 6.16 – p-values from the paired t-tests comparing the results obtained from the various methods of measurement on the 'L-shaped' footprint of the project having multiple lifts with same mix design (OR 18)

	SSD (Method A)	SSD (Method C)	Troxler 3430 - 38806 (2)	Troxler 3430 - 38994 (2)
CoreLok	0.0046	0.0043	0.1412	0.0655
SSD (Method A)		0.0004	0.5347	0.2002
SSD (Method C)			0.0651	0.0262
Troxler 3430 - 38806 (2)				0.3882

Based on the results summarized in tables 6.15 and 6.16, a number of observations can be made as follows:

- For the ‘cross-shaped’ and ‘L-shaped’ footprint investigation, a majority of the comparisons between the densities derived from the various tests on cores indicate that they were significantly different from each other at a 95% confidence level. However, on the ‘cross-shaped’ footprint, the densities derived from the CoreLok method were not significantly different from the densities derived from the SSD method C at a 95% confidence level.

- A majority of the comparisons indicate that there was no significant difference (at a 95% confidence level) between the densities derived from the various tests on cores and the densities measured using the nuclear gauges.

- Also, in both the tables, results from the comparisons indicate that there was no significant difference (at a 95% confidence level) between the various nuclear gauge density measurements.

- For the ‘cross-shaped’ footprint, there was no significant difference (at a 95% confidence level) between the densities obtained using the average of 4 nuclear gauge measurements at every location and the densities obtained using the average of 2 nuclear gauge measurements. On the ‘L-shaped’ footprint, only 2 measurements were taken by a nuclear gauge at every location.

6.2.3 Summary

To identify if an adjustment factor (average difference) must be obtained between the densities derived from the tests on cores and the densities measured using the nuclear gauges, the primary objective of interest is the results of the comparisons between the densities from the two methods of measurements.

For the project having multiple lifts with same mix design (OR 18), the comparisons on 7 of 8 various lifts (i.e., tables 6.1 through 6.8 except 6.5) indicate that the densities obtained from various tests on cores are significantly different from densities obtained from nuclear gauges. For the project having two lifts with different mix designs (I-5), the comparisons on the bottom lift (table 6.9) indicate that the densities obtained from various tests on cores are significantly different from the densities obtained from nuclear gauges, and comparisons made on the top lift (table 6.10) indicate that there is no significant difference between densities from tests on cores and densities from nuclear gauges. For the mill and fill project (OR 140) having an inlay and an overlay lift, a majority of the comparisons indicate that the densities derived from CoreLok and SSD method C tests were not significantly different from the density measurements from the nuclear gauges Also, tables 6.12 and 6.14 indicate that the densities derived from the SSD method A were not significantly different from the densities measured from nuclear gauges. However, tables 6.11 and 6.13, whose results carry more weight because measurements here were taken in accordance with WAQTC TM8 procedure, indicate that densities derived from SSD method A were significantly different from densities measured using nuclear gauges. On the ‘L-shaped’ and ‘cross-shaped’ footprints, the results indicate that there was no significant difference (at 95% confidence level) between the densities derived from tests on cores and densities measured using the nuclear gauges.

Results from a majority of the lifts provide evidence to suggest that it is probably better to obtain an average difference and use it to adjust gauge densities to core densities. This is also supported by finding from the literature studies which indicate that core results give us a more accurate density than nuclear gauge results (26, 27, and 34).

Also, an appropriate method should be selected to test cores since all the various methods of testing cores produced significantly different results.

6.3 Pooled t-tests to compare the methods of measurement between lifts

This section presents the methodology and results of statistical analyses conducted to determine if the adjustment factors from one lift of a project can be used on another lift on that project. These adjustment factors can be described as the average differences between the densities derived from the tests on cores and the densities measured from the nuclear gauges.

6.3.1 Methodology

The ‘adjustment factor’ or ‘average difference’ was first calculated. To do this, the difference in densities between any two particular methods of measurements was obtained at each of the 10 given locations on each lift section. For example, at each of the 10 locations, the difference between the density derived from the tests on cores using the CoreLok device and the density measured using the average of 4 density readings per location using the control nuclear gauge was calculated. Then, the average of the 10 differences was calculated to determine the ‘adjustment factor’ for the nuclear gauge on that lift. In this example, this adjustment factor was the “CoreLok – Control nuclear gauge (average of 4 readings)” adjustment factor, wherein the ‘average of 4 readings’ indicates that the density from the control nuclear gauge at each location was measured using the average of 4 density readings from the gauge. Thus, in this manner, various ‘adjustment factors’ or ‘average differences’ were calculated for each lift as follows:

- SSD (Method A) – Nuclear gauge (average of 4 readings),
- SSD (Method A) – Nuclear gauge (average of 2 readings),
- SSD (Method C) – Nuclear gauge (average of 4 readings),
- SSD (Method C) – Nuclear gauge (average of 2 readings),

- CoreLok – Nuclear gauge (average of 4 readings),
- CoreLok – Nuclear gauge (average of 2 readings),
- SSD (Method A) – Electromagnetic gauge (average of 5 readings; un-sanded condition),
- SSD (Method A) – Electromagnetic gauge (average of 5 readings; sanded condition),
- SSD (Method C) – Electromagnetic gauge (average of 5 readings; un-sanded condition),
- SSD (Method C) – Electromagnetic gauge (average of 5 readings; sanded condition),
- CoreLok - Electromagnetic gauge (average of 5 readings; un-sanded condition),
- CoreLok - Electromagnetic gauge (average of 5 readings; sanded condition),

It should be emphasized that there were 3 nuclear gauges and 2 electromagnetic gauges that were commonly used on more than one lift; all these gauges were the ones whose results were used for the purpose of obtaining the 'average differences' to answer the research questions. Since they were used on more than one lift, densities obtained from these gauges were used in calculating adjustment factors and in the pooled t-tests. The remaining gauges, which were not used on more than one lift, could only be used in the paired t-tests (which compare the various methods of measurements on a given lift) to derive various other conclusions as indicated in section 6.2. Also, most of the SSD tests which were used for the purpose of obtaining the 'average differences' were performed multiple times using the AASHTO T 166 procedure. They were designated as follows: SSD Method A (before APAO testing), SSD Method A (APAO testing), SSD Method A (after APAO testing), and SSD Method C.

In this manner, the ‘average differences’ described above were calculated on every lift. Then, statistical comparisons (pooled t-tests) was performed to verify if the various adjustment factors or average differences from one lift were significantly different from those from another lift. Thus, the statistical tests were always used to compare the adjustment factors between any two lifts.

Since the densities obtained at each location on the various lifts were from the same stretch of pavement, they were considered a part of the same population for statistical purposes. Similarly, since the locations were randomly selected and were at different spots with each successive lift, they were considered two separate samples. Thus, in this case, pooled t-tests were performed between the various data sets.

Pooled t-tests were conducted to determine if the adjustment factors (i.e., differences between average densities derived from tests on cores and average densities measured using nuclear gauges) on one lift were equal to the adjustment factors on another lift (i.e., to determine if the difference in adjustment factors from the two lifts was zero). Any non-zero differences found would provide evidence to suggest the adjustment factors on one lift cannot be used on another lift. Thus, there would then be a need for obtaining separate adjustment factors or ‘average differences’ between core and nuclear gauge densities on each individual lift for a given project.

More specifically, for a particular method of testing cores and a particular gauge commonly used on two given lifts, the difference between the density derived from the test on cores and the density measured from the gauge was calculated at every location on both the lifts. The average and standard deviation of these differences between the densities from the two methods of measurement was also calculated. The average and standard deviation of the differences were then utilized in the pooled t-test to determine if the difference between the adjustment factors (average of differences) was equal to zero at a specified confidence level. Thus, the hypotheses and test statistic were as follows:

Hypotheses:

- μ_1 = True mean adjustment factor on one lift (i.e., difference between the true mean density of one pavement lift as determined by a particular method of measurement on cores and the true mean density on the same lift as determined by measurements from a particular nuclear gauge)

- μ_2 = True mean adjustment factor on another lift (i.e., difference between the true mean density of the other pavement lift as determined by the same particular method of measurement on cores and the true mean density on the same lift as determined by measurements from the same particular nuclear gauge)

- Null hypothesis, H_0: $\mu_1 - \mu_2 = 0$ (i.e., no difference between the true mean adjustment factors from the two lifts)

- Alternate hypothesis, H_a: $\mu_1 - \mu_2 \neq 0$ (i.e., difference between the true mean adjustment factors from the two lifts)

Test statistic:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - 0}{\sqrt{\frac{s_1^{\ 2}}{n_1} + \frac{s_2^{\ 2}}{n_2}}} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^{\ 2}}{n_1} + \frac{s_2^{\ 2}}{n_2}}}$$

Where:

t = Pooled t-statistic

$\overline{X}_1$ = Estimate of the true mean adjustment factor on one lift (i.e., mean of differences between core densities and nuclear gauge densities)

$\overline{X}_2$ = Estimate of the true mean adjustment factor on the other lift (i.e., mean of differences between core densities and nuclear gauge densities)

s_1 = Standard deviation of the differences on one lift (i.e., standard deviation of the differences between core densities and nuclear gauge densities)

s_2 = Standard deviation of the differences on the other lift (i.e., standard deviation of the differences between core densities and nuclear gauge densities)

n_1 = Number of observations of the differences on one lift

n_2 = Number of observations of the differences on the other lift

The nature of the alternate hypothesis (H_a) implies that a two-tailed rejection region should be used, meaning that the difference could be positive or negative. That is, the two-tailed t-test correctly accounts for the possibility that the adjustment factor on one lift could be less than or greater than the adjustment factor on the other lift. The t-test was conducted using a confidence level of 95% (significance level $\alpha = 0.05$, or 5%). Figure 6.1 in section 6.2.1, which illustrates the interpretation of paired t-tests, can be used to illustrate the interpretation of pooled t-tests. Figure 6.1a illustrates a result that would not allow rejection of the null hypothesis at a 95% confidence level. That is, the 95% confidence interval centered about the mean difference obtained from test results captures the value of zero indicating that there is insufficient evidence to suggest a difference in the adjustment factors, as determined on each of the lifts, at this confidence level. Conversely, Figure 6.1b illustrates a result that would allow rejection of the null hypothesis (i.e., the value of zero falls outside of the 95% confidence interval). The results of the various comparisons of the adjustment factors between the various lifts are indicated using p-values. The explanation and interpretation of p-values has been described in section 6.2.1.

6.3.2 Results

The tables contained in this section (tables 6.17 through 6.22) summarize the results of pooled t-tests utilized to determine whether or not the null hypothesis of each comparison (i.e., difference between the adjustment factors of two given lifts is equal to zero) can be rejected at the 95% confidence level (5% significance level). The p-values from these pooled t-tests are provided in the tables in the Appendix. However, in this section, each cell in each of the tables contains information to indicate the number of comparisons wherein the adjustment factors from one lift were not significantly different from those of another lift at a 95% confidence level. Cells highlighted with yellow indicate that only a few of the adjustment factors on one lift were not significantly different from the adjustment factors on the other lift. In other words, a majority of the null hypotheses between those two lifts can be rejected (i.e., sufficient evidence at the 95% confidence level to indicate that the adjustment factor on one lift was significantly different from the adjustment factor on the other lift for a majority of the adjustment factors compared between the two lifts). Cells highlighted with green indicate that the null hypothesis cannot be rejected in a majority of the instances. Each cell also contains a reference to a particular table in appendix B and from which more inferences can be drawn. For example, in table 6.17, the first cell reads ‘18 of 18 (table B4)’. This indicates that of the 18 comparisons between the average differences for those two lifts (detailed in table B4), all 18 of them were not significantly different at 95% confidence level. Also, the blue diagonal line cutting across the table is an indication that the values on one side of the line are a transpose of those on the other side of the line (i.e., like a mirror reflection); it was included to easily draw interpretations from the findings.

6.3.2.1 Summary of comparisons across lifts on the OR 18 project

Section 3.2 provides information on the experiment design for this project i.e., to answer the first research question which is as follows:

“Is it necessary to obtain the average difference between densities from tests on cores and densities from nuclear gauges for each lift of hot mix asphalt (HMA) placed in multiple lifts with the same mix design for each lift?”

As indicated in Section 3.3.1, the base course along this project was constructed in three lifts, all having the same mix design. For the purpose of answering this research question, densities and core samples were obtained from all three lifts of the base course on this project. The results presented in this section were derived from comparisons of the various adjustment factors between the various lifts on this project. Again, the adjustment factors were the averages of differences between densities derived from various tests on cores and densities measured from various gauges. The purpose of these comparisons was to determine if the adjustment factors on one lift could be used on another lift on the same project. This was done by performing statistical analysis between the adjustment factors from two given lifts to determine if they were significantly different or not at a 95% confidence level.

Table 6.17 presents the summary of the results of the pooled t-tests comparing the various adjustment factors between the various lifts. From table 6.17, it can be observed that a majority of the comparisons of adjustment factors across a majority of the lifts were not significantly different from each other. However, a majority of the adjustment factors obtained on the third set of the first lift (18-3 lift 1) were significantly different (at 95% confidence level) from those on the remaining lifts. Also, a majority of the adjustment factors on the composite pavement comprised of all three lifts (18-3) were significantly different (at a 95% confidence level) from those of most other lifts. Another important observation is that a majority of the adjustment factors on the first set of second lift (18-2 lift 2) were significantly different (at a 95% confidence level) from those of the third lift of the composite pavement comprised of all three lifts (18-3 lift 3).

Further, a closer observation of the results from the various tables in the Appendix (used to create the summary values in table 6.17) indicate that the various methods of testing cores to derive the densities provide similar results when used to calculate the adjustment factor and when compared between two lifts. In a majority of instances, the results from comparisons of adjustment factors between two given lifts were the same regardless of the core testing method used. In other words, irrespective of whether the CoreLok method was used or whether the SSD methods were used, similar results were obtained at 95% confidence level. Thus, no conclusive evidence was obtained to suggest that a particular testing method on cores was preferable in comparison to other testing methods on cores. Another observation indicates that, irrespective of the nuclear gauge used in the comparison of adjustment factors between lifts, similar results were obtained at 95% confidence level.

Table 6.17 - Summary of results from pooled t-tests comparing average differences (difference between densities obtained from cores and nuclear gauges) across lifts on the project having multiple lifts with the same mix design (OR 18)

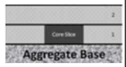
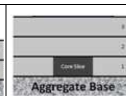
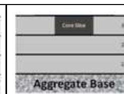
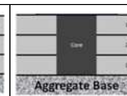
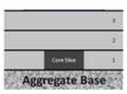
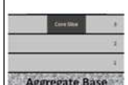
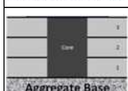
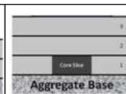
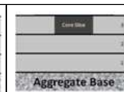
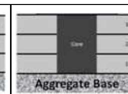
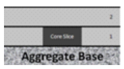
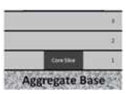
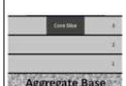
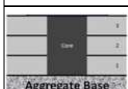
								
		16 of 16 (table B4)	10 of 16 (table B1)	6 of 8 (table B5)	0 of 8 (table B6)	8 of 8 (table B7)	7 of 8 (table B2)	2 of 4 (table B8)
	16 of 16 (table B4)		30 of 30 (table B9)	12 of 12 (table B10)	0 of 20 (table B11)	20 of 20 (table B12)	20 of 20 (table B13)	5 of 8 (table B14)
	10 of 16 (table B1)	30 of 30 (table B9)		12 of 12 (table B18)	0 of 20 (table B15)	20 of 20 (table B16)	5 of 20 (table B3)	3 of 8 (table B17)
	6 of 8 (table B5)	12 of 12 (table B10)	12 of 12 (table B18)		0 of 8 (table B26)	8 of 8 (table B27)	5 of 8 (table B28)	3 of 8 (table B25)
	0 of 8 (table B6)	0 of 20 (table B11)	0 of 20 (table B15)	0 of 8 (table B26)		4 of 20 (table B19)	0 of 20 (table B20)	4 of 8 (table B21)
	8 of 8 (table B7)	20 of 20 (table B12)	20 of 20 (table B16)	8 of 8 (table B27)	4 of 20 (table B19)		20 of 20 (table B22)	8 of 8 (table B23)
	7 of 8 (table B2)	20 of 20 (table B13)	7 of 20 (table B3)	5 of 8 (table B28)	0 of 20 (table B20)	20 of 20 (table B22)		4 of 8 (table B24)
	2 of 4 (table B8)	5 of 8 (table B14)	3 of 8 (table B17)	3 of 8 (table B25)	4 of 8 (table B21)	8 of 8 (table B23)	4 of 8 (table B24)	

Table 6.18 presents the summary of the results of the pooled t-tests comparing the various adjustment factors between the various lifts. From the table, it is observed that a majority of the comparisons of adjustment factors across a majority of the lifts were not significantly different from each other. However, many of the adjustment factors obtained on the third set of the first lift (18-3 lift 1) were significantly different (at a 95% confidence level) from those on the remaining lifts. Thus, for all lift sections other than third set of first lift (18-3 lift 1), the adjustment factors obtained on one lift could be used on the other lifts as well.

Further, in contrast to the observations from table 6.17, a closer observation of the results from the various tables in the Appendix (used to create the summary values in table 6.18) indicates that the CoreLok method provided better results than the SSD methods. In other words, when the densities from the CoreLok method were used with those from either of the electromagnetic gauges to obtain the adjustment factors, then the comparison of these adjustment factors between lifts were not significantly different at a 95% confidence level in more instances than when the densities from the SSD methods were used with electromagnetic gauges to obtain the adjustment factors. As seen from table B11 and table B20, when the densities from the SSD methods were used with those from the 2 electromagnetic gauges to obtain the adjustment factors, the comparison of many of these adjustment factors between the lifts were significantly different at a 95% confidence level. Another observation is that, irrespective of which electromagnetic gauge was used while comparing adjustment factors between lifts, similar results were obtained at a 95% confidence level. Even with the presence or absence of sand while using either of the electromagnetic gauges, similar results of comparisons of adjustment factors between lifts were obtained at a 95% confidence level (as seen in tables B21 to B24).

Table 6.18 - Summary of results from pooled t-tests comparing average differences (difference between densities obtained from cores and electromagnetic gauges) across lifts on the project having multiple lifts with the same mix design (OR 18)

								
		N/A	N/A	N/A	N/A	N/A	N/A	N/A
	N/A		10 of 10 (table B9)	4 of 4 (table B10)	4 of 10 (table B11)	10 of 10 (table B12)	9 of 10 (table B13)	4 of 4 (table B14)
	N/A	10 of 10 (table B9)		4 of 4 (table B18)	1 of 10 (table B15)	10 of 10 (table B16)	10 of 10 (table B3)	3 of 4 (table B17)
	N/A	4 of 4 (table B10)	4 of 4 (table B18)		3 of 4 (table B26)	4 of 4 (table B27)	3 of 4 (table B28)	4 of 4 (table B25)
	N/A	4 of 10 (table B11)	1 of 10 (table B15)	3 of 4 (table B26)		12 of 20 (table B19)	4 of 20 (table B20)	8 of 8 (table B21)
	N/A	10 of 10 (table B12)	10 of 10 (table B16)	4 of 4 (table B27)	12 of 20 (table B19)		20 of 20 (table B22)	8 of 8 (table B23)
	N/A	9 of 10 (table B13)	10 of 10 (table B3)	3 of 4 (table B28)	4 of 20 (table B20)	20 of 20 (table B22)		8 of 8 (table B24)
	N/A	4 of 4 (table B14)	3 of 4 (table B17)	4 of 4 (table B25)	8 of 8 (table B21)	8 of 8 (table B23)	8 of 8 (table B24)	

6.3.2.2 Summary of comparisons across lifts on the I-5 project

Section 3.2 provides information on the experiment design for this project; i.e., to answer the second research question which is as follows:

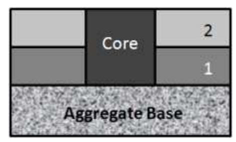
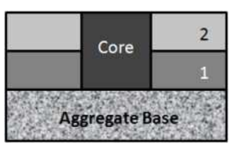
"For projects incorporating different mix designs for individual lifts (e.g., use of a rich binder base), is there a significant difference between nuclear gauge calibration factors for the different mixtures?"

As indicated in Section 3.3.2, the first lift had a different mix design than the other lifts of the base course. For the purpose of answering the above research question, densities and samples were only obtained from the first and second lift of the base course of this project. The results presented in this section were derived from comparisons of the various adjustment factors between the two lifts on this project. The adjustment factors were the averages of differences between densities derived from various tests on cores and densities measured from various gauges. Again, the purpose of these comparisons was to determine if the adjustment factors on one lift could be used on another lift with a different mix design on the same project. This was done by performing statistical analysis between the adjustment factors from the two lifts to determine if they were significantly different or not at a 95% confidence level.

Table 6.19 presents the summary of the results of the pooled t-tests comparing the various adjustment factors between the two lifts. From the table, observations indicate that a majority of the comparisons of adjustment factors across the two lifts were not significantly different from each other at a 95% confidence level.

Further, a closer observation of the results from the table B29 in the Appendix B (used to create the summary values in table 6.17) indicated that the various nuclear gauges used on this project produced differing results when comparing the adjustment factors between the two lifts. For the adjustment factors calculated using the densities derived from the SSD tests and the densities measured using the nuclear gauges, the results from 2 of the 3 nuclear gauges indicated that there was a significant difference between the adjustment factors from the two lifts at a 95% confidence level, whereas the results from the other nuclear gauge indicated that there was no significant difference between the adjustment factors at 95% confidence level. For the adjustment factors calculated using the densities derived from the CoreLok tests and the densities measured using the nuclear gauges, the results from all the 3 nuclear gauges indicated that there was no significant difference between the adjustment factors from the two lifts at a 95% confidence level. Thus, these results indicate that on this project, if densities from CoreLok (in comparison to SSD) are used to obtain the average differences on one lift, then those average differences can be used on the other lift as well.

Table 6.19 - Summary of results from pooled t-tests comparing average differences (difference between densities obtained from cores and nuclear gauges) across lifts on the project having two lifts with different mix designs (I-5)

		
		10 of 18 (table B29)
	10 of 18 (table B29)	

6.3.2.3 Summary of comparisons across lifts on the OR 140 project

Section 3.2 provides information on the experiment design for this project; i.e., to answer the third research question which is as follows:

"On mill & fill (inlay) plus overlay projects, is there a significant difference between the nuclear gauge calibration factors for the fill (inlay) material and the overlay material?"

As indicated in Section 3.3.3, this repaving project had two lifts (the inlay and the overlay) both having the same mix design. Densities and samples were obtained from these two lifts along this project. The results presented in this section were derived from comparisons of the various adjustment factors between the inlay and overlay lifts. The adjustment factors are the averages of differences between densities derived from various tests on cores and densities measured from various gauges. Again, the purpose of these comparisons was to determine if the adjustment factors on one lift could be used on the other lift on mill and inlay projects. This was done by performing statistical analysis between the adjustment factors from the two lifts to determine if they were significantly different or not at a 95% confidence level.

Table 6.20 presents the summary of the results of the pooled t-tests comparing the various adjustment factors between the various lift sections. In the table, since all the cells are green, it is observed that a majority of the comparisons of adjustment factors between any two lift sections were not significantly different from each other at a 95% confidence level.

Further, a closer observation of the results from the various tables in the Appendix B (used to create the summary values in table 6.20) indicates that the CoreLok method provided better results than the SSD methods. In other words, when the differences between the densities from the CoreLok method and the densities from any of the 3 nuclear gauges were used to obtain the adjustment factors, then the comparison of these adjustment factors between lifts were not significantly different at a 95% confidence level in all the instances. However, as seen in table B31 and B32, when the differences between the densities from the SSD method A and the densities from the 3 nuclear gauges were used to obtain the adjustment factors, the comparison of these adjustment factors between lifts were not significantly different (at a 95% confidence level) for 2 of the 3 nuclear gauges used, whereas there was a significant difference for the third nuclear gauge. Also, the p-values in the tables indicated that all three nuclear gauges provided similar results in most of the comparisons.

Table 6.20 - Summary of results from pooled t-tests comparing average differences (difference between densities obtained from cores and nuclear gauges) across lifts on the mill and fill (inlay) plus overlay project having same mix design on both lifts (OR 140)

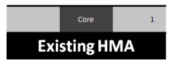
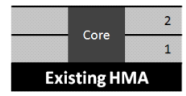
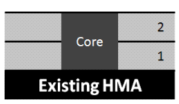
				
		24 of 24 (table B30)	20 of 24 (table B31)	10 of 12 (table B32)
	24 of 24 (table B30)		30 of 30 (table B33)	12 of 12 (table B34)
	20 of 24 (table B31)	30 of 30 (table B33)		12 of 12 (table B35)
	10 of 12 (table B32)	12 of 12 (table B34)	12 of 12 (table B35)	

Table 6.21 presents the summary of the results of the pooled t-tests comparing the adjustment factors between the various lifts. A majority of the comparisons indicated that there is no significant difference in adjustment factors between the various lifts. However, the yellow cell in the table 6.21 indicates that a majority of the comparisons of adjustment factors between the first set of inlay lift (140-1) and the overlay lift section (140-2 lift 2) were significantly different from each other at a 95% confidence level.

Further, a closer observation of the results from the various related tables (table B30 to B35) in the Appendix B used to create the summary values in table 6.21 indicates that the CoreLok method provided better results than the SSD method. In other words, when the densities from the CoreLok method were used with those from either of the electromagnetic gauges to obtain the adjustment factors, then the comparison of these adjustment factors between lifts were not significantly different at a 95% confidence level in all the instances. However, as seen from table B30 to table B32, when the densities from the SSD methods were used with those from the 2 electromagnetic gauges to obtain the adjustment factors, the comparison of many of these adjustment factors were significantly different at a 95% confidence level. Another observation is that, in certain lift comparisons (as seen in table B30) the SSD method C provided better results than the SSD method A.

Table 6.21 - Summary of results from pooled t-tests comparing average differences (difference between densities obtained from cores and electromagnetic gauges) across lifts on the mill and fill (inlay) plus overlay project having same mix design on both lifts (OR 140)

	Core 1 Existing HMA	2 Core Slice 1 Existing HMA	Core Slice 2 1 Existing HMA	Core 2 1 Existing HMA
Core 1 Existing HMA		6 of 8 (table B30)	2 of 8 (table B31)	3 of 4 (table B32)
2 Core Slice 1 Existing HMA	6 of 8 (table B30)		10 of 10 (table B33)	4 of 4 (table B34)
Core Slice 2 1 Existing HMA	2 of 8 (table B31)	10 of 10 (table B33)		4 of 4 (table B35)
Core 2 1 Existing HMA	3 of 4 (table B32)	4 of 4 (table B34)	4 of 4 (table B35)	

6.3.2.4 Summary of comparisons based on the gauge footprint and core extraction location on the OR 18 project

Section 3.2 provides information on the experiment design for this project; i.e., to answer the fourth research question which is as follows:

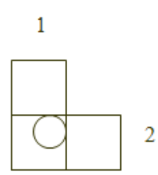
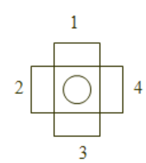
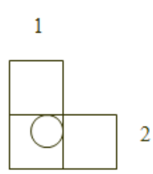
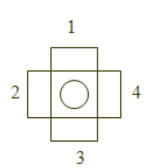
“Is there a significant difference between nuclear gauge calibration factors derived from cores obtained from the corner of the overlapping portion of the nuclear gauge measurement footprint (Figure 3.1) and cores obtained from the center of the overlapping portion of the footprint (Figure 3.2)?”

As indicated in Section 3.2, samples and measurements were taken from the project with three lifts having the same mix design (OR 18). The samples and measurements were taken from the ‘L-shaped’ and ‘cross-shaped’ footprints on a single lift. The results presented in this section were derived from comparisons of the various adjustment factors between the ‘L-shaped’ and the ‘cross-shaped’ footprints on this project. The adjustment factors are the averages of differences between densities derived from various tests on cores and densities measured from various gauges. Again, the purpose of these comparisons was to determine if there was a difference between the adjustment factors measured using the ‘L-shaped’ and ‘cross-shaped’ footprints. This was done by performing statistical analysis between the adjustment factors from the ‘L-shaped’ and ‘cross-shaped’ footprints to determine if they were significantly different or not at a 95% confidence level.

Table 6.22 presents the summary of the results of the pooled t-tests comparing the various adjustment factors between the ‘L-shaped’ and the ‘cross-shaped’ footprints. In the table, since all the cells are green, it is observed that a majority of the adjustment factors between any two lift sections were not significantly different from each other at a 95% confidence level. This suggests that the adjustment factor is not influenced by the location of the core provided that it is obtained within the overlapped region of the nuclear gauge measurements.

Further, a closer observation of the results from the related table (table B36) in the Appendix B used to create the summary values in table 6.22 indicated that the various methods of testing cores to derive the densities provide similar results when used to calculate the adjustment factor and when compared between two lifts. The results from comparisons of adjustment factors between the 'L-shaped' and the 'cross-shaped' footprints were the same (at a 95% confidence level) regardless of which method of testing cores was used to derive the density of cores. In other words, irrespective of whether the CoreLok method was used or whether the SSD methods were used, all the results were not significantly different at a 95% confidence level. Thus, no conclusive evidence was obtained to suggest that a particular testing method on cores was preferable to the other testing methods. Another observation is that irrespective of which nuclear gauge was used, a majority of the results were not significantly different between nuclear gauges.

Table 6.22 - Summary of results from pooled t-tests comparing average differences (difference between densities obtained from cores and nuclear gauges) between the cross-shaped and the L-shaped footprint on the project having multiple lifts with the same mix design (OR 18)

		
		12 of 12 (table B36)
	12 of 12 (table B36)	

6.3.3 Summary

An adjustment factor is the average difference between the densities from a particular nuclear gauge and the densities from a particular core testing method. This adjustment factor is used to adjust the nuclear gauge density measurements on that lift. To identify if these adjustment factors can be used to adjust density measurements made on another lift on the same project, the primary objective of interest is the results of the comparisons between the various adjustment factors among the various lifts.

Overall, comparisons of most of the adjustment factors between most of the lifts indicated that no significant difference is observed at a 95% confidence level. However, comparisons of adjustment factors between certain lifts on some projects indicated that there were significant differences at a 95% confidence level. For the project having multiple lifts with same mix design (OR 18), a majority of the adjustment factors between the third set of the first lift (18-3 lift 1) and the remaining lifts were significantly different at a 95% confidence level. Also, a majority of the adjustment factors (obtained using densities from nuclear gauges but not electromagnetic gauges) between the first set of second lift (18-2 lift 2) and the third lift section (18-3 lift 3) were significantly different at a 95% confidence level. Another important observation on OR 18 indicates that the results of the comparisons of the various adjustment factors between lifts were the same (at a 95% confidence level) irrespective of which method of testing was used on OR 18. This is also true when comparing the various adjustment factors between the 'L-shaped' and 'cross-shaped' footprints on OR 18 project. However, on the OR 140 and I-5 projects, although a majority of the comparisons of adjustment factors between the various lifts indicated no significant differences at a 95% confidence level, a closer observation revealed that the adjustment factors obtained using the CoreLok method were not significantly different between lifts in more cases than those obtained using the SSD methods. While the adjustment factors between the CoreLok method and the gauges were not significantly different between lifts irrespective of which nuclear or electromagnetic gauge was used, the adjustment factors between the SSD methods and the gauges were not significantly different between lifts only when certain nuclear or electromagnetic gauges were used. This was clearly observed in tables B29 to table B32 (i.e., summary of p-value results comparing lifts on the I-5 project and also on the OR 140 project).

6.4 Summary of observation from all data analyses performed

From the paired t-tests comparing the densities from the various methods of measurements on the various projects, certain common observations were made as follows:

1. Overall, on most of the lifts on all the projects, a majority of the densities from the various tests on cores were significantly different from each other at a 95% confidence level.

2. On the new pavement construction project having multiple lifts with same mix design (OR 18), a majority of the densities derived from the various tests on cores were significantly different from those of nuclear gauge measurements at a 95% confidence level. This is also true for the bottom lift of the project with different mix designs (I-5); but, this is not true for the top lift. This is also true for the inlay and overlay lifts when the SSD method A was used for obtaining the densities of the lifts. However, on the top lift of the I-5 project, on the repaving project (OR 140) when CoreLok method and SSD method C were used, and on the project having ‘L-shaped’ and ‘cross-shaped’ measurement footprints, there is no significant difference between the densities derived from the various tests on cores and the densities measured from the nuclear gauges at a 95% confidence level.

3. On OR 18, a majority of the comparisons on a majority of the lifts indicated that the densities derived from the CoreLok or SSD methods were not significantly different from the densities measured using the electromagnetic gauges at a 95% confidence level. However, on the OR 140 project, only the densities from the CoreLok method and those from the nuclear gauges were not significantly

different at a 95% confidence level; for the SSD method, the results were inconclusive.

4. Overall, a majority of the comparisons on the OR 18, I-5 projects, and on the inlay lift of OR 140 indicated that the densities from the various nuclear gauges were significantly different from each other at a 95% confidence level. However, a majority of the comparisons on the overlay lift on OR 140 and on the 'L-shaped' and 'cross-shaped' footprint indicated that the densities from the various nuclear gauges were not significantly different from each other at a 95% confidence level.

5. On all projects, it was commonly observed that the average density of any lift obtained using the average of 4 nuclear gauge measurements at every location was not significantly different from the average density of the lift obtained using the average of 2 nuclear gauge measurements.

6. A majority of the comparisons on OR 18 indicate that the densities from the various nuclear gauges were significantly different from the densities obtained from the electromagnetic gauges at a 95% confidence level. On OR 140, the results from the comparisons on the overlay lift indicated that there was significant difference between the densities from the nuclear gauges and the PQI 300 but not the PQI 301; whereas, on the inlay lift, there was significant difference between the densities from the nuclear gauges and the PQI 301 but not the PQI 300.

7. On the third (top) lift on the OR 18 project and on both lifts on the OR 140 project, the density measurements from the PQI 301 were always significantly different from those of the PQI 300 electromagnetic gauge at a 95% confidence level. However, on the second (middle) lift of the OR 18 project, the density measurements from the PQI 301 were not significantly different from those of the PQI 300 electromagnetic gauge at a 95% confidence level.

8. On the OR 18 project and the OR 140 project, the moisture measurements from the PQI 301 were significantly different from those of the PQI 300 at a 95% confidence level.

9. For the PQI 301, the density and moisture measurements on a sanded surface were not significantly different (at 95% confidence level) from the density and moisture measurements without sand. However, for the PQI 300, the density and moisture measurements on a sanded surface were significantly different (at 95% confidence level) from the density and moisture measurements without sand.

From the pooled t-tests comparing the average differences (between the densities from the various tests on cores and the densities from the various gauges) between the various lifts for all the projects, certain common observations were made as follows:

1. Comparison of adjustment factors between almost all lifts on the OR 18, OR 140, and I-5 projects and also between ‘L-shaped’ and ‘cross-shaped’ footprints on the OR 18 project indicated that a majority of the adjustment factors were not significantly different at a 95% confidence level. Some of the exceptions are as follows:

    a. On OR 18, a majority of the adjustment factors compared between the third set of first lift (18-3 lift 1) and the remaining lifts were significantly different at a 95% confidence level.

    b. Also, a majority of the adjustment factors (differences between core densities and nuclear gauge densities) compared between the first set of second lift (18-2 lift 2) and the third lift section (18-3 lift 3) were significantly different at 95% confidence level.

    c. In contrast, the adjustment factors (differences between core densities and electromagnetic gauge densities) compared between the first set of second lift (18-2 lift 2) and the third lift (18-3 lift 3) were not significantly different at a 95% confidence level.

2. Another common observation on all projects was that, in a vast majority of instances, the results of the comparisons of adjustment factors were similar irrespective of the gauge used. In other words, all gauges (whether nuclear or electromagnetic) produced similar results most of the time. So, if particular results were obtained from the comparisons of adjustment factors between two lifts on a project using a Troxler nuclear gauge, most of the time, similar results will be observed from the comparisons of adjustment factors between those two lifts using some other nuclear gauge or even an electromagnetic gauge. However, for the electromagnetic gauges, better results were obtained when measurements were taken in an un-sanded condition than when measurements were taken in a sanded condition (as seen in tables B19 and B20 in Appendix B). In these tables, the adjustment factors that utilized electromagnetic gauge densities taken in an un-sanded condition were not significantly different between lifts, whereas the adjustment factors that utilized electromagnetic gauge measurements in sanded condition were significantly different between lifts.

3. Another interesting observation from the results of the comparisons of adjustment factors across lifts on the various projects was that the adjustment factors calculated utilizing CoreLok densities were not significantly different (at 95% confidence level) in more cores than adjustment factors that utilized densities from SSD methods. More specifically:

    a. On the OR 140 project and the I-5 project, while the adjustment factors (difference between CoreLok density and gauge density) were not significantly different between lifts irrespective of which nuclear or electromagnetic gauge was used, the adjustment factors between the SSD methods and the gauges were not significantly different between lifts only when certain nuclear or electromagnetic gauges were used.

    b. Even on the OR 18 project, for the adjustment factors obtained between core densities and electromagnetic gauge densities, there were a greater number of instances where the adjustment factors that utilized the CoreLok densities were not significantly different as compared to instances where the SSD methods were used.

7.0 DISCUSSION OF RESULTS

In accordance with the experimental design that was developed to answer the research questions, field and laboratory studies were conducted and statistical analyses were performed on the data collected. This data primarily included the bulk specific gravity and density obtained from the field and laboratory studies. Various analyses were performed on the data and the findings were presented in chapters four through six; the majority of the findings were presented in chapter six. This chapter provides a discussion of the results obtained from the field and laboratory studies in continuation with the results from the various statistical analyses.

7.1 Results obtained to answer the research questions

In summary, the first research question was aimed at understanding if there is a significant difference in adjustment factors between the various lifts of a pavement with the same mix design. The second research question was aimed at understanding if there is a significant difference in adjustment factors between the various lifts of a pavement with different mix designs. The third research question was aimed at understanding if there is significant difference in adjustment factors between the lifts (both with the same mix design) of a pavement that underwent a mill and inlay operation. The fourth research question was aimed at understanding if there is a significant difference in adjustment factors when cores were extracted from the cores of an ‘L-shaped’ footprint as opposed to the center of a ‘cross-shaped’ footprint.

Before discussing the results, it should be recalled that the adjustment factors, which were compared between the various lifts, were not just between the nuclear gauges and cores but also between the electromagnetic gauges and cores for each of the lifts. Besides that, it must also be noted that the nuclear gauge measurements used to determine the adjustment factors were the densities using the average of 2 nuclear gauge readings per location and densities using the average of 4 nuclear gauge readings per location. The densities from the electromagnetic gauges, used in the calculation of adjustment factors, were the densities from the average of five readings per location. The densities from the cores used to calculate the adjustment factors were obtained from the CoreLok method, SSD Method A, and SSD Method C tests. The SSD method A test was performed three times on the cores from most of the lifts; first, it was performed at Oregon State University, then it was performed by APAO personnel in their lab, and finally it was performed again at Oregon State University so that the final SSD mass and submerged mass could be obtained before SSD method C tests were performed.

A majority of the results from the comparisons of adjustment factors between lifts are similar irrespective of the project from which the results were obtained; this indicates that there is a common answer to all the research questions. However, some of the adjustment factors were better than others (i.e., they were not significantly different in a greater number of cases.

7.1.1 Results obtained to answer the first research question

The discussion of results to answer the first research question (i.e., “Is it necessary to determine nuclear gauge calibration factors for each lift of hot mix asphalt (HMA) placed in multiple lifts with the same mix design for each lift?”) is provided in the following paragraphs.

Based on the analysis results, there is substantial evidence to suggest that it is not necessary to obtain the adjustment factors (average difference between the nuclear gauge densities and various core densities) on each lift of a pavement with the same mix design provided that, the adjustment factors are obtained on the first lift (bottom lift). The adjustment factors obtained from the first lift could be used on the second lift as well as the third lift provided that the nuclear gauge was in contact with the first lift for the purposes of determining the adjustment factor. However, adjustment factors obtained from the second lift should not be used on the third lift of the combined lifts and vice versa because a majority of the adjustment factors compared between these two lifts were significantly different at a 95% confidence level. Also, the adjustment factors obtained from the cores of the first lift when the nuclear gauges measurements were made on the top (third) lift should not be used on any other lift because these adjustment factors were significantly different from the adjustment factors of many of the remaining lift sections at a 95% confidence level.

This is evidence to indicate that the gauge measurements and in turn, the adjustment factors are possibly influenced by the depth of the pavement. It also indicates that the gauge measurements and adjustment factors are possibly influenced by the effect of underlying layers. On a side note, the measurements and samples taken from the second and third lifts were within the stations from which measurements and samples were taken from the first lift; whereas there was no overlap in stations between the measurements and samples taken from the second lift and those taken from the third lift (as seen in figure 3.2).

Also, it was inconclusive as to whether the adjustment factors utilizing CoreLok method with nuclear gauges were better or worse than the adjustment factors utilizing SSD methods with nuclear gauges. In other words, the results from the comparison between lifts of adjustment factors utilizing the CoreLok method were similar to the results from the comparisons between lifts of adjustment factors utilizing the SSD methods at a 95% confidence level.

In contrast to the adjustment factors that utilized nuclear gauge densities, the adjustment factors that utilized electromagnetic gauges (average difference between the electromagnetic gauge densities and various core densities) obtained on the second lift were not significantly different from similar adjustment factors obtained on the third lift at a 95% confidence level. Even the adjustment factors from the first lift were not significantly different from those of the second or third lift at 95% confidence level. This is evidence to suggest that, on this project, the adjustment factors that utilized electromagnetic gauge densities could be used on more lifts than the adjustment factors that utilized nuclear gauge densities. Only the adjustment factors from the third set of first lift (18-3 lift 1) could not be used on any other lift because these adjustment factors were significantly different from the adjustment factors of many of the remaining lift sections at 95% confidence level.

Another observation from the results of comparisons of adjustment factors on this project was that, the adjustment factors utilizing densities from the CoreLok method with densities from electromagnetic gauges could be used on more number of lifts than the adjustment factors utilizing densities from the SSD methods with densities from electromagnetic gauges. This is because the adjustment factors utilizing the CoreLok method were not significantly different between more number of lifts in comparison to adjustment factors utilizing SSD method (particularly the lifts compared in table B11, B20, B26, and B28 in Appendix B).

Also, the electromagnetic gauge measurements taken in un-sanded condition produced better results than those taken in sanded condition (i.e., adjustment factors that utilized electromagnetic gauge measurements taken when the pavement locations were un-sanded were not significantly different in more number of lift comparisons at a 95% confidence level than adjustment factors that utilized electromagnetic gauge measurements taken when the pavement locations were sanded). This is observed in tables B19 and B20 in Appendix B. Thus, this is evidence to suggest that electromagnetic gauges produce more desirable readings when the pavement locations are not sanded. In other words, when electromagnetic gauge measurements are taken in the absence of sand, the adjustment factors that are obtained on one lift (using average of differences between these measurements and densities from tests on cores) can be used to adjust the densities measured by these electromagnetic gauges on other lifts as well.

Finally, no evidence was found to suggest that one electromagnetic gauge provided better results than another. The results from the paired t-tests indicated that the measurements from the PQI 300 were significantly different from the PQI 301 on third lift but not on second lift and that the measurements made in sanded condition using an electromagnetic gauge were significantly different from measurements made in un-sanded condition at a 95% confidence level. However, the results of comparisons of adjustment factors obtained from the PQI 300 were almost always equal to the results of comparisons of adjustment factors obtained from the PQI 301. In fact, although the paired t-tests indicated that the PQI 301 was unaffected by sand and the PQI 300 was affected by sand, it was the presence or absence of sand that caused variability in results of comparisons of adjustment factors as seen in tables B19 and B20 in Appendix B.

7.1.2 Results obtained to answer the second research question

The discussion of results to answer the second research question “For projects incorporating different mix designs for individual lifts (e.g., use of a rich binder base), is there a significant difference between nuclear gauge calibration factors for the different mixtures?” is provided in the following paragraphs.

Based on the analysis results, there is substantial evidence to suggest that there is no significant difference (at a 95% confidence level) between a majority of the nuclear gauge adjustment factors (average difference between the nuclear gauge densities and various core densities) from the different mixtures i.e., two lifts having different mix designs.

However, a closer observation of the results from the table B29 in the Appendix B (used to create the summary values in table 6.17 to answer this research question) indicated that the various nuclear gauges used on this project produced differing results when comparing the adjustment factors between the two lifts.

For the adjustment factors calculated using the difference between the densities derived from the CoreLok tests and the densities measured using the nuclear gauges, the analyses results from all the 3 nuclear gauges indicated that there was no significant difference (at 95% confidence level) between these adjustment factors from the bottom lift and similar adjustment factors from the top lift.

However, for the adjustment factors calculated using the difference between the densities derived from the SSD tests and the densities measured using the nuclear gauges, the analyses results from 2 of the 3 nuclear gauges indicated that there was a significant difference (at 95% confidence level) between these adjustment factors from the bottom lift and similar adjustment factors from the top lift; while, 1 nuclear gauge indicated that there was no significant difference (at 95% confidence level) between these adjustment factors from the bottom lift and similar adjustment factors from the top lift.

Thus, these results indicated that on this project, if densities from CoreLok (in comparison to SSD) were used to obtain the adjustment factors on one lift, then those adjustment factors could be used on the other lift as well, despite the two lifts having different mix designs. In other words, the adjustment factors utilizing CoreLok method with nuclear gauges was better than the adjustment factors utilizing SSD method with nuclear gauges.

7.1.3 Results obtained to answer the third research question

The discussion of results to answer the third research question, “On mill & fill (inlay) plus overlay projects, is there a significant difference between the nuclear gauge calibration factors for the fill (inlay) material and the overlay material?” is provided in the following paragraphs.

Based on the analysis results, there is substantial evidence to suggest that there is no significant difference (at 95% confidence level) between the adjustment factors for the inlay lift sections and the adjustment factors for the overlay lift sections.

The adjustment factors (average difference between the nuclear gauge densities and various core densities) obtained from the inlay lift sections could be used on the overlay lift sections and vice versa (as seen in table 6.20 which summarizes tables B30 to B35 in Appendix B). A majority of the comparisons of adjustment factors between any two lift sections were not significantly different from each other at 95% confidence level. Thus, it is not necessary to obtain the adjustment factors (average difference between the densities from nuclear gauges and densities from tests on cores) on every lift of reconstructed pavement (mill and inlay with overlay) with same mix design provided that, the adjustment factors are obtained on at least one lift section.

A closer observation indicates that when the adjustment factors (which are calculated using the differences between the densities from the CoreLok method tests and the densities from the nuclear gauges) are compared to the adjustment factors (which are calculated using the differences between the densities from the SSD methods and the densities from the nuclear gauges), the former adjustment factors provide better results than the latter adjustment factors. In other words, when the differences between the densities from the CoreLok method and the densities from any of the 3 nuclear gauges were used to obtain the adjustment factors, then the comparison of these adjustment factors between lifts were not significantly different at 95% confidence level in all the instances. However, as seen in table B31 and B32, when the differences between the densities from the SSD methods and the densities from the 3 nuclear gauges were used to obtain the adjustment factors, the comparison of these adjustment factors between lifts were not significantly different at 95% confidence level only for 2 of the 3 nuclear gauges used; the adjustment factors obtained using the difference between the densities from the SSD method A and those from the third nuclear gauge were significantly different between the lifts compared in tables B31 and B32. Another observation from the related tables in the Appendix B (comparing the various adjustment factors between the various lift sections) indicated that all three nuclear gauges provided similar results in most of the comparisons.

Similar to the nuclear gauge adjustment factors, the electromagnetic gauge adjustment factors (average difference between the densities from the tests on cores and the densities from the electromagnetic gauges) obtained from the inlay lift sections could be used on the overlay lift sections and vice versa (as seen in table 6.21 which summarizes tables B30 to B35 in Appendix B). The only exception is that the adjustment factors (calculated using the differences between the densities from the SSD methods and the densities from the electromagnetic gauges) obtained from the first set of first lift (140-1) were significantly different from similar adjustment factors obtained on the second lift section (140-2 Lift 2) at 95% confidence level. However, the adjustment factors (calculated using the differences between the densities from the CoreLok method and the densities from the electromagnetic gauges) obtained from the first set of first lift (140-1) were not significantly different from similar adjustment factors obtained on the second lift section (140-2 Lift 2) at 95% confidence level.

When the adjustment factors (which are calculated using the differences between the densities from the CoreLok and the densities from the nuclear gauges) were compared to the adjustment factors (which are calculated using the differences between the densities from the SSD methods and the densities from the nuclear gauges), the former adjustment factors provided better results than the latter adjustment factors. In other words, when the differences between the densities from the CoreLok method and the densities from both the electromagnetic gauges were used to obtain the adjustment factors, then the comparison of these adjustment factors between lifts were not significantly different at 95% confidence level in all the instances between all the lift comparisons (tables B30 to B35). However, as seen in tables B30 to B32, when the differences between the densities from the SSD method A and the densities from both the electromagnetic gauges were used to obtain the adjustment factors, the comparison of these adjustment factors between lifts were significantly different at 95% confidence level. Another observation from the related tables in Appendix B is that the comparisons of the adjustment factors between the various lift sections indicated that both the electromagnetic gauges provided similar results in most of the comparisons, especially when the CoreLok method was used (instead if the SSD methods). This is evidence to suggest that the CoreLok method is more suitable than the SSD method since the adjustment factors obtained from the CoreLok density values were not significantly different between more number of lift comparisons; thus, these adjustment factors could be used to adjust the electromagnetic gauge densities on more number of lift sections.

7.1.4 Results obtained to answer the fourth research question

The discussion of results to answer the fourth research question, “Is there a significant difference between nuclear gauge calibration factors derived from cores obtained from the corner of the overlapping portion of the nuclear gauge measurement footprint (Figure 3.1) and cores obtained from the center of the overlapping portion of the footprint (Figure 3.2)?" is provided in the following paragraphs.

Based on the analysis results, there is strong evidence to suggest that there was no significant difference (at 95% confidence level) between nuclear gauge adjustment factors derived from cores obtained from the corner of the overlapping portion of a nuclear gauge measurement footprint (Figure 3.1) and cores obtained from the center of the overlapping portion of another nuclear gauge measurement footprint (Figure 3.2)

The results from the comparison of adjustment factors (utilizing CoreLok method) between the footprint measurements as indicated in figure 3.1 were similar to the results of the comparison of adjustment factors (utilizing SSD methods) between the footprint measurements as indicated in figure 3.2.

Another observation was that both the nuclear gauges produced similar results in the comparisons of their respective adjustment factors between the footprint measurements as indicated in figure 3.1 and figure 3.2.

7.1.5 Results obtained to answer the fifth research question

The discussion of results to answer the fifth research question, "ODOT currently allows use of the saturated surface-dry (SSD) method for determining the bulk specific gravity of a core even if the water absorption of the core is greater than 2%. Is it appropriate to continue to allow this?" is provided in the following paragraphs.

Based on the literature studies, for fine and coarse mixes (with water absorption greater than 0.4%), the density results from the Automatic Vacuum Sealing method are more accurate than those from the saturated surface-dry method. At high water absorption levels (greater than 0.4%), low densities, low levels of compaction, and lower level of gyration, the AASHTO T 166 overestimates the density. This answers the fifth research question and provides evidence to suggest that the SSD method must not be used at water absorption levels of 2% and instead, the CoreLok method should be used as recommended.

7.2 Summary comparing results obtained from the various projects

Every project was selected for answering a particular research question. It is interesting to note the various similarities and differences that were observed between the various results obtained from the various projects as discussed in the following sections.

7.2.1 Similar results obtained from the various projects

The results can be summarized as follows:

1. Results from all the projects indicated that there were no significant differences (at a 95% confidence level) for most of the adjustment factors compared between most of the lifts.

2. A majority of the nuclear and electromagnetic gauges provided similar results in the comparison of their adjustment factors between lifts. In other words, the result obtained from one particular nuclear gauge adjustment factor compared between two lifts is similar to the result obtained from another nuclear gauge adjustment factor compared between two lifts. The results from the comparisons of nuclear gauge adjustment factors between most of the lifts are also similar to the results from the comparisons of electromagnetic gauge adjustment factors.

3. On all the projects, the adjustment factors utilizing the densities from the CoreLok method could be used on more lifts than the adjustment factors utilizing the densities from the SSD methods. This is because the former adjustment factors were not significantly different (at 95% confidence level) between more lift comparisons than the latter adjustment factors.

4. On all the projects, the average densities obtained from 4 nuclear gauge measurements at every location on a lift were not significantly different (at a 95% confidence level) from the average densities obtained from 2 nuclear gauge measurements at every location on the same lift.

5. A majority of the densities from the nuclear gauges were almost always significantly different from the densities from the electromagnetic gauges at 95% confidence level.

7.2.2 Differing results obtained from the various projects

The results can be summarized as follows:

1. While the nuclear gauge adjustment factors utilizing CoreLok method densities provided similar results as the nuclear gauge adjustment factors utilizing SSD method densities on the OR 18 project, the electromagnetic gauge adjustment factors utilizing CoreLok method densities provided better results (i.e., not significantly different at 95% confidence level in more number of lift comparisons) than the electromagnetic gauge adjustment factors utilizing SSD method densities on the OR 18 project.

2. On the I-5 project and the OR 140 project, the nuclear or electromagnetic gauge adjustment factors utilizing CoreLok method densities provided better results (i.e., not significantly different at 95% confidence level in more number of lift comparisons) than the gauge adjustment factors utilizing SSD method densities.

3. Although the overall results comparing the densities from the various tests on cores indicate that they are significantly different from each other at a 95% confidence level, a closer observation reveals the following:

    a. On OR 18, on almost all the lifts, the densities from the CoreLok method were not significantly different from the densities from the SSD method A at a 95% confidence level.

    b. On I-5 and on the cross-shaped footprint (as seen in figure 3.2), a majority of the densities from the various methods of tests on cores were significantly different from each other at a 95% confidence level.

c. On OR 140 and on the ‘L-shaped’ footprint (as seen in figure 3.1), a majority of the densities from the CoreLok method were not significantly different from the densities from the SSD method C at a 95% confidence level.

d. Also, on OR 140, the results from the overlay and composite pavements indicate that the densities from the CoreLok method were not significantly different from the densities from the SSD method A at 95% confidence level. However, the results from the inlay lift sections (140-1 and 140-2 lift 1) indicate that the densities from the CoreLok method were significantly different from the densities from the SSD method A at a 95% confidence level.

4. On the OR 18 project and on the first lift of I-5 project, the densities from the tests on cores were significantly different from the densities from the nuclear gauges at 95% confidence level. On the second lift of the I-5 project, the OR 140 project, and on the footprints (figures 3.1 and 3.2) on the OR 18 project, the densities from the various tests on cores were not significantly different from the densities from the nuclear gauges at a 95% confidence level.

5. On the OR 18 project, the densities from the CoreLok and SSD methods on cores were not significantly different from the densities from the electromagnetic gauges at a 95% confidence level. Also, on the overlay lift of the OR 140 project, the densities from the CoreLok method on cores were not significantly different from the densities from the electromagnetic gauges at 95% confidence level. However, on the inlay lift of the OR 140 project, the densities from the CoreLok method were significantly different from the densities from the PQI method at a 95% confidence level.

6. On the OR 18 project, I-5 project, the inlay lift of the OR 140 project, the densities from the various nuclear gauges were significantly different from each other at a 95% confidence level. However, on the overlay lift of the OR 140

project and on the measurement from the two footprints on OR 18 project, the densities from the nuclear gauges were not significantly different from each other at a 95% confidence level.

7. On the third (top) lift of the OR 18 project and on the OR 140 project, the densities from the PQI 301 electromagnetic gauge were significantly different from the densities from the PQI 300 electromagnetic gauge at a 95% confidence level. However, on the second (middle) lift of the OR 18 project, the densities from the PQI 301 electromagnetic gauge were significantly different from the densities from the PQI 300 electromagnetic gauge at a 95% confidence level.

7.2.3 Other significant results obtained from the data analyses

These results can be summarized as follows:

1. The density and moisture measurements obtained by the PQI 300 in the sanded condition were significantly different (at 95% confidence level) from those in the un-sanded condition. However, the density and moisture measurements obtained by the PQI 301 in the sanded condition were not significantly different from those in the un-sanded condition. In other words, the PQI 300 is influenced by sand whereas; the PQI 301 is not influenced by sand.

2. When comparing the electromagnetic gauge adjustment factors between lifts, the influence of sand on the results of the comparisons of adjustment factors between lifts is more than the influence of the type of electromagnetic gauge (whether PQI 300 or PQI 301) on the results of the comparisons of adjustment factors between lifts. This is observed in tables B19 and B20.

8.0 CONCLUSIONS AND RECOMMENDATIONS

The objectives of this study (as described in section 1.2) were to develop recommendations for improved HMA density measurement using nuclear gauges, recommendations for alternate ways to measure HMA density, and recommendations for the optimal system for quantifying dense-graded HMA density (which could be a combination of measurements using nuclear gauges and other means). In order to accomplish these objectives certain research questions (provided in section 3.1) were developed. Some of the objectives were addressed through these research questions. The various findings from the literature studies and from the analysis of the data collected were used to answer these research questions and to fulfill these objectives.

To fulfill the first objective, the various methods of measurement of in-place density were investigated. Oregon department of transportation (ODOT) uses nuclear gauges for QC and QA purposes and, to verify nuclear gauge results and for core correlation purposes the AASHTO T 166 method is sometimes used. To fulfill the second objective, current practices utilized by ODOT to measure in-place density were described as a part of the literature studies. To fulfill the third objective, the results from the analyses of field and lab data were summarized in the previous chapters and five of the eight research questions were answered. Based on these findings, the conclusions and recommendations have been written to fulfill the fourth and fifth objectives.

It should be noted that the findings from this research project are limited by various factors and are bounded in their scope and application. The findings only apply to dense-graded hot mix asphalt pavements. Data was obtained from only one project each for three types of pavement construction projects; they were new pavement construction project with multiple lifts having same mix design, new pavement construction project having different mix designs, and rehabilitation project (mill and inlay with overlay) having same mix design. Based on the findings from the data, conclusions and recommendations have been made for those pavement construction project types; if more projects were investigated for each of those types of pavement construction projects, stronger evidence could have been obtained to support the findings. The findings are based on data obtained from certain nuclear gauges (Troxler) and electromagnetic gauges (Pavement Quality Indicator) used by Oregon Department of Transportation and could thus be limited in their scope to these gauges only. Another limitation is that these findings may not apply to gauges from other manufacturers. Thus, the conclusions and recommendations are only applicable to dense-graded hot mix asphalt pavements, the projects investigated, and the gauges used.

8.1 Conclusions

Based on the various findings, the conclusions are as follows:

1. Based on substantial evidence, nuclear gauge measurements need to be adjusted to core density values to obtain more accurate results. In other words, nuclear gauge measurements not adjusted to core density values will most likely produce inaccurate results.

2. It is sufficient to take two nuclear gauge density readings rather than four readings at a given location. The readings should be taken as indicated in figure 3.1.

3. No changes need to be made to the procedures used by ODOT for calibration and operation of a nuclear gauge (which are described in ODOT TM 304 and WAQTC TM 8). In other words, the nuclear gauge must be operated in backscatter mode, the density readings must be of one minute duration each, and the gauge must be rotated 90° keeping the probe at a constant location.

4. Cores extracted for the purpose of obtaining an adjustment factor to adjust nuclear gauge density values, should be extracted from the overlapping portion of the gauge footprint. The location of the core can be anywhere within the overlapping portion of the gauge footprints.

5. While taking two readings at a given location using a nuclear gauge, there is no significant difference in results when the gauge footprints are as indicated in figure 3.1 and when they are as indicated in figure 3.2. Thus, either method can be used to obtain two readings.

6. Literature studies indicate that the CoreLok method provides a more accurate measure of density than the SSD methods (AASHTO T 166) at high air void content (greater than 8%), high water absorption levels (greater than 0.4%), low gyration levels, and coarse mixes. Thus, under these conditions, the CoreLok method should be used instead of the SSD methods.

7. Electromagnetic gauges can possibly replace nuclear gauges to measure in-place density of pavements for quality control purposes. This conclusion is based on the evidence from the results from the data analyses (paired and pooled t-tests).

8. The densities obtained from electromagnetic gauges and adjusted using the CoreLok densities should be used because these adjustment factors provided the best results in comparison to adjustment factors utilizing nuclear gauges or AASHTO T 166 (SSD methods). In other words, the former adjustment factors were not significantly different in more lift comparisons than the latter adjustment factors.

9. Electromagnetic gauge densities might not even require adjustment to core densities (i.e., unlike nuclear gauges) because the results from these two methods of measurement were not significantly different at 95% confidence level.

10. The electromagnetic gauges should be used in the un-sanded condition rather than in the sanded condition.

11. The PQI 301 is not significantly affected by sand or moisture whereas; the PQI 300 is significantly affected by sand and moisture.

12. The PQI 301 should be used instead of the PQI 300 while measuring in-place density of dense-graded hot mix asphalt pavements.

13. On new pavement construction project having multiple lifts with same mix design, the core adjustment factors that are obtained to adjust the nuclear gauge densities need not be obtained on every lift. The adjustment factors should only be obtained on the bottom (first) lift wherein, the nuclear gauge is in direct contact with the bottom lift core being extracted. These adjustment factors can be used on at least the next two overlying lifts which have thickness of two to three inches each. The adjustment factors obtained from the second or third lift cannot be used on overlying lifts. This is because the results indicate that the gauge densities are possibly affected by the depth of measurement and the underlying layers.

14. On new pavement construction project having two lifts with different mix designs, the core adjustment factors should only be obtained on one lift wherein, the gauge is in direct contact with the lift whose core densities will be used to obtain adjustment factors.

15. On mill and fill pavement construction project with two lifts having same mix design, the core adjustment factors should only be obtained on one lift wherein, the gauge is in direct contact with the lift whose core densities will be used to obtain adjustment factors.

16. The CoreLok procedure for obtaining core density should be used instead of the AASHTO T 166 (SSD methods) while obtaining core adjustment factors because it provides stronger evidence for use on multiple lifts. In other words, the CoreLok adjustment factors were not significantly different in more lift comparisons than the AASHTO T 166 adjustment factors.

8.2 Recommendations

Based on the various findings from the literature studies and the results and conclusions from the research project, the following recommendations can be made:

1. The nuclear gauge calibration procedure implemented by ODOT (which are described in ODOT TM 304) can continue to be used without it undergoing any changes.

2. It is strongly recommended that the nuclear gauge density measurements should be adjusted to core densities (based on an adjustment factor) to obtain a more accurate in-place density.

3. The adjustment factor utilized to adjust the gauge densities should be calculated according to the procedure described in ODOT TM 327 or WAQTC TM 8 except that the following changes are recommended to the procedures:

    a. Densities from at least ten cores should be used to obtain the adjustment factor on every lift,

    b. The CoreLok method should be used to test the cores instead of the AASHTO T 166 (SSD methods),

    c. The high and low values of the ratio (calculated in accordance with ODOT TM 327) or the highest average difference value (calculated as described in WAQTC TM 8) if the standard deviation is greater than 2.5 pcf should not be discarded while calculating the core adjustment factor to adjust the gauge densities.

4. The adjustment factors that are calculated utilizing one nuclear gauge should not be used to adjust the densities from any other nuclear gauge on the same lift or on another lift.

5. ODOT should strongly consider adopting the CoreLok method of measuring density instead of the SSD methods.

6. In contrast to literature studies, the electromagnetic gauges show significant promise and could possibly entirely replace nuclear gauges for measuring in-place pavement density. A more extensive study on electromagnetic gauges is strongly recommended so that it can be used for quality control and quality assurance in the measurement of in-place density of dense-graded hot mix asphalt pavements.

7. It is recommended that ODOT adopt the electromagnetic gauges as the viable option to nuclear gauges (pending further study) and especially if they adopt CoreLok densities to adjust electromagnetic gauge densities.

8. The use of the PQI 300 electromagnetic gauge should be discontinued because it is affected by sand and moisture and instead, later models like the PQI 301 should be used as a replacement.

9. It should be made mandatory to obtain the core adjustment factors on at least the first lift if not the next two subsequent lifts. It should be emphasized that these core adjustment factors can be used on at least the next two subsequent lifts (having thickness of two to three inches each) irrespective of the mix design or type of the dense graded pavement being constructed.

BIBLIOGRAPHY

1. Sanders, S.R., Rath, D. and Parker Jr, F. *Comparison of nuclear and core pavement density measurements*, Journal of Transportation Engineering, Vol. 120 No. 6, p. 953-966, 1994

2. Schmitt, R.L., Hanna, A.S., Russel, J.S., and Nordheim, E.V. *Pavement Density Measurement Comparative Analysis using Core and Nuclear Methods*, Asphalt Paving Technology, Vol. 66, 1997

3. Parker, F. and Hossain, M.S. *An Analysis of hot mix asphalt Mat Density Measurements*, Journal of Testing and Evaluation, JTEVA, Vol.23, No. 6, p. 415-423, Alabama, Nov 1995

4. Minchin, R.E. and Thomas, H.R. *Validation of vibration-based onboard asphalt density measuring system*, Journal of Construction Engineering and Management, Vol. 129 No.1, pp1-7, Pennsylvania, Jan/Feb 2003

5. Liao, Y., Sargand, S.M. and Kim, S. *Electromagnetic density gauge comparative study for Quality Control/Quality Assurance in hot mix asphalt construction*, Proceedings of the 2006 Airfield and Highway Pavement Specialty Conference, American Society of Civil Engineers, pp365-376, Ohio, 2006.

6. Brown, E.R. and Greene, J. *Issues in mix design and construction of hot mix asphalt for airfields*, Proceedings of the XXIInd PIARC World Road Congress, World Road Association – PIARC, 2003

7. Hurley, G.C., Prowell, B.D. and Cooley Jr, L.A. *Evaluating electromagnetic measurement devices to determine in-place pavement density (with discussion and closure)*, Transportation Research Record No.1900, Transportation Research Board, pp56-64, Washington DC, 2004

8. Allen, D.L., Schultz Jr, D.B. and Willett, D.A. *Evaluation of electromagnetic density gauges*, Kentucky Transportation Cabinet, Kentucky, 2003

9. Larsen, D.A. and Henault, J.W. *Quantifying Segregation in hot mix asphalt Pavements Using Non-Nuclear Density Devices: Data Collection Report for Connecticut*, Connecticut Department of Transportation, Connecticut, 2006

10. Sebesta, S., Zeig, M. and Scullion, T. *Evaluation of Non-Nuclear Density Gauges for HMAC: Year 1 Report*, Texas Transportation Institute, Texas Department of Transportation, 2003

11. Spellerberg, P. and Savage, D. An Investigation of the Cause of Variation in hot mix asphalt Bulk Specific Gravity Test Results Using Non-Absorptive Aggregates, Transportation Research Board, AASHTO Materials Reference Laboratory, Maryland, July, 2004

12. Cooley Jr, L.A., Prowell, B.D., Hainin, M.R., Buchanan, M.S. and Harrington, J. *Bulk Specific Gravity Round-Robin using the CoreLok Vacuum Sealing Device*, NCAT Report, Alabama, November, 2002

13. Choubane, B., Upshaw, P.B., Sholar, G.A., Page, G.C. and Musselman, J.A. *Nuclear Density Readings and Core Densities: A Comparative Study;* TRR no.1654; pp70-78; Florida, 1999

14. Stroup-Gardiner, M. and Newcomb, D. *Statistical Evaluation of Nuclear Density Gauges under Field Conditions*, TRR No. 1178, 1988

15. Gedney, D.S., Shirley, E.G., Agent, K.R., Fuca, T.A. *Nuclear density gauge monitoring of asphalt concrete compaction*, Transportation Research Circular No. 321, Transportation Research Board, Washington, DC, 1987

16. Rogge, D.F. and Jackson, M.A. *Compaction and measurement of field density for Oregon open-graded (f-mix) asphalt pavement*, Oregon State University and Oregon Department of Transportation, Oregon, 1999

17. Sargand, S.M., Kim, S. and Farrington, S.P. *A Working Review of Available Non-Nuclear Equipment for Determining In-Place Density of Asphalt*, Ohio Department of Transportation, Ohio, 2005

18. Troxler Jr., W.F., Dep, L. Measurement of Longitudinal Joint Density in Asphalt pavements Using Nuclear and Nonnuclear Gauges, Transportation Research E-Circular No.E-C105, pp100-119, TRB, Washington DC, 2006

19. Huot, G.L., Verrette, Y. and Henault, G.G. *Nuclear measurement of bituminous pavement densities*, Canadian Tech Asphalt Assoc Proc, Vol 11, pp 115-134, 1966

20. Rao, C., Von Quintus, H.L. and Schmitt, R.L. *Calibration of Nonnuclear Density Gauge Data for Accurate In-Place Density Prediction*, Transportation Research Board 86th Annual Meeting, November, 2006

21. Esch, D.C. *Field Differences Testing: Nuclear vs. Pavement Core Density Tests*, Alaska Department of Transportation and Public Facilities, 1972

22. Romero, P. and Kuhnow, F. Evaluation of New Nonnuclear Pavement Density Gauges with Data from Field projects, TRR no.1813, Utah, 2002

23. Turochy, R.E. and Parker, F. *Comparisons of Contractor and State Transportation Agency Quality Assurance Test Results on Mat Density of Hot-Mix Asphalt Concrete*, Transportation Research Record no.2040, Alabama, 2007

24. Williams, S.G., Hall, K.D. *Critical Factors Affecting the Field Determination of hot mix asphalt Density Using Non-Nuclear Devices*, Transportation Research Board, 2008

25. Willoughby, K.A., Mahoney, J.P. *An Assessment of WSDOT's hot mix asphalt Quality Control and Assurance Requirements*, Washington State Department of Transportation, Washington DC, 2007

26. Padlo, P.T., Mahoney, J., Aultman-hall, L. *Differences of Nuclear Density Readings with Cores Cut from Compacted Roadways*, Final Report; Connecticut DOT, November, 2005

27. Smith, B.C., Diefender, B.K. *Comparing Nuclear and Nonnuclear Pavement Density Testing Devices*, TRB, Virginia, 2008

28. Sargand, S.M., Kim, S., and Farrington, S.P. *Non-Nuclear Density Gauge Comparative Study*, Ohio Department of Transportation, Ohio, August 2005

29. Prowell, B.D. and Dudley, M.C. *Evaluation of Measurement Techniques for Asphalt Pavement Density and Permeability*, TRR no.1789, TRB, Bituminous Paving Mixtures, Virginia, 2002

30. Parker, F. *Alabama DOT Implementation of hot mix asphalt Quality Control/Quality Assurance Procedures*

31. Schmitt, R. L., Whited, G.C.; Waelti, G.J., Hanna, A.S., Russell, J.S., Bahia, H.U. *Current initiatives in quality asphaltic pavement construction: The Wisconsin Experience*, Transportation Research Record No. 1543, p. 120-124, Transportation Research Board, Wisconsin, 1996

32. Hausman, J.J. and Buttlar, W.G. *Analysis of TransTech Model 300 Pavement Quality Indicator: Laboratory and Field Studies for determining asphalt pavement density*, Transportation Research Record No. 1813, Transportation Research Board, pp191-200, Illinois, 2002

33. Parikh, J. *Compaction control of bituminous concrete using nuclear density gauges: Final report*, Maryland Department of Transportation; Federal Highway Administration, 1990

34. Andrewski, D.H. *Density Measurements of hot mix asphalt (hot mix asphalt) Pavements*, Final Report; Indiana DOT, December, 2003

35. Burati, J.L. and Elzoghbi, G.B. *Differences of Nuclear Density Results With Core Densities*, Transportation Research Record No. 1126, Transportation Research Board, Washington D.C., 1987

36. Kvasnak, A.N., Williams, R.C., Ceylan, H. and Gopalakrishnan, K. *Investigation of Electromagnetic Gauges for Determining In-Place hot mix asphalt Density*, Iowa Department of Transportation, Iowa, 2007

37. Schmitt, R., Chetana, R. and Von Quintos, H.L. Non-Nuclear Density Testing Devices and Systems to Evaluate In-Place Asphalt Pavement Density, Wisconsin, May, 2006

38. TransTech Systems. *Assessment of Field Asphalt Density Gauge Data When Compared With Cores Processed According to AASHTO T-166*, TransTech Technical Note 0302, Rev 03, Schenectady, NY; March 24, 2004

39. Romero, P. *Evaluation of Non-Nuclear Gauges to Measure Density of Hot-Mix Asphalt Pavements*, University of Utah, July 18, 2002

40. Andrea Kvasnak, "*Non-Nuclear Gauge Study*"41.

41. Standard Method of Test for Bulk Specific Gravity of Compacted Asphalt Mixtures using Saturated Surface-Dry Specimens (AASHTO T 166)

42. Standard Test Method for Bulk Specific Gravity and Density of Non-Absorptive Compacted Bituminous Mixtures (ASTM D 2726)

43. Standard Specification for In-Place Density and Moisture Content of Soil and Soil-Aggregate by Nuclear Methods (Shallow Depth) (AASHTO T 310)

44. Standard Test Method for Density of Bituminous Concrete in Place by Nuclear Methods (ASTM D 2950)

45. Standard Method of Test for Theoretical Maximum Specific Gravity and Density of Bituminous Paving Mixtures (AASHTO T 209)

46. Standard Test Method for Theoretical Maximum Specific Gravity and Density of Bituminous Paving Mixtures (ASTM D 2041-03a)

47. Standard method of Test for Bulk Specific Gravity of Compacted Bituminous Mixtures using Paraffin-Coated Specimens (AASHTO T 275)

48. Standard Test method for Bulk Specific Gravity and Density of Compacted Bituminous Mixtures Using Coated Samples (ASTM D 1188)

49. Standard Method of Test for Bulk Specific Gravity and Density of Compacted Asphalt Mixtures using Automatic Vacuum Sealing Method (AASHTO T 331)

50. Standard Test Method for Bulk Specific Gravity and Density of Compacted Bituminous Mixtures Using Automatic Vacuum Sealing Method (ASTM D 6752)

51. Standard Method of Test for Sampling of Bituminous Paving mixtures (AASHTO T 168)

52. Standard Practice for Random Sampling of Construction Materials (ASTM D 3665-07)

53. CoreLok Operator’s Guide, InstroTek, Inc., July, 2003 Version 20

54. Oregon Standard Specifications for Construction, APWA Oregon Chapter, Oregon Department of Transportation, Vol 1, 2008

55. In-Place Density of Bituminous Mixes Using the Nuclear Moisture-Density Gauge FOP for WAQTC TM 8, Test Procedure WAQTC, October, 2005

56. Method of Test for Nuclear Density/Moisture gauge Calibration and Affect of Hot Substrate, ODOT TM 304

57. Method of Test for Differences of Nuclear Gauge Readings with Pavement Cores, ODOT TM 327

58. Test Method for Density of Bituminous Paving Mixtures in Place by the Electromagnetic Surface Contact Methods, ASTM D 7113-05, Vol-04.03, 2005

59. Standard Method of Test for Density of In-Place Hot Mix Asphalt (HMA) Pavement by Electronic Surface Contact Devices, AASHTO TP 68-04

APPENDIX

(Results from pooled t-tests)

The p-values in the following table suggest that the different nuclear gauges provide differing results when the average differences (between the densities from the SSD methods and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the control gauge and the densities from other methods of measurement) do not differ significantly between the two lifts at 90% level of confidence whereas, the average differences (between the SSD results and the other nuclear gauge used) differ significantly between the two lifts even at 97% level of confidence. Another observation is that the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 56%; and this result stays consistent irrespective of the nuclear gauge used. Overall, at 95% level of confidence, a majority of the p-values (10 of 16) from the table indicate that there is no significant difference between the average differences compared between the two lifts.

Hypothesis for the pooled t-test at 95% level of confidence:
True mean 'average difference' from the first set of first lift (μ_{1a}) and true mean 'average difference' from the first set of second lift (μ_{2a}):

H_0: $\mu_{1a} - \mu_{2a} = 0$

H_a: $\mu_{1a} - \mu_{2a} \neq 0$

Test: $t = \dfrac{\bar{X}_{1a} - \bar{X}_{2a}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A1 – p-values from the pooled t-tests comparing the various average differences across lifts 18-1 and 18-2 Lift 2 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
CoreLok - gauge (avg of 4 readings)		0.7962	0.5941				
CoreLok - gauge (avg of 2 readings)		0.6527	0.4424				
SSD Method A (before APAO) - gauge (avg of 4 readings)		0.2398	0.0250				
SSD Method A (before APAO) - gauge (avg of 2 readings)		0.3997	0.0139				
SSD Method A (after APAO) - gauge (avg of 4 readings)		0.1162	0.0003				
SSD Method A (after APAO) - gauge (avg of 2 readings)		0.2048	0.0003				
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)		0.3829	0.0050				
SSD Method C - gauge (avg of 2 readings)		0.5075	0.0035				

The p-values in the following table suggest that the average differences (between the densities from the SSD method the densities from the nuclear gauge) do not differ significantly between the two lifts at 76% level of confidence. Another observation is that the average differences (between the densities from the CoreLok and the densities from the nuclear gauge) do not differ significantly between the two lifts only at 97% level of confidence. Overall, at 95% level of confidence, a majority of the p-values (7 of 8) from the table indicate that there is no significant difference between the average differences compared between the two lifts.

Hypothesis for the pooled t-test at 95% level of confidence:

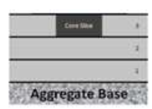
True mean 'average difference' from the first set of first lift (μ1a) and true mean 'average difference' from the third lift (μ3):

H$_0$: $\mu_{1a} - \mu_3 = 0$

H$_a$: $\mu_{1a} - \mu_3 \neq 0$

Test: $t = \dfrac{\bar{X}_{1a} - \bar{X}_3}{s_p/\sqrt{n}}$; where s_p=pooled standard deviation

Table A2 – p-values from the pooled t-tests comparing the various average differences across lifts 18-1 and 18-3 Lift 3 on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
CoreLok - gauge (avg of 4 readings)		0.0323					
CoreLok - gauge (avg of 2 readings)		0.0646					
SSD Method A (before APAO) - gauge (avg of 4 readings)		0.2470					
SSD Method A (before APAO) - gauge (avg of 2 readings)		0.5548					
SSD Method A (after APAO) - gauge (avg of 4 readings)		0.5848					
SSD Method A (after APAO) - gauge (avg of 2 readings)		0.9606					
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)		0.1542					
SSD Method C - gauge (avg of 2 readings)		0.3397					

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the cores and the densities from the nuclear gauges) are significantly different between the two lifts at 94% level of confidence in 17 of 20 comparisons. However, when the average of two nuclear gauge readings are taken instead of four to calculate the average differences (between the densities from the control gauge and the densities from the other methods of measurements on the cores), it can be observed that there is no significant difference between the average differences of the two lifts at 95% level of confidence.

For the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at a level of confidence as low as 65% (in 9 of 10 comparisons). For the same average differences compared between the two lifts, it must be noted that although both the electromagnetic gauges provide similar results, the PQI 301 provides stronger evidence than the PQI 300 of not being significantly different between the two lifts i.e., the level of confidence was as low as 63%.

Hypothesis for the pooled t-test at 95% level of confidence:

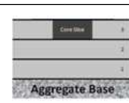
True mean correlation factor from first set of second lift (μ_{2a}) and true mean correlation factor from third lift (μ_3):

H_0: $\mu_{2a} - \mu_3 = 0$

H_a: $\mu_{2a} - \mu_3 \neq 0$

Test: $t = \dfrac{\overline{X}_{2a} - \overline{X}_3}{s_p \big/ \sqrt{n}}$; where s_p=pooled standard deviation

Table A3 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 2 and 18-3 Lift 3 on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0029	0.0071		0.7448		0.6352	
CoreLok - gauge (avg of 2 readings)	0.0158	0.0617					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0108	0.0222		0.0827		0.3796	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0170	0.1530					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.0042	0.0081		0.3519		0.9958	
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.0274	0.0656					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.0156	0.0281		0.3674		0.9406	
SSD Method A (APAO) - gauge (avg of 2 readings)	0.0546	0.1415					
SSD Method C - gauge (avg of 4 readings)	0.0028	0.0051		0.4723		0.8541	
SSD Method C - gauge (avg of 2 readings)	0.0192	0.0380					

The p-values in the following table are evidence to indicate that the average differences (between the densities from the various tests on cores and the densities from the two nuclear gauges) do not differ significantly at a level of confidence as low as 86%. This is true for all the 16 comparisons made. Another observation is that both the nuclear gauges provide similar results when the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) are compared between the two lifts.

Hypothesis for the pooled t-test at 95% level of confidence:

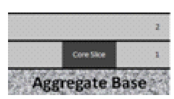
True mean correlation factor from first set of first lift (μ_{1a}) and true mean correlation factor from second set of first lift (μ_{1b}):

H_0: $\mu_{1a} - \mu_{1b} = 0$

H_a: $\mu_{1a} - \mu_{1b} \neq 0$

Test: $t = \dfrac{\bar{X}_{1a} - \bar{X}_{1b}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A4 – p-values from the pooled t-tests comparing the various average differences across lifts 18-1 and 18-2 Lift 1 on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)		0.2822	0.6973				
CoreLok - gauge (avg of 2 readings)		0.2465	0.8136				
SSD Method A (before APAO) - gauge (avg of 4 readings)		0.9237	0.4282				
SSD Method A (before APAO) - gauge (avg of 2 readings)		0.9486	0.3514				
SSD Method A (after APAO) - gauge (avg of 4 readings)		0.4022	0.1661				
SSD Method A (after APAO) - gauge (avg of 2 readings)		0.5108	0.1418				
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)		0.4739	0.2091				
SSD Method C - gauge (avg of 2 readings)		0.5811	0.1773				

The p-values from the following table suggest that the different nuclear gauges provide differing results when the average differences (between the densities from the SSD methods and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the control gauge and the densities from other methods of measurement) do not differ significantly between the two lifts at 86% level of confidence whereas, the average differences (between the SSD results and the other nuclear gauge used) differ significantly between the two lifts even at 97% level of confidence. Another observation is that the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 86%; and this result stays consistent irrespective of the nuclear gauge used. Overall, at 95% level of confidence, a majority of the p-values (6 of 8) from the table indicate that there is no significant difference between the average differences compared between the two lifts.

Hypothesis for the pooled t-test at 95% level of confidence:

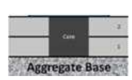
True mean correlation factor from first set of first lift (μ_{1a}) and true mean correlation factor from combination of first and second lifts (μ_{12}):

H_0: $\mu_{1a} - \mu_{12} = 0$

H_a: $\mu_{1a} - \mu_{12} \neq 0$

Test: $t = \dfrac{\bar{X}_{1a} - \bar{X}_{12}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A5 – p-values from the pooled t-tests comparing the various average differences across lifts 18-1 and 18-2 uncut on the project having multiple lifts with same mix design (OR 18)

 							
Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)		0.1911	0.7661				
CoreLok - gauge (avg of 2 readings)		0.1406	0.9259				
SSD Method A (before APAO) - gauge (avg of 4 readings)		0.2359	0.0236				
SSD Method A (before APAO) - gauge (avg of 2 readings)		0.4005	0.0217				
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

The p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauge) differ significantly between the two lifts even at 99% level of confidence. This is true for all the 8 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

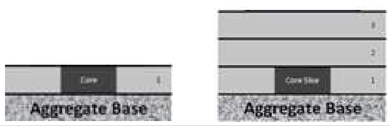
True mean correlation factor from first set of first lift (μ_{1a}) and true mean correlation factor from third set of first lift (μ_{1c}):

H_0: $\mu_{1a} - \mu_{1c} = 0$

H_a: $\mu_{1a} - \mu_{1c} \neq 0$

Test: $t = \dfrac{\bar{X}_{1a} - \bar{X}_{1c}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A6 – p-values from the pooled t-tests comparing the various average differences across lifts 18-1 and 18-3 Lift 1 on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)		0.0006					
CoreLok - gauge (avg of 2 readings)		0.0020					
SSD Method A (before APAO) - gauge (avg of 4 readings)		0.0052					
SSD Method A (before APAO) - gauge (avg of 2 readings)		0.0085					
SSD Method A (after APAO) - gauge (avg of 4 readings)		0.0001					
SSD Method A (after APAO) - gauge (avg of 2 readings)		0.0005					
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)		0.0002					
SSD Method C - gauge (avg of 2 readings)		0.0007					

The p-values in the following table suggest that the average differences (between the densities from the SSD method the densities from the nuclear gauge) do not differ significantly between the two lifts at 67% level of confidence. This is true for all the 8 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from first set of first lift (μ_{1a}) and true mean correlation factor from first set of second lift (μ_{2a}):

H_0: $\mu_{1a} - \mu_{2a} = 0$

H_a: $\mu_{1a} - \mu_{2a} \neq 0$

Test: $t = \dfrac{\bar{X}_{1a} - \bar{X}_{2a}}{s_p/\sqrt{n}}$; where s_p=pooled standard deviation

Table A7 – p-values from the pooled t-tests comparing the various average differences across lifts 18-1 and 18-3 Lift 2 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)		0.3303					
CoreLok - gauge (avg of 2 readings)		0.4659					
SSD Method A (before APAO) - gauge (avg of 4 readings)		0.5296					
SSD Method A (before APAO) - gauge (avg of 2 readings)		0.7181					
SSD Method A (after APAO) - gauge (avg of 4 readings)		0.7699					
SSD Method A (after APAO) - gauge (avg of 2 readings)		0.9511					
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)		0.7849					
SSD Method C - gauge (avg of 2 readings)		0.9606					

The p-values in the following table suggest that the CoreLok and the SSD tests on cores provide differing results when the average differences (between the densities from the tests on cores and the densities from the nuclear gauge) are compared between the two lifts. In other words, all the average differences (between the densities from the CoreLok and the densities from the nuclear gauge) differ significantly between the two lifts even at 99% level of confidence; whereas, the average differences (between the densities from the SSD test and the densities from the nuclear gauge) do not differ significantly between the two lifts at a level of confidence as low as 62%.

Hypothesis for the pooled t-test at 95% level of confidence:

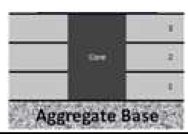
True mean correlation factor from first set of first lift (μ_{1a}) and true mean correlation factor from combination of first, second, and third lifts (μ_{123}):

H_0: $\mu_{1a} - \mu_{123} = 0$

H_a: $\mu_{1a} - \mu_{123} \neq 0$

Test: $t = \dfrac{\bar{X}_{1a} - \bar{X}_{123}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A8 – p-values from the pooled t-tests comparing the various average differences across lifts 18-1 and 18-3 uncut on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)		0.0004					
CoreLok - gauge (avg of 2 readings)		0.0013					
SSD Method A (before APAO) - gauge (avg of 4 readings)		0.3835					
SSD Method A (before APAO) - gauge (avg of 2 readings)		0.6426					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 58%. This is true for all the 30 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at a level of confidence as low as 64%. This is true for all the 10 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from second set of first lift (μ_{1b}) and true mean correlation factor from first set of second lift (μ_{2a}):

H_0: $\mu_{1b} - \mu_{2a} = 0$

H_a: $\mu_{1b} - \mu_{2a} \neq 0$

Test: $t = \frac{\bar{X}_{1b} - \bar{X}_{2a}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A9 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 1 and 18-2 Lift 2 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.3075	0.3074	0.3622	0.2646		0.3230	
CoreLok - gauge (avg of 2 readings)	0.2863	0.3400	0.3824				
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.4620	0.4541	0.5119	0.3683		0.4348	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.4218	0.4894	0.5314				
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.9100	0.8687	0.8167	0.8670		0.9451	
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.9799	0.8405	0.7927				
SSD Method A (APAO) - gauge (avg of 4 readings)	0.8194	0.7731	0.7326	0.9984		0.9461	
SSD Method A (APAO) - gauge (avg of 2 readings)	0.8751	0.7612	0.7135				
SSD Method C - gauge (avg of 4 readings)	0.8194	0.8531	0.9312	0.5912		0.6892	
SSD Method C - gauge (avg of 2 readings)	0.7386	0.8998	0.9654				

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 55%. This is true for all the 12 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at a level of confidence as low as 60%. This is true for all the 4 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from second set of first lift (μ_{1b}) and true mean correlation factor from combination of first and second lifts (μ_{12}):

H_0: $\mu_{1b} - \mu_{12} = 0$

H_a: $\mu_{1b} - \mu_{12} \neq 0$

Test: $t = \dfrac{\overline{X}_{1b} - \overline{X}_{12}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A10 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 1 and 18-2 uncut on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.8176	0.8153	0.8421	0.7828		0.8080	
CoreLok - gauge (avg of 2 readings)	0.8009	0.8256	0.8509				
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.4938	0.4926	0.5611	0.4046		0.4787	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.4532	0.5156	0.5867				
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) differ significantly between the two lifts even at 97% level of confidence. This is true for all the 20 comparisons made.

For the electromagnetic gauges, the p-values in the following table suggest that the CoreLok and the SSD tests on cores provide differing results when the average differences (between the densities from the tests on cores and the densities from the electromagnetic gauge) are compared between the two lifts. In other words, all the average differences (between the densities from the CoreLok and the densities from the nuclear gauge) do not differ significantly between the two lifts even at 92% level of confidence; whereas, the average differences (between the densities from the SSD test and the densities from the nuclear gauge) differ significantly between the two lifts even at a level of confidence as high as 98%.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from second set of first lift (μ_{1b}) and true mean correlation factor from third set of first lift (μ_{1c}):

H_0: $\mu_{1b} - \mu_{1c} = 0$

H_a: $\mu_{1b} - \mu_{1c} \neq 0$

Test: $t = \dfrac{\bar{X}_{1b} - \bar{X}_{1c}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A11 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 1 and 18-3 Lift 1 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0082	0.0141		0.1768		0.1019	
CoreLok - gauge (avg of 2 readings)	0.0100	0.0347					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0046	0.0093		0.1736		0.0848	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0061	0.0242					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.0002	0.0005		0.0065		0.0044	
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.0002	0.0021					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.0005	0.0009		0.0086		0.0056	
SSD Method A (APAO) - gauge (avg of 2 readings)	0.0005	0.0033					
SSD Method C - gauge (avg of 4 readings)	0.0004	0.0007		0.0104		0.0068	
SSD Method C - gauge (avg of 2 readings)	0.0004	0.0028					

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 73%. This is true for all the 20 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at a level of confidence as low as 87%. This is true for all the 10 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

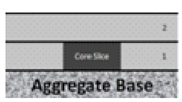
True mean correlation factor from second set of first lift (μ_{1b}) and true mean correlation factor from second set of second lift (μ_{2b}):

H_0: $\mu_{1b} - \mu_{2b} = 0$

H_a: $\mu_{1b} - \mu_{2b} \neq 0$

Test: $t = \dfrac{\overline{X}_{1b} - \overline{X}_{2b}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A12 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 1 and 18-3 Lift 2 on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.8243	0.9955		0.1766		0.4169	
CoreLok - gauge (avg of 2 readings)	0.9002	0.7807					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.3828	0.5397		0.4029		0.7891	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.4197	0.7860					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.2849	0.3996		0.6869		0.9299	
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.3043	0.6019					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.2712	0.3867		0.6866		0.9120	
SSD Method A (APAO) - gauge (avg of 2 readings)	0.2904	0.5869					
SSD Method C - gauge (avg of 4 readings)	0.3338	0.4528		0.6467		0.9896	
SSD Method C - gauge (avg of 2 readings)	0.3557	0.6562					

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts at a 93% level of confidence. This is true for all the 20 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at a level of confidence as low as 85%. This is true for 9 of 10 comparisons made. The average difference between the densities from the SSD method A and the densities from the PQI 300 differs significantly between the 2 lifts even at 97% level of confidence.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

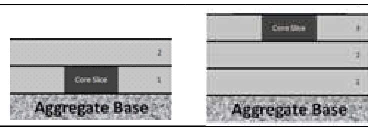
True mean correlation factor from second set of first lift (μ_{1b}) and true mean correlation factor from third lift (μ_3):

H_0: $\mu_{1b} - \mu_3 = 0$

H_a: $\mu_{1b} - \mu_3 \neq 0$

Test: $t = \dfrac{\bar{X}_{1b} - \bar{X}_3}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A13 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 1 and 18-3 Lift 3 on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.4860	0.6859		0.1673		0.5123	
CoreLok - gauge (avg of 2 readings)	0.5691	0.9857					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.2475	0.4164		0.0264		0.1524	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.3019	0.7627					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.1445	0.2279		0.5162		0.9481	
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.1703	0.4523					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.2099	0.3230		0.3981		0.8965	
SSD Method A (APAO) - gauge (avg of 2 readings)	0.2483	0.5887					
SSD Method C - gauge (avg of 4 readings)	0.0743	0.1143		0.9557		0.5961	
SSD Method C - gauge (avg of 2 readings)	0.0830	0.2463					

The CoreLok and the SSD tests on cores provide differing results when the average differences (between the densities from the tests on cores and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) differ significantly between the two lifts at 89% level of confidence; whereas, the average differences (between the densities from the SSD test and the densities from the nuclear gauges) do not differ significantly between the two lifts at a level of confidence as low as 70%.

For the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at a level of confidence as low as 71%. This is true for all 4 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

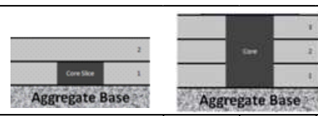
True mean correlation factor from second set of first lift (μ_{1b}) and true mean correlation factor from combination of first, second, and third lifts (μ_{123}):

H_0: $\mu_{1b} - \mu_{123} = 0$

H_a: $\mu_{1b} - \mu_{123} \neq 0$

Test: $t = \dfrac{\bar{X}_{1b} - \bar{X}_{123}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A14 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 1 and 18-3 uncut on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0246	0.0404		0.5203		0.2955	
CoreLok - gauge (avg of 2 readings)	0.0306	0.1089					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.3029	0.4533		0.3981		0.8037	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.3732	0.7572					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) differ significantly between the two lifts even at a level of confidence as high as 99%. This is true for all the 20 comparisons made.

For the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) differ significantly between the two lifts at a level of confidence of 92%. This is true for all the 10 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from first set of second lift (μ_{2a}) and true mean correlation factor from third set of first lift (μ_{1c}):

H_0: $\mu_{1c} - \mu_{2a} = 0$

H_a: $\mu_{1c} - \mu_{2a} \neq 0$

Test: $t = \dfrac{\bar{X}_{1c} - \bar{X}_{2a}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A15 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 2 and 18-3 Lift 1 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0006	0.0011		0.0442		0.0187	
CoreLok - gauge (avg of 2 readings)	0.0010	0.0029					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0003	0.0008		0.0588		0.0232	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0006	0.0032					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.0000	0.0000		0.0049		0.0010	
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.0000	0.0000					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.0000	0.0000		0.0071		0.0018	
SSD Method A (APAO) - gauge (avg of 2 readings)	0.0000	0.0001					
SSD Method C - gauge (avg of 4 readings)	0.0000	0.0000		0.0154		0.0033	
SSD Method C - gauge (avg of 2 readings)	0.0000	0.0001					

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence of 89%. This is true for all the 20 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at a level of confidence as low as 68%. This is true for all the 10 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from first set of second lift (μ_{2a}) and true mean correlation factor from second set of second lift (μ_{2b}):

H_0: $\mu_{2a} - \mu_{2b} = 0$

H_a: $\mu_{2a} - \mu_{2b} \neq 0$

Test: $t = \dfrac{\overline{X}_{2a} - \overline{X}_{2b}}{s_p/\sqrt{n}}$; where s_p=pooled standard deviation

Table A16 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 2 and 18-3 Lift 2 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.2360	0.3677		0.5625		0.9637	
CoreLok - gauge (avg of 2 readings)	0.2736	0.6110					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.1141	0.2058		0.8344		0.6971	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.1251	0.3886					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.1735	0.2633		0.5855		0.9666	
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.2205	0.4411					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.1354	0.2127		0.6833		0.8574	
SSD Method A (APAO) - gauge (avg of 2 readings)	0.1727	0.3706					
SSD Method C - gauge (avg of 4 readings)	0.3226	0.4513		0.3684		0.7327	
SSD Method C - gauge (avg of 2 readings)	0.4042	0.6692					

The CoreLok and the SSD tests on cores provide differing results when the average differences (between the densities from the tests on cores and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) differ significantly between the two lifts at 99% level of confidence; whereas, the average differences (between the densities from the SSD test and the densities from the nuclear gauges) do not differ significantly between the two lifts at a 96% level of confidence.

For the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at a 94% level of confidence. This is true for all 4 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

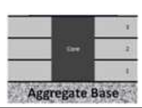
True mean correlation factor from first set of second lift (μ_{2a}) and true mean correlation factor from combination of first, second, and third lifts (μ_{123}):

H_0: $\mu_{2a} - \mu_{123} = 0$

H_a: $\mu_{2a} - \mu_{123} \neq 0$

Test: $t = \dfrac{\bar{X}_{2a} - \bar{X}_{123}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A17 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 2 and 18-3 uncut on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0000	0.0001		0.1019		0.0309	
CoreLok - gauge (avg of 2 readings)	0.0002	0.0007					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0419	0.0935		0.8249		0.6714	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0700	0.2790					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 55%. This is true for all the 12 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at a level of confidence as low as 48%. This is true for all the 4 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from first set of second lift (μ_{2a}) and true mean correlation factor from combination of first and second lifts (μ_{12}):

H_0: $\mu_{2a} - \mu_{12} = 0$

H_a: $\mu_{2a} - \mu_{12} \neq 0$

Test: $t = \dfrac{\bar{X}_{2a} - \bar{X}_{12}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A18 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 Lift 2 and 18-2 uncut on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.1701	0.1312	0.1708	0.2580		0.2108	
CoreLok - gauge (avg of 2 readings)	0.1472	0.1793	0.2099				
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.8866	0.8800	0.8980	0.8703		0.8851	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.8736	0.8977	0.9035				
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) differ significantly between the two lifts at a 92% level of confidence. This is true for all the 20 comparisons made.

For the electromagnetic gauges, the p-values in the following table suggest that the presence or absence of sand on the surface of the locations provides differing results when the average differences are compared between the two lifts. In sanded condition, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges when the locations were sanded) differ significantly between the two lifts at 97% level of confidence. This is true for all the 10 comparisons made. In un-sanded condition, the p-values suggest that the average differences do not differ significantly at 92% level of confidence. This is true for all the 10 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

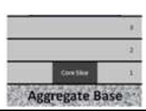
True mean correlation factor from third set of first lift (μ_{1c}) and true mean correlation factor from second set of second lift (μ_{2b}):

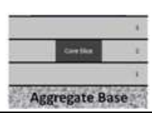
H_0: $\mu_{1c} - \mu_{2b} = 0$

H_a: $\mu_{1c} - \mu_{2b} \neq 0$

Test: $t = \dfrac{\bar{X}_{1c} - \bar{X}_{2b}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A19 – p-values from the pooled t-tests comparing the various average differences across lifts 18-3 Lift 1 and 18-3 Lift 2 on the project having multiple lifts with same mix design (OR 18)





Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0179	0.0212		0.0298	0.1477	0.0296	0.1482
CoreLok - gauge (avg of 2 readings)	0.0181	0.0273					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0523	0.0636		0.0697	0.2545	0.0682	0.2578
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0541	0.0742					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.0155	0.0201		0.0143	0.2959	0.0151	0.2960
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.0135	0.0275					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.0161	0.0224		0.0137	0.0903	0.0140	0.0870
SSD Method A (APAO) - gauge (avg of 2 readings)	0.0146	0.0288					
SSD Method C - gauge (avg of 4 readings)	0.0125	0.0168		0.0133	0.0933	0.0136	0.0915
SSD Method C - gauge (avg of 2 readings)	0.0109	0.0225					

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) differ significantly between the two lifts even at 98% level of confidence. This is true for all the 20 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) differ significantly between the two lifts at a level of confidence of 94%. This is true for 18 of 20 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from third set of first lift (μ_{1c}) and true mean correlation factor from third lift (μ_3):

H_0: $\mu_{1c} - \mu_3 = 0$

H_a: $\mu_{1c} - \mu_3 \neq 0$

Test: $t = \dfrac{\overline{X}_{1c} - \overline{X}_3}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A20 – p-values from the pooled t-tests comparing the various average differences across lifts 18-3 Lift 1 and 18-3 Lift 3 on the project having multiple lifts with same mix design (OR 18)

							
Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0099	0.0107		0.0317	0.0547	0.0304	0.0564
CoreLok - gauge (avg of 2 readings)	0.0133	0.0138					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0116	0.0142		0.0094	0.0170	0.0088	0.0183
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0168	0.0169					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.0000	0.0000		0.0009	0.2795	0.0010	0.2797
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.0001	0.0003					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.0001	0.0001		0.0013	0.0014	0.0014	0.0010
SSD Method A (APAO) - gauge (avg of 2 readings)	0.0002	0.0007					
SSD Method C - gauge (avg of 4 readings)	0.0002	0.0009		0.0043	0.0081	0.0046	0.0091
SSD Method C - gauge (avg of 2 readings)	0.0008	0.0022					

The CoreLok and the SSD tests on cores provide differing results when the average differences (between the densities from the tests on cores and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 71%; whereas, the average differences (between the densities from the SSD test and the densities from the nuclear gauges) differ significantly between the two lifts at a 95% level of confidence.

For the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at a 93% level of confidence. This is true for all 8 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

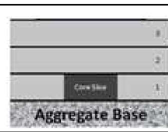
True mean correlation factor from third set of first lift (μ_{1c}) and true mean correlation factor from combination of first, second, and third lifts (μ_{123}):

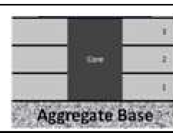
H_0: $\mu_{1c} - \mu_{123} = 0$

H_a: $\mu_{1c} - \mu_{123} \neq 0$

Test: $t = \dfrac{\overline{X}_{1c} - \overline{X}_{123}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A21 – p-values from the pooled t-tests comparing the various average differences across lifts 18-3 Lift 1 and 18-3 uncut on the project having multiple lifts with same mix design (OR 18)





Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.2952	0.3175		0.4326	0.4888	0.4404	0.5090
CoreLok - gauge (avg of 2 readings)	0.3197	0.3306					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0252	0.0374		0.0760	0.1205	0.0753	0.1443
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0360	0.0416					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 36%. This is true for all the 20 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at a level of confidence as low as 87%. This is true for all the 20 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from second set of second lift (μ_{2b}) and true mean correlation factor from third lift (μ_3):

H_0: $\mu_{2b} - \mu_3 = 0$

H_a: $\mu_{2b} - \mu_3 \neq 0$

Test: $t = \dfrac{\bar{X}_{2b} - \bar{X}_3}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A22 – p-values from the pooled t-tests comparing the various average differences across lifts 18-3 Lift 2 and 18-3 Lift 3 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.7191	0.7286		0.7137	0.7439	0.7190	0.7404
CoreLok - gauge (avg of 2 readings)	0.7169	0.7389					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.9177	0.9228		0.3126	0.1363	0.3197	0.1337
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.9171	0.9261					
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.9638	0.9651		0.9639	0.5373	0.9643	0.5299
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.9634	0.9664					
SSD Method A (APAO) - gauge (avg of 4 readings)	0.8085	0.8175		0.8102	0.4368	0.8118	0.4269
SSD Method A (APAO) - gauge (avg of 2 readings)	0.8072	0.8236					
SSD Method C - gauge (avg of 4 readings)	0.6495	0.6609		0.6498	0.8442	0.6531	0.8408
SSD Method C - gauge (avg of 2 readings)	0.6448	0.6708					

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) differ significantly between the two lifts at 91% level of confidence. However, the average differences (between the densities from the SSD and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 3%.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at 92% level of confidence. This is true for all the 8 comparisons made.

Overall, at 95% level of confidence, all the comparisons made in the following table indicate that there is no significant difference between the average differences of the two lifts.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from second set of second lift (μ_{2b}) and true mean correlation factor from combination of first, second, and third lifts (μ_{123}):

H_0: $\mu_{2b} - \mu_{123} = 0$

H_a: $\mu_{2b} - \mu_{123} \neq 0$

Test: $t = \dfrac{\bar{X}_{2b} - \bar{X}_{123}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A23 – p-values from the pooled t-tests comparing the various average differences across lifts 18-3 Lift 2 and 18-3 uncut on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0606	0.0654		0.0838	0.3295	0.0832	0.3141
CoreLok - gauge (avg of 2 readings)	0.0609	0.0847					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.9857	0.9867		0.9943	0.6577	0.9848	0.6936
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.9779	0.9896					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

The CoreLok and the SSD tests on cores provide differing results when the average differences (between the densities from the tests on cores and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) differ significantly between the two lifts at 97% level of confidence; whereas, the average differences (between the densities from the SSD test and the densities from the nuclear gauges) do not differ significantly between the two lifts even at a level of confidence as low as 13%.

For the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at 94% level of confidence. This is true for all 8 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

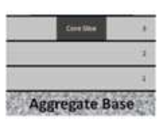
True mean correlation factor from third lift (μ_3) and true mean correlation factor from combination of first, second, and third lifts (μ_{123}):

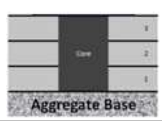
H_0: $\mu_3 - \mu_{123} = 0$

H_a: $\mu_3 - \mu_{123} \neq 0$

Test: $t = \dfrac{\bar{X}_3 - \bar{X}_{123}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A24 – p-values from the pooled t-tests comparing the various average differences across lifts 18-3 Lift 3 and 18-3 uncut on the project having multiple lifts with same mix design (OR 18)





Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0086	0.0069		0.0638	0.0835	0.0622	0.0792
CoreLok - gauge (avg of 2 readings)	0.0215	0.0105					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.9051	0.8741		0.3128	0.2419	0.2981	0.2725
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.9277	0.9164					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

The CoreLok and the SSD tests on cores provide differing results when the average differences (between the densities from the tests on cores and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) differ significantly between the two lifts even at 99% level of confidence; whereas, the average differences (between the densities from the SSD test and the densities from the nuclear gauges) do not differ significantly between the two lifts at 97% level of confidence.

For the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at 87% level of confidence. This is true for all 4 comparisons made.

Overall, the results are similar irrespective of the gauge used.

Hypothesis for the pooled t-test at 95% level of confidence:

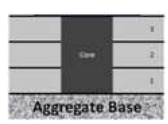
True mean correlation factor from combination of first and second lift (μ_{12}) and true mean correlation factor from combination of first, second, and third lifts (μ_{123}):

H_0: $\mu_{12} - \mu_{123} = 0$

H_a: $\mu_{12} - \mu_{123} \neq 0$

Test: $t = \dfrac{\overline{X}_{12} - \overline{X}_{123}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A25 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 uncut and 18-3 uncut on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0008	0.0012		0.3260		0.1338	
CoreLok - gauge (avg of 2 readings)	0.0014	0.0085					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0330	0.0880		0.7514		0.7316	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0611	0.2743					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) differ significantly between the two lifts even at 98% level of confidence. This is true for all the 8 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) differ significantly between the two lifts at a level of confidence of 88%. However, at 95% level of confidence 3 out of 4 comparisons indicate that there is no significant difference.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from combination of first and second lifts (μ_{12}) and true mean correlation factor from third set of first lift (μ_{1c}):

H_0: $\mu_{1c} - \mu_{12} = 0$

H_a: $\mu_{1c} - \mu_{12} \neq 0$

Test: $t = \dfrac{\overline{X}_{1c} - \overline{X}_{12}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A26 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 uncut and 18-3 Lift 1 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860)	(Troxler 3430 # 38806)	(Troxler 3440 # 25714)	PQI 300 (sand)	PQI 300 (no sand)	PQI 301 (sand)	PQI 301 (no sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0021	0.0038		0.1152		0.0504	
CoreLok - gauge (avg of 2 readings)	0.0035	0.0102					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0006	0.0014		0.0638		0.0252	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0013	0.0039					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts at 88% level of confidence.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts even at 82% level of confidence. This is true for all the 8 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from combination of first and second lifts (μ_{12}) and true mean correlation factor from second set of second lift (μ_{2b}):

H_0: $\mu_{2b} - \mu_{12} = 0$

H_a: $\mu_{2b} - \mu_{12} \neq 0$

Test: $t = \dfrac{\overline{X}_{2b} - \overline{X}_{12}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A27 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 uncut and 18-3 Lift 2 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.6316	0.8322		0.1897		0.4241	
CoreLok - gauge (avg of 2 readings)	0.7057	0.8827					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.1238	0.2190		0.7534		0.7600	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.1362	0.4120					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

The CoreLok and the SSD tests on cores provide differing results when the average differences (between the densities from the tests on cores and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) do not differ significantly between the two lifts even at 88% level of confidence; whereas, the average differences (between the densities from the SSD test and the densities from the nuclear gauges) differ significantly between the two lifts at 89% level of confidence.

For the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at 96% level of confidence. This is true for all 4 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from third lift (μ_3) and true mean correlation factor from first and second lifts (μ_{12}):

H_0: $\mu_3 - \mu_{12} = 0$

H_a: $\mu_3 - \mu_{12} \neq 0$

Test: $t = \dfrac{\bar{X}_3 - \bar{X}_{12}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A28 – p-values from the pooled t-tests comparing the various average differences across lifts 18-2 uncut and 18-3 Lift 3 on the project having multiple lifts with same mix design (OR 18)

Average difference compared	(Troxler 3430 # 30860) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 25714) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 300 (no sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)	PQI 301 (no sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.1202	0.2386		0.1367		0.4840	
CoreLok - gauge (avg of 2 readings)	0.1893	0.6892					
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0027	0.0111		0.0456		0.2843	
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0079	0.1018					
SSD Method A (after APAO) - gauge (avg of 4 readings)							
SSD Method A (after APAO) - gauge (avg of 2 readings)							
SSD Method A (APAO) - gauge (avg of 4 readings)							
SSD Method A (APAO) - gauge (avg of 2 readings)							
SSD Method C - gauge (avg of 4 readings)							
SSD Method C - gauge (avg of 2 readings)							

The p-values in the following table suggest that the different nuclear gauges provide differing results when the average differences (between the densities from the SSD methods and the densities from the nuclear gauges) are compared between the two lifts. In other words, all the average differences (between the densities from the Troxler 3430 # 38996 and the densities from other methods of measurement) do not differ significantly between the two lifts even at 59% level of confidence whereas, the average differences (between the SSD results and the other nuclear gauge used) differ significantly between the two lifts even at 97% level of confidence. Another observation is that the average differences (between the densities from the CoreLok and the densities from the nuclear gauges) do not differ significantly between the two lifts even at 89% level of confidence; and this result stays consistent irrespective of the nuclear gauge used. Overall, at 95% level of confidence, a majority of the p-values (10 of 18) from the table indicate that there is no significant difference between the average differences compared between the two lifts.

Hypothesis for the pooled t-test at 95% level of confidence:

True mean correlation factor from first set of first lift (μ_{1a}) and true mean correlation factor from combination of first and second lifts (μ_{12}):

H_0: $\mu_{1a} - \mu_{12} = 0$

H_a: $\mu_{1a} - \mu_{12} \neq 0$

Test: $t = \dfrac{\overline{X}_{1a} - \overline{X}_{12}}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A29 – p-values from the pooled t-tests comparing the various average differences across lifts 5-1 and 5-2 on the project having two lifts with different mix designs (I-5)

Pooled t-tests comparing correlation factors between lifts 1 and 2

Average difference compared	(Troxler 3430 # 38996) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(CPN MC3 Portaprobe - 7443) Pooled t-test (p-value)
SSD Method A-nuclear gauge (avg of 4 readings)	0.4119	0.0006	0.0194
SSD Method A-nuclear gauge (avg of 2 readings)	0.6126	0.0014	0.0056
CoreLok-nuclear gauge (avg of 4 readings)	0.8442	0.1090	0.1896
CoreLok-nuclear gauge (avg of 2 readings)	0.9911	0.1466	0.1084
SSD Method C-nuclear gauge (avg of 4 readings)	0.4165	0.0009	0.0272
SSD Method C-nuclear gauge (avg of 2 readings)	0.5965	0.0018	0.0084

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts at 93% level of confidence.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at 95% level of confidence. This is observed in 4 of 6 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

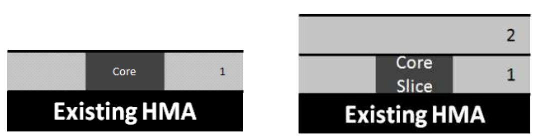
True mean correlation factor from first set of inlay (μ_{Ia}) and true mean correlation factor from second set of inlay (μ_{Ib}):

H$_O$: $\mu_{Ia} - \mu_{Ib} = 0$

H$_a$: $\mu_{Ia} - \mu_{Ib} \neq 0$

Test: $t = \dfrac{\bar{X}_{Ia} - \bar{X}_{Ib}}{s_p / \sqrt{n}}$; where s_p = pooled standard deviation

Table A30 – p-values from the pooled t-tests comparing the various average differences across lifts 140-1 and 140-2 Lift 1 on the mill and fill (inlay) plus overlay project having one mix design (OR 140)



Average difference compared	(Troxler 3440 # 38996) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 35601) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0954	0.4320	0.3462	0.9130	0.0947
CoreLok - gauge (avg of 2 readings)	0.0978	0.3776	0.3901		
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0752	0.4136	0.3307	0.8163	0.0190
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0780	0.3510	0.3818		
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.0699	0.4049	0.3278	0.6271	0.0337
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.0680	0.3176	0.3785		
SSD Method A (APAO) - gauge (avg of 4 readings)					
SSD Method A (APAO) - gauge (avg of 2 readings)					
SSD Method C - gauge (avg of 4 readings)	0.0505	0.2433	0.2172	0.8327	0.0772
SSD Method C - gauge (avg of 2 readings)	0.0505	0.2196	0.2896		

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts at 95% level of confidence. This is observed in 20 of 24 comparisons.

However, for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) differ significantly between the two lifts at 95% level of confidence. This is observed in 6 of 8 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

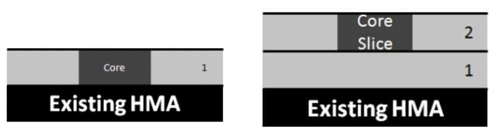
True mean correlation factor from first set of inlay (μ_{Ia}) and true mean correlation factor from overlay (μ_O):

H_O: $\mu_{Ia} - \mu_O = 0$

H_a: $\mu_{Ia} - \mu_O \neq 0$

Test: $t = \dfrac{\bar{X}_{Ia} - \bar{X}_O}{s_p / \sqrt{n}}$; where s_p = pooled standard deviation

Table A31 – p-values from the pooled t-tests comparing the various average differences across lifts 140-1 and 140-2 Lift 2 on the mill and fill (inlay) plus overlay project having one mix design (OR 140)



Average difference compared	(Troxler 3440 # 38996) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 35601) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.1250	0.4318	0.3589	0.9404	0.1844
CoreLok - gauge (avg of 2 readings)	0.1226	0.3729	0.3938		
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0023	0.9454	0.4516	0.0359	0.0011
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0065	0.6522	0.5932		
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.0019	0.8799	0.4297	0.0263	0.0039
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.0038	0.5489	0.5788		
SSD Method A (APAO) - gauge (avg of 4 readings)					
SSD Method A (APAO) - gauge (avg of 2 readings)					
SSD Method C - gauge (avg of 4 readings)	0.0055	0.9500	0.5635	0.0207	0.0029
SSD Method C - gauge (avg of 2 readings)	0.0107	0.6839	0.6877		

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts at 95% level of confidence. This is observed in 10 of 12 comparisons.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at 95% level of confidence. This is observed in 3 of 4 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

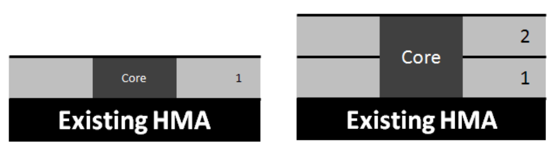
True mean correlation factor from first set of inlay (μ_{Ia}) and true mean correlation factor from both lifts i.e., inlay and overlay (μ_{IO}):

H_O: $\mu_{Ia} - \mu_{IO} = 0$

H_a: $\mu_{Ia} - \mu_{IO} \neq 0$

Test: $t = \dfrac{\overline{X}_{Ia} - \overline{X}_{IO}}{s_p / \sqrt{n}}$; where s_p = pooled standard deviation

Table A32 – p-values from the pooled t-tests comparing the various average differences across lifts 140-1 and 140-2 uncut on the mill and fill (inlay) plus overlay project having one mix design (OR 140)



Average difference compared	(Troxler 3440 # 38996)	(Troxler 3430 # 38806)	(Troxler 3440 # 35601)	PQI 300 (sand)	PQI 301 (sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.0865	0.6706	0.3147	0.6341	0.1912
CoreLok - gauge (avg of 2 readings)	0.1022	0.6318	0.3458		
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.0083	0.5240	0.1389	0.2930	0.0042
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.0116	0.4510	0.1300		
SSD Method A (after APAO) - gauge (avg of 4 readings)					
SSD Method A (after APAO) - gauge (avg of 2 readings)					
SSD Method A (APAO) - gauge (avg of 4 readings)					
SSD Method A (APAO) - gauge (avg of 2 readings)					
SSD Method C - gauge (avg of 4 readings)					
SSD Method C - gauge (avg of 2 readings)					

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at 77% level of confidence. This is observed in all 30 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at 87% level of confidence. This is observed in all 10 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

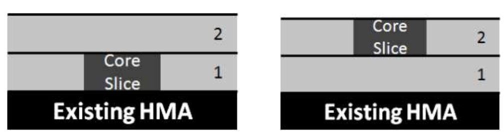
True mean correlation factor from second set of inlay (μ_{Ib}) and true mean correlation factor from both lifts i.e., inlay and overlay (μ_O):

H_O: $\mu_{Ib} - \mu_O = 0$

H_a: $\mu_{Ib} - \mu_O \neq 0$

Test: $t = \dfrac{\bar{X}_{Ib} - \bar{X}_O}{s_p / \sqrt{n}}$; where s_p = pooled standard deviation

Table A33 – p-values from the pooled t-tests comparing the various average differences across lifts 140-2 Lift 1 and 140-2 Lift 2 on the mill and fill (inlay) plus overlay project having one mix design (OR 140)



Average difference compared	(Troxler 3440 # 38996) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 35601) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.9283	0.9275	0.9336	0.8889	0.8695
CoreLok - gauge (avg of 2 readings)	0.9270	0.9291	0.9371		
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.4203	0.4105	0.4751	0.2075	0.2105
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.4161	0.4250	0.5164		
SSD Method A (after APAO) - gauge (avg of 4 readings)	0.4300	0.4249	0.4812	0.2448	0.2528
SSD Method A (after APAO) - gauge (avg of 2 readings)	0.4249	0.4403	0.5211		
SSD Method A (APAO) - gauge (avg of 4 readings)	0.5422	0.5333	0.5830	0.3814	0.3863
SSD Method A (APAO) - gauge (avg of 2 readings)	0.5411	0.5468	0.6162		
SSD Method C - gauge (avg of 4 readings)	0.2529	0.2313	0.2920	0.1423	0.1440
SSD Method C - gauge (avg of 2 readings)	0.2558	0.2757	0.3636		

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at 40% level of confidence. This is observed in all 12 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at 33% level of confidence. This is observed in all 4 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

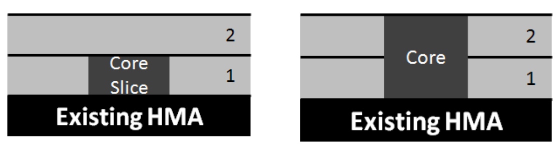
True mean correlation factor from second set of inlay (μ_{Ib}) and true mean correlation factor from both lifts i.e., inlay and overlay (μ_{IO}):

H_O: $\mu_{Ib} - \mu_{IO} = 0$

H_a: $\mu_{Ib} - \mu_{IO} \neq 0$

Test: $t = \dfrac{\bar{X}_{Ib} - \bar{X}_{IO}}{s_p / \sqrt{n}}$; where s_p = pooled standard deviation

Table A34 – p-values from the pooled t-tests comparing the various average differences across lifts 140-2 Lift 1 and 140-2 uncut on the mill and fill (inlay) plus overlay project having one mix design (OR 140)



Average difference compared	(Troxler 3440 # 38996)	(Troxler 3430 # 38806)	(Troxler 3440 # 35601)	PQI 300 (sand)	PQI 301 (sand)
	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)	Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.7059	0.6620	0.8009	0.7384	0.7818
CoreLok - gauge (avg of 2 readings)	0.6826	0.6442	0.8701		
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.7424	0.6499	0.8548	0.6737	0.7678
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.6721	0.6058	0.9271		
SSD Method A (after APAO) - gauge (avg of 4 readings)					
SSD Method A (after APAO) - gauge (avg of 2 readings)					
SSD Method A (APAO) - gauge (avg of 4 readings)					
SSD Method A (APAO) - gauge (avg of 2 readings)					
SSD Method C - gauge (avg of 4 readings)					
SSD Method C - gauge (avg of 2 readings)					

For the nuclear gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at 67% level of confidence. This is observed in all 12 comparisons made.

Even for the electromagnetic gauges, the p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the electromagnetic gauges) do not differ significantly between the two lifts at 82% level of confidence. This is observed in all 4 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

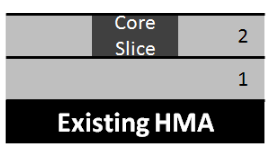
True mean correlation factor from overlay (μ_O) and true mean correlation factor from both lifts i.e., inlay and overlay (μ_{IO}):

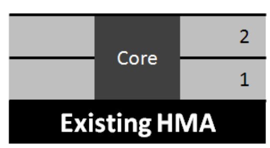
H$_O$: $\mu_O - \mu_{IO} = 0$

H$_a$: $\mu_O - \mu_{IO} \neq 0$

Test: $t = \dfrac{\overline{X}_O - \overline{X}_{IO}}{s_p / \sqrt{n}}$; where s_p = pooled standard deviation

Table A35 – p-values from the pooled t-tests comparing the various average differences across lifts 140-2 Lift 2 and 140-2 uncut on the mill and fill (inlay) plus overlay project having one mix design (OR 140)





Average difference compared	(Troxler 3440 # 38996) Pooled t-test (p-value)	(Troxler 3430 # 38806) Pooled t-test (p-value)	(Troxler 3440 # 35601) Pooled t-test (p-value)	PQI 300 (sand) Pooled t-test (p-value)	PQI 301 (sand) Pooled t-test (p-value)
CoreLok - gauge (avg of 4 readings)	0.6771	0.6370	0.7595	0.6698	0.6995
CoreLok - gauge (avg of 2 readings)	0.6528	0.6146	0.8211		
SSD Method A (before APAO) - gauge (avg of 4 readings)	0.4457	0.5132	0.3382	0.1896	0.3194
SSD Method A (before APAO) - gauge (avg of 2 readings)	0.3881	0.5834	0.3425		
SSD Method A (after APAO) - gauge (avg of 4 readings)					
SSD Method A (after APAO) - gauge (avg of 2 readings)					
SSD Method A (APAO) - gauge (avg of 4 readings)					
SSD Method A (APAO) - gauge (avg of 2 readings)					
SSD Method C - gauge (avg of 4 readings)					
SSD Method C - gauge (avg of 2 readings)					

The p-values in the following table suggest that the average differences (between the densities from the various tests on cores and the densities from the nuclear gauges) do not differ significantly between the two lifts even at 44% level of confidence. This is observed in all 12 comparisons made.

Hypothesis for the pooled t-test at 95% level of confidence:

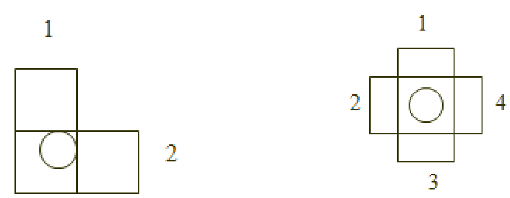
True mean correlation factor from 'L-shaped' footprint (μ_L) and true mean correlation factor from 'cross-shaped' footprint (μ_C):

H_0: $\mu_L - \mu_C = 0$

H_a: $\mu_L - \mu_C \neq 0$

Test: $t = \dfrac{\bar{X}_L - \bar{X}_C}{s_p / \sqrt{n}}$; where s_p=pooled standard deviation

Table A36 – p-values from the pooled t-tests comparing the various average differences across 'L-shaped' and 'cross-shaped' footprint on the project having multiple lifts with same mix design (OR 18)



Average difference compared	(Troxler 3430-38806) Pooled t-test (p-value)	(Troxler 3430-38994) Pooled t-test (p-value)
CoreLok-nuclear gauge (avg of 4 readings)	0.9019	0.7924
CoreLok-nuclear gauge (avg of 2 readings)	0.9116	0.9716
SSD Method A-nuclear gauge (avg of 4 readings)	0.6629	0.5691
SSD method A-nuclear gauge (avg of 2 readings)	0.8220	0.7169
SSD Method C-nuclear gauge (avg of 4 readings)	0.8576	0.7492
SSD Method C-nuclear gauge (avg of 2 readings)	0.9825	0.9128