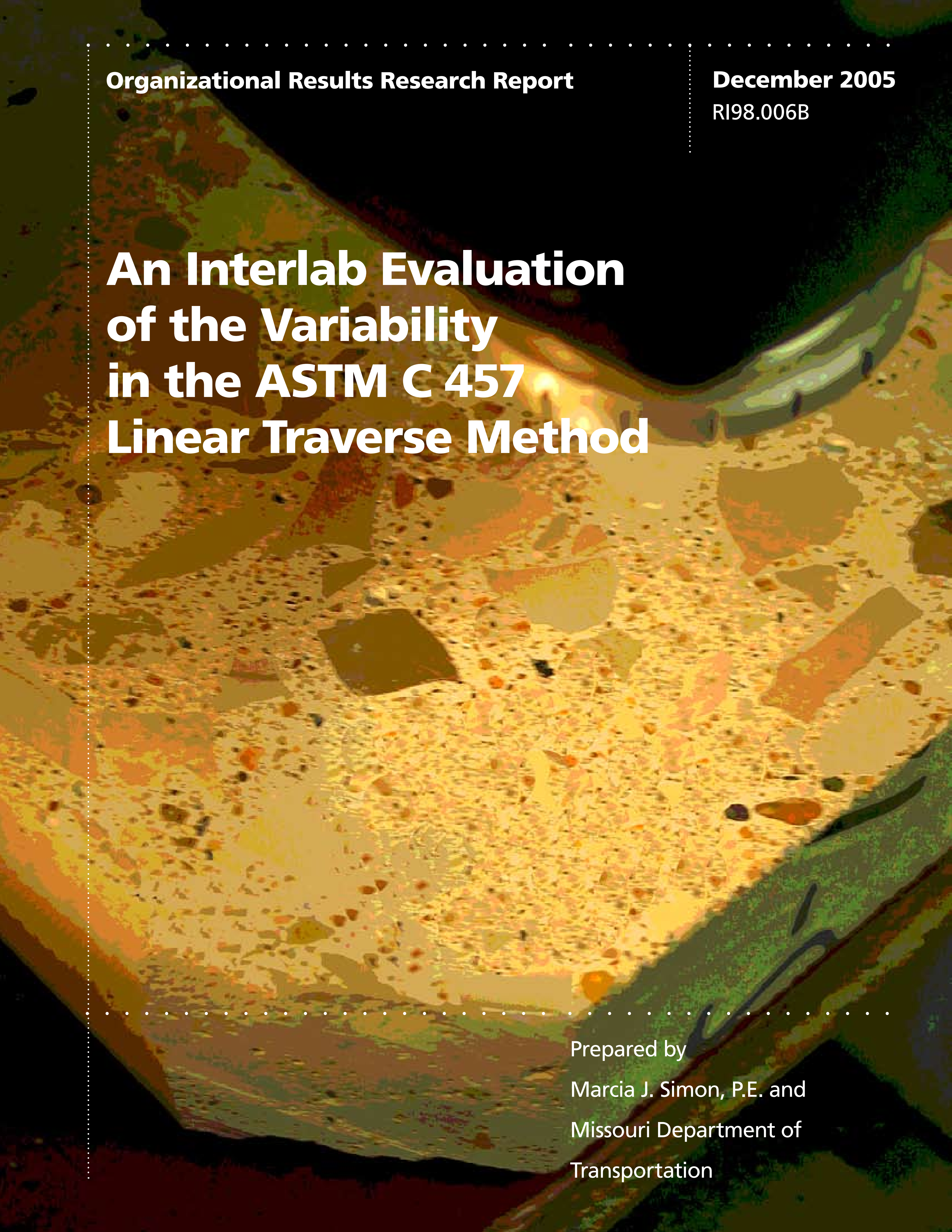

Organizational Results Research Report

December 2005
RI98.006B

An Interlab Evaluation of the Variability in the ASTM C 457 Linear Traverse Method

Prepared by
Marcia J. Simon, P.E. and
Missouri Department of
Transportation

FINAL REPORT

RI98-006B

An Interlaboratory Evaluation of Variability in the ASTM C 457 Linear Traverse Method

Prepared for the Missouri Department of Transportation

By:

Marcia J. Simon, P.E.

December 2005

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. OR 06 – 009	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle **An Interlaboratory Evaluation of Variability in the ASTM C 457 Linear Traverse Method**		5. Report Date December 2005 6. Performing Organization Code
7. Author(s) Marcia J. Simon, P.E.		8. Performing Organization Report No. OR 06 – 009 / RI 98 – 006 B
9. Performing Organization Name and Address Marcia J. Simon, P.E.		10. Work Unit No. 11. Contract or Grant No. RI 98 – 006 B
12. Sponsoring Agency Name and Address Missouri Department of Transportation Organizational Results P. O. Box 270-Jefferson City, MO 65102		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code MoDOT

15. Supplementary Notes
The investigation was conducted in cooperation with the U. S. Department of Transportation, Federal Highway Administration.

16. Abstract
The vital role of air entrainment in preventing freeze-thaw damage in concrete is well known and well documented [Powers 1949]. Through the action of an air entraining agent (AEA) added to fresh concrete, an air void system comprised of various microscopic voids is established. There are several parameters of the air void system, which are considered important indicators of freeze-thaw resistance. Measuring these parameters and their adequacy provides extremely useful information on concrete freeze-thaw resistance. This report documents the results of a round-robin study of the ASTM C 457 linear traverse method, which was initiated to access the typical variability associated with the linear traverse test when performed by a human operator. The round robin was performed as part of a national pooled-fund study, led by the Missouri Department of Transportation (MoDOT), and sponsored by 13 states, entitled "Advanced Research…of a Fully Automated Image analysis system." The goal of the pooled-fund study was to refine and complete the development of a fully automated, computer-based linear traverse system, which could provide results equal to or better than those of a linear traverse performed by a human operator. Thus, the round robin study was undertaken to access accuracy and precision of a human-based linear traverse for which the automated system could be measured against.

17. Key Words Concrete, freeze-thaw, air void, air void parameters, linear traverse, round robin, precision, variability, air content, specific surface, spacing factor	18. Distribution Statement No restrictions. This document is available to the public through National Technical Information Center, Springfield, Virginia 22161

19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 83	22. Price

Form DOT F 1700.7 (06/98)

Executive Summary

The vital role of air entrainment in preventing freeze-thaw damage in concrete is well known and well documented [Powers 1949]. Through the action of an air entraining agent (AEA) added to fresh concrete, an air void system comprised of various microscopic voids is established. There are several parameters of the air void system, which are considered important indicators of freeze-thaw resistance. Measuring these parameters and their adequacy provides extremely useful information on concrete freeze-thaw resistance.

ASTM C 457 describes two methods for assessing the adequacy of the air void system in hardened concrete: the linear traverse and modified point count [ASTM 2004]. Both procedures in ASTM C 457 require a human operator to use a microscope, or a video monitor and camera attached to a microscope, to make the necessary observations and measurements. The process has been long recognized as tedious and time-consuming, requiring a skilled and experienced operator to ensure reliable, consistent results. In addition to operator training, ability, and bias, factors that contribute to variation in air void measurements include differences in procedures and equipment, differences in specimen preparation quality, and inherent statistical variability of the test method itself. This report documents the results of a round-robin study of the ASTM C 457 linear traverse method, which was initiated to access the typical variability associated with the linear traverse test when performed by a human operator. The round robin was performed as part of a national pooled-fund study, led by the Missouri Department of Transportation (MoDOT), and sponsored by 13 states, entitled "Advanced Research…of a Fully Automated Image analysis system." The goal of the pooled-fund study was to refine and complete the development of a fully automated, computer-based linear traverse system, which could provide results equal to or better than those of a linear traverse performed by a human operator. Thus, the round robin study was undertaken to access accuracy and precision of a human-based linear traverse for which the automated system could be measured against.

The round robin study entailed five concrete samples, representing a broad range of concrete specimens for system validation, analyzed by laboratories and operators among eight of the pooled-fund states, the FHWA, and a private consulting firm in accordance with the study's experimental plan. The experimental plan was executed to access both within-laboratory variability (repeatability) and between-laboratory variability (reproducibility). Multiple statistical analyses were performed on the round robin data to access variability. Air void parameters evaluated included air content, voids per inch, spacing factor, and specific surface. The following study findings were noted:

- Within-laboratory variation, as expected, was less than between-laboratory variation; however, although within-laboratory variation provides a means of assessing operator consistency, it does not provide a useful indication of accuracy. Large within-laboratory variation may or may not be coupled with poor accuracy.

- Between-laboratory results exhibited wide variation, even when the data were confined to a subset of the five best labs. The differences between laboratories most likely reflect differences in operator ability, because the study was designed to minimize the effects of

different equipment by requiring that all laboratories use the same magnification for all tests. Other equipment differences, such as use of a monitor instead of a microscope for viewing may have had a minor effect; however, this effect would not be expected to cause such wide variation.

- The results of the MoDOT round robin were more variable than a previous study reported in ASTM C 457, Table 4; however, the MoDOT results were less variable or comparable to the results of another study reported in ASTM C 457, Table 6.
- The wide variation between laboratories strongly supports the development of an automated image analysis system for performing ASTM C 457 measurements. A properly designed automated system should be able to provide more consistent results than a human operator and thus reduce the considerable variation noted in this study.
- The key to a successful air void measurement system, whether human-based or automated, is its ability to make accurate measurements. At present, there is no way to assess the true accuracy of air void parameter estimates; therefore, the accuracy of an automated system should be assessed by comparing its results with the results of a study such as this one.
- The comparison of the results between Phase 1 and Phase 2 of the round robin indicated that, with the exception of specimen RR3, specimen preparation appeared to be very good and did not appear to have a significant effect on the results.

Based upon analysis of the round robin data, this study also covers recommendations for statistical measures of acceptable variability for both precision (within-laboratory variability) and accuracy (between-laboratory variability) of an automated testing system conducting the ASTM C 457 linear traverse method.

Table of Contents

List of Figures

List of Tables

List of Tables (cont'd.)

Acknowledgements

The Missouri Department of Transportation would like to acknowledge the following organizations for their participation in the round robin testing:

Arkansas Department of Highways
CTLGroup
FHWA – Turner Fairbank Highway Research Center
Indiana Department of Transportation
Illinois Department of Transportation
Minnesota Department of Transportation
Missouri Department of Transportation
Nebraska Department of Highways
Virginia Department of Transportation

The following states provided funding for the image analysis research project under which this round robin study was performed:

Arkansas
California
Colorado
Illinois
Indiana
Iowa
Minnesota
Missouri
Montana
Nebraska
Ohio
Virginia
Wisconsin

Chapter 1
Project Introduction

The vital role of air entrainment in preventing freeze-thaw damage in concrete is well-known and well-documented [Powers 1949]. The air void system comprises billions of microscopic voids, ranging in size from several micrometers to a millimeter or more, which are incorporated into the fresh concrete as bubbles through the action of an air entraining agent (AEA). The AEA reduces the surface tension at the air-water interface, stabilizing air incorporated during mixing into small, stable bubbles. These bubbles remain as voids in the hardened concrete. [Mielenz et al 1958].

There are several parameters of the air void system that are considered important indicators of freeze-thaw resistance. The most commonly known and specified parameter is the air content (the volume fraction of air in the concrete). More important than the air content, however, is the number and size distribution of the air voids. These cannot be measured directly, but instead can be described using various parameters: the mean chord length (also called average chord length, or average chord intercept), the specific surface (area of bubble surface per unit air volume), and the spacing factor (an indicator of the distance water must travel to reach an air void boundary) [Powers 1949].

ASTM C 457 describes two methods for assessing the adequacy of the air void system in hardened concrete: the linear traverse and modified point count [ASTM 2004]. Both procedures in ASTM C 457 require a human operator to use a microscope, or a video monitor and camera attached to a microscope[1], to make the necessary observations and measurements. The operator must distinguish among the various concrete constituents (air, paste, aggregate) on a prepared plane surface of concrete and record measurements (counts in the modified point count, chord length measurements in the linear traverse). This process has been long recognized as tedious and time-consuming, requiring a skilled and experienced operator to ensure reliable, consistent results[2]. In addition to operator training, ability, and bias, factors that contribute to variation in air void measurements include differences in procedures and equipment, differences in specimen preparation quality, and inherent statistical variability of the test method itself.

This report documents the results of a round-robin study of the ASTM C 457 linear traverse method. The round-robin was performed as part of a national pooled-fund study, led by the Missouri Department of Transportation (MoDOT), and sponsored by 13 states[3], entitled "Advanced Research…. of a Fully Automated Image Analysis System." The goal of the pooled-fund study was to refine and complete the development of a fully automated, computer-based linear traverse system that could provide results equal to or better than those of a linear traverse

[1] ASTM C 457 does not specifically mention the use of a video monitor, but many laboratories use one (see Appendix A.5, question 2).

[2] Operator skill and dedication are especially important in seeing and including the smallest air void chords. These chords have little influence on air content but have a significant influence on other parameters such as spacing factor.

[3] The state sponsors include Arkansas, California, Colorado, Illinois, Indiana, Iowa, Minnesota, Missouri, Montana, Nebraska, Ohio, Virginia, and Wisconsin.

performed by a human operator. A technical advisory committee (TAC) comprising representatives from each participating state, the Federal Highway Administration (FHWA), and the contractor was formed to provide technical oversight and input as the study progressed.

One of the project's goals was to assess the accuracy and precision of the automated system relative to that of a human operator, with the ultimate goal that the automated system would provide results at least as accurate and precise as the results of tests performed by human operators. In order to meet this objective, the researchers working on the automated system needed to assess the typical variability associated with the linear traverse test when performed by a human operator. This round-robin study was undertaken to address this task. The TAC collected concrete samples from all participating states and elsewhere in order to provide a broad range of concrete specimens for system validation, comprising different aggregate types, paste characteristics, and air void systems. The round-robin study was then planned and executed to provide the needed information on variability. Eight of the sponsoring states, the FHWA, and a private consulting firm participated in the round robin testing (the consulting firm participated only in Phase 2).

Chapter 2 of this report provides brief background on the linear traverse test method and its potential sources of variability. Chapter 3 describes the experimental plan and the methods used, and Chapter 4 presents the results obtained. Chapter 5 presents a summary of general findings and provides recommendations for precision and accuracy standards that could be used for an automated air void measurement system. Finally, the Appendix contains a comprehensive tabulation of the data collected in the study, as well as results of a survey conducted among potential study participants.

Chapter 2
Background

2.1 Background on linear traverse testing

2.1.1 Description of test method

The linear traverse and modified point count are the two methods prescribed in ASTM C 457 for estimating the air void parameters of hardened concrete [ASTM 2004]. This interlaboratory study deals exclusively with the linear traverse (Procedure A in ASTM C 457). In a linear traverse, parallel lines are superimposed on a polished plane surface of concrete, and the chords formed by the intersection of these lines with exposed air void sections are counted and measured. Air void parameters such as the air content, specific surface, and spacing factor, are calculated using equations set forth in the test method.

Modern linear traverse equipment includes an X-Y stage (motorized or hand-cranked), a microscope or microscope and video monitor for magnified viewing of the plane concrete surface, and some type of length measuring and counting device. Stage movement is accomplished either manually or through computer software control. As the stage moves linearly, a human operator views a cross-hair superimposed on the concrete surface, which traces a virtual line on the surface. As the cross-hair moves over various phases (e.g., air, paste, aggregate), the operator measures chord lengths of the different constituents by pressing a button (in a computerized system) or recording a value from a length counter for each measured constituent.

The data obtained from a linear traverse vary with equipment type. All systems must provide the minimum information needed to calculate the air void parameters: the total traverse length, T_l, the length traversed on air T_a, the number of air void chords intercepted N_a, and the length traversed on paste T_p (or, alternatively, the paste content estimated from mix proportions or a point count) [Snyder et al 2002].

2.1.2 Estimation of air void parameters

The air void parameters that are estimated[4] from C 457 test data include air content, mean chord length, voids per inch, specific surface, and spacing factor. Equations for estimating each of these parameters from linear traverse data are shown in Table 1.

2.2 Variability in the Linear Traverse Method

The round robin testing program was designed to estimate the expected precision, or variability[5] of the ASTM C 457 linear traverse method as performed by a human operator. Both within-

---

[4] technically, the values calculated using the equations in Table 1 are “sample statistics,” which are estimates of the “true” parameters of the volume of concrete from which the sample is taken. See Section 2.2.4.
[5] the terms “precision” ,“variability”, and “uncertainty” are used interchangeably in this report.

laboratory variability[6] and between-laboratory variability[7] were of interest. The following aspects of linear traverse testing contribute to the variability of the method, and are discussed further below:

- operator characteristics (e.g., training, ability, dedication, and bias)
- differences in equipment and procedure
- surface preparation quality
- inherent statistical variation of the method

Table 1. Equations for ASTM C 457 air void parameters

Parameter	Units	Equation
Air content, *A*	%	$A = 100 \times \frac{T_a}{T_l}$
Paste content, *p*	%	$p = 100 \times \frac{T_p}{T_l}$
Mean chord length, $\bar{l}$	in	$\bar{l} = \frac{T_a}{N_a}$
Voids per inch, *n*	in	$n = \frac{N_a}{T_l}$
Specific surface, α	in^{-1}	$\alpha = \frac{4}{\bar{l}}$
Spacing factor, $\bar{L}$	in	$\bar{L} = \frac{3}{\alpha}\left[1.4\left(1+\frac{p}{A}\right)^{\frac{1}{3}} - 1\right], \frac{p}{A} > 4.342$ $\bar{L} = \frac{p}{A\alpha}$, otherwise

2.2.1 Operator characteristics

Operator characteristics that contribute to variable test results for the same specimen surface include ability, training, experience and dedication. Conducting a linear traverse involves microscopical examination of a prepared concrete surface by a human operator. Based on his observations, either directly through a microscope or from a microscopical image

[6] within-laboratory variability is synonymous with “single-operator precision” as defined in ASTM C 670. This term is also referred to as repeatability.
[7] between-laboratory variability is synonymous with “multi-laboratory precision” as defined in ASTM C 670. This term is also referred to as reproducibility.

projected on a monitor, the operator measures chord lengths of the components of the concrete (air, paste, and aggregate). The work is time-consuming and tedious, requiring judgment, hand-eye coordination, and attentiveness. Therefore, the operator's skill (or lack thereof) in performing the linear traverse will inevitably influence the results.

Another component of operator variability is test-to-test variation for a particular operator. An individual operator cannot be expected to be perfectly consistent from test to test, even if he could run two tests over identical traverse lines. However, this variability is likely to be far less than the variability due to differences in ability among different operators.

2.2.2 Differences in equipment and procedure

Equipment differences such as measurement resolution, measurement accuracy, method of viewing (direct viewing through microscope viewing via video monitor and camera attached to microscope), magnification, and lighting can contribute to variability of linear traverse results. Assuming that equipment is functioning properly, the most significant element of variability associated with equipment is probably the magnification level. Differences in magnification affect the operator's ability to discern the smallest circular sections of voids on the concrete surface, and thus the smallest chord lengths he can measure. These smallest chords do not greatly affect air content estimates, but can contribute significantly to the estimates of other parameters such as specific surface and spacing factor. According to ASTM C 457, the allowable magnification for a test may range from 50X to 125X [ASTM C 457 2004].

2.2.3 Surface preparation quality

ASTM C 457 requires that the surface of the concrete specimen to be examined must be ground and polished to obtain an acceptably smooth, plane surface for microscopical observation. However, ASTM C 457 does not specify any particular preparation method or equipment. Therefore, the preparation methods used and the definitions of what constitutes an "acceptable" surface are often unique to individual laboratories. Moreover, operators often mentally "reconstruct" a surface when they encounter minor surface damage. The accuracy of such reconstructions is another factor that may contribute to variability.

Round robin test results from phase I (differing surface preparation methods) compared to those from Phase 2 (single surface preparation method) should allow a better determination of this influence due to the different surface preparation methods used in this study. A description of the various surface preparation methods used in this study is included in Section A.4 of the Appendix.

2.2.4 Inherent variability of the method

Because the linear traverse is essentially a statistical sampling process, the test results will have some inherent uncertainty due to sampling error. In a linear traverse, the sample is the set of chord lengths obtained from the intersection of a set of regularly spaced traverse lines superimposed on a two-dimensional concrete surface with the circular air void sections on the surface. The air void system parameters (air content, specific surface, spacing factor, etc.) are functions of the chord length measurements and the number of chords measured. For each distinct set of linear traverse lines that can be superimposed on the surface, a different sample of chord lengths will be obtained; therefore, the estimates of air void system parameters based on the sample will vary from one set of lines to another. Thus, the inherent variation is what would be expected if a perfect operator, using perfect equipment, performed multiple traverses (on different sets of traverse lines) on the same ideally prepared specimen surface.

The air void system parameters are estimates for a concrete volume (three-dimensional) based on a sample of chord length measurements (one-dimensional) of circles on a plane surface. The chord lengths are measured on one particular sample of circular air void sections (two-dimensional) exposed when the three-dimensional concrete specimen (e.g., a core or cylinder) is cut to expose a plane surface. The circular sections of voids visible on this surface represent only a fraction of the three-dimensional voids existing throughout the entire sample.

Equations for the expected minimum uncertainty of the linear traverse method have been derived from the equations used to calculate the air void parameters from linear traverse data [Snyder et al 2002]. Table 2 summarizes the equations for the uncertainties in air content, specific surface, and spacing factor, expressed in terms of the squared coefficient of variation. The coefficient of variation, C_X, of a parameter *X* is the standard deviation of *X* divided by the expected value (mean) of X.[8]

Table 2. Equations for minimum expected uncertainties in linear traverse air void parameters

Parameter	Uncertainty equation	Notes
A	$C_A^2 = \frac{1}{N_a}\left[1 + C_l^2\right] = \frac{1}{N_a} + C_{\bar{l}}^2$	1. individual air chord lengths required 2. N_a = number of chords counted
n	$C_n = \frac{1}{\sqrt{N_a}}$	
α	$C_\alpha^2 = \frac{1}{N_a} C_l^2 = C_{\bar{l}}^2$	3. individual air chord lengths required 4. assumes $C_{\bar{l}} << 1$ (reasonable if $N_a \geq 1000$)
$\bar{L}$	$C_{\bar{L}}^2 = \frac{(1-\beta)^2}{N_a} C_l^2 + \beta^2\left(C_p^2 + \frac{1}{N_a}\right)$ where $\beta = 0.35\frac{\bar{l}}{\bar{L}}\left(\frac{p}{A}\right)\left(1 + \frac{p}{A}\right)^{-2/3}$	5. individual air and paste chord lengths required 6. if point count is used to estimate paste content, C_p can be calculated from point count data using the following equation where S_p = stops on paste and S_t = total stops: $C_p = \frac{1}{S_p} - \frac{1}{S_t}$ 7. the uncertainty equation shown is a simplified version in which terms with negligible contribution are eliminated

[8] The coefficients of variation in the uncertainty equations listed in Table 2 are expressed as fractions, not as percentages.

In order to calculate the expected uncertainties, individual air chord length measurements are required to calculate C_l. Likewise, to calculate the uncertainty in the spacing factor, individual paste lengths must be measured to calculate C_P, the coefficient of variation of the paste content[9]. In addition, the standard data obtained from a linear traverse is needed: N_a, the number of chords on air, N_p, the number of chords on paste, T_l, the total length of traverse, and T_a, the traverse length on air.

The mean chord length, $\bar{l}$, is calculated by dividing the traverse length on air by the number of chords on air, or equivalently, by averaging the individual chord measurements. The number of air chords and paste chords are assumed to be distributed as a Poisson distribution. Thus, the standard deviation of the number of air chords, N_a, is equal to $s_{N_a} = \sqrt{N_a}$, and similarly, the standard deviation of the number of paste chords is equal to $s_{N_p} = \sqrt{N_p}$.

The coefficient of variation of the chord length, l, is related to the coefficient of variation of the mean chord length, $\bar{l}$. By the central limit theorem, the expected value (mean) of the mean chord length $\bar{l}$ is the same as the expected value of the chord length l, $E[l] = E[\bar{l}] = \bar{l}$. Similarly, the standard deviation of the mean chord length is given by

$$s_{\bar{l}} = \sqrt{\frac{s_l^2}{N_a}}$$

Therefore, the coefficient of variation of the chord length is related to the coefficient of variation of the mean chord length as follows:

$$C_l^2 = \frac{s_l^2}{\bar{l}^2} = \frac{N_a \cdot s_{\bar{l}}^2}{\bar{l}^2} = N_a \cdot C_{\bar{l}}^2$$

In this round robin study, several laboratories provided individual chord length data from the linear traverses they ran. For those laboratories, the above equations were used to calculate minimum uncertainties for the air void parameters. The results are presented in Section 4.4 of this report.

---

[9]The paste chords are needed to estimate the paste volume for calculation of spacing factor. Individual paste chord measurements are required to estimate the uncertainty in the spacing factor. Some laboratories use the linear traverse for air chord measurements and the point count to estimate the paste volume for spacing factor calculations. In this case, C_p can be calculated as described in Note 6 in Table 2.

Chapter 3
Experimental Plan

3.1 Selection of concrete specimens

The round robin testing program involved five, pre-selected concrete specimens obtained from six-inch concrete cylinders or cores. Potential specimens were submitted to MoDOT by pooled-fund participant states and FHWA, and MoDOT selected five specimens to be used in the round robin. These specimens, shown in Figure 1, represented a sampling of concretes and aggregates from around the country.

Laboratories submitting potential specimens for the study were asked to send a concrete specimen with approximate dimensions of 6 in x 6 in x 1 in, with surfaces prepared using the laboratory's equipment and standard grinding and polishing procedure. After selecting the round robin specimens, MoDOT sawed specimens RR1 through RR4 into an octagonal shape (this step was required to enable the specimens to fit in MoDOT's linear traverse equipment and did not affect other laboratories' ability to perform the tests). Specimen RR5 was left in its original rectangular shape. MoDOT did not perform any additional surface preparation on the specimens prior to starting Phase 1 of the study.

3.2 Selection of round-robin participants

The initial round-robin participants were selected from among the states participating in the pooled fund study. States were initially contacted regarding their willingness to participate. Those who responded affirmatively completed a questionnaire regarding their linear traverse testing equipment and procedure, as well as their specimen preparation equipment and technique. Potential participants were asked for the following information:

- general description of equipment used for linear traverse
- usual specimen size and shape
- minimum and maximum specimen sizes
- method of viewing (directly through microscope or via camera/monitor)
- usual magnification used
- available magnification range
- method of distinguishing entrained/entrapped air (if any)
- ability to record and report individual chord lengths
- specimen preparation equipment and usual procedure (make and model, auto or manual, type of grinding material used, grit sizes used, time on each grit, etc).

The responses to this questionnaire are provided in Section A.5 of the Appendix.

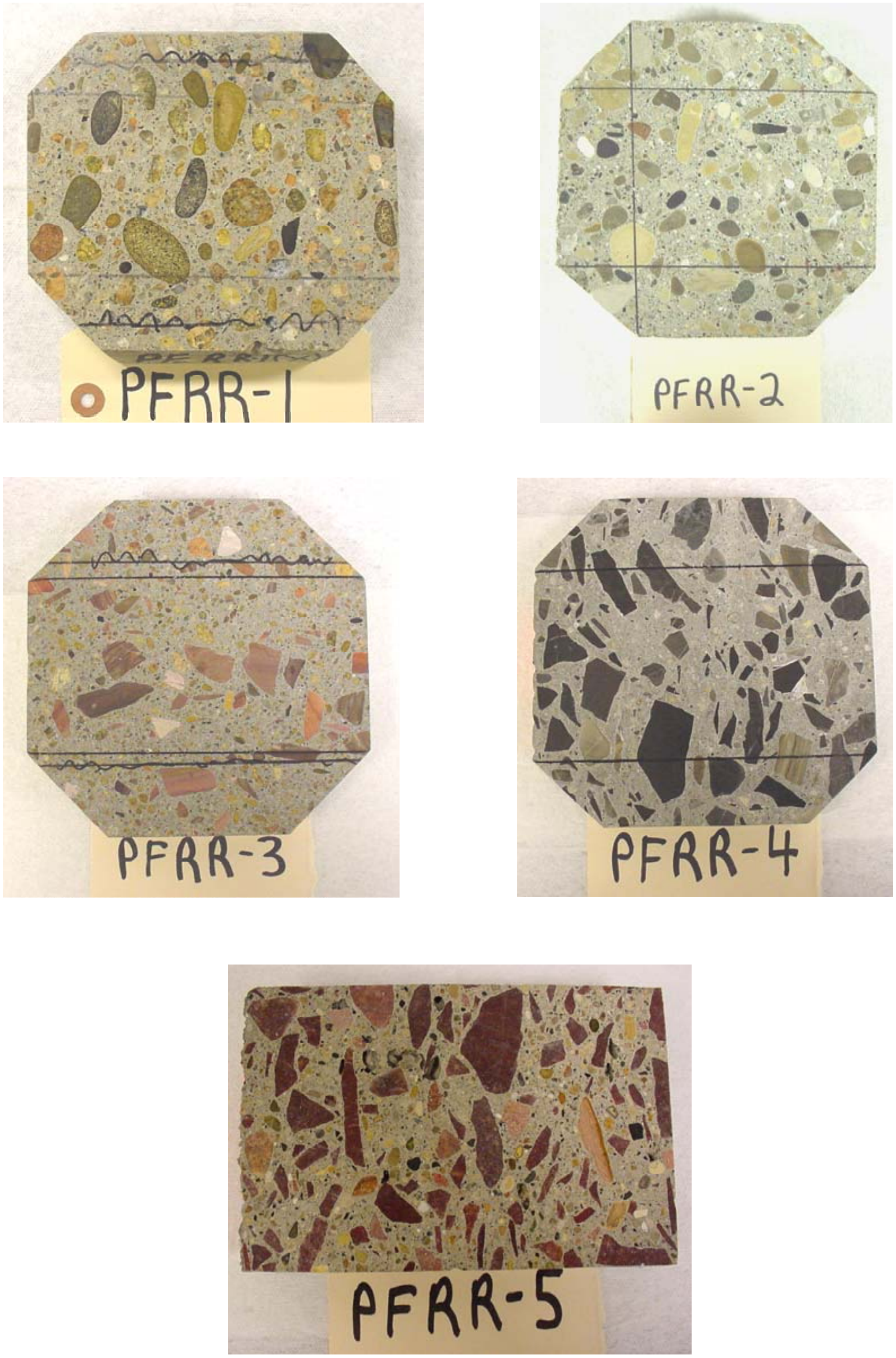

Figure 1. Photographs of the five RR specimens

3.3 General requirements

The round robin study comprised 18 ASTM C 457 linear traverse tests. The study was divided into two phases. Nine laboratories participated in Phase 1, and ten laboratories participated in Phase 2[10]. In each phase, each lab received five concrete specimens for testing. In Phase 1, the specimens were prepared at five different participating labs using the standard preparation equipment and procedure of the laboratory (see Section A.4 of the Appendix for a summary of specimen preparation methods). A testing schedule was created and specimens were shipped from lab to lab and tested in a designated order. When the Phase 1 tests were complete, the specimens were shipped back to MoDOT. MoDOT re-prepared the specimen surfaces using their standard equipment and method (to eliminate any differences associated with sample preparation), and the specimens were shipped out again, in the same order as before, for Phase 2 testing.

In both phases, each laboratory performed nine linear traverses – one traverse each for specimens RR1, RR3, RR4, and RR5, and five traverses, each on a distinct set of traverse lines, for specimen RR2. After testing was completed, each specimen was sent back to MoDOT. Laboratories were instructed not to share their results with other participating laboratories during the testing process.

3.4 Experimental details

The following specific instructions regarding the conduct of the tests were provided to each laboratory:

- The same operator was to perform all of the linear traverse tests (18 total – 9 in Phase 1 and 9 in Phase 2).
- Each linear traverse was to have a length of 100 in and was to cover an area of 24 in^2. If specimen size and shape limited the area to less than 24 in^2, the maximum possible area was to be covered.
- Each linear traverse was to be performed at a magnification of 100X.
- On the octagonal specimens (RR1 to RR4), black lines were drawn on the sample to show a rectangular area to be traversed (see Figure 1). On specimen RR5, the rectangular specimen, the entire surface area was used. Two sides of each specimen were marked with "X" and "Y" to indicate the orientation to be used (x-axis parallel to the "X", with the "X" facing the operator). Using this approach, the X and Y directions were the same for each test in each laboratory. However, no specific starting point was indicated, and the actual sets of traverse lines used were expected to be different for each laboratory.
- In the case where five linear traverses were to be performed on the same specimen (RR2), the sides of the specimen were marked A, B, C, D, and E. As illustrated in Figure 2, A and E were oriented in the same direction, while B, C, and D were oriented at 90, 180, and 270

[10] The laboratories in Phase 1 included seven state DOT laboratories and FHWA. One of the state laboratories had two operators perform the tests – each operator was considered a separate "laboratory" in the analysis of the results. In Phase 2, the same labs from Phase 1 participated, along with a private consulting laboratory.

degrees, respectively, from side A/E. The first test was performed with side A facing the operator, the second with side B, and so forth. In each case, the x-axis was parallel to the side facing the operator. An additional black line was drawn on specimen RR2 (see Figure 1) to indicate a boundary of the area to be traversed in the B and D orientations.

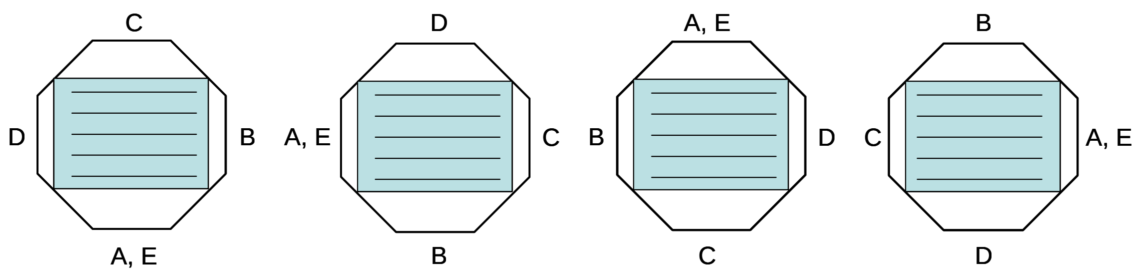

Figure 2. Illustration of orientations for RR2.

Chapter 4
Results and Analysis

4.1 Introduction

This chapter presents and discusses the results of the round robin testing program. The round robin results are presented in graphical and tabular form in Section 4.2. First, the results for within-laboratory variability (repeatability) are presented, followed by the results for between-laboratory variability (reproducibility). In Sections 4.3 and 4.4, ASTM C 670 precision limits and theoretical minimum uncertainties are calculated and compared with the round robin data. In Section 4.5, the results of Phase 1 and Phase 2 are compared to assess effects of specimen preparation. Finally, Section 4.6 presents and compares chord distributions obtained from several participating laboratories that provided individual chord measurements along with their test results.

Complete tabulations of the data collected during this study can be found in Sections A.1 through A.4 of the Appendix. Section A.1 contains the within-laboratory data. Sections A.2 and A.3 contain between-laboratory data arranged by air void parameter (Section A.2) and by specimen (Section A.3). Section A.4 contains tables of multi-laboratory (between-laboratory) precision limits according to ASTM C 670.

4.2 Results of round robin testing

4.2.1 Repeatability (within-laboratory variability)

The multiple traverses performed on specimen RR2 provide data for estimating the repeatability, or within-laboratory variability, of the linear traverse for each participating operator. The within-laboratory variability is the same as the single-operator precision as defined by ASTM C 670. The results from RR2 provide some information regarding reproducibility (between-laboratory variability) as well.

Figures 3 through 5 show comprehensive summary plots of the results from all participating laboratories for the five traverses performed on specimen RR2 in both phases of the study. Each figure presents one of the air void parameters: air content in Figure 3, specific surface in Figure 4, and spacing factor in Figure 5. In each figure, the five individual test results for each lab are shown as open diamonds (Phase 1) or open triangles (Phase 2). The mean of the five test results is shown as a filled diamond (Phase 1) or filled triangle (Phase 2). Error bars representing ± 1 standard deviation are shown for each lab as well. Finally, the overall mean (based on all data from all labs) is denoted by the horizontal line (dot-dash for Phase 1, dashed for Phase 2).

Figure 3 shows considerable variation in air content among laboratories. In Phase 1, the measured air contents ranged from about 2.5 percent (lab 7) to nearly 9 percent (lab 8). In Phase 2, the range was somewhat smaller – about 4.6 to 8 percent. Except for labs 3 and 7, there was fairly close agreement between Phase 1 and Phase 2, and the overall means for Phase 1 and Phase 2 are fairly close. The error bars indicate a range of variability within laboratories.

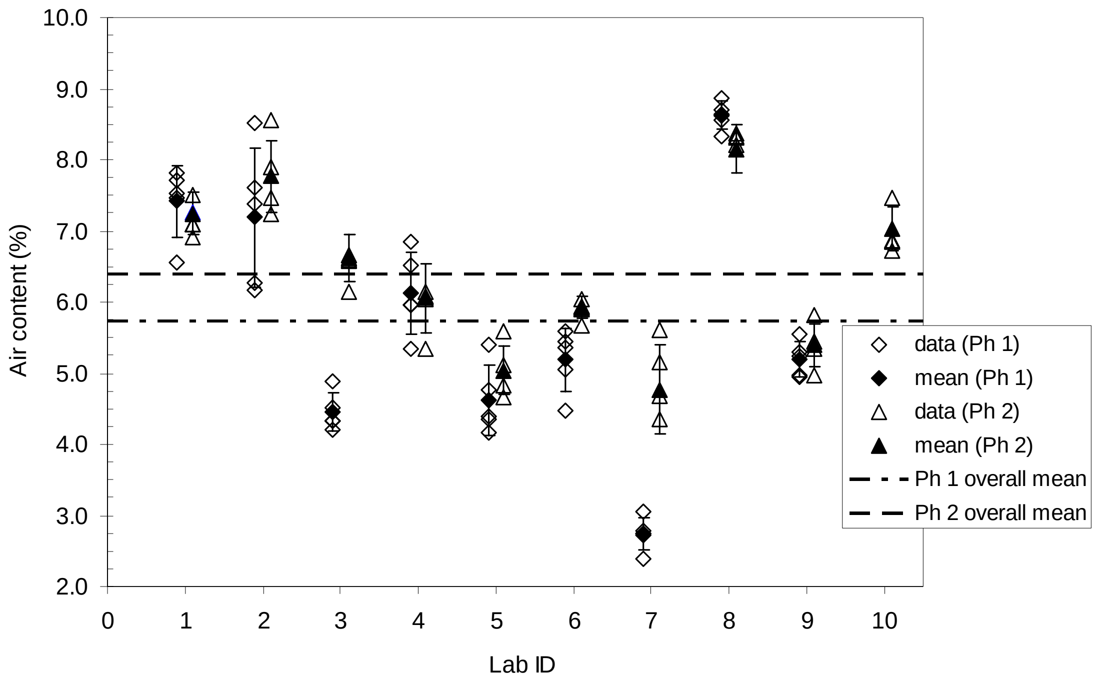

Figure 3. Within and between-laboratory air content, RR2/RR2X

For specific surface (Figure 4), the labs fell in several general groups. Labs 1, 4, 6, 8, 9, and 10 had values ranging from about 1000 in^{-1} to 1300 in^{-1} for both phases. These labs were consistent between phases. Lab 3 had an average of about 600 in^{-1} for Phase 1 and 1000 in^{-1} for Phase 2. Labs 2 and 5 had significantly lower results in both phases – around 500 to 650. Lab 7 was the lowest, with average values of 200 in^{-1} and 400 in^{-1} for phases 1 and 2, respectively. Again, the error bars indicate the within-laboratory variability for each lab.

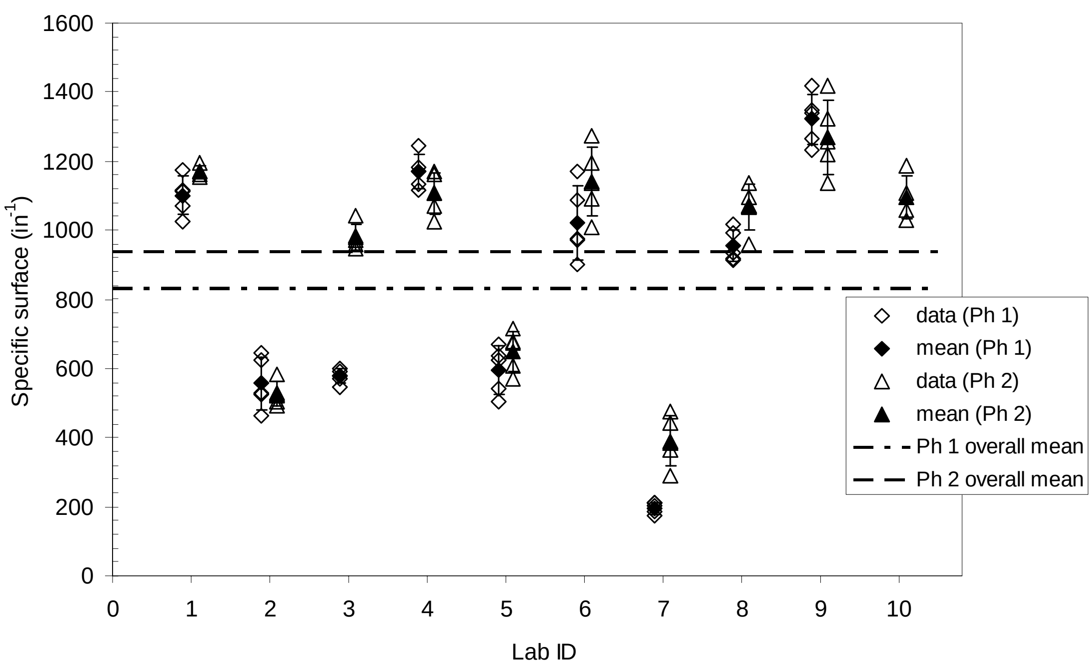

Figure 4. Within and between-laboratory specific surface, RR2/RR2X

Figure 5 shows that one lab (lab 7) had a substantially different spacing factor than the rest of the labs. Therefore, Figure 6 shows a second plot of spacing factor, excluding lab 7 and making the variability among the remaining labs easier to discern. Again, some groupings are apparent. Labs 1, 4, 6, 8, 9, and 10 have spacing factors ranging from 0.003 to 0.005 (approximately). All of those labs, except for lab 6, had very consistent results from Phase 1 to Phase 2. Lab 2 had a spacing factor of about 0.006 for both phases. Labs 3 and 5 had the highest results for Phase 1

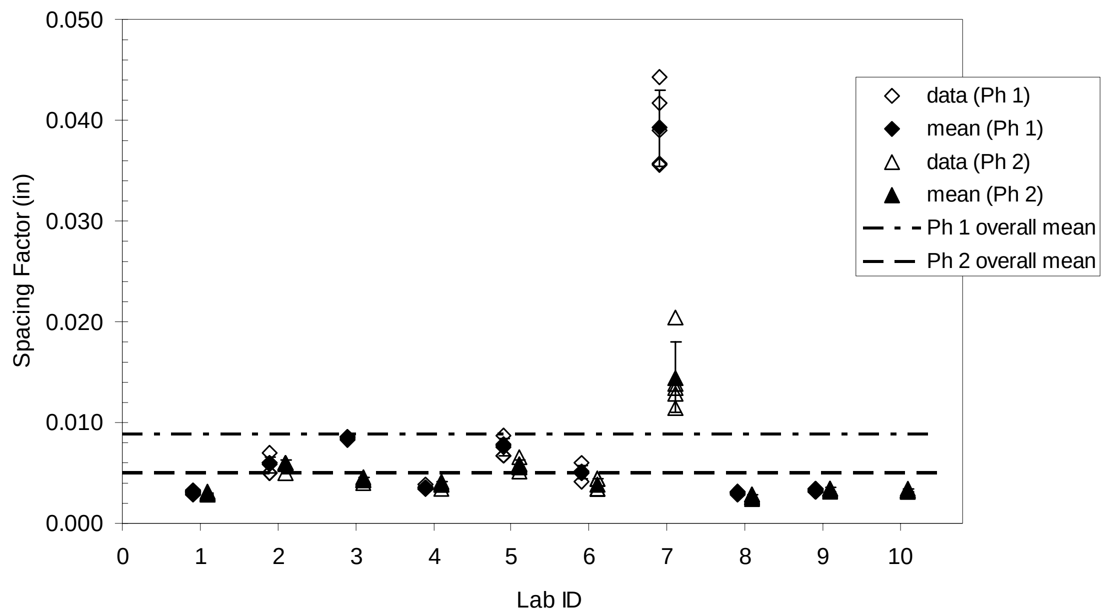

Figure 5. Within and between-laboratory spacing factor, RR2/RR2X

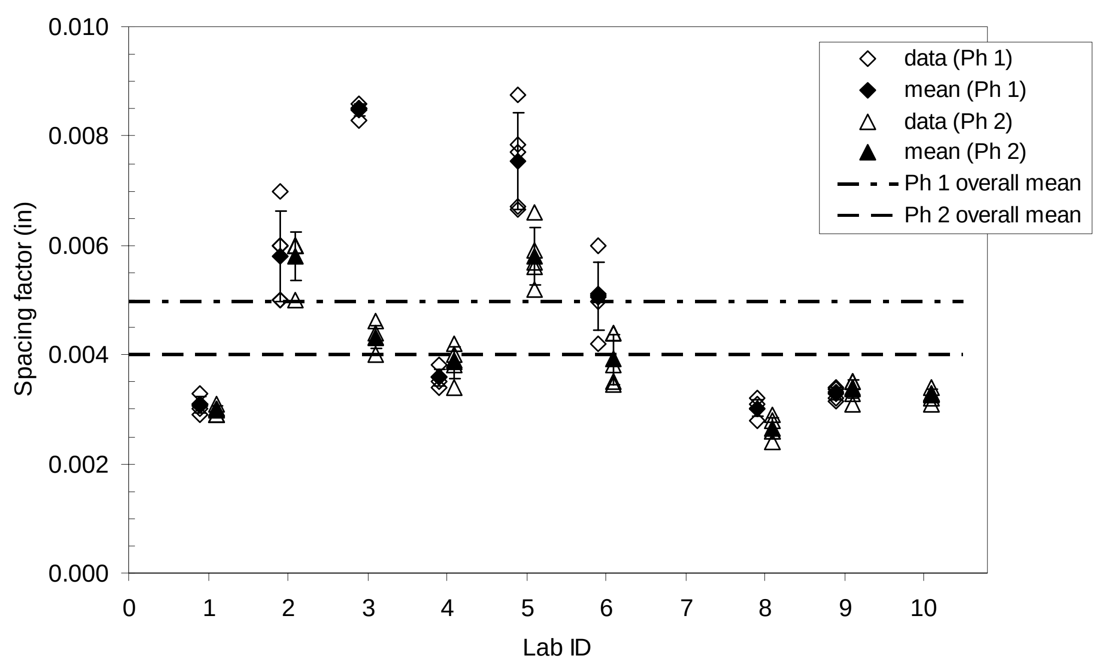

Figure 6. Within and between-laboratory spacing factor, RR2/RR2X (excluding lab 7)

(0.0085 and 0.0075, respectively), but substantially lower results for Phase 2 (0.0045 and 0.006, respectively). Overall, the within-laboratory variability for spacing factor seems to be smaller than for the other air void parameters.

Figures 7 through 9 summarize the within-laboratory coefficients of variation for air content, specific surface, and spacing factor based on five linear traverse results. Results for Phase 1 (RR2) and Phase 2 (RR2X) are shown. For air content, the coefficient of variation ranged from about 3 percent to about 13 percent. Most labs were below 10 percent in both phases. Several labs were around 5 percent or less. Labs 1, 3, 4, 8, and 9 were very consistent from Phase 1 to Phase 2. For specific surface, the coefficients of variation ranged from about 2 percent to nearly 20 percent. Six labs were below 10 percent in both phases, with several at or near 5 percent. Labs 2 and 7 had noticeably different coefficients of variation between phases 1 and 2. For spacing factor, results were generally similar to those for specific surface.

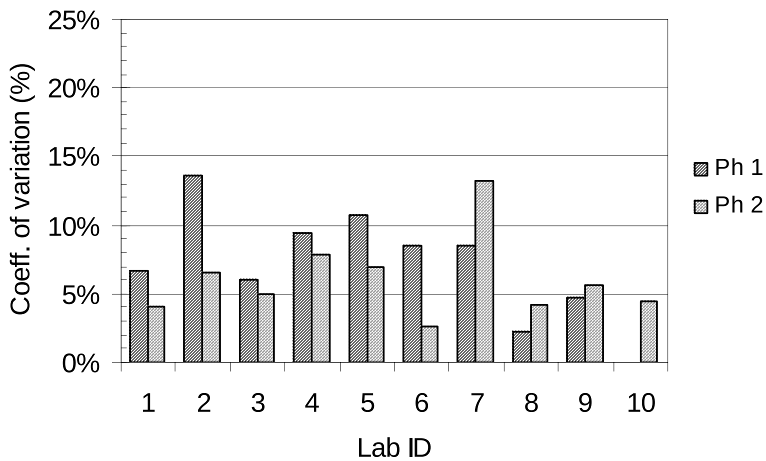

Figure 7. Within-laboratory coefficient of variation for air content (RR2/RR2X)

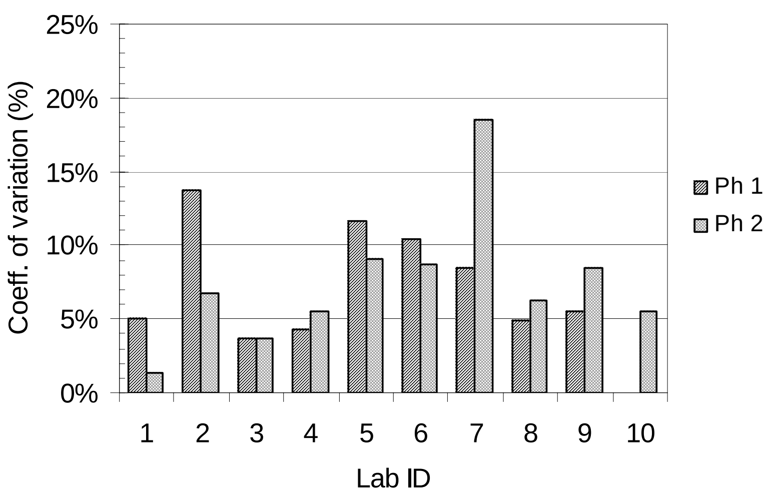

Figure 8. Within-laboratory coefficient of variation for specific surface (RR2/RR2X)

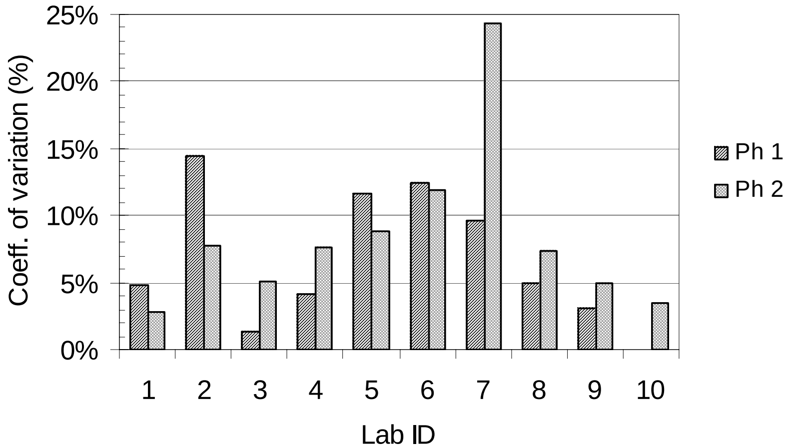

Figure 9. Within-laboratory coefficient of variation for spacing factor (RR2/RR2X)

4.2.2 Reproducibility (between-laboratory variability)

Figures 10 through 13 summarize the Phase 1 between-laboratory air void parameter results for specimens RR1, RR3, RR4 and RR5. Each figure contains three pairs of graphs. The graphs on the left side show the between-laboratory variation in air content, spacing factor, and specific surface for all nine participating labs in Phase 1, while the graphs on the right show corresponding plots for all labs except lab 7. Figures 14 through 17 summarize the Phase 2 between-laboratory air void parameters for RR1X, RR3X, RR4X, and RR5X in a similar fashion.

Each graph in Figures 10 through 17 shows the data points, the mean as a dashed line, and the 99 percent confidence intervals as solid lines. The 99 percent confidence intervals were chosen as the range outside of which data would be considered outliers. When all of the data were considered, lab 7 had significantly more outlier values than any other lab (for example, compare the summary statistics for air content in Table 14 in the Appendix with the data in Table 13). Therefore, a second analysis of the results was performed excluding lab 7.

For this second analysis, referred to as “Subset 1,” the overall means and 99 percent confidence limits were recalculated using the data from the remaining eight labs (nine in Phase 2). In a number of cases, particularly for spacing factor, this resulted in a shifting of the mean value and a narrowing of the confidence limits, which can be seen by comparing the graphs on the left with those on the right for each specimen. The changes for air content and specific surface were typically less dramatic. When the numbers of outliers for each lab were tallied again based on the new Subset 1 values, five labs (labs 1, 4, 8, 9 and 10) had significantly fewer outliers than the others. Thus, the data for these labs, referred to as “Subset 2,” were again analyzed separately. In summary, three distinct analyses were performed – all of the data, Subset 1 (all labs except lab 7), and Subset 2 (labs 1, 4, 8, 9 and 10).

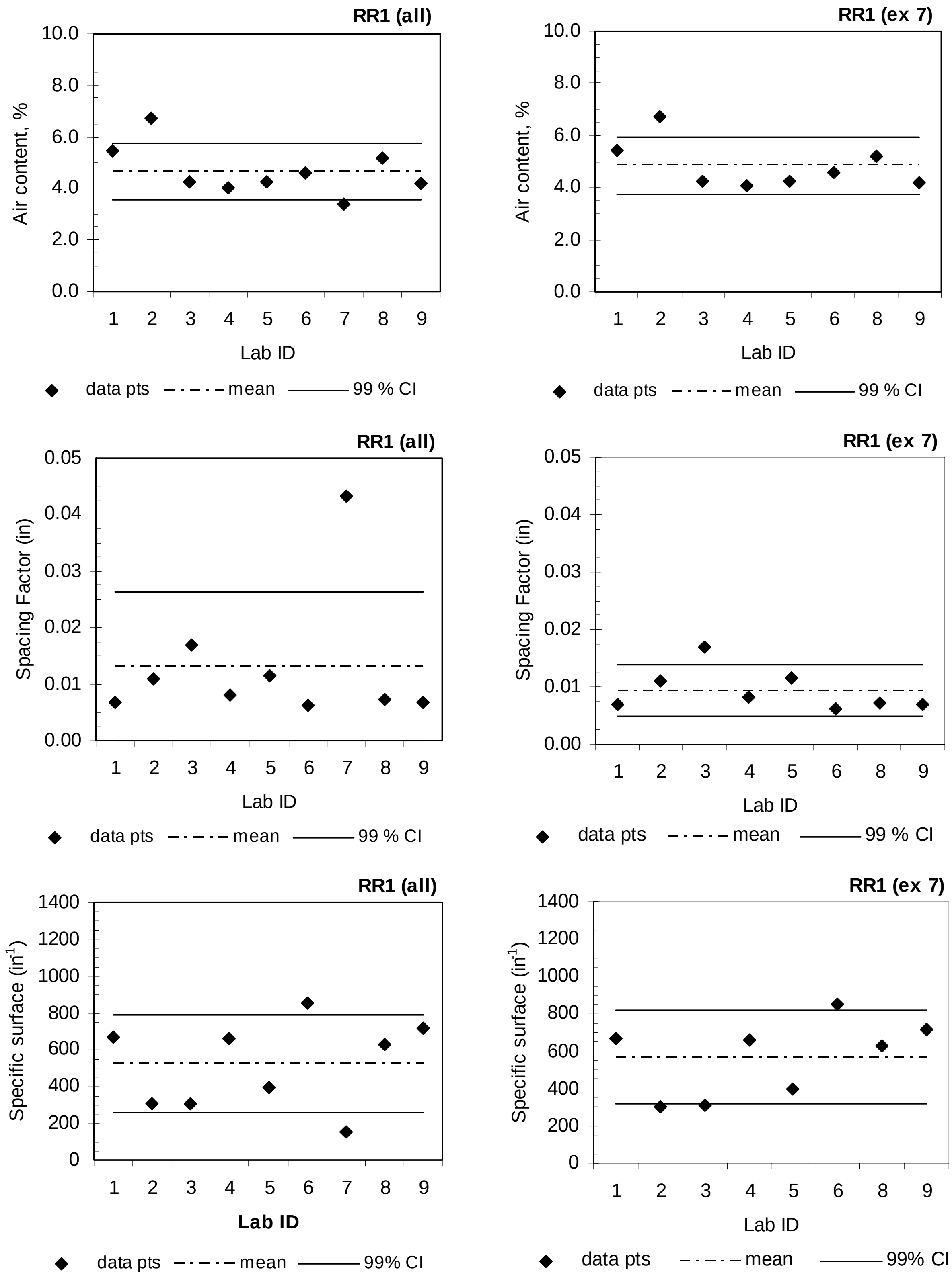

Figure 10. Between-laboratory comparison of air void parameters, RR1

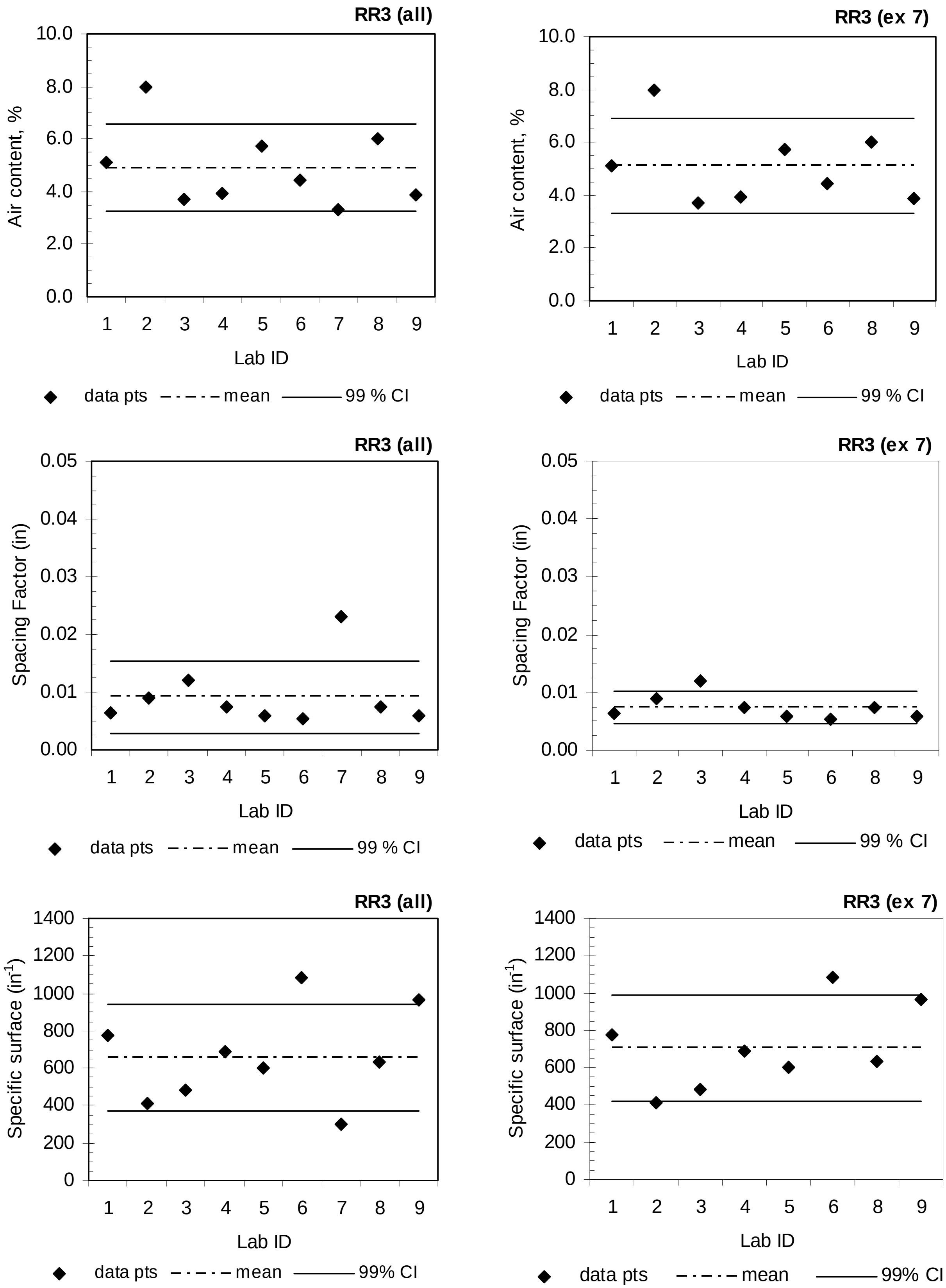

Figure 11. Between-laboratory comparison of air void parameters, RR3

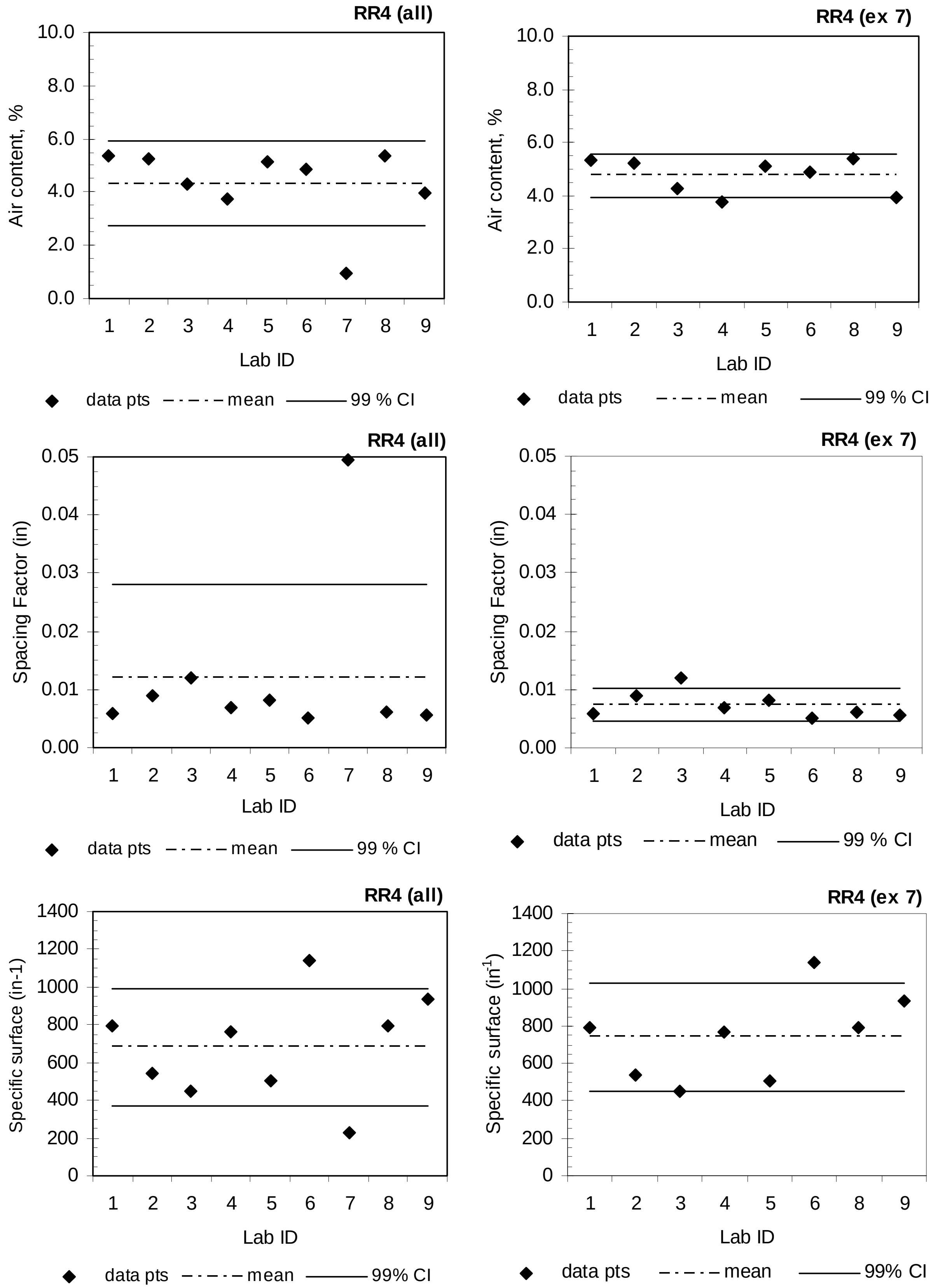

Figure 12. Between-laboratory comparison of air void parameters, RR4

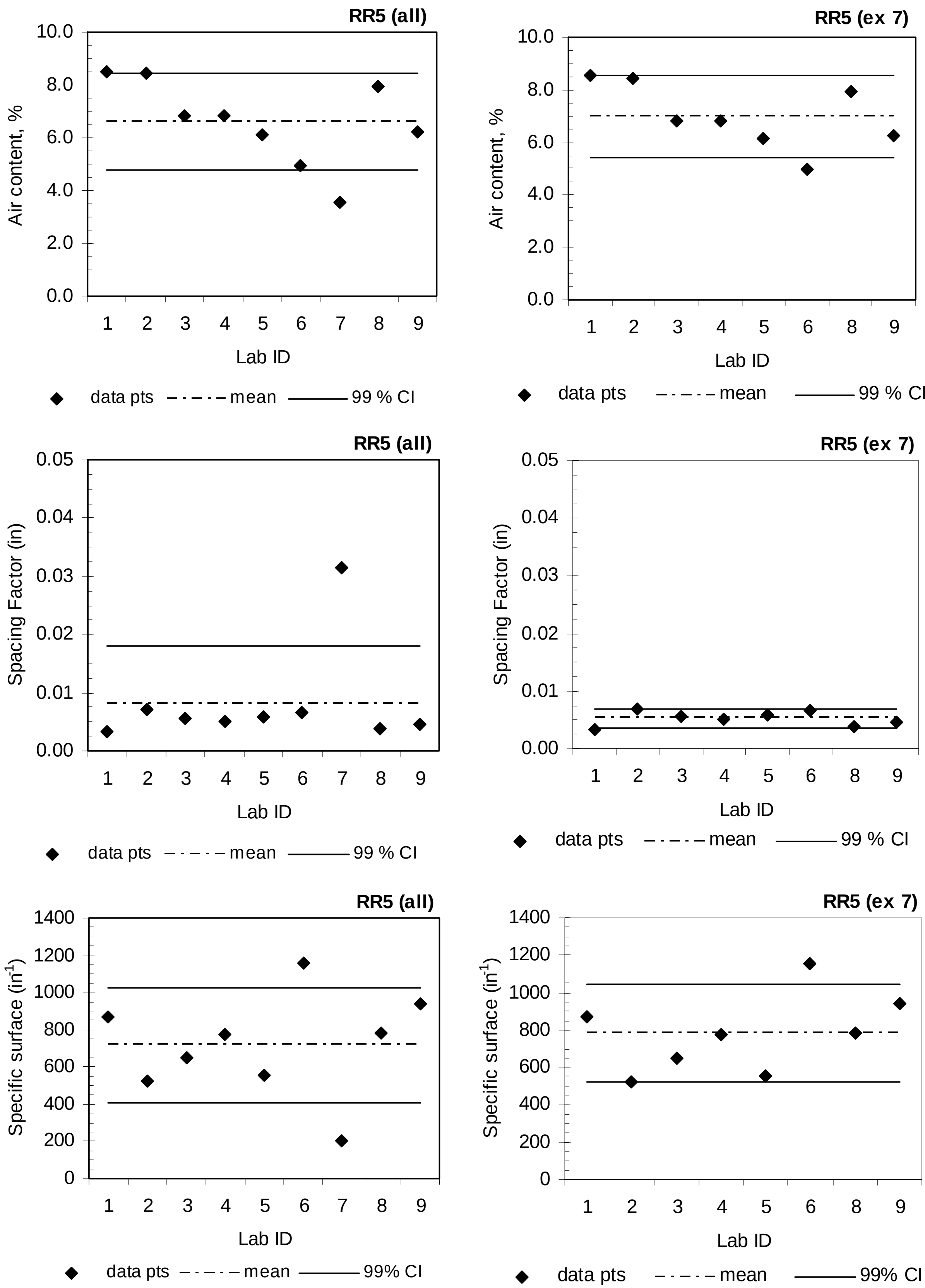

Figure 13. Between-laboratory comparison of air void parameters, RR5

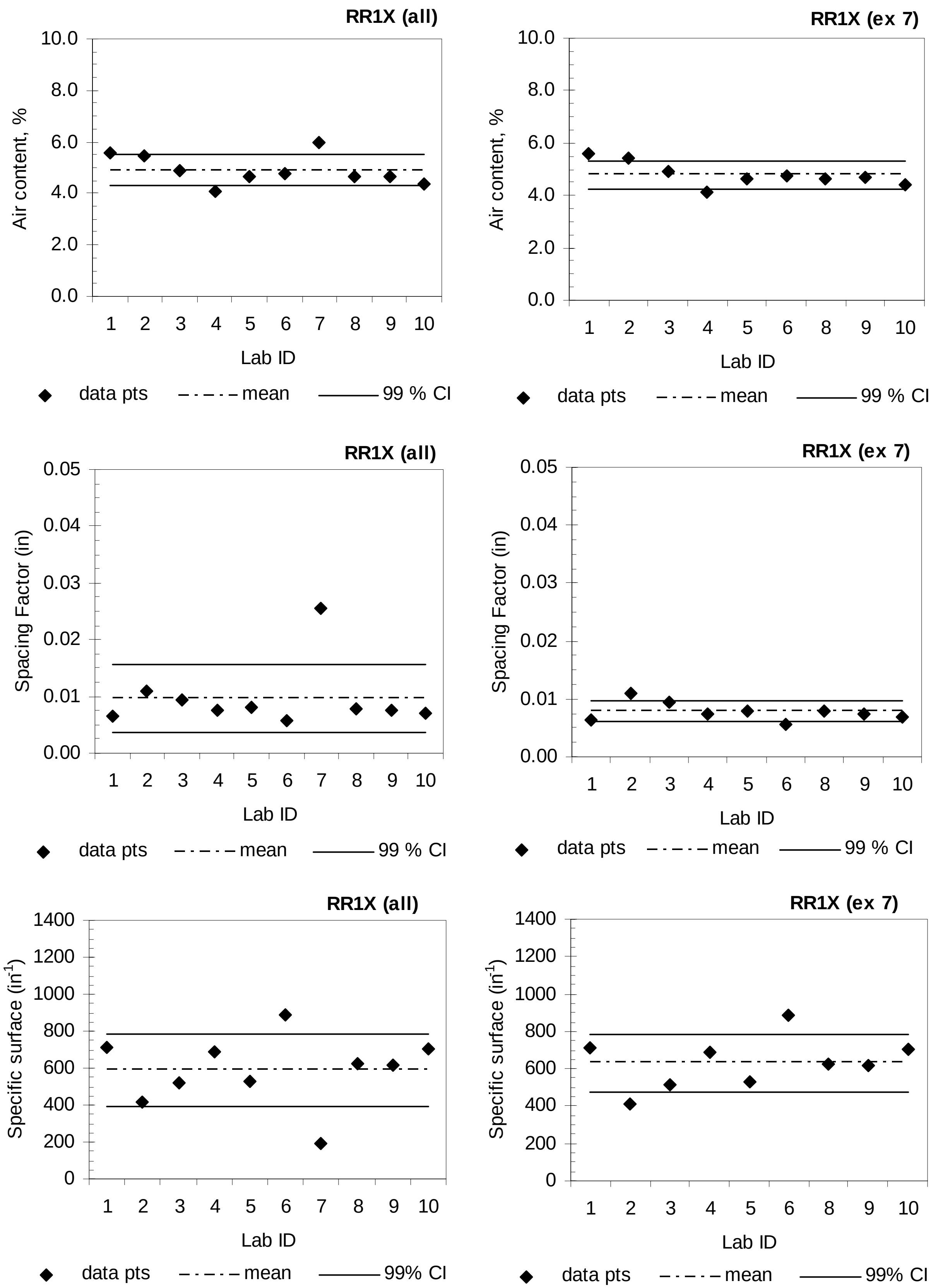

Figure 14. Between-laboratory comparison of air void parameters, RR1X

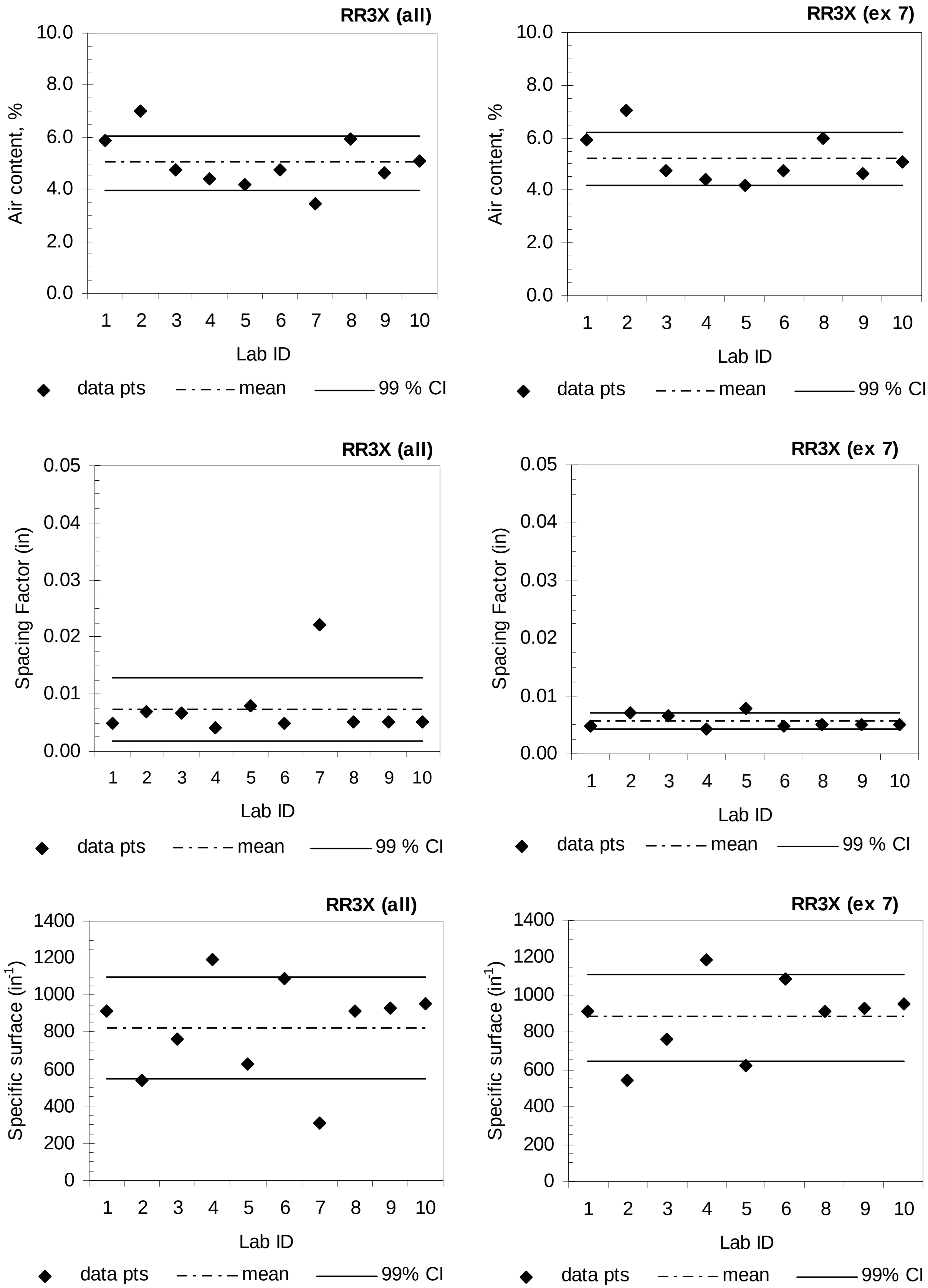

Figure 15. Between-laboratory comparison of air void parameters, RR3X

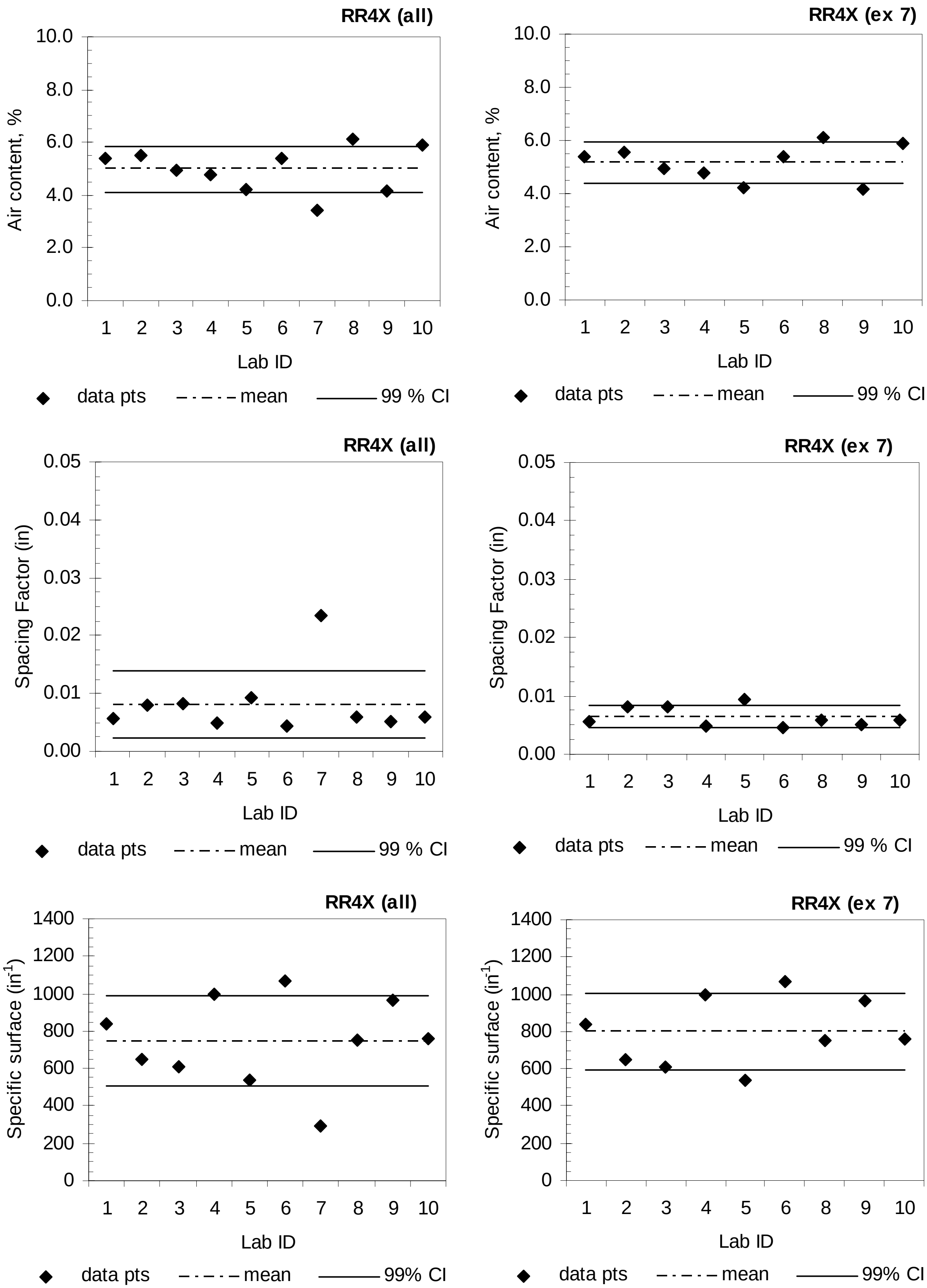

Figure 16. Between-laboratory comparison of air void parameters, RR4X

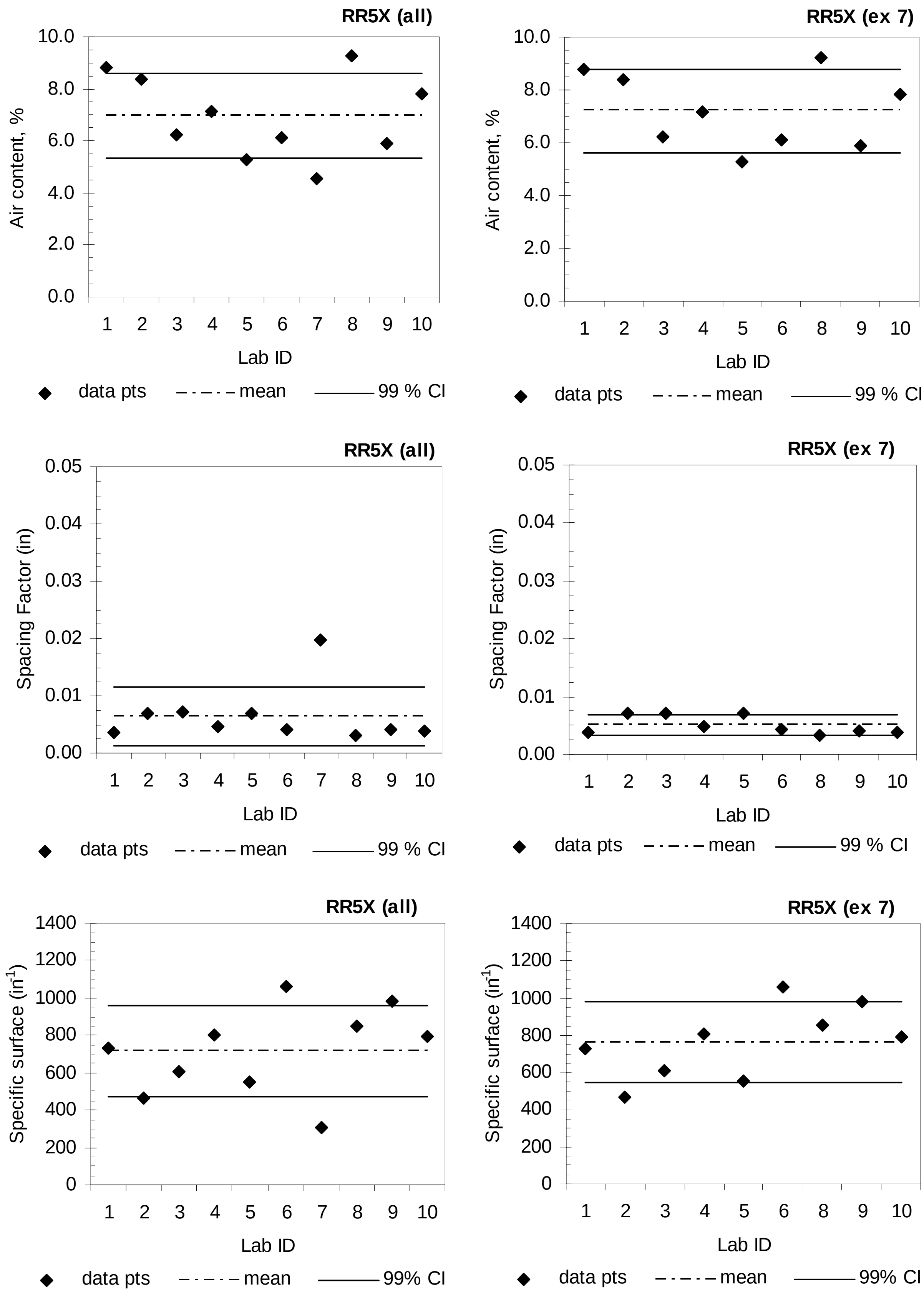

Figure 17. Between-laboratory comparison of air void parameters, RR5X

Between-laboratory standard deviations and coefficients of variation were calculated for air content, specific surface, and spacing factor. These values are plotted in Figures 18 to 20 for Phase 1 and Figures 21 to 23 for Phase 2. The plots show coefficients of variation for the three analysis groups – all data, Subset 1 and Subset 2 – for each of the five specimens. For air content, the coefficients of variation ranged from about 20 to 30 percent for all data. There was generally a slight decrease for Subset 1, and another slight decrease for Subset 2, with values ranging from about 15 to 20 percent for Subset 2. For specific surface, the coefficients of variation for all data ranged from about 35 to 45 percent. They decreased to 25 to 35 percent for Subset 1, and decreased significantly, to 5 to 10 percent for RR1, RR4, and RR5, and 15 to 20 percent for the other two specimens. For spacing factor (Figure 20), the coefficient of variation was quite large for all data, ranging from 60 to 130 percent. For Subset 1, the values dropped dramatically, to 25 to 40 percent, and for Subset 2, the values were generally in the range of 5 to 15 percent (slightly more for RR5). The data from Phase 2 (Figures 21 to 23) generally follow similar trends.

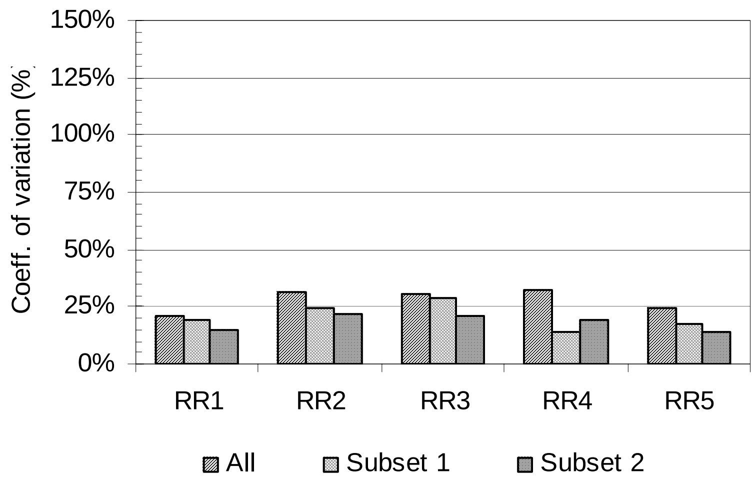

Figure 18. Between-laboratory coefficients of variation for air content (Phase 1)

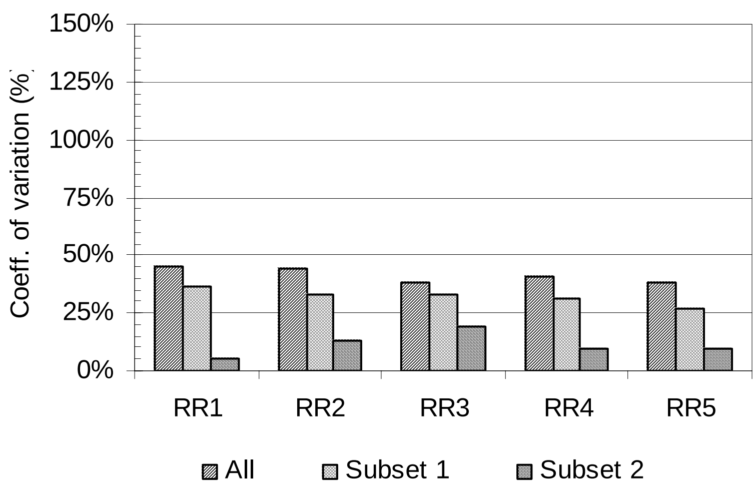

Figure 19. Between-laboratory coefficients of variation for specific surface (Phase 1)

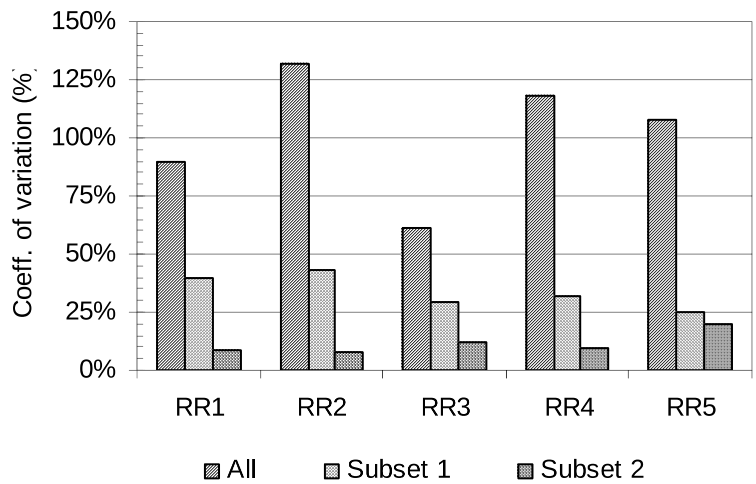

Figure 20. Between-laboratory coefficients of variation for spacing factor (Phase 1)

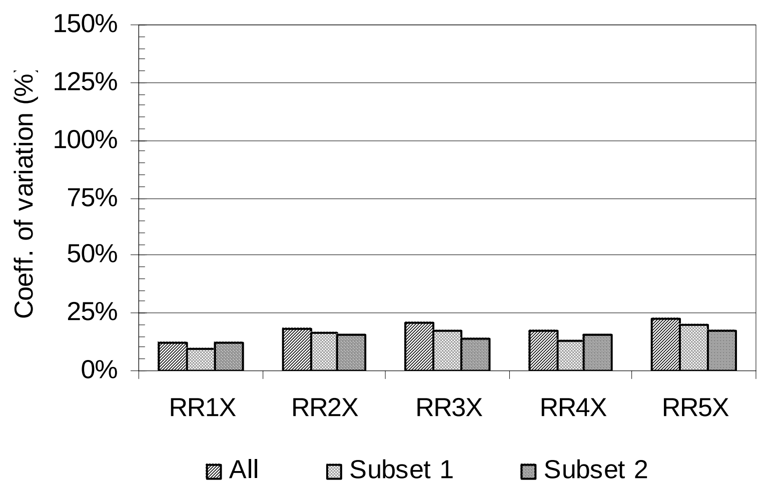

Figure 21. Between-laboratory coefficients of variation for air content (Phase 2)

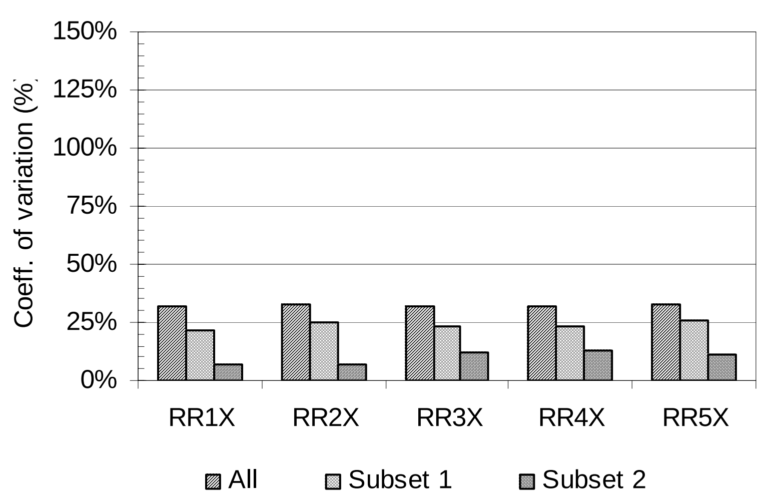

Figure 22. Between-laboratory coefficients of variation for specific surface (Phase 2)

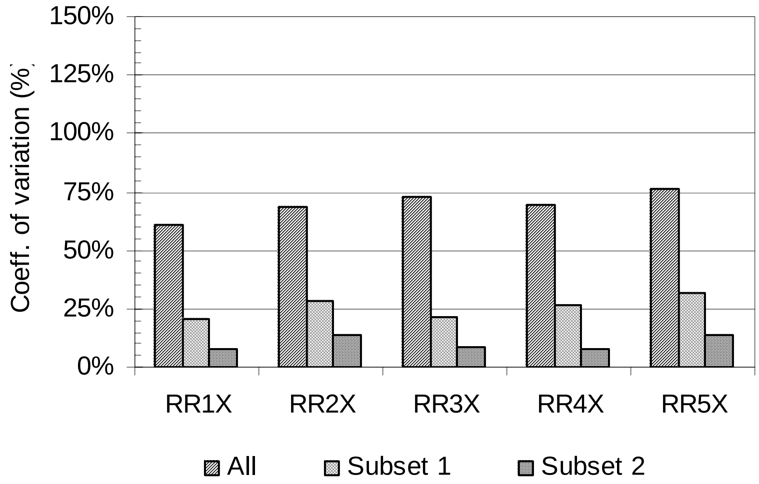

Figure 23. Between-laboratory coefficients of variation for spacing factor (Phase 2)

4.3 Minimum expected uncertainties

The minimum expected uncertainty for an air void parameter reflects the variability inherent in the linear traverse method itself, assuming perfect test conditions and a perfect operator. Table 3 shows the number of chords counted and the minimum expected uncertainties in air content, voids per inch, spacing factor, and specific surface. The values shown are coefficients of variation. Calculations were performed for all of the Phase 1 results and for RR1X in Phase 2[11]. For each parameter, the row labeled "calc" were calculated using the equations in Table 2 (Section 2.2.4). The rows labeled "actual" are the between-laboratory Subset 2 results from the MoDOT round robin.

Table 3. Minimum expected uncertainties in air void parameters (selected specimens)

		RR1	RR1X	RR2	RR3	RR4	RR5
Number of chords*	N_a	856	809	1887	1001	976	1687
Air content, C_a	calc	12.8	12.7	6.2	6.4	7.7	8.4
	actual	14.9	**12.1**	21.9	21.3	18.9	14.1
Voids per inch, C_n	calc	3.4	3.5	2.3	3.2	3.2	2.5
	actual	13.1	16.0	9.2	15.5	17.3	14.4
Spacing factor, $C_{\bar{L}}$	calc	7.0	7.0	3.8	3.9	4.5	4.5
	actual	8.8	7.9	7.8	12.0	9.4	19.4
Specific surface, C_α	calc	12.3	12.2	5.8	5.6	6.9	8.0
	actual	**5.5**	**6.9**	13.4	19.1	9.5	9.3

*equals average number of chords counted for all data sets considered. RR1 – two data sets; RR2 – three data sets; all others – four data sets. Data from labs 1, 4, 8 and 9 were used in these calculations.

[11]Because of the close agreement between RR1 and RR1X, calculations were not performed for the remaining Phase 2 results; they were assumed be similar to the corresponding Phase 1 results.

The calculated minimum uncertainties for air content, spacing factor, and specific surface are relatively consistent for each specimen with the exception of RR1. In the case of RR1 (and RR1X), it is possible that several very large chords contributed to the greater uncertainty, because the chord length distribution has a substantial influence on the uncertainty. A few very large chord lengths (3 mm or more) could contribute disproportionately to the calculated minimum uncertainty. Using a maximum cutoff value for chord lengths might help alleviate this problem.

A comparison of the results for specimens RR1 and RR4 suggest that the minimum uncertainty is not necessarily inversely proportional to the number of chords counted, except in the case of voids per inch.

Table 3 indicates that for all but three cases (highlighted in bold in the table), the between-laboratory variability exceeds the minimum expected uncertainty for all air void parameters. The exceptions occurred for RR1. As explained above, the minimum uncertainties for RR1 were atypical compared with the other specimens, possibly due to the contribution of very large measured chord lengths.

4.4 ASTM C 670 precision limits

ASTM C 670[12] is intended to provide guidance in preparing precision and bias statements for ASTM test methods related to construction materials. ASTM C 457 discusses precision according to ASTM C 670 using precision limits based on two round robin studies. In this section of the report, ASTM C 670 precision limits are calculated for the MoDOT round robin data and are compared with the precision limits reported in ASTM C 457 and with the minimum expected uncertainties calculated in Section 4.3.

ASTM C 670 defines the population standard deviation, σ, to be "the fundamental statistic underlying all indexes of precision" [ASTM C 670 2004]. ASTM defines two indices of precision based on σ: the single-operator one-sigma limit and the multi-laboratory one-sigma limit, both denoted "1s".[13] The former, synonymous with within-laboratory precision or repeatability, estimates the variability of a large group of individual test results made on the same material by a single operator using the same apparatus in the same laboratory. The latter, synonymous with between-laboratory precision or reproducibility, estimates the variability of a large group of individual test results, each made in a different laboratory, using test portions of material that are as identical as possible.

4.4.1 Single-operator precision – 1s limit

Tables 4 and 5 summarize the single-operator one-sigma limits (1s) for air content and the one-sigma limits in percent (1s%) for voids per inch, spacing factor, and specific surface. Table 4 summarizes the Phase 1 results based on five traverses per lab on specimen RR2. Table 5 summarizes the Phase 2 results based on five traverses per lab on specimen RR2X. In both

[12] "Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials"
[13] ASTM C 670 also designates the "one-sigma limit in percent" (denoted "1s%"), equal to the standard deviation divided by the mean (i.e., the coefficient of variation), to be used in certain situations.

cases, the reported one-sigma limit for each parameter is the pooled standard deviation[14] for that parameter, and the reported one-sigma limit in percent (coefficient of variation) is the pooled standard deviation divided by the overall mean for all laboratories. As shown in the tables, the one-sigma limits were calculated for all laboratories, Subset 1, and Subset 2.

The fourth column under "MoDOT Round Robin," labeled "Min expected," shows the minimum expected uncertainties from Table 3 for the listed air void parameters. The last two columns under the heading "ASTM C 457" contain the values from other round robin studies reported in ASTM C 457 Tables 4 and 6 [ASTM C 457 2004].

Table 4. ASTM C 670 single operator precision – Phase 1 (RR2)

Parameter		MoDOT Round Robin				ASTM C-457	
		All labs	Subset 1	Subset 2	Min. expected	Table 4	Table 6
Air	1s	0.50	0.52	0.41	–	0.29	0.57
	1s%	8.7%	8.5%	6.0%	6.2%	–	–
Voids per inch	1s	1.02	1.08	0.97	–	–	–
	1s%	8.1%	7.7%	5.1%	2.3%	3.7%	–
Spacing factor	1s	0.0013	0.00049	0.00014	–	–	–
	1s%	15.3%	9.9%	4.3%	3.8%	–	8.0%
Specific surface	1s	63	67	57	–	–	–
	1s%	7.6%	7.3%	5.0%	5.8%	–	–

Table 5. ASTM C 670 single operator precision – Phase 2 (RR2X)

Parameter		MoDOT Round Robin				ASTM C-457	
		All labs	Subset 1	Subset 2	Min. expected	Table 4	Table 6
Air	1s	0.39	0.35	0.35	–	0.29	0.57
	1s%	6.1%	5.4%	5.2%	6.2%	–	–
Voids per inch	1s	1.12	1.13	1.17	–	–	–
	1s%	7.4%	6.9%	6.1%	2.3%	3.7%	–
Spacing factor	1s	0.0011	0.00032	0.00019	–	–	–
	1s%	22.8%	7.9%	5.8%	3.8%	–	8.0%
Specific surface	1s	67	66	69	–	–	–
	1s%	7.1%	6.6%	6.0%	5.8%	–	–

As expected, the coefficients of variation generally decrease as the analysis proceeds from all labs to Subset 1 to Subset 2. The change is especially noticeable from all labs to Subset 1, where

[14]The pooled standard deviation is the square root of the weighted average of the variances for each specimen. The equation at the bottom of page 42 of this report illustrates how the pooled standard deviation is calculated.

the data from lab 7 were excluded. Since the minimum expected value is the theoretical minimum, assuming perfect test conditions, the experimental data should not be less than the minimum expected value. This was the generally the case for all parameters in Phase 1, and all parameters except air in Phase 2. Similarly, the values from the previous ASTM studies are also greater than the theoretical minimum.

Comparing the Phase 1 round robin results to the previous ASTM values, the round robin within-laboratory standard deviations (1s limit) for air content were 0.52 for Subset 1 and 0.41 for Subset 2. These values exceed the "average" standard deviation of 0.29 reported in Table 4 of ASTM C 457; however, they are less than the reported value of 0.57 cited in Table 6 of ASTM C 457 (based on results of a European study for specimens prepared and tested by the same laboratory). In Phase 2, the MoDOT round robin within-laboratory standard deviations for air content (0.35 for Subset 1 and 0.35 for Subset 2) were less than in Phase 1 and closer to the reported value of 0.29 from ASTM C 457 Table 4.

The within-laboratory 1s% limits (coefficients of variation) for voids per inch ranged from 5.1 to 8.1 percent for Phase 1 and 6.1 to 7.4 percent for Phase 2. These values are higher than the reported value of 3.7 percent in ASTM C 457 Table 4. For spacing factor, the results for Subset 1 in Phase 1 is slightly higher than the result of 8.0 percent reported in ASTM C 457 Table 6. The results for Subset 1 in Phase 2 and Subset 2, however, are less than the reported value of 8.0 percent.

Overall, the round robin results for Subset 2 compare reasonably well to the reported results in ASTM C 457. The within-laboratory variations of the air void parameters (reflected Tables 4 and 5) were greater than the minimum expected uncertainties (Table 3), as expected.

4.4.2 Multi-laboratory precision – 1s limit

Figures 24 to 27 show the multi-laboratory one-sigma limits for air content, voids per inch, spacing factor, and specific surface. Each figure shows the results for each of the five round robin specimens.[15] Four sets of results are shown: Phase 1 Subset 1 (■), Phase 1 Subset 2 (▲), Phase 2 Subset 1 (□) and Phase 2 Subset 2 (△). Data from ASTM C 457 Tables 4 and 6, where available, are shown in the figures as dashed or solid lines, respectively.

In Figure 24, the one-sigma limits (1s limit) for air content were generally greater for Subset 1 than for Subset 2. No consistent differences were apparent between Phase 1 and Phase 2. All of the round robin results were greater than the limit of 0.41 reported in ASTM C 457, Table 4. Several results, especially for Subset 2 (both phases), approached the limit of 0.73 reported in ASTM C 457, Table 6.

In Figures 25, 26 and 27, the precision limits for Subset 2 were significantly less than those of Subset 1 for both phases. Again, there were no consistent differences between Phase 1 and Phase 2. In Figure 25, the Phase 2 results for voids per inch were close to the reported value

[15] These figures show only the results for subsets 1 and 2. Tabulations of ASTM C 670 multi-laboratory limits for all laboratories, Subset 1, and Subset 2 can be found in Section A.4 of the Appendix.

from ASTM C 457, Table 4. In Figure 26, the Phase 2 results for spacing factor fell below the reported values in ASTM C 457, Table 6.

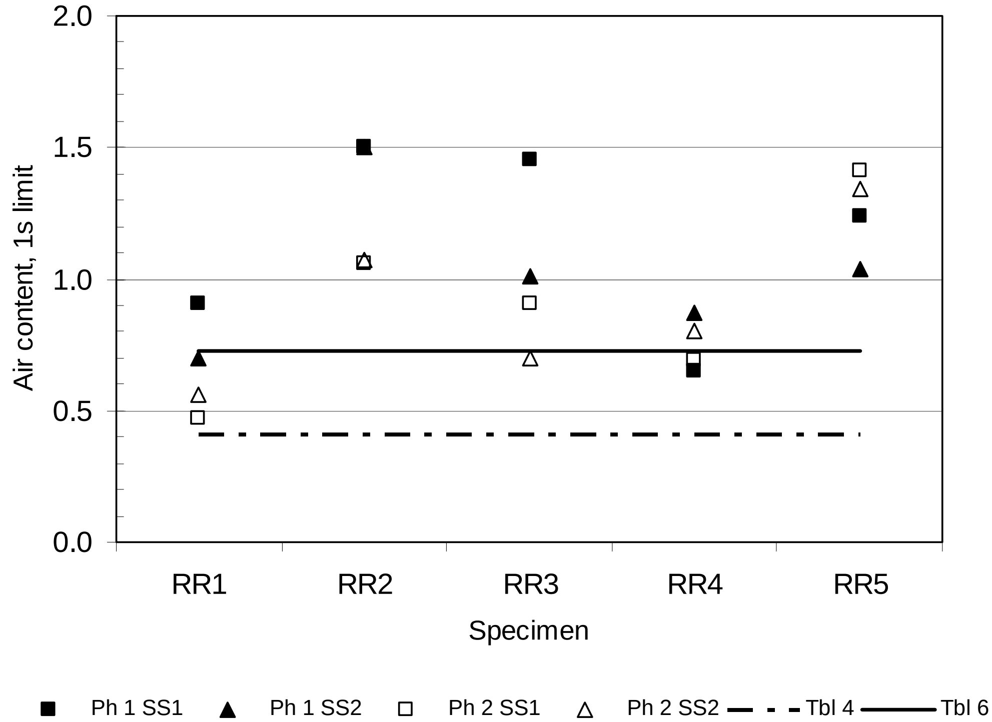

Figure 24. Multilaboratory precision for air content

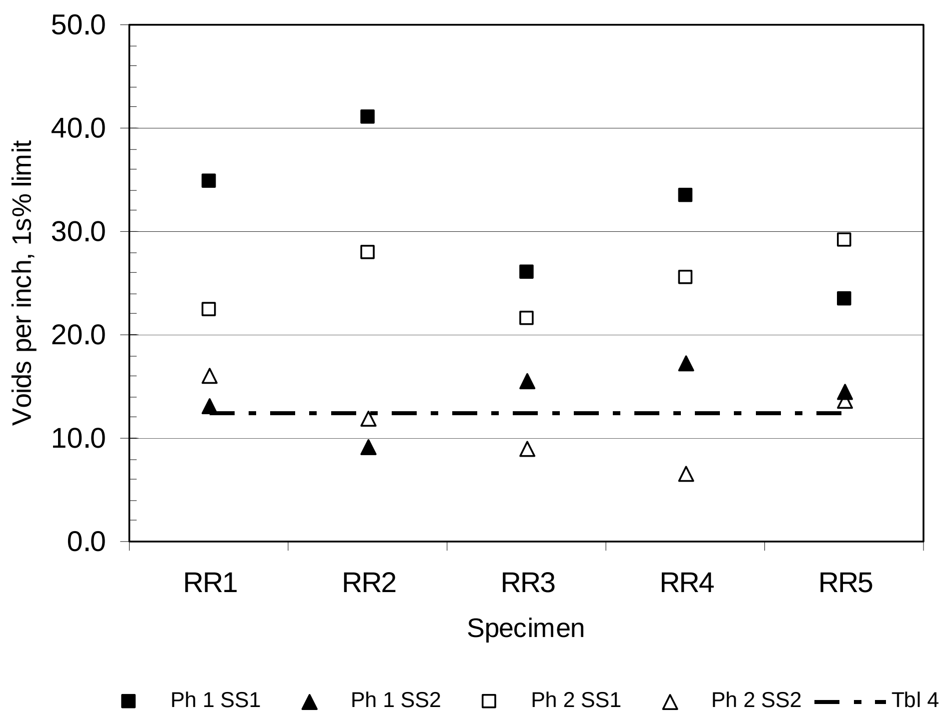

Figure 25. Multilaboratory precision for voids per inch

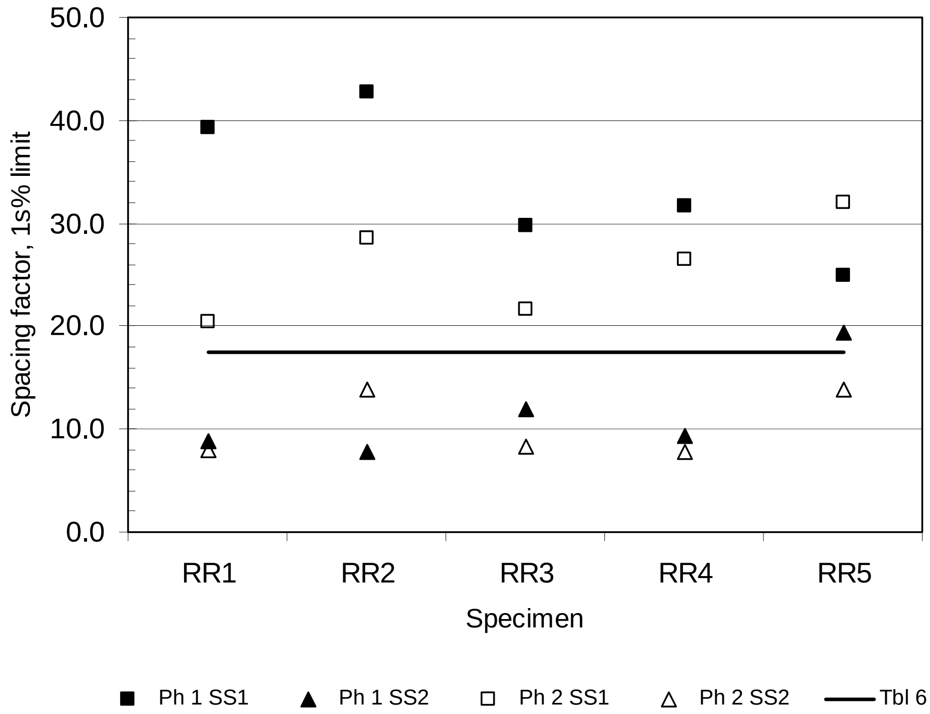

Figure 26. Multilaboratory precision for spacing factor

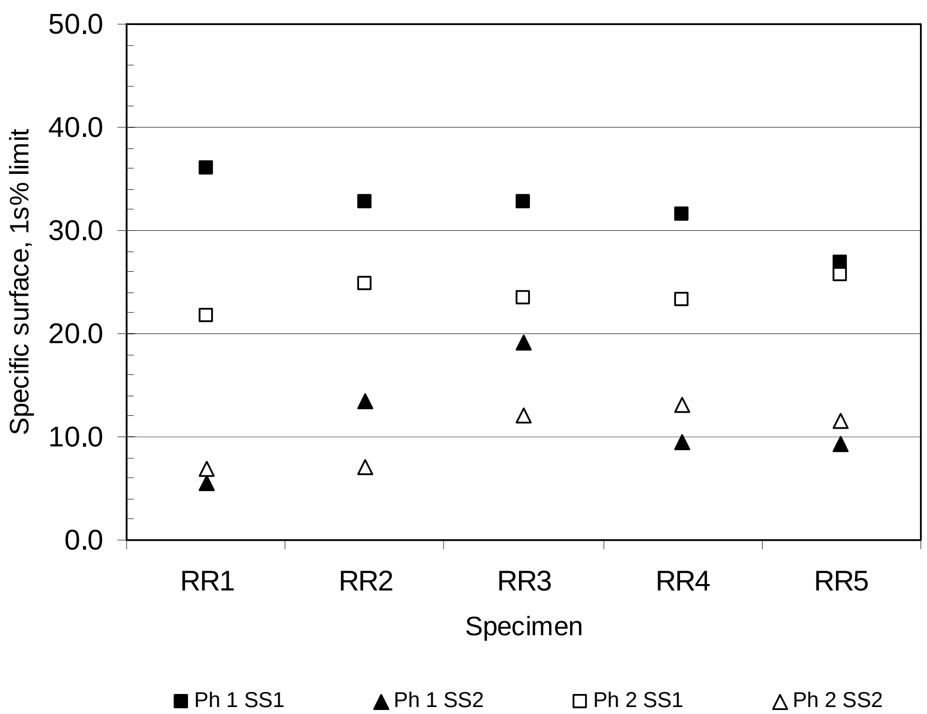

Figure 27. Multilaboratory precision for specific surface

ASTM C 670 defines a precision limit for the acceptable range between two test results. This limit, d2s[16], represents "the difference between two individual test results that would be equaled or exceeded in the long run in only 1 case in 20 in the normal and correct operation of the method." [ASTM C 670]. The d2s (or d2s%) precision limit is defined as $2.83 \times 1s$ (or 1s%). For situations where more than two results are reported, ASTM C 670 Table 1 provides an appropriate multiplier to be substituted for 2.83 in order to calculate the acceptable range. For example, the multiplier is 3.6 for four results and 3.9 for five results.

Table 6 compares the range of results for each air void parameter from Phase 2, Subset 2, with the ASTM C 670 acceptable range precision limits. For each specimen, the range of the between-laboratory results for air content are compared with the calculated d2s limit, and the range as percent of the mean of the between-laboratory results for voids per inch, spacing factor,

Table 6. Between-laboratory results compared with ASTM C 670 d2s or d2s% limits

Parameter	ID	Mean	Std dev (1s)	Range	d2s	Range (% of mean)	d2s%
Air content	RR1	4.68	0.56	1.50	1.58	–	–
	RR2	6.77	1.07	2.75	3.03	–	–
	RR3	5.20	0.70	1.54	1.98	–	–
	RR4	5.27	0.80	1.96	2.26	–	–
	RR5	7.78	1.34	3.36	3.79	–	–
Voids per inch	RR1	7.82	1.25	2.95	–	37.7	45.3
	RR2	19.20	2.29	5.00	–	26.0	33.7
	RR3	12.59	1.14	2.69	–	21.4	25.5
	RR4	11.20	0.74	2.00	–	17.9	18.7
	RR5	16.03	2.18	5.35	–	33.4	38.5
Spacing factor	RR1	0.0073	0.00058	0.0015	–	20.7	22.4
	RR2	0.0032	0.00045	0.0012	–	37.2	39.3
	RR3	0.0049	0.00041	0.0010	–	20.4	23.5
	RR4	0.0055	0.00042	0.0010	–	18.2	21.8
	RR5	0.0039	0.00054	0.0015	–	38.5	39.3
Specific surface	RR1	668	46	94	–	14.1	19.5
	RR2	1142	82	205	–	17.9	20.1
	RR3	980	119	278	–	28.4	34.2
	RR4	863	113	245	–	28.4	37.1
	RR5	832	95	255	–	30.6	32.5

[16]ASTM C 670 also defines d2s% which is analogous to 1s%.

and specific surface are compared with the calculated d2s% limits[17]. In all cases, the range of the round robin results for the five laboratories is less than the calculated d2s or d2s% limit. Because the round robin results fall within the d2s or d2s% limits, the round robin will also fall within the acceptable range for five test results according to ASTM C 670. Therefore, there is no reason to believe that any of the tests in Phase 2, Subset 2 were improperly conducted.

4.5 Phase 1 versus Phase 2

The Phase 1 and Phase 2 data were compared to ascertain the effects of sample preparation on the test results. Because the data for each lab are paired observations, statistical tests are based on the differences in each air void parameter [Box et al 1978]. For example, if A_1 is the air content measured by lab 1 in Phase 1, and A_2 is the air content measured by lab 1 in Phase 2, the difference in air content for lab 1 is $D_1 = A_1 - A_2$. A t-test is performed to see if the mean of the differences for all labs is statistically significant from zero. A significant result would indicate a difference between Phase 1 and Phase 2. A significance level of 0.05 was used in the analysis, which is summarized in Tables 7 through 11. In the tables, significant p-values (less than 0.05) are shown in bold type.

Table 7. Significance tests for differences in mean air void parameters, RR1

Parameter	Mean diff.	S. D.	$t_{(.05,n-1)}$	Std. Error	p-value
Air content	-0.29	1.04	2.306	0.35	0.42
Mean chord length	0.0019	0.0024	2.306	0.0008	**0.05**
Voids per inch	-0.85	1.23	2.306	0.41	0.07
Spacing factor	0.0032	0.0060	2.306	0.0020	0.14
Spec surface	-54	89	2.306	30	0.11

Table 8. Significance tests for differences in mean air void parameters, RR2

Parameter	Mean diff.	S. D.	$t_{(.05,n-1)}$	Std. Error	p-value
Air content	-0.60	0.93	2.306	0.31	0.09
Mean chord length	0.0015	0.0033	2.306	0.0011	0.20
Voids per inch	-2.14	3.25	2.306	1.08	0.08
Spacing factor	0.0036	0.0081	2.306	0.0027	0.22
Spec surface	-90	146	2.306	49	0.10

[17] The d2s limits are used for comparison even though the results of Phase 2, Subset 2 contain five test results (where each test result is the average of the results of the five laboratories). Because the multiplier increases with the number of test results, the d2s is a more conservative limit. If the data fall below the d2s limit they will automatically fall within the broader limits that apply to a greater number of test results.

Table 9. Significance tests for differences in mean air void parameters, RR3

Parameter	Mean diff.	S. D.	$t_{(.05,n-1)}$	Std. Error	p-value
Air content	-0.12	0.86	2.306	0.29	0.70
Mean chord length	0.0012	0.0012	2.306	0.0004	**0.02**
Voids per inch	-2.27	2.58	2.306	0.86	**0.03**
Spacing factor	0.0016	0.0020	2.306	0.0007	**0.05**
Spec surface	-148	177	2.306	59	**0.04**

Table 10. Significance tests for differences in mean air void parameters, RR4

Parameter	Mean diff.	S. D.	$t_{(.05,n-1)}$	Std. Error	p-value
Air content	-0.56	0.91	2.306	0.30	0.10
Mean chord length	0.0010	0.0014	2.306	0.0005	0.07
Voids per inch	-1.49	1.62	2.306	0.54	**0.02**
Spacing factor	0.0037	0.0085	2.306	0.0028	0.23
Spec surface	-62	95	2.306	32	0.08

Table 11. Significance tests for differences in mean air void parameters, RR5

Parameter	Mean diff.	S. D.	$t_{(.05,n-1)}$	Std. Error	p-value
Air content	-0.25	0.78	2.306	0.26	0.36
Mean chord length	0.0005	0.0023	2.306	0.0008	0.55
Voids per inch	-0.25	2.11	2.306	0.70	0.73
Spacing factor	0.0014	0.0040	2.306	0.0013	0.32
Spec surface	10	79	2.306	26	0.72

Four of the five air void parameters for RR3 were significantly different between Phase 1 and Phase 2, suggesting that surface preparation may have contributed to the variability for RR3. In contrast, the other four specimens had either no parameters different (RR2 and RR5) or only one parameter different (RR1 and RR4) between Phases 1 and 2.

When the significance tests were repeated for Subset 1 (results not shown in this report), there were three statistically significant differences for RR3 (mean chord length, voids per inch, and spacing factor) while there was only one other significant difference among the remaining specimens (voids per inch for RR4). Again, this indicates that specimen preparation may have contributed to the variability for RR3 but contributed little to the variability for the other specimens.

4.6 Chord Distributions

Several laboratories[18] provided raw data files containing individual chord length data for the linear traverse tests they performed. These data were sorted into bins to create histograms of chord length distributions. Figures 28 through 30 show chord distributions from laboratories 1, 4 and 8, respectively, for the five traverses on specimen RR2. These figures show generally consistent results in terms of chord distributions within each laboratory. Some variation is to be expected for the chords measured over five distinct sets of traverse lines.

The figures also show distinct differences between laboratories – for example, lab 1 consistently found very few chord lengths in the first bin (size range 0 to 25 microns), while lab 8 found the most in the smallest bin. Thus, each laboratory appears to have its own inherent bias (based on equipment, operator, or both) in terms of its chord distribution results. Nevertheless, in many cases these laboratories showed excellent agreement in terms of air void parameters, and these labs were all part of Subset 2 – the most consistent labs in the study.

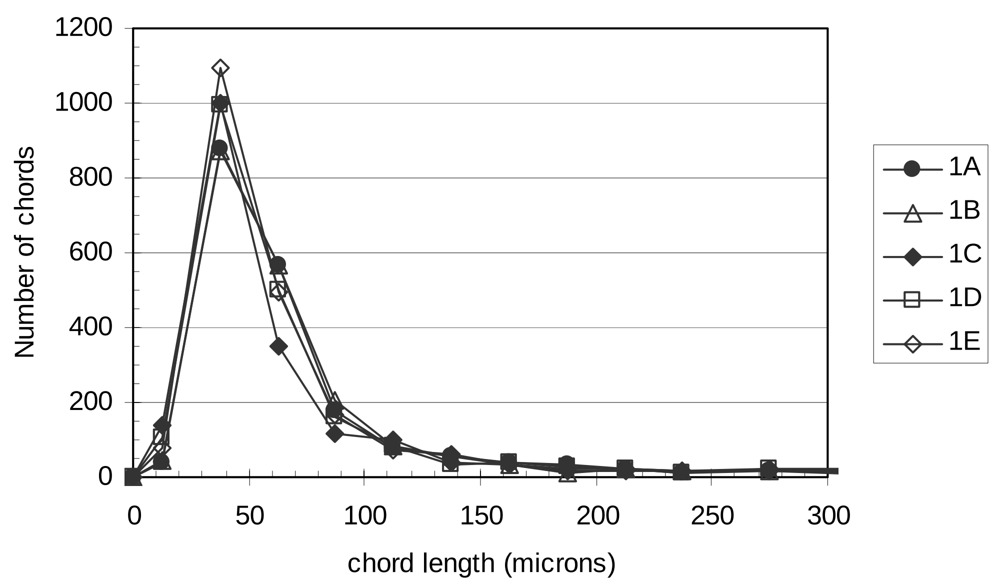

Figure 28. Chord distributions for RR2, Lab 1

[18]Laboratories 1, 4, 8 and 9 provided individual chord length data; however, the data from lab 9 for RR2 were not used because problems were encountered processing the raw data files.

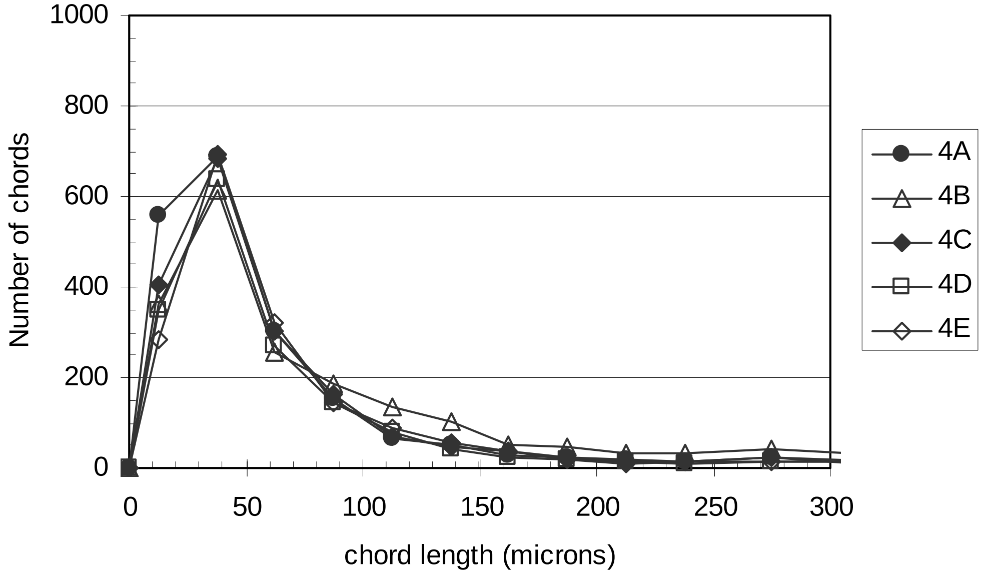

Figure 29. Chord distributions for RR2, Lab 4

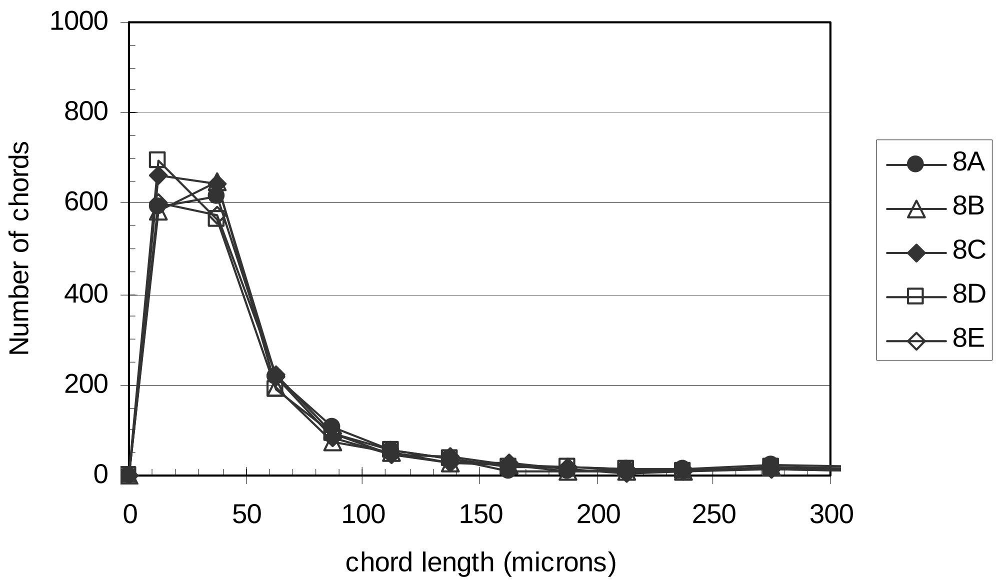

Figure 30. Chord distributions for RR2, Lab 8

Figures 31 and 32 compare the chord distributions obtained in four laboratories – labs 1, 4, 8, and 9 – for specimen RR5. Figure 27 shows the distributions over the range 0 to 1000 micrometers, while Figure 28 shows the lower end of the distribution, ranging from 0 to 300 micrometers

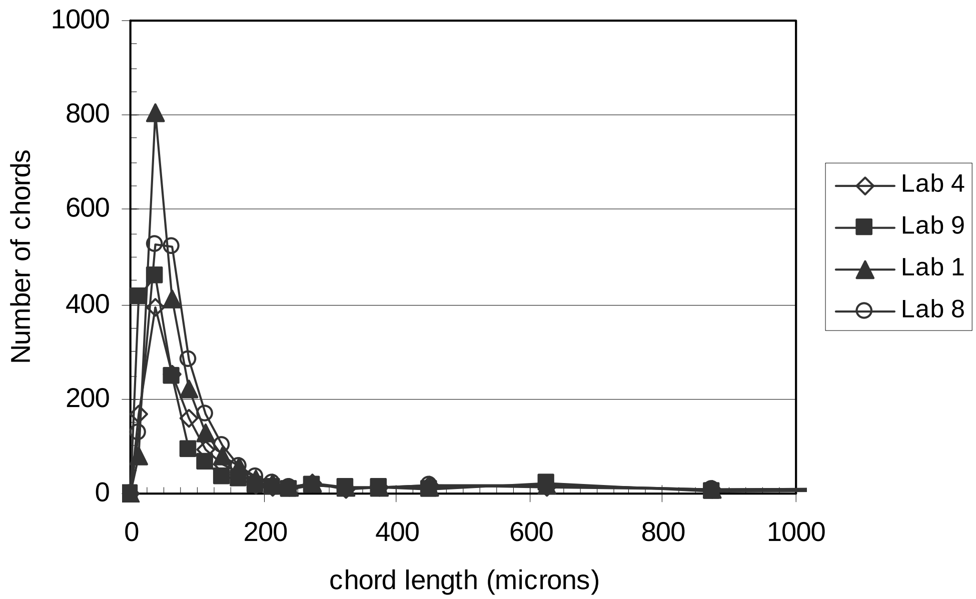

Figure 31. Chord distributions for RR5

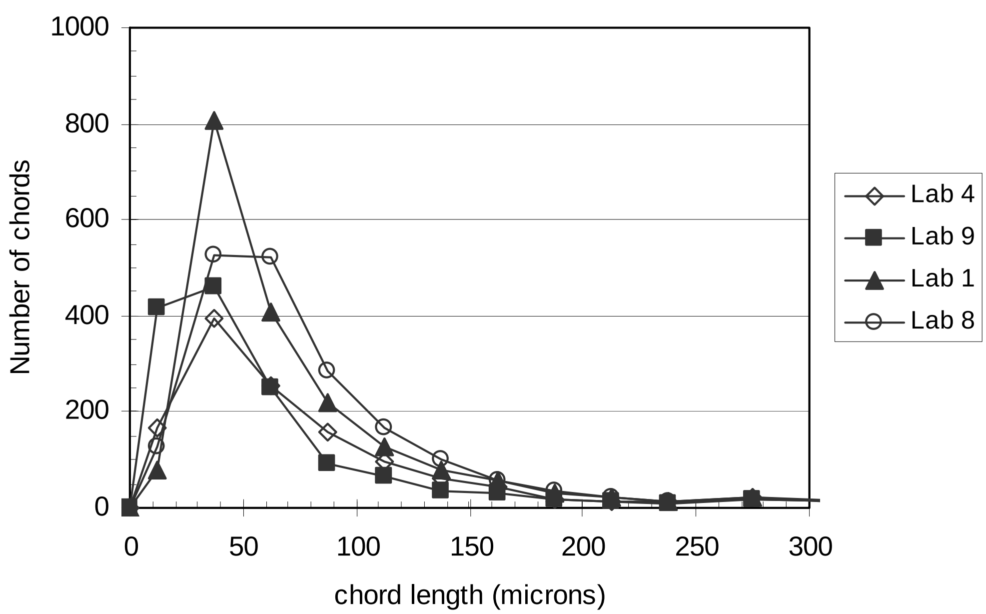

Figure 32. Lower end of chord distributions for RR5

Chapter 5
Conclusions and Recommendations

5.1 Summary of findings

The round-robin study results suggest the following general findings and conclusions:

- Within-laboratory variation, as expected, was less than between-laboratory variation; however, although within-laboratory variation provides a means of assessing operator consistency, it does not provide a useful indication of accuracy. Large within-laboratory variation may or may not be coupled with poor accuracy.
- Between-laboratory results exhibited wide variation, even when the data were confined to the subset of the five best labs. The differences between laboratories most likely reflect differences in operator ability, because the study was designed to minimize the effects of different equipment by requiring that all laboratories use the same magnification for all tests. Other equipment differences, such as use of a monitor instead of a microscope for viewing, may have had a minor effect; however, this effect would not be expected to cause such wide variation.
- The results of the MoDOT round robin were more variable than a previous study reported in ASTM C 457, Table 4; however, the MoDOT results were less variable or comparable to the results of another study reported in ASTM C 457, Table 6.
- The wide variation between laboratories strongly supports the development of an automated image analysis system for performing ASTM C 457 measurements. A properly designed automated system should be able to provide more consistent results than a human operator and thus reduce the considerable variation noted in this study.
- The key to a successful air void measurement system, whether human-based or automated, is its ability to make accurate measurements. At present, there is no way to assess the true accuracy of air void parameter estimates; therefore, the accuracy of an automated system should be assessed by comparing its results with the results of a study such as this one.
- The comparison of the results between Phase 1 and Phase 2 indicated that, with the exception of specimen RR3, specimen preparation appeared to be very good and did not appear to have a significant effect on the results.

5.2 Recommendations – precision and accuracy standard for automated system

One of the primary goals of this study was to quantify a standard that an automated testing system would have to meet to be considered comparable to and an acceptable substitute for a human operator-based system; however, there is no one standard that is clearly proper for this purpose. The ASTM C 670 one-sigma limits for single operator and multi-laboratory precision, discussed in Section 4.4, can be used as measures of precision and accuracy for an automated system. Several common statistical measures, such as 95 percent confidence intervals, could also be used. Regardless of the type of data analysis employed, however, the ultimate decision regarding the acceptability criteria for an automated system (or any system) is a matter of engineering judgment. Is an estimated air content of 6 percent acceptable if the true air content is 5 percent? Is an estimated spacing factor of 0.008 inches acceptable if the true value is 0.010 inches? Whether a system can provide results within one standard deviation, or within a 95 percent confidence interval, or within a 95 percent prediction interval, is irrelevant if the preceding questions are not considered.

The remainder of this section discusses several possible measures of acceptability for precision and accuracy based on the round robin results. The Phase 2, Subset 2 data (laboratories 1, 4, 8, 9, 10) were used as the basis for the following analyses.[19]

5.2.1. Precision

For the purposes of this report, "precision" means expected variability of the estimated parameter in multiple tests on the same specimen and is synonymous with within-laboratory variability. In a properly operating automated system, the only source of variation in repeated tests on the same specimen (different traverse lines) should be the variability inherent in the test method itself. In contrast, a human operator-based system may also have operator-related variability associated with its within-laboratory results. Thus, it is reasonable to expect an automated system to achieve a within-laboratory variability less than or equal to the best human operator-based system, and ideally, the automated system's precision should approach the minimum expected uncertainty.

Table 12 summarizes the within-laboratory results for standard deviations of air content and coefficients of variation of voids per inch, spacing factor, and specific surface. The second column from the right contains the minimum values for each parameter's variability measure.

Table 12. Summary of within-laboratory variability (Phase 2, SS2, RR2X)

Parameter		Laboratory ID					Min	Pooled
		1	4	8	9	10		
Air content	std dev	0.30	0.48	0.34	0.30	0.31	0.30	0.35
Voids per inch	C.V.	4.8%	9.8%	6.9%	3.4%	3.6%	3.4%	6.2%
Spacing factor	C.V.	2.8%	7.7%	7.3%	5.0%	3.5%	2.8%	5.6%
Specific surface	C.V.	1.3%	5.5%	6.2%	8.5%	5.5%	1.3%	5.9%

A pooled value (rightmost column) can also be calculated. The pooled within-laboratory standard deviation is calculated using the following equation:

$$s_{pooled} = \sqrt{\frac{n_1 s_1^{\ 2} + n_2 s_2^{\ 2} + \cdots + n_k s_k^{\ 2}}{N}}$$

[19]These data represent the most consistent laboratories, and in Phase 2 any variability associated with different methods of specimen preparation was eliminated.

where the s_k^2 are the variances for each laboratory, the n_k are the number of results obtained for each laboratory, and N is the total number of test results (sum of the n_k). An analogous equation, substituting the coefficients of variation for the s_i, can be used to calculate a pooled coefficient of variation.

The following options can be used as measures of acceptability for the precision of an automated system:

- individual within-laboratory standard deviation (single laboratory)
- pooled within-laboratory standard deviation (multiple laboratories)
- multiple of within-laboratory standard deviation (individual or pooled) based on a 95 percent confidence interval[20]
- another multiple of within-laboratory standard deviation (individual or pooled)

5.2.2 Accuracy

Accuracy means proximity to a true value. In the context of this study, and of the linear traverse in general, it is difficult to identify a measure of accuracy because the true values of the air void parameters for a concrete specimen are unknown. Therefore, the suggested approach for specifying accuracy is to use an interval estimate – a range of values whose midpoint is the mean of a number of test results. A test result from an automated system falling within that range would be considered acceptable. Several possible ways to define this range are discussed below.

As with precision, the standard deviation can be used to define the required accuracy range. If test results from multiple laboratories on a single specimen are available, the between-laboratory standard deviation can be used to define the allowable range. If results from multiple laboratories on several specimens are available, as in this study, the range can also be defined using a pooled between-laboratory standard deviation:

$$s_{pooled} = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2 + \cdots + n_m s_m^2}{N}}$$

where the s_m^2 are the variances (or the squared coefficients of variation) for each specimen, the n_m are the number of results obtained for each specimen, and N is the total number of test results (sum of the n_m). In either case, the standard deviation itself, or some multiple of the standard deviation, can be used to define the limits of the allowable range. An arbitrary multiple (e.g., ± 2 standard deviations) or a statistically-based measure such as a confidence interval can be used.

[20]The multiple used to calculate a confidence interval is $t(1-\frac{\alpha}{2}, n-1)\frac{s}{\sqrt{n}}$, where α= desired significance level (e.g., 0.05 for a 95% confidence interval), n = degrees of freedom, and s = standard deviation, and t is the Student t-distribution.

Two possibilities are shown in Table 13, which summarizes the between-laboratory variability for air content, voids per inch, spacing factor and specific surface using the Phase 2, Subset 2 data. The first possibility is to use the standard deviation (or coefficient of variation) to define the acceptable range for a given air void parameter. The second possibility is to define the range using a 95 percent confidence interval (see footnote 20 on page 43).

Table 13. Summary of between-laboratory variability (Phase 2, SS2)

Parameter	ID	Mean	Std dev	C.V.	95% low	95% high
Air content	RR1	4.68	0.56	12.1%	3.98	5.38
	RR2	6.77	1.07	15.8%	5.44	8.11
	RR3	5.20	0.70	13.6%	4.32	6.07
	RR4	5.27	0.80	15.2%	4.27	6.26
	RR5	7.78	1.34	17.2%	6.12	9.45
Voids per inch	RR1	7.82	1.25	16.0%	6.27	9.37
	RR2	19.20	2.29	11.9%	16.36	22.05
	RR3	12.59	1.14	9.0%	11.18	14.00
	RR4	11.20	0.74	6.6%	10.28	12.12
	RR5	16.03	2.18	13.6%	13.32	18.74
Spacing factor	RR1	0.0073	0.00058	7.9%	0.0065	0.0080
	RR2	0.0032	0.00045	13.9%	0.0027	0.0038
	RR3	0.0049	0.00041	8.3%	0.0044	0.0054
	RR4	0.0055	0.00042	7.7%	0.0050	0.0060
	RR5	0.0039	0.00054	13.9%	0.0032	0.0046
Specific surface	RR1	668	46	6.9%	611	725
	RR2	1142	82	7.1%	1041	1244
	RR3	980	119	12.1%	833	1127
	RR4	863	113	13.1%	723	1003
	RR5	832	95	11.5%	713	951

Pooled between-laboratory standard deviations or coefficients of variation are shown in Table 14. For each specimen, Table 14 shows the individual (i.e., for each specimen) standard deviations for air content and the coefficients of variation of voids per inch, spacing factor, and specific surface. Each entry represents the between-laboratory variability of five test results (for the five laboratories in Phase 2, Subset 2).[21] In the two rightmost columns, the table shows

[21]For specimen RR2, each individual laboratory's result is the mean from the five traverses conducted on that specimen. These means are treated as individual results in this analysis.

pooled estimates of standard deviation (for air content) and coefficient of variation (for voids per inch, spacing factor, and specific surface).

Table 14. Pooled estimates of standard deviation or coefficient of variation (Phase 2, SS2)

Parameter	Std. dev or C.V.*					Pooled std. dev.	Pooled C.V.*
	RR1	RR2	RR3	RR4	RR5		
Air content	0.56	1.07	0.70	0.80	1.34	0.94	–
Voids per inch	*16.0*	*11.9*	*9.0*	*6.6*	*13.6*	–	*11.9*
Spacing factor	*7.9*	*13.9*	*8.3*	*7.7*	*13.9*	–	*10.7*
Specific surface	*6.9*	*7.1*	*12.1*	*13.1*	*11.5*	–	*10.5*
*Note: values in italics are C.V. in percent							

Strictly speaking, a confidence interval is only appropriate for predicting a range in which the mean of an air void parameter would fall. To predict a range for a single newly-observed individual outcome, a prediction interval should be used [Neter et al 1985]. A prediction interval is wider than a confidence interval for a given level of significance. A prediction interval can be based on either the individual or pooled standard deviation.

Tables 15 to 19 show 95 percent prediction intervals for the air void parameters for each of the five test specimens, based on Phase 2, Subset 2 data (using individual standard deviations). For some parameters, particularly air content, the prediction intervals are quite wide, perhaps too wide to be useful in deciding whether an automated system is acceptable.

Table 15. 95 percent prediction intervals for RR1

Parameter	95% PI low	95% PI high
Air content (%)	2.96	6.40
Mean chord length (in)	0.0047	0.0073
Voids per inch	4.02	11.62
Spacing factor (in)	0.0055	0.0090
Specific surface (in-1)	529	808

Table 16. 95 percent prediction intervals for RR2

Parameter	95% PI low	95% PI high
Air content (%)	3.51	10.03
Mean chord length (in)	0.0028	0.0042
Voids per inch	12.23	26.18
Spacing factor (in)	0.0019	0.0046
Specific surface (in^{-1})	894	1390

Table 17. 95 percent prediction intervals for RR3

Parameter	95% PI low	95% PI high
Air content (%)	3.05	7.34
Mean chord length (in)	0.0029	0.0054
Voids per inch	9.14	16.05
Spacing factor (in)	0.0037	0.0061
Specific surface (in^{-1})	619	1340

Table 18. 95 percent prediction intervals for RR4

Parameter	95% PI low	95% PI high
Air content (%)	2.83	7.71
Mean chord length (in)	0.0026	0.0067
Voids per inch	8.94	13.46
Spacing factor (in)	0.0042	0.0068
Specific surface (in^{-1})	520	1206

Table 19. 95 percent prediction intervals for RR5

Parameter	95% PI low	95% PI high
Air content (%)	3.71	11.86
Mean chord length (in)	0.0034	0.0065
Voids per inch	9.39	22.67
Spacing factor (in)	0.0022	0.0056
Specific surface (in^{-1})	542	1122

In summary, the following options may be used to assess the accuracy of an automated system compared with the results of the round robin study:

- mean ± individual between-laboratory standard deviation (multiple laboratories, single specimen)
- mean ± pooled between-laboratory standard deviation (multiple laboratories, multiple specimens)
- mean ± arbitrary multiple of between-laboratory standard deviation (individual or pooled)
- 95 percent confidence interval centered on the mean, based on between-laboratory standard deviation (individual or pooled)

- 95 percent prediction interval centered on the mean, based on between-laboratory standard deviation (individual or pooled)

In each case, the estimated value from the automated system would have to fall within the range of values defined by the selected option.

5.3 Other recommendations

One of the limitations of the round-robin study is that the true parameters of the air void systems of the specimens could not be known with certainty. Because there was no benchmark against which accuracy could be gauged directly, the mean values based on test results had to be used as estimates of the true "population" parameters. Therefore, a method to determine directly the true population values is needed.

Microtomography is a method that might be useful in providing such a benchmark in future studies. Microtomography is similar to a medical CAT scan where x-ray images are taken at many angles around an object and a tomogram (projection of internal structures) is computed from these images. In microtomography, resolutions as low as 2 microns can be achieved. Microtomography at a resolution of 5 to 10 microns could provide a significantly better baseline for assessing accuracy. In the future, microtomography could even become the method of choice for air void system evaluation, although at present the resolution may not be sufficient, and the cost would be prohibitive.

The between-laboratory measurements in this study were quite variable. The variability appears to be largely a result of differences in operator experience and ability. Therefore, training of operators is an important issue. When training a new operator, it would be beneficial to have some kind of standard specimen with a known two-dimensional circle distribution on the surface. This could be, for example, a concrete specimen whose surface has been painstakingly analyzed using an image analysis system (with a human operator measuring each circular air void section). Such a specimen would be ideal for operator training because it would require a new operator to learn under realistic conditions, with the actual colors and features of a concrete specimen. Alternatively, an artificial "specimen" (for example, a surface with multi-sized black circles superimposed on a white background) could be used, although such a specimen might be more appropriate for ensuring that equipment is making accurate measurements.

REFERENCES

1. American Society for Testing and Materials (ASTM), "Standard Practice for Microscopical Determination of Air Void Content and Parameters of the Air Void System in Hardened Concrete," (ASTM C 457). In *2004 Annual Book of ASTM Standards*, Volume 4.02. ASTM, West Conshohocken, PA, 2004.

2. American Society for Testing and Materials (ASTM), "Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials," (ASTM C 670). In *2004 Annual Book of ASTM Standards*, Volume 4.02. ASTM, West Conshohocken, PA, 2004.

3. Box, G.E.P., Hunter, W.G., and J.S. Hunter, *Statistics for Experimenters*. John Wiley & Sons, New York, 1978.

4. Mielenz, R.C., Backstrom, J.E., Burrows, R.W. and V.E. Wolkodoff, "Origin, Evolution and Effects of the Air Void System in Concrete. Part 1 – Entrained Air in Unhardened Concrete." *Journal of the American Concrete Institute*, Vol. 55, July, 1958, pp. 95-121.

5. Neter, J., Wasserman, W., and M.K. Kutner, *Applied Linear Statistical Models, 2nd Edition*. Richard D. Irwin, Inc., Homewood, Illinois, 1985, pp. 76-81.

6. Powers, T.C., "The Air Requirement of Frost Resistant Concrete." In *Proceedings of the Highway Research Board*, Vol. 29, 1949, pp. 184-211.

7. Snyder, K.A., Natesaiyer, K.C., and K.C. Hover, "The Stereological and Statistical Properties of Entrained Air Voids in Concrete: A Mathematical Basis for Air Void System Characterization." In *Materials Science of Concrete VI (*S. Mindess and J. Skalny, Eds.). The American Ceramic Society, Westerville, Ohio, 2002, pp. 129-21.

Appendix

A.1 Data tables – within-laboratory results

Tables 20 through 24 contain the within-laboratory data from Phase 1 for air content, mean chord length, voids per inch, spacing factor, and specific surface, respectively. Tables 25 through 29 contain the analogous within-laboratory data from Phase 2.

Table 20. Within-laboratory air content, RR2

Lab ID	2A	2B	2C	2D	2E	Mean (% air)	Std dev (% air)	C.V. (%)
1	7.82	7.52	6.56	7.46	7.71	7.41	0.50	6.73
2	6.17	7.60	7.38	8.52	6.27	7.19	0.98	13.67
3	4.33	4.51	4.89	4.33	4.21	4.45	0.27	5.98
4	6.84	5.95	6.51	5.35	5.95	6.12	0.57	9.39
5	4.35	4.16	5.41	4.40	4.77	4.62	0.49	10.72
6	5.36	5.06	5.58	4.48	5.45	5.19	0.44	8.46
7	2.79	2.72	2.40	3.05	2.74	2.74	0.23	8.45
8	8.64	8.33	8.56	8.71	8.86	8.62	0.20	2.27
9	5.54	4.96	4.95	5.24	5.29	5.20	0.25	4.77

Table 21. Within-laboratory mean chord length, RR2

Lab ID	2A	2B	2C	2D	2E	Mean (in)	Std dev (in)	C.V. (%)
1	0.0039	0.0037	0.0034	0.0036	0.0036	0.0036	0.0002	4.99
2	0.0062	0.0076	0.0064	0.0087	0.0076	0.0073	0.0010	13.97
3	0.0067	0.0070	0.0073	0.0069	0.0068	0.0069	0.0002	3.32
4	0.0034	0.0036	0.0035	0.0032	0.0034	0.0034	0.0001	4.34
5	0.0079	0.0063	0.0064	0.0060	0.0074	0.0068	0.0008	11.90
6	0.0041	0.0044	0.0041	0.0037	0.0034	0.0039	0.0004	9.93
7	0.0190	0.0231	0.0191	0.0197	0.0217	0.0205	0.0018	8.80
8	0.0039	0.0040	0.0044	0.0044	0.0043	0.0042	0.0002	5.58
9	0.0032	0.0030	0.0028	0.0030	0.0032	0.0030	0.0002	5.49

Table 22. Within-laboratory voids per inch, RR2

Lab ID	2A	2B	2C	2D	2E	Mean	Std dev	C.V. (%)
1	20.08	20.09	19.28	20.77	21.50	20.34	0.83	4.10
2	9.98	10.00	11.50	9.84	8.28	9.92	1.14	11.50
3	6.48	6.43	6.66	6.26	6.22	6.41	0.18	2.77
4	20.02	16.58	18.46	16.64	17.59	17.86	1.43	8.03
5	5.49	6.61	8.46	7.38	6.45	6.88	1.11	16.15
6	13.04	11.43	13.56	12.18	15.94	13.23	1.72	13.00
7	1.46	1.18	1.25	1.55	1.26	1.34	0.16	11.72
8	21.94	20.66	19.62	19.88	20.71	20.56	0.91	4.41
9	17.08	16.64	17.55	17.64	16.74	17.13	0.46	2.66

Table 23. Within-laboratory spacing factor, RR2

Lab ID	2A	2B	2C	2D	2E	Mean (in)	Std dev (in)	C.V. (%)
1	0.0031	0.0031	0.0033	0.0030	0.0029	0.0031	0.0001	4.82
2	0.0060	0.0050	0.0050	0.0060	0.0070	0.0058	0.0008	14.43
3	0.0083	0.0085	0.0086	0.0085	0.0085	0.0085	0.0001	1.29
4	0.0034	0.0038	0.0036	0.0035	0.0036	0.0036	0.0001	4.14
5	0.0088	0.0078	0.0067	0.0067	0.0077	0.0075	0.0009	11.63
6	0.0051	0.0060	0.0050	0.0051	0.0042	0.0051	0.0006	12.46
7	0.0357	0.0443	0.0390	0.0356	0.0417	0.0393	0.0038	9.66
8	0.0028	0.0030	0.0032	0.0031	0.0030	0.0030	0.0001	4.91
9	0.0034	0.0033	0.0032	0.0032	0.0034	0.0033	0.0001	3.12

Table 24. Within-laboratory specific surface, RR2

Lab ID	2A	2B	2C	2D	2E	Mean (in^{-1})	Std dev (in^{-1})	C.V. (%)
1	1027	1069	1175	1114	1116	1100	56	5.05
2	647	526	623	462	528	557	76	13.69
3	599	571	545	579	591	577	21	3.62
4	1171	1115	1134	1244	1183	1169	50	4.27
5	505	636	625	671	541	596	70	11.68
6	974	903	972	1088	1169	1021	106	10.38
7	210	173	209	203	184	196	16	8.42
8	1015	992	917	913	935	954	46	4.85
9	1234	1341	1418	1348	1264	1321	73	5.53

Table 25. Within-laboratory air content, RR2X

Lab ID	2A	2B	2C	2D	2E	Mean (%)	Std dev (%)	C.V. (%)
1	7.60	7.50	6.90	7.10	7.10	7.24	0.30	4.10
2	7.69	7.46	7.23	8.56	7.89	7.77	0.51	6.54
3	7.07	6.15	6.67	6.60	6.57	6.61	0.33	4.95
4	6.67	5.33	6.08	6.14	6.04	6.05	0.48	7.89
5	5.01	5.12	4.67	5.58	4.82	5.04	0.35	6.90
6	6.08	5.93	6.04	5.68	5.89	5.92	0.16	2.65
7	4.03	4.35	5.15	4.69	5.61	4.77	0.63	13.19
8	7.56	8.37	8.20	8.33	8.30	8.15	0.34	4.13
9	5.41	5.35	5.82	4.97	5.44	5.40	0.30	5.60
10	7.25	6.86	7.46	6.85	6.73	7.03	0.31	4.41

Table 26. Within-laboratory mean chord length, RR2X

Lab ID	2A	2B	2C	2D	2E	Mean (in)	Std dev (in)	C.V. (in)
1	0.0033	0.0034	0.0034	0.0034	0.0035	0.0034	0.0001	2.08
2	0.0069	0.0079	0.0076	0.0081	0.0077	0.0076	0.0005	5.97
3	0.0042	0.0041	0.0042	0.0038	0.0041	0.0041	0.0002	4.03
4	0.0034	0.0034	0.0036	0.0037	0.0039	0.0036	0.0002	5.89
5	0.0066	0.0070	0.0059	0.0059	0.0056	0.0062	0.0006	9.34
6	0.0033	0.0040	0.0035	0.0037	0.0031	0.0035	0.0003	9.92
7	0.0138	0.0091	0.0104	0.0084	0.0110	0.0105	0.0021	19.85
8	0.0037	0.0037	0.0037	0.0042	0.0035	0.0038	0.0003	6.88
9	0.0033	0.0030	0.0035	0.0028	0.0032	0.0032	0.0003	8.55
10	0.0038	0.0036	0.0039	0.0034	0.0036	0.0037	0.0002	5.33

Table 27. Within-laboratory voids per inch, RR2X

Lab ID	2A	2B	2C	2D	2E	Mean	Std dev	C.V. (%)
1	22.60	21.90	20.10	20.90	20.60	21.22	1.01	4.78
2	11.21	9.42	9.57	10.51	10.27	10.20	0.73	7.15
3	16.97	14.93	15.78	17.14	16.14	16.19	0.90	5.58
4	19.48	15.48	16.86	16.41	15.48	16.74	1.64	9.82
5	7.60	7.28	7.89	9.39	8.61	8.15	0.85	10.40
6	18.16	14.96	17.16	15.52	18.76	16.91	1.64	9.71
7	2.91	4.80	4.98	5.57	5.08	4.67	1.02	21.92
8	20.32	22.35	22.35	20.07	23.62	21.74	1.51	6.93
9	16.48	17.71	16.52	17.64	17.08	17.09	0.59	3.44
10	19.15	18.98	19.23	20.35	18.45	19.23	0.69	3.61

Table 28. Within-laboratory spacing factor, RR2X

Lab ID	2A	2B	2C	2D	2E	Mean (in)	Std dev (in)	C.V. (%)
1	0.0029	0.0029	0.0030	0.0030	0.0031	0.0030	0.0001	2.81
2	0.0060	0.0060	0.0060	0.0050	0.0060	0.0058	0.0004	7.71
3	0.0044	0.0046	0.0043	0.0040	0.0043	0.0043	0.0002	5.02
4	0.0034	0.0039	0.0038	0.0040	0.0042	0.0039	0.0003	7.69
5	0.0056	0.0066	0.0057	0.0052	0.0059	0.0058	0.0005	8.88
6	0.0035	0.0044	0.0038	0.0044	0.0035	0.0039	0.0005	11.87
7	0.0205	0.0128	0.0139	0.0115	0.0135	0.0144	0.0035	24.30
8	0.0028	0.0026	0.0026	0.0029	0.0024	0.0027	0.0002	7.33
9	0.0035	0.0033	0.0035	0.0031	0.0034	0.0034	0.0002	4.98
10	0.0033	0.0033	0.0031	0.0032	0.0034	0.0033	0.0001	3.50

Table 29. Within-laboratory specific surface, RR2X

Lab ID	2A	2B	2C	2D	2E	Mean (in^{-1})	Std dev (in^{-1})	C.V. (%)
1	1196	1163	1172	1172	1155	1172	15	1.31
2	583	505	529	491	520	526	35	6.70
3	959	971	947	1040	983	980	36	3.69
4	1168	1162	1109	1069	1025	1107	61	5.52
5	607	569	676	673	715	648	59	9.07
6	1194	1009	1137	1093	1273	1141	100	8.76
7	289	441	386	475	362	391	72	18.47
8	1072	1069	1095	958	1138	1066	67	6.24
9	1219	1324	1135	1420	1257	1271	108	8.47
10	1057	1107	1030	1188	1096	1096	60	5.49

A.2 Data tables – between-laboratory results tabulated by air void parameter

The tables in this section are organized by air void parameter. There are four tables for each air void parameter: an overall summary table containing the test result[22] for each laboratory (e.g., Table 30 below), followed by three tables of summary statistics -- for all laboratories, for Subset 1, and for Subset 2 (e.g., Tables 31, 32, and 33 below).

[22] NOTE: In this section, each laboratory's result for RR2 is the mean of the five within-laboratory tests performed on that specimen in each phase of the study.

Table 30. Between-laboratory results – air content

Lab ID	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
1	5.45	5.60	7.41	7.24	5.09	5.90	5.34	5.40	8.52	8.80
2	6.70	5.44	7.19	7.77	7.99	7.02	5.24	5.53	8.42	8.37
3	4.25	4.89	4.45	6.61	3.70	4.77	4.28	4.94	6.84	6.24
4	4.05	4.10	6.12	6.05	3.96	4.41	3.76	4.80	6.81	7.15
5	4.25	4.63	4.62	5.04	5.73	4.20	5.14	4.20	6.12	5.30
6	4.58	4.75	5.19	5.92	4.42	4.76	4.88	5.41	4.95	6.11
7	3.38	5.97	2.74	4.77	3.31	3.44	0.94	3.41	3.53	4.54
8	5.20	4.63	8.62	8.15	5.99	5.95	5.38	6.11	7.96	9.25
9	4.20	4.68	5.20	5.40	3.88	4.66	3.96	4.15	6.25	5.89
10	–	4.38	–	7.03	–	5.06	–	5.88	–	7.83

Table 31. Air content summary statistics – all labs

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	4.67	4.91	5.73	6.40	4.90	5.02	4.32	4.98	6.60	6.95
Std dev	0.98	0.58	1.80	1.15	1.48	1.02	1.41	0.85	1.64	1.58
C.V.	20.9%	11.9%	31.4%	18.0%	30.3%	20.4%	32.6%	17.1%	24.8%	22.7%
99% L	3.58	4.31	3.72	5.21	3.24	3.96	2.75	4.11	4.77	5.33
99% H	5.77	5.51	7.74	7.58	6.55	6.07	5.90	5.86	8.43	8.57

Table 32. Air content summary statistics – Subset 1

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	4.84	4.79	6.10	6.58	5.10	5.19	4.75	5.16	6.98	7.22
Std dev	0.91	0.47	1.50	1.06	1.45	0.91	0.65	0.69	1.24	1.41
C.V.	18.8%	9.9%	24.6%	16.1%	28.5%	17.6%	13.7%	13.4%	17.8%	19.6%
99% L	3.71	4.26	4.24	5.39	3.30	4.17	3.94	4.39	5.44	5.64
99% H	5.96	5.32	7.96	7.77	6.89	6.21	5.55	5.93	8.52	8.80

Table 33. Air content summary statistics – Subset 2

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	4.73	4.68	6.84	6.77	4.73	5.20	4.61	5.27	7.39	7.78
Std dev	0.70	0.56	1.50	1.07	1.01	0.70	0.87	0.80	1.04	1.34
C.V.	14.9%	12.1%	21.9%	15.8%	21.3%	13.6%	18.9%	15.2%	14.1%	17.2%
99% L	2.67	3.52	2.47	4.57	1.79	3.74	2.07	3.62	4.35	5.03
99% H	6.78	5.84	11.21	8.98	7.67	6.65	7.15	6.92	10.42	10.54

Table 34. Between-laboratory results – specific surface

Lab ID	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
1	670	710	1100	1172	772	915	792	839	868	728
2	306	413	557	526	408	542	541	645	525	467
3	307	518	577	980	482	764	447	610	647	608
4	660	688	1169	1107	691	1190	764	999	775	804
5	396	530	596	648	604	625	505	536	552	551
6	853	888	1021	1141	1086	1086	1137	1070	1154	1063
7	153	195	196	391	299	310	230	294	204	305
8	630	622	954	1066	632	912	794	754	784	853
9	718	616	1321	1271	967	931	937	962	941	983
10	–	706	–	1096	–	951	–	762	–	792

Table 35. Specific surface summary statistics – all labs

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	521	589	832	940	660	823	683	747	717	715
Std dev	236	190	367	304	255	265	277	236	275	234
C.V.	45.3%	32.2%	44.1%	32.4%	38.6%	32.2%	40.6%	31.6%	38.3%	32.7%
99% L	258	394	422	627	375	550	373	504	409	475
99% H	785	784	1243	1252	945	1095	993	990	1024	956

Table 36. Specific surface summary statistics – Subset 1

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	568	632	912	1001	705	880	740	797	781	761
Std dev	205	138	298	249	231	207	234	185	210	196
C.V.	36.0%	21.8%	32.7%	24.9%	32.7%	23.5%	31.6%	23.2%	26.9%	25.7%
99% L	314	478	544	722	420	649	450	591	521	542
99% H	821	786	1280	1279	990	1111	1029	1004	1040	980

Table 37. Specific surface summary statistics – Subset 2

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	670	668	1136	1142	766	980	822	863	842	832
Std dev	37	46	152	82	146	119	78	113	78	95
C.V.	5.5%	6.9%	13.4%	7.1%	19.1%	12.1%	9.5%	13.1%	9.3%	11.5%
99% L	563	574	691	974	339	736	594	631	614	635
99% H	776	763	1581	1310	1192	1224	1050	1096	1070	1029

Table 38. Between-laboratory results – spacing factor

Lab ID	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
1	0.0068	0.0064	0.0031	0.0030	0.0065	0.0049	0.0058	0.0056	0.0034	0.0037
2	0.0110	0.0110	0.0058	0.0058	0.0090	0.0070	0.0090	0.0080	0.0070	0.0070
3	0.0170	0.0094	0.0085	0.0043	0.0120	0.0067	0.0120	0.0082	0.0057	0.0071
4	0.0082	0.0075	0.0036	0.0039	0.0075	0.0042	0.0069	0.0049	0.0052	0.0047
5	0.0115	0.0080	0.0075	0.0058	0.0058	0.0079	0.0083	0.0093	0.0059	0.0070
6	0.0061	0.0056	0.0051	0.0039	0.0054	0.0049	0.0051	0.0045	0.0067	0.0042
7	0.0431	0.0255	0.0393	0.0144	0.0232	0.0222	0.0496	0.0235	0.0314	0.0198
8	0.0073	0.0079	0.0030	0.0027	0.0074	0.0051	0.0061	0.0058	0.0037	0.0032
9	0.0069	0.0075	0.0033	0.0034	0.0058	0.0052	0.0056	0.0052	0.0045	0.0040
10	–	0.0070	–	0.0033	–	0.0051	–	0.0059	–	0.0039

Table 39. Spacing factor summary statistics – all labs

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	0.0131	0.0096	0.0088	0.0050	0.0092	0.0073	0.0120	0.0081	0.0082	0.0065
Std dev	0.0118	0.0058	0.0116	0.0035	0.0056	0.0054	0.0143	0.0056	0.0088	0.0049
C.V.	89.8%	60.5%	131.9%	68.9%	61.6%	73.1%	118.4%	69.8%	107.8%	76.2%
99% L	-0.0001	0.0036	-0.0042	0.0015	0.0029	0.0018	-0.0039	0.0023	-0.0017	0.0014
99% H	0.0263	0.0155	0.0218	0.0086	0.0155	0.0128	0.0280	0.0139	0.0180	0.0115

Table 40. Spacing factor summary statistics – Subset 1

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	0.0094	0.0078	0.0050	0.0040	0.0074	0.0057	0.0073	0.0064	0.0053	0.0050
Std dev	0.0037	0.0016	0.0021	0.0011	0.0022	0.0012	0.0023	0.0017	0.0013	0.0016
C.V.	39.2%	20.5%	42.7%	28.5%	29.7%	21.6%	31.6%	26.4%	24.9%	32.0%
99% L	0.0048	0.0060	0.0023	0.0027	0.0047	0.0043	0.0045	0.0045	0.0036	0.0032
99% H	0.0139	0.0096	0.0076	0.0053	0.0101	0.0070	0.0102	0.0083	0.0069	0.0068

Table 41. Spacing factor summary statistics – Subset 2

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	0.0073	0.0073	0.0032	0.0032	0.0068	0.0049	0.0061	0.0055	0.0042	0.0039
Std dev	0.00064	0.00058	0.00025	0.00045	0.00082	0.00041	0.00057	0.00042	0.00082	0.00054
C.V.	8.8%	7.9%	7.8%	13.9%	12.0%	8.3%	9.4%	7.7%	19.4%	13.9%
99% L	0.0054	0.0062	0.0025	0.0023	0.0044	0.0041	0.0044	0.0046	0.0018	0.0028
99% H	0.0092	0.0084	0.0040	0.0041	0.0092	0.0057	0.0078	0.0063	0.0066	0.0050

Table 42. Between-laboratory results – voids per inch (VPI)

Lab ID	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
1	9.13	10.00	20.34	21.22	9.83	13.40	10.58	11.30	18.50	16.10
2	5.13	5.62	9.92	10.20	8.16	9.51	7.09	8.92	11.06	9.78
3	3.27	6.32	6.41	16.19	4.46	9.10	4.79	7.55	11.06	9.49
4	6.68	7.05	17.86	16.74	6.84	13.12	7.18	11.99	13.20	14.37
5	4.21	6.13	6.88	8.15	8.65	6.57	6.48	5.63	8.44	7.29
6	9.76	10.54	13.23	16.91	12.00	12.90	13.87	14.46	14.27	16.24
7	1.29	2.90	1.34	4.67	2.48	2.67	0.54	2.50	1.80	3.46
8	8.20	7.11	20.56	21.74	9.47	13.55	10.68	11.52	15.60	19.72
9	7.55	7.20	17.13	17.09	9.38	10.86	9.28	9.99	14.71	14.48
10	–	7.72	–	19.23	–	12.03	–	11.20	–	15.49

Table 43. VPI summary statistics – all labs

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	6.14	7.06	12.63	15.21	7.92	10.37	7.83	9.51	12.07	12.64
Std dev	2.85	2.16	6.86	5.68	2.92	3.53	3.86	3.49	4.84	4.95
C.V.	46.5%	30.6%	54.3%	37.4%	36.9%	34.1%	49.3%	36.7%	40.1%	39.2%
99% L	2.95	4.84	4.95	9.37	4.65	6.74	3.51	5.92	6.66	7.55
99% H	9.33	9.28	20.31	21.06	11.18	14.00	12.15	13.09	17.49	17.73

Table 44. VPI summary statistics – Subset 1

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	6.74	7.52	14.04	16.39	8.60	11.23	8.74	10.28	13.36	13.66
Std dev	2.35	1.69	5.78	4.57	2.23	2.41	2.92	2.62	3.14	3.98
C.V.	34.9%	22.4%	41.1%	27.9%	26.0%	21.5%	33.4%	25.5%	23.5%	29.1%
99% L	3.83	5.63	6.89	11.27	5.84	8.53	5.13	7.36	9.47	9.21
99% H	9.65	9.41	21.19	21.50	11.36	13.92	12.35	13.21	17.24	18.12

Table 45. VPI summary statistics – Subset 2

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	7.89	7.82	18.97	19.20	8.88	12.59	9.43	11.20	15.50	16.03
Std dev	1.03	1.25	1.74	2.29	1.37	1.14	1.63	0.74	2.23	2.18
C.V.	13.1%	16.0%	9.2%	11.9%	15.5%	9.0%	17.3%	6.6%	14.4%	13.6%
99% L	4.87	5.24	13.90	14.48	4.87	10.25	4.67	9.67	8.99	11.54
99% H	10.91	10.39	24.04	23.93	12.89	14.93	14.19	12.73	22.02	20.53

Table 46. Between-laboratory results – mean chord length (MCL)

Lab ID	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
1	0.0061	0.0056	0.0036	0.0034	0.0052	0.0044	0.0051	0.0048	0.0046	0.0055
2	0.0131	0.0097	0.0073	0.0076	0.0098	0.0074	0.0074	0.0062	0.0076	0.0086
3	0.0130	0.0077	0.0069	0.0041	0.0083	0.0052	0.0090	0.0066	0.0062	0.0066
4	0.0061	0.0058	0.0034	0.0036	0.0058	0.0034	0.0052	0.0037	0.0052	0.005
5	0.0101	0.0076	0.0068	0.0062	0.0066	0.0064	0.0079	0.0075	0.0073	0.0073
6	0.0047	0.0045	0.0039	0.0035	0.0037	0.0037	0.0035	0.0040	0.0035	0.0038
7	0.0261	0.0205	0.0205	0.0105	0.0134	0.0129	0.0174	0.0136	0.0196	0.0131
8	0.0063	0.0064	0.0042	0.0038	0.0063	0.0044	0.0050	0.0053	0.0051	0.005
9	0.0056	0.0065	0.0030	0.0032	0.0041	0.0043	0.0043	0.0042	0.0043	0.0041
10	–	0.0057	–	0.0037	–	0.0042	–	0.0052	–	0.0051

Table 47. MCL summary statistics – all labs

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	0.0101	0.0080	0.0066	0.0050	0.0070	0.0056	0.0072	0.0061	0.0070	0.0064
Std dev	0.0068	0.0046	0.0055	0.0024	0.0031	0.0028	0.0042	0.0029	0.0049	0.0028
C.V.	67.2%	57.8%	82.2%	49.1%	43.5%	50.3%	58.8%	47.4%	69.6%	43.2%
99% L	0.0025	0.0032	0.0005	0.0025	0.0036	0.0027	0.0025	0.0031	0.0016	0.0036
99% H	0.0177	0.0128	0.0128	0.0075	0.0104	0.0085	0.0119	0.0091	0.0125	0.0093

Table 48. MCL summary statistics – Subset 1

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	0.0081	0.0066	0.0049	0.0043	0.0062	0.0048	0.0059	0.0053	0.0055	0.0057
Std dev	0.0034	0.0015	0.0018	0.0015	0.0020	0.0013	0.0019	0.0013	0.0015	0.0016
C.V.	42.2%	23.1%	36.3%	35.3%	32.8%	27.0%	32.5%	24.2%	26.5%	27.4%
99% L	0.0039	0.0049	0.0027	0.0026	0.0037	0.0034	0.0035	0.0038	0.0037	0.0039
99% H	0.0124	0.0083	0.0071	0.0060	0.0088	0.0063	0.0083	0.0067	0.0073	0.0074

Table 49. MCL summary statistics – Subset 2

Statistic	RR1		RR2		RR3		RR4		RR5	
	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2	Ph 1	Ph 2
Mean	0.0060	0.0060	0.0036	0.0035	0.0054	0.0041	0.0049	0.0046	0.0048	0.0049
Std dev	0.0003	0.0004	0.0005	0.0002	0.0009	0.0004	0.0004	0.0007	0.0004	0.0005
C.V.	5.2%	7.0%	13.6%	6.8%	17.3%	10.2%	8.6%	14.7%	9.3%	10.4%
99% L	0.0051	0.0051	0.0022	0.0030	0.0026	0.0033	0.0037	0.0032	0.0035	0.0039
99% H	0.0069	0.0069	0.0050	0.0040	0.0081	0.0050	0.0061	0.0060	0.0061	0.0060

A.3 Data tables – between-laboratory results tabulated by specimen

This section contains ten tables of between-laboratory test results, organized by specimen. Tables 50 through 54 contain the Phase 1 results and Tables 55 through 59 contain the Phase 2 results. In Tables 51 (RR2) and 56 (RR2X), the table entries for each laboratory represent the average of the five test results on those specimens.

Table 50. Between-laboratory results, RR1

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in^{-1})
1	5.45	0.0061	9.13	0.0068	670
2	6.70	0.0131	5.13	0.0110	306
3	4.25	0.0130	3.27	0.0170	307
4	4.05	0.0061	6.68	0.0082	660
5	4.25	0.0101	4.21	0.0115	396
6	4.58	0.0047	9.76	0.0061	853
7	3.38	0.0261	1.29	0.0431	153
8	5.20	0.0063	8.20	0.0073	630
9	4.20	0.0056	7.55	0.0069	718
Mean	4.67	0.0101	6.14	0.0131	521
Std dev	0.98	0.0068	2.85	0.0118	236
C.V.	20.9%	67.2%	46.5%	89.8%	45.3%

Table 51. Between-laboratory results, RR2

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in-1)
1	7.41	0.0036	20.34	0.0031	1100
2	7.19	0.0073	9.92	0.0058	557
3	4.45	0.0069	6.41	0.0085	577
4	6.12	0.0034	17.86	0.0036	1169
5	4.62	0.0068	6.88	0.0075	596
6	5.19	0.0039	13.23	0.0051	1021
7	2.74	0.0205	1.34	0.0393	196
8	8.62	0.0042	20.56	0.0030	954
9	5.20	0.0030	17.13	0.0033	1321
Mean	5.73	0.0066	12.63	0.0088	832
Std dev	1.80	0.0055	6.86	0.0116	367
C.V.	31.4%	82.2%	54.3%	131.9%	44.1%

Table 52. Between-laboratory results, RR3

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in-1)
1	5.09	0.0052	9.83	0.0065	772
2	7.99	0.0098	8.16	0.0090	408
3	3.70	0.0083	4.46	0.0120	482
4	3.96	0.0058	6.84	0.0075	691
5	5.73	0.0066	8.65	0.0058	604
6	4.42	0.0037	12.00	0.0054	1086
7	3.31	0.0134	2.48	0.0232	299
8	5.99	0.0063	9.47	0.0074	632
9	3.88	0.0041	9.38	0.0058	967
Mean	4.90	0.0070	7.92	0.0092	660
Std dev	1.48	0.0031	2.92	0.0056	255
C.V.	30.3%	43.5%	36.9%	61.6%	38.6%

Table 53. Between-laboratory results, RR4

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in-1)
1	5.34	0.0051	10.58	0.0058	792
2	5.24	0.0074	7.09	0.0090	541
3	4.28	0.0090	4.79	0.0120	447
4	3.76	0.0052	7.18	0.0069	764
5	5.14	0.0079	6.48	0.0083	505
6	4.88	0.0035	13.87	0.0051	1137
7	0.94	0.0174	0.54	0.0496	230
8	5.38	0.0050	10.68	0.0061	794
9	3.96	0.0043	9.28	0.0056	937
Mean	4.32	0.0072	7.83	0.0120	683
Std dev	1.41	0.0042	3.86	0.0143	277
C.V.	32.6%	58.8%	49.3%	118.4%	40.6%

Table 54. Between-laboratory results, RR5

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in-1)
1	8.52	0.0046	18.50	0.0034	868
2	8.42	0.0076	11.06	0.0070	525
3	6.84	0.0062	11.06	0.0057	647
4	6.81	0.0052	13.20	0.0052	775
5	6.12	0.0073	8.44	0.0059	552
6	4.95	0.0035	14.27	0.0067	1154
7	3.53	0.0196	1.80	0.0314	204
8	7.96	0.0051	15.60	0.0037	784
9	6.25	0.0043	14.71	0.0045	941
Mean	6.60	0.0070	12.07	0.0082	717
Std dev	1.64	0.0049	4.84	0.0088	275
C.V.	24.8%	69.6%	40.1%	107.8%	38.3%

Table 55. Between-laboratory results, RR1X

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in-1)
1	5.60	0.0056	10.00	0.0064	710
2	5.44	0.0097	5.62	0.0110	413
3	4.89	0.0077	6.32	0.0094	518
4	4.10	0.0058	7.05	0.0075	688
5	4.63	0.0076	6.13	0.0080	530
6	4.75	0.0045	10.54	0.0056	888
7	5.97	0.0205	2.90	0.0255	195
8	4.63	0.0064	7.11	0.0079	622
9	4.68	0.0065	7.20	0.0075	616
10	4.38	0.0057	7.72	0.0070	706
Mean	4.91	0.0080	7.06	0.0096	589
Std dev	0.58	0.0046	2.16	0.0058	190
C.V.	11.9%	57.8%	30.6%	60.5%	32.2%

Table 56. Between-laboratory results, RR2X

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in^{-1})
1	7.24	0.0034	21.22	0.0030	1172
2	7.77	0.0076	10.20	0.0058	526
3	6.61	0.0041	16.19	0.0043	980
4	6.05	0.0036	16.74	0.0039	1107
5	5.04	0.0062	8.15	0.0058	648
6	5.92	0.0035	16.91	0.0039	1141
7	4.77	0.0105	4.67	0.0144	391
8	8.15	0.0038	21.74	0.0027	1066
9	5.40	0.0032	17.09	0.0034	1271
10	7.03	0.0037	19.23	0.0033	1096
Mean	6.40	0.0050	15.21	0.0050	940
Std dev	1.15	0.0024	5.68	0.0035	304
C.V.	18.0%	49.1%	37.4%	68.9%	32.4%

Table 57. Between-laboratory results, RR3X

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in^{-1})
1	5.90	0.0044	13.40	0.0049	915
2	7.02	0.0074	9.51	0.0070	542
3	4.77	0.0052	9.10	0.0067	764
4	4.41	0.0034	13.12	0.0042	1190
5	4.20	0.0064	6.57	0.0079	625
6	4.76	0.0037	12.90	0.0049	1086
7	3.44	0.0129	2.67	0.0222	310
8	5.95	0.0044	13.55	0.0051	912
9	4.66	0.0043	10.86	0.0052	931
10	5.06	0.0042	12.03	0.0051	951
Mean	5.02	0.0056	10.37	0.0073	823
Std dev	1.02	0.0028	3.53	0.0054	265
C.V.	20.4%	50.3%	34.1%	73.1%	32.2%

Table 58. Between-laboratory results, RR4X

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in^{-1})
1	5.40	0.0048	11.30	0.0056	839
2	5.53	0.0062	8.92	0.0080	645
3	4.94	0.0066	7.55	0.0082	610
4	4.80	0.0037	11.99	0.0049	999
5	4.20	0.0075	5.63	0.0093	536
6	5.41	0.0040	14.46	0.0045	1070
7	3.41	0.0136	2.50	0.0235	294
8	6.11	0.0053	11.52	0.0058	754
9	4.15	0.0042	9.99	0.0052	962
10	5.88	0.0052	11.20	0.0059	762
Mean	4.98	0.0061	9.51	0.0081	747
Std dev	0.85	0.0029	3.49	0.0056	236
C.V.	17.1%	47.4%	36.7%	69.8%	31.6%

Table 59. Between-laboratory results, RR5X

Lab ID	Air (%)	MCL (in)	Voids per inch	Spacing factor (in)	Specific surface (in^{-1})
1	8.80	0.0055	16.10	0.0037	728
2	8.37	0.0086	9.78	0.0070	467
3	6.24	0.0066	9.49	0.0071	608
4	7.15	0.0050	14.37	0.0047	804
5	5.30	0.0073	7.29	0.0070	551
6	6.11	0.0038	16.24	0.0042	1063
7	4.54	0.0131	3.46	0.0198	305
8	9.25	0.0050	19.72	0.0032	853
9	5.89	0.0041	14.48	0.0040	983
10	7.83	0.0051	15.49	0.0039	792
Mean	6.95	0.0064	12.64	0.0065	715
Std dev	1.58	0.0028	4.95	0.0049	234
C.V.	22.7%	43.2%	39.2%	76.2%	32.7%

A.4 Data tables – ASTM C 670 Multi-laboratory precision

Tables 60 through 64 contain the calculated 1s and 1s% limits for multi-laboratory precision according to ASTM C 670.

Table 60. Multi-laboratory precision, RR1/RR1X

Parameter		Phase 1			Phase 2		
		All labs	Subset 1	Subset 2	All labs	Subset 1	Subset 2
Air	1s	0.98	0.91	0.70	0.58	0.47	0.56
	1s%	20.9%	18.8%	14.9%	11.9%	9.9%	12.1%
Voids per inch	1s	2.85	2.35	1.03	2.16	1.69	1.25
	1s%	46.5%	34.9%	13.1%	30.6%	22.4%	16.0%
Spacing factor	1s	0.0118	0.0037	0.00064	0.0058	0.0016	0.00058
	1s%	89.8%	39.2%	8.8%	60.5%	20.5%	7.9%
Specific surface	1s	236	205	37	190	138	46
	1s%	45.3%	36.0%	5.5%	32.2%	21.8%	6.9%

Table 61. Multi-laboratory precision, RR2/RR2X

Parameter		Phase 1			Phase 2		
		All labs	Subset 1	Subset 2	All labs	Subset 1	Subset 2
Air	1s	1.80	1.50	1.50	1.15	1.06	1.07
	1s%	31.4%	24.6%	21.9%	18.0%	16.1%	15.8%
Voids per inch	1s	6.86	5.78	1.74	5.68	4.57	2.29
	1s%	54.3%	41.1%	9.2%	37.4%	27.9%	11.9%
Spacing factor	1s	0.0116	0.0021	0.00025	0.0035	0.0011	0.00045
	1s%	131.9%	42.7%	7.8%	68.9%	28.5%	13.9%
Specific surface	1s	367	298	152	304	249	82
	1s%	44.1%	32.7%	13.4%	32.4%	24.9%	7.1%

Table 62. Multilaboratory precision, RR3/RR3X

Parameter		Phase 1			Phase 2		
		All labs	Subset 1	Subset 2	All labs	Subset 1	Subset 2
Air	1s	1.48	1.45	1.01	1.02	0.91	0.70
	1s%	30.3%	28.5%	21.3%	20.4%	17.6%	13.6%
Voids per inch	1s	2.92	2.23	1.37	3.53	2.41	1.14
	1s%	36.9%	26.0%	15.5%	34.1%	21.5%	9.0%
Spacing factor	1s	0.0056	0.0022	0.00082	0.0054	0.0012	0.00041
	1s%	61.6%	29.7%	12.0%	73.1%	21.6%	8.3%
Specific surface	1s	255	231	146	265	207	119
	1s%	38.6%	32.7%	19.1%	32.2%	23.5%	12.1%

Table 63. Multilaboratory precision, RR4/RR4X

Parameter		Phase 1			Phase 2		
		All labs	Subset 1	Subset 2	All labs	Subset 1	Subset 2
Air	1s	1.41	0.65	0.87	0.85	0.69	0.80
	1s%	32.6%	13.7%	18.9%	17.1%	13.4%	15.2%
Voids per inch	1s	3.86	2.92	1.63	3.49	2.62	0.74
	1s%	49.3%	33.4%	17.3%	36.7%	25.5%	6.6%
Spacing factor	1s	0.0143	0.0023	0.00057	0.0056	0.0017	0.00042
	1s%	118.4%	31.6%	9.4%	69.8%	26.4%	7.7%
Specific surface	1s	277	234	78	236	185	113
	1s%	40.6%	31.6%	9.5%	31.6%	23.2%	13.1%

Table 64. Multilaboratory precision, RR5/RR5X

Parameter		Phase 1			Phase 2		
		All labs	Subset 1	Subset 2	All labs	Subset 1	Subset 2
Air	1s	1.64	1.24	1.04	1.58	1.41	1.34
	1s%	24.8%	17.8%	14.1%	22.7%	19.6%	17.2%
Voids per inch	1s	4.84	3.14	2.23	4.95	3.98	2.18
	1s%	40.1%	23.5%	14.4%	39.2%	29.1%	13.6%
Spacing factor	1s	0.0088	0.0013	0.00082	0.0049	0.0016	0.00054
	1s%	107.8%	24.9%	19.4%	76.2%	32.0%	13.9%
Specific surface	1s	275	210	78	235	196	95
	1s%	38.3%	26.9%	9.3%	32.7%	25.7%	11.5%

A.5 Summary of specimen preparation procedures

Table 65. Specimen preparation procedures, Phase 1

	Equipment	Abrasive	Size	Time	Lubricant
RR1	Lapmaster 24F	silicon carbide grit	22.5 μm	?	?
RR2	Struers Abrapol-10	diamond-embedded disc	#120	2 min	water
			#220	1 min	
			#600	1 min	
			#`1200	1 min	
		diamond grit solution	3 μm	?	polishing lubricant
RR3	Lapidary polisher	silicon carbide grit	#120	30 min	water
			#120	4 hr	
			#240	8 hr	
			#400	8 hr	
			#600	24 hr	

RR4	Struers Abramin	diamond-embedded disc	#220	2 min	water
			#600	3 min	
			#1200	3 min	
		diamond grit solution	3 µm	4 min	polishing lubricant
RR5	Metaserv 2000	diamond-embedded disc	125 µm	?	water
			15 µm	?	
		silicon carbide grit	#800	?	
		aluminum oxide grit	#1000	?	
		diamond paste	3 µm	?	

Table 66. Specimen preparation procedures, Phase 2

	Equipment	Abrasive	Size	Time	Lubricant
ALL	Struers Abrapol	diamond-embedded disc	125 µm	2 min	water
			40 µm	3 min	
		diamond grit solution	15 µm	2 min	polishing lubricant
				4 min	
			3 µm	3 min	

A.6 Results of survey of potential round-robin participants

This section of the Appendix contains responses to a questionnaire sent by MODOT to laboratories that expressed an interest in participating in the round robin study. The results of this questionnaire were used to select the final participants and to develop the experimental plan.

1) *What's the maximum specimen size or dimensions that your current ASTM C 457 system can analyze? (FYI - The new automated system's maximum size is a 6"x6" specimen. Missouri's polishing machine can handle a maximum 7" diameter surface area.)*

Lab	Response
Wisconsin	Maximum specimen size which will physically fit in our table is 5"x5". Maximum analysis area (limits of stage travel) is 4"x4".
Minnesota	Maximum specimen size for analysis is 7"x 6", but automated sample prep equipment can prepare samples only 3.5"x 3.5". Use manual preparation for larger samples
Nebraska	We can handle a 6" x 12" specimen and possibly larger. We will be trying a 14" specimen in the near future.
Virginia	The table has 7.5 x 7.5-in travel. Our largest lapping rings are 6-in diameter. Normally we look at slabs that are 4 x 6 inches (max)

Lab	Response
Indiana	The maximum specimen diameter is about 5 inches. That is based on a maximum specimen length of about 5 inches. The width of the specimen is not an issue. The maximum specimen thickness is about 1 inch.
Illinois	Our ASTM C 457 system has a maximum movement of 7 inches in each direction. However, our polisher limits us to a maximum of 3.75 inches by 3.75 inches.
Missouri	For 6" X 6" samples we cut the corners off so the sample will fit inside the 7" circle. Ideally the sample should be sawed 1" + - 1/10" thick so the finished sample is 3/4" min. to just over 1" max. thick. Uniform thickness is not a requirement but is highly recommended.
FHWA	Maximum specimen size is approximately 7 x 7 inches.
CTLGroup	Just under 7.5 x 7.5 inches

2) *Provide a general description of your current equipment or system used to conduct ASTM C 457. For example, do you conduct analysis directly from viewing a microscope or do you use a microscopic image projected to a monitor?*

Lab	Response
Wisconsin	The basic air void analysis system, including stage and movement hardware and analysis software is a model 602 from Trilogy Systems Corp. We later added a new Bausch and Lomb microscope and a Hitachi video camera and monitor. Our operator currently performs analysis using a video image on a monitor.
Minnesota	Stage is a computer driven "Parker positioning system – Daedal Division". The computer software is "AV2000, Air Void Analysis System" Version 2.02 (January 1995). This is a DOS program. Updated Windows based software is available but we do not have ... yet. The system includes a control box with 9 buttons used to control the stage and to record what the material is as it traverses under the crosshairs (eg: void, paste, etc.). We have a binocular microscope (see magnifications below) and a high intensity fiber-optics light we direct at a low angle onto the specimen.
Nebraska	We use a microscope image projected to a monitor.
Virginia	The C 457 equipment is a CAS-2000 system. The microscope is fitted with a video camera and the analysis is conducted viewing a monitor.
Indiana	See attached pictures. You can view the image from the monitor or through the microscope. The image can be recorded on video tape through the monitor. We rarely record the test.
Illinois	Our ASTM C 457 system uses a semi-automated system. It consists of a motorized stage which is controlled by special software. We count and measure voids and aggregates by observing them on a TV screen and pressing appropriate

Lab	Response
	buttons. Our system uses a 486 based computer, a closed circuit TV camera mounted on the microscope, and a TV screen for viewing.
Missouri	We do all viewing from the microscope at 100 X. A variable speed electric motor drives the stage under the microscope and a electronic counter measures the distance traversed. I verify the distance traversed by a mechanical counter. The traverse data is recorded in a chord file and a summary is computed by software using the ASTM C 457 formulas.
FHWA	System is semi-automated with a computer-controlled motorized X-Y stage. Chords are counted and measured using keyboard and special joystick buttons. All stage movements and calculations are performed by software. Operator can view specimen through the microscope or via a camera and video monitor.
CTLGroup	Stereomicroscope set up with a mechanical, computer driven stage. We do not use a monitor to observe the traverse, but look directly through the eyepieces of the stereomicroscope.

3) *Does your current microscope offer a range of magnification (e.g. 10x to 100x)? If so, what is that range? What magnification do you currently use for ASTM C 457 analysis?*

Lab	Response
Wisconsin	Range is 26.8X to 160X. Standard magnification used for air analysis is 80X.
Minnesota	Microscope has zoom capabilities with magnification ranging from 30x to 180x. Usually use approximately 100x for a traverse, zooming in on questionable areas when necessary.
Nebraska	We are trying to determine the magnification range of our microscope. It has been modified a few times, but we know it meets the requirements of C 457 which is 50x to 125x. We believe it is probably of the order of 10x to 125x.
Virginia	The microscope has an magnification range of 3.5x - 140x. Linear traverse analysis is normally conducted at 90x.
Indiana	The range of magnification is 20-400x. Perform test at 120x.
Illinois	The microscope currently has a zoom lens. Viewing the specimen on the TV screen gives a magnification range of approximately 100:1 up to 600:1. We normally view the specimen at a magnification of approximately 100:1.
Missouri	Our microscope offers a range of magnification range of 32x to 160x. We use 100x for all testing.
FHWA	Yes, the range is approximately 10x to 300x through the microscope. Usual magnification for C 457 is 100x (microscope or camera/monitor).
CTLGroup	Our microscope offers a range of magnifications from 20 to135X. We typically use a magnification of 90X for ASTM C 457.

4) Is your system able to measure individual chord lengths?

Lab	Response
Wisconsin	Yes, we measure individual chord lengths by marking coordinates of beginning and end points.
Minnesota	If I understand this question correctly - yes we measure individual chord lengths. Our stage automatically moves the specimen and we hold down a different button, depending on whether we are traversing over paste, aggregate, entrained air or entrapped air (or filling in a void). The computer program then interprets these into chord lengths and reports out spacing factor, %air, specific surface, etc. To get a report of the individual chord lengths we have to dump the data into a spreadsheet.
Nebraska	Chord lengths can be measured, but not automatically. The machine has to be stopped at each edge of the void and the difference in positions calculated.
Virginia	The system does collect individual chord lengths
Indiana	Yes
Illinois	Our system is not able to measure individual chord lengths.
Missouri	Yes
FHWA	Yes
CTLGroup	Yes

5) When conducting analysis, do you distinguish between entrapped and entrained air voids? How do you accomplish this (e.g. what limit do you set?)?

Lab	Response
Wisconsin	Yes, we distinguish between entrained and entrapped air voids, using a 1 mm cutoff threshold.
Minnesota	Yes - we look at the void and determine from size and shape if it is entrained or entrapped. If it is smaller, and round, a coalescing of round voids, or a slightly stretched round shape it is entrained. If irregular, especially if larger and irregular, it is entrapped. We do not use a specific size criteria to distinguish the difference.
Nebraska	We consider any voids with diameters greater than 1mm or irregular in shape to be entrapped. Anything else is counted as entrained.
Virginia	Normally, no. When consolidation is an issue, we have categorized voids as entrapped (over 1 mm/irregular shape) or entrained (under 1 mm/normal shape).
Indiana	No
Illinois	We do not distinguish between entrapped and entrained air voids

Lab	**Response**
Missouri	Our Summary program reports % of air for four categories of void sizes. 0.00001 to 0.00599, 0.00600 to 0.03999, 0.04000 to 0.08000, and >0.08000 inches. We report % air in concrete, and % air in mortar, in all four void size ranges.
FHWA	Operator has option to set a cutoff value (default = 1 mm) to distinguish entrapped air void chords by length
CTLGroup	When performing linear traverse method, we are unable to specify if the void is entrapped or entrained, but the analysis produces a void-size breakdown ranging from less than 0.001 in. to 0.039 in. that can be applied to the specimen's air content. When performing point count method, a void greater than 1 mm or irregularly shaped or both is typically considered entrapped. Voids less than 1 mm and spherical is shape are typically considered entrained.