Hindawi
Advances in Civil Engineering
Volume 2019, Article ID 4539549, 11 pages
https://doi.org/10.1155/2019/4539549

Research Article

Quantifying Repeatability Reproducibility Sources of Error and Capacity of a Measurement: Demonstrated Using Laboratory Soil Plasticity Tests

Cheng Li,[1,2] **Jeramy C. Ashlock**,[3] **and Xuhao Wang**[1,2]

[1]*School of Highway, Chang'an University, Xi'an, Shanxi Province 710064, China*
[2]*Research and Development Center of Transport Industry of Technologies, Materials and Equipment of Highway Construction and Maintenance, Gansu Road & Bridge Construction Group, China*
[3]*Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, IA 50010, USA*

Correspondence should be addressed to Xuhao Wang; wangxh@chd.edu.cn

Received 19 April 2019; Accepted 10 June 2019; Published 2 July 2019

Academic Editor: Timo Saksala

The repeatability, reproducibility, and sources of error inherent in a given measurement are important considerations for potential users. To quantify errors arising from a single operator or multiple laboratories, most testing standards uses a one-way analysis of variance- (ANOVA-) based method, which utilizes a simple standard deviation across all measurements. However, this method does not allow users to quantify the sources of error and capacity (i.e., the precision to tolerance ratio). In this study, an innovative two-way ANOVA-based analysis method is selected to quantify the relative contributions of different sources of error and determine whether a measurement can be used to check conformance of a measured characteristic to engineering specifications. In this study, the standardized Atterberg limits tests, fall-cone device Atterberg limits tests, and bar linear shrinkage tests widely used for determining the soil plasticity were selected for evaluation and demonstration. Comparisons between results of the various testing methods are presented, and the error sources contributing to the overall variations between tests are discussed. Based on the findings of this study, the authors suggest use of two-way ANOVA-based R&R analysis to quantify the sources of measurement error and capacity and also recommend using the fall cone device and ASTM standardized thread rolling device for determining liquid and plastic limits of soils, respectively.

1. Introduction

The repeatability (i.e., the single operator or intralaboratory precision) and reproducibility (the interlaboratory precision) of a measurement are important characteristics which can be quantified to enable users to understand the variability of test results. The ASTM E691 standard practice on interlaboratory testing states "ASTM standard regulations require precision statements in all test methods in terms of repeatability and reproducibility" and specifies a one-way ANOVA (i.e., a simple standard deviation across all measurements) to quantify the single-operator or multilaboratory errors [1]. However, the repeatability and reproducibility (R&R) statistics from the one-way ANOVA analysis cannot quantify the contributions of multiple error sources to the overall variation in a measurement. Such information would be useful for identifying potential ways to further improve the test methods. For example, the design of the testing device may need to be improved or training of operators may need to be enhanced. The capacity of a measurement, defined as the ratio of precision to tolerance, is also an important parameter for determining whether a measurement is useful for checking conformance of a measured characteristic to engineering specifications. The ASTM standards (i.e., ASTM E177) use the 95% limits on the difference between two test results, referred to as the "d2s" limit (i.e., $1.960\sqrt{2} \cdot 1s$, where 1s is one standard deviation) to determine the acceptable range of two test results [2]. However, for a given testing method, the acceptable ranges calculated using the method are usually different between the single-operator and multilaboratory test

results due to their different standard deviations. In recent years, more and more new laboratory and field testing devices and methods have been developed for evaluating soil characteristics [3–5]. However, very few of them report the repeatability, reproducibility, and capacity of newly developed testing methods, which considers errors arising from both the device and different operators.

To address these issues, the authors propose to use a two-way ANOVA-based R&R analysis to evaluate the repeatability, reproducibility, sources of error, and capacity of a measurement. In this study, several widely used Atterberg limits testing devices are employed for comparisons and demonstration of the statistical analysis. The ASTM-standardized Atterberg limits tests [6], fall cone test [7], and bar linear shrinkage test [8] were conducted on specimens prepared by incorporating different percentages of pure bentonite into the minus No. 40 fraction of crushed limestone samples from granular roadways. The R&R and capacity of the testing methods were determined using the two-way ANOVA-based analysis and compared to those determined using the one-way ANOVA methods described in ASTM E691-15; ASTM E177-13 [1, 2]. Based on the results of the laboratory tests and statistical analyses, the method for quantifying repeatability, reproducibility, sources of error, and capacity of a given measurement is demonstrated, the testing methods with the best R&R are identified, and the correlations between the different tests are also discussed.

2. Various Tests for Determining the Soil Consistency

Swedish soil scientist Atterberg defined moisture content limits to delineate transitions in the consistency of fine-grained soils [9]. Atterberg initially set up five limits to describe the consistency of a soil at different water contents: (1) the upper limit of fluidity, (2) the lower limit of fluidity (flow limit), (3) the sticky limit, (4) the roll-out limit, and (5) the cohesion limit. Based on his laboratory evaluations, Atterberg established that a soil is plastic between the flow limit (liquid limit) and roll-out limit (plastic limit) and the plasticity number (plasticity index), which is the difference between the flow and roll-out limits, is the best measure of the plasticity of soils [10].

Since Terzaghi introduced Atterberg's limits into modern soil mechanics practice and Casagrande standardized the testing devices, the liquid and plastic limit tests have been extensively performed in geotechnical engineering and soil science fields worldwide [11, 12]. To date, Atterberg limits remain a requirement for most soil classification systems, and they are used in many empirical models for predicting soil engineering properties and specifications for control material properties [12–14]. However, many previous studies have demonstrated that the conventional Atterberg limits tests are highly operator dependent and thus produce significant variations in the test results [11, 15–18]. To provide more repeatable and reproducible test results, different devices and testing methods including the fall cone and bar linear shrinkage tests have been proposed and evaluated [8, 17, 18].

2.1. Liquid Limit Test. In 1932, Casagrande developed a device to standardize the liquid limit (LL) test and in 1949 further refined the design to overcome inherent shortcomings [19]. The later design of the device is standardized in the current ASTM D4318-10 [6]. Although the Casagrande device has become ubiquitous in geotechnical testing, many previous studies have demonstrated that the device yields large variations in LL values. Some of the factors responsible for the large variation are a strong dependency of the results on operator judgment, wear of the grooving tool, and variations in hardness of the base materials of different devices [15–18]. Since late 1950s, many studies have focused on alternative LL measurement methods, and several have concluded that the fall cone device originally developed for testing bitumen materials can eliminate most of the shortcomings of the Casagrande device and provide more consistent test results [17, 18]. Sowers et al. [18] evaluated the effects of cone angle, cone mass, and penetration time on the test results and concluded that the fall cone test is a promising method for measurement of LL. Haigh [16] concluded that the Casagrande cup and fall cone devices actually measure different physical properties of soils and reported that the fall cone test is a measure of specific strength which corresponds to a soil shear strength of ~1.7 kPa, but the Casagrande cup test corresponds to a mean specific strength of ~1.07 m^2/s^2. However, many studies have reported strong correlations between LL values determined by the Casagrande and fall cone test devices for a range of material types [15, 17, 20–25].

2.2. Plastic Limit Test. The fall cone test device was also evaluated for determination of the plastic limit (PL) of soils by [11, 26]. The data interpretations used in these studies were developed based on three assumptions: (1) the undrained shear strength (C_u) of a soil at its PL is approximately 100 times that at its LL [27]; (2) the relationship between moisture content (w) and $\ln(C_u)$ is linear based on critical state soil mechanics concepts [11, 28]; (3) $C_u \cdot d_2/W$ is constant for the same cone geometry, where d is the fall cone penetration depth and W is the weight of the fall cone [11, 28]. Based on these three assumptions, Wroth and Wood [11] proposed to determine the plasticity index (PI) of a soil by conducting fall cone tests with two different cone weights (W_1 and W_2) to determine the water content separation (Δ) of the two parallel flow lines as shown in Figure 1, from which the PI of the specimen can be calculated using the equation shown in the same figure.

2.3. Bar Linear Shrinkage Test. Paige-Green and Ventura concluded that the bar linear shrinkage (BLS) test result is a good indicator of the plasticity of soils [8]. The BLS specimen is prepared by mixing minus No. 40 material at a water content close to its LL, then transferring the material to a 150 mm long by 10 mm square trough and oven drying at 110°C until shrinkage stops. The shrinkage of the specimen is then measured and expressed as a percentage of the original specimen length, which is defined as the BLS value. Paige-Green and Ventura observed a linear correlation between

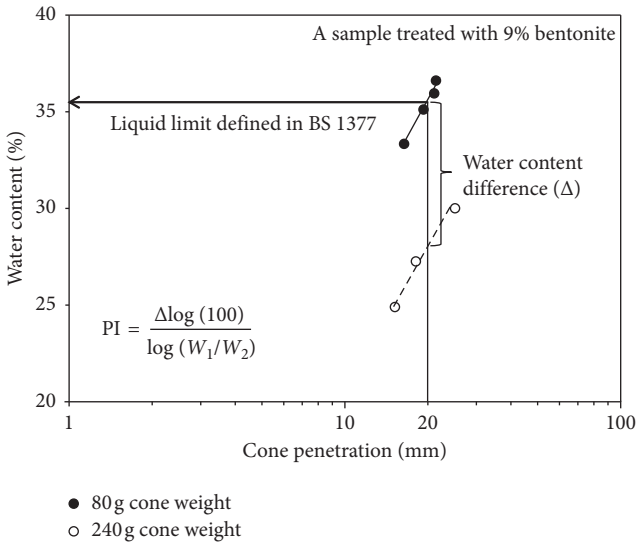

Figure 1: An example of using fall cone test results to determine LL and PI.

BLS values and PI, with the PI values approximately two times the corresponding BLS values [8]. Paige-Green and Ventura [8] found that the BLS test is less susceptible to operator error and is much quicker and easier to learn and perform than the conventional Atterberg limits tests. To improve repeatability and reduce uneven shrinkage leading to bowing (bending) of the specimens, Sampson et al. recommended using a mold with openings on two sides instead of a trough and placing the specimen into the oven immediately after filling to reduce cracking [29]. Similar bar shrinkage tests can be found in several different testing standards, which are summarized in Table 1 along with their various trough dimensions, drying methods, and oven temperatures.

3. Materials and Testing Methods

3.1. Materials. In this study, a total of five samples were prepared by incorporating different percentages of pure bentonite powder into the minus No. 40 fraction of existing crushed limestone fines. The incorporated bentonite content by dry mass of the minus No. 40 material was increased from 0% to 12% in 3% increments. The sieve analysis, hydrometer analysis, and Atterberg limits test results of the initial full granular-surface material gradation are shown in Figure 2.

The chemical composition and mineralogy of the bentonite were determined using X-ray fluorescence (XRF) X-ray diffraction (XRD) analyses, respectively. The XRD results showed that the bentonite was sodium montmorillonite ($Na_{0.3}(Al,Mg)_2Si_4O_{10}(OH)_2{\cdot}4H_2O$) with calcite ($CaCO_3$) and quartz ($SiO_2$). The chemical composition determined by the XRF results is shown in Table 2, indicating that the primary chemical components are SiO_2 and Al_2O_3.

The liquid and plastic limits of the bentonite determined using the methods of ASTM D4318 [6] were 297% and 35%, respectively. Following the recommendation from Bergeson and Wahbeh [30], a 0.5% sodium carbonate (i.e., soda ash) solution was used to increase the water content of the bentonite-treated samples, in order to disperse the bentonite particles and reach a more uniform consistency.

3.2. Testing Methods. The conventional LL and PL tests were performed in accordance with ASTM D4318-10 [6]. The testing devices used in this study are shown in Figures 3(a) and 3(b).

The fall cone LL test was conducted in accordance with the British Standard 1377-2 [7] using the fall cone device shown in part (c) of the figure. The cone weights 80 g and has an apex angle of 30°. The LL is determined as the moisture content corresponding to a cone penetration of 20 mm using a best-fit straight line through the data points of moisture content vs. cone penetration, plotted on linear scales.

To determine the PI using the fall cone device, the testing and calculation methods recommended by Wroth and Wood were followed [11]. As illustrated in Figure 1, the PI of the sample can be calculated based on two known cone weights (W_1 and W_2) and the water content separation (Δ) of the two flow lines at a cone penetration of 20 mm. Along with the fall cone LL tests using the 80 g cone, another set of tests were performed on the same samples using a 240 g cone for this purpose.

The BLS test specimens were also prepared during the fall cone LL tests, because the initial water content of the BLS specimens should be close to the LL that results in a cone penetration of 20 mm. As recommended by Sampson et al., the aluminum BLS molds custom-fabricated for this study are open on two sides, with a length of 150 mm and a 10 mm by 10 mm square cross section (Figures 3(d) and 3(e)) [29]. The molds were first oven-heated at 110°C, then lubricated using a wax bar to reduce friction between the inside walls of the mold and soil specimen, which helps eliminate cracking and uneven or incomplete shrinkage. The wax-lined molds were filled with the soil specimens at moisture contents close to their LL and immediately placed in the oven at 110°C. After drying for 24 hours, the lengths of the specimens were measured using calipers. If a specimen was bowed, the arc height and chord length of the specimen were measured to calculate the average specimen length.

In this study, a total of five samples were prepared at the same time to minimize possible variations caused by sample preparation. For each of the samples, three well-trained operators performed three replicate tests each. The three operators were all trained on all the different tests at the same time in order to minimize errors associated with the interoperator variability.

4. Correlations and Variations between the Various Consistency Tests

4.1. Liquid Limit by Fall Cone vs. Casagrande Cup. The correlation between the liquid limits determined using the Casagrande cup (LL_{cup}) and the fall cone (LL_{cone}) was determined using a total of 45 tests for each device (five bentonite contents times three operators times three

Table 1: Trough designs and drying methods for various BLS testing standards.

Parameter	South Africa TMH1 A4	British standard BS1377	Australia P6A/1	Texas 107-E	California CTM-228
Cross-section shape	Square	Semicircular	Semicircular	Square	Tapered
Cross-section dimension (mm)	10 × 10	25 diameter	25 diameter	19 × 19	Top width 19.05, bottom width 17.48
Length (mm)	150	140	135 or 250	127	127
Drying method	Oven dry (110°C)	Air dry + oven dry (65 and 110°C)[a]	Air dry (24 h) + oven dry (110°C)	Air dry[b] + oven dry (110°C)	Air dry[b] + oven dry (110°C)

[a]Place the mold where the soil/water can air-dry slowly in a position free from drafts until the soil has shrunk away from the walls of the mold. Then complete the drying, first at a temperature not exceeding 65°C until shrinkage has largely ceased and then at 105°C to 110°C to complete the drying. [b]Air-dry the soil bar at room temperature until color changes slightly.

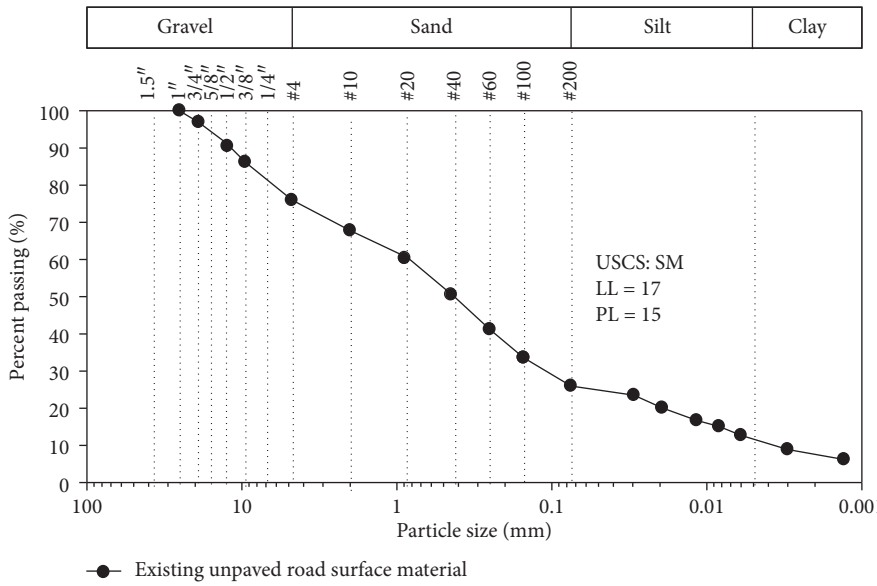

Figure 2: Particle size distribution of the granular-surface material.

Table 2: Chemical composition of the bentonite used in the lab study.

Chemical component	Percent
SiO_2	58.77
Al_2O_3	20.66
Fe_2O_3	3.81
SO_3	0.86
CaO	2.42
MgO	3.61
Na_2O	2.45
K_2O	0.62
P_2O_5	0.08
TiO_2	0.18
SrO	0.03
BaO	0.02
Total	93.50
LOI	6.15
Bulk moisture	7.60

replicates per operator). For each bentonite content, the average LL values from the replicate tests are shown in Figure 4, with error bars indicating the maximum and minimum values.

A strong linear correlation can be observed between the two testing methods. The best-fit line is very close to the 1 : 1 line, but on average, the fall cone test yields higher LL_{cone} values for LL_{cup} values below 33, and lower LL_{cone} values above LL_{cup} equals 33. Both tests yield progressively larger variations with the increasing LL values that result from increasing the bentonite content. However, the variations in the fall cone test results are much smaller than those of the Casagrande cup, as clearly demonstrated by the smaller range of the vertical error bars compared to the horizontal ones. The standard deviation of the test results determined using the two test methods are summarized in Table 3.

The linear correlations determined in the previous and present studies for different types of materials are summarized in Table 4. These linear correlations indicate that using the fall cone test to determine LL is promising, and the test results are very close to those of the conventional Casagrande cup test despite the different mechanisms of the two testing methods.

4.2. Plastic Limit and Plasticity Index by Fall Cone vs. Conventional Method. Fall cone tests were also performed on the five bentonite-treated samples to determine the PI and thereby

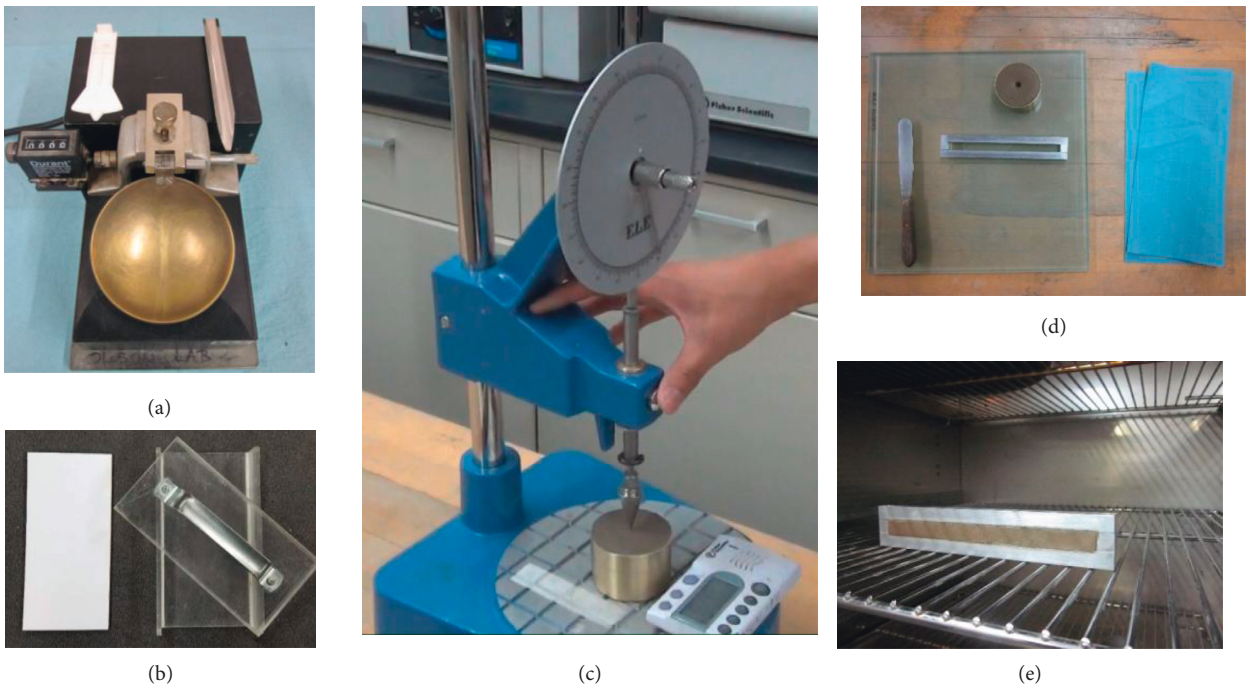

Figure 3: Test devices used in this study: (a) Casagrande LL test device, (b) PL rolling device, (c) fall cone test device, (d) BLS mold, and (e) BLS test specimen in the oven.

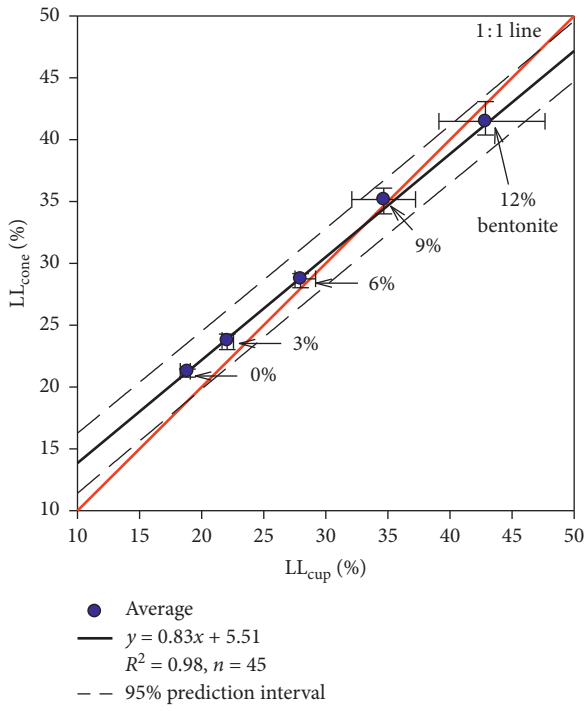

Figure 4: Correlation between LL values determined using Casagrande cup and fall cone device (3 operators × 3 replicates for each bentonite content). *Note.* Error bars show the maximum or minimum values.

the PL, using the previously described method of Wroth and Wood [11]. The relationships between the PI determined using the fall cone (PI_{cone}) and the ASTM-standardized testing methods (PI_{ASTM}) involving rolling specimens into 3.2 mm diameter threads are shown in Figure 5(a).

A linear correlation can be observed between the two testing methods, but the PI_{cone} values are approximately 40% greater than those determined by the ASTM test method using the ASTM plastic roller device. For the fall cone device, the plastic limit (PL_{cone}) was calculated by subtracting the PI_{cone} values from LL_{cone} values. The resulting plastic limits are compared with those from the conventional ASTM-standardized rolling device in Figure 5(b). Interestingly, the PL_{roller} values are approximately the same for the samples treated with different percentages of bentonite (from 0% to 12%), whereas the PL_{cone} values vary over a much wider range. This phenomenon may indicate that the PL determined using the conventional method is governed by the dominant material of the samples (i.e., the minus No. 40 sieved granular limestone material), and the fall cone test is more sensitive to the bentonite content. This observation warrants further study.

4.3. Plasticity Index vs. Bar Linear Shrinkage Values. In this study, the BLS test was also conducted on the five samples with bentonite contents varying from 0 to 12%. The BLS test results compared with the PI values determined using the ASTM standard tests are shown in Figure 6. A linear correlation can be observed between the two parameters, and the variation of the BLS test results generally increases as the bentonite content increases.

However, as the bentonite content increases from 0% to 12%, the PI determined by the ASTM methods varies from

Table 3: Summary of the standard deviation of the liquid limit test results determined using the two tests.

Bentonite content (%)	Standard deviation (%)	
	Casagrande cup	Fall cone
0	0.24	0.21
3	0.28	0.38
6	0.57	0.4
9	1.67	0.67
12	2.93	0.92

Table 4: Correlations between the Casagrande cup and fall cone liquid limit test results.

Reference	Material	LL range (%)	Number of specimens	Correlations
Sherwood and Ryley [17]	Various clays	30–76	25	$LL_{cone} = 0.95\ LL_{cup} + 0.95$
Belviso et al. [21]	Natural soils, Southern Italy	34–134	16	$LL_{cone} = 0.97\ LL_{cup} + 1.19$
Wasti and Bezirci [22]	Turkey natural soils	27–110	15	$LL_{cone} = 1.01\ LL_{cup} + 4.92$
Dragoni et al. [24]	Clayey soils, Central Italy	28–74	41	$LL_{cone} = 1.02\ LL_{cup} + 2.87$
Özer [25]	Natural soils, Turkey	29–104	32	$LL_{cone} = 0.90\ LL_{cup} + 6.04$
Fojtová et al. [23]	Ostrava Basin clay, Czek Republic	20–50	52	$LL_{cone} = 1.00\ LL_{cup} + 2.44$
Di Matteo [15]	Database of various soils	24–50	>50	$LL_{cone} = 1.00\ LL_{cup} + 2.20$
Spagnoli [20]	Kaolinite and illitic clay	20–61	50	$LL_{cone} = 0.99\ LL_{cup} + 1.05$
Present study	Crushed limestone material plus bentonite	20–45	45	$LL_{cone} = 0.85\ LL_{cup} + 5.51$

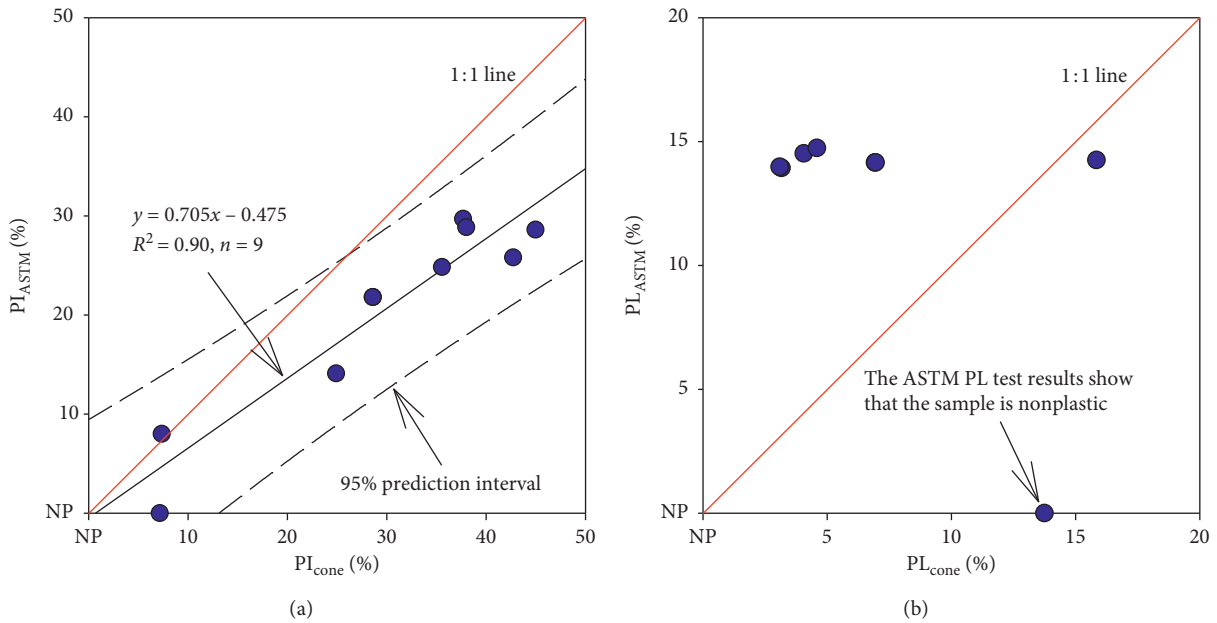

Figure 5: The (a) PI and (b) PL values determined using the fall cone and ASTM standardized methods.

0 to 28%, whereas the BLS values vary over a much smaller range of 2% to 8%. This indicates that BLS values are much less sensitive than PI values to changes in plasticity. More importantly, the ranges of maximum and minimum values of PI (vertical error bars) for the different bentonite contents do not overlap, whereas most of the BLS ranges (horizontal error bars) do overlap. This means that a BLS measurement on the high end of the range for a bentonite content of 3%, for example, could have the same value as the BLS measurement on the low end of the range for a bentonite content of 12%. In both cases, plugging in the BLS value into the linear equation for converting BLS to PI in Figure 6 would result in a significant error in the estimated PI. For this reason, use of BLS as an alternative test method to directly obtain PI instead of measuring LL and PL separately is not recommended in this study.

5. Two-Way ANOVA-Based Repeatability and Reproducibility Analysis

5.1. Description of the Statistical Analysis Method. In this study, a two-way ANOVA-based R&R analysis was used to

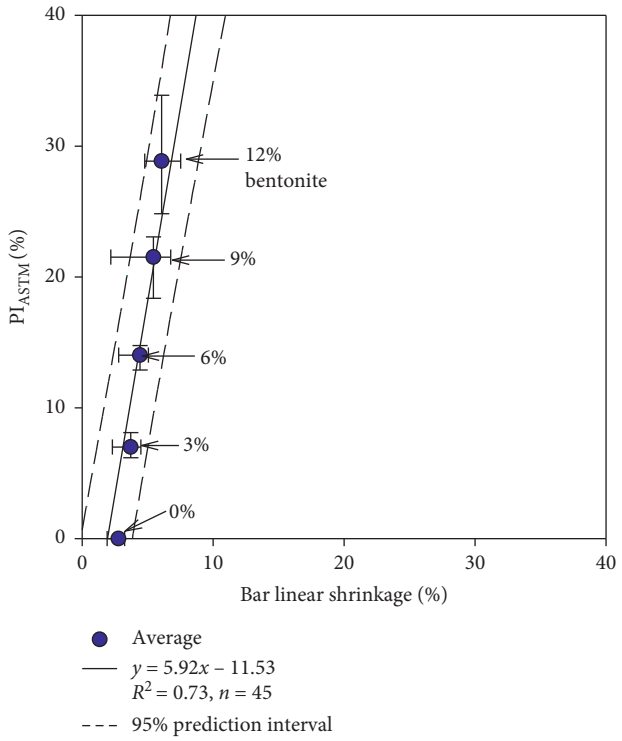

Figure 6: Correlation between BLS values and PI determined by ASTM methods (Casagrande cup for LL and plastic limit roller for PL), for five testing samples. *Note.* Error bars show the maximum or minimum values.

statistically quantify the repeatability, reproducibility, overall variability, and error sources of the various laboratory plasticity tests. This statistical analysis method is detailed in Vardeman and Jobe [31]. The input data for the analysis require J different operators to measure each of I different parts a total of m times. The two-way random effects model is represented by

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}, \tag{1}$$

where y_{ijk} is the k^{th} measurement made by operator j on part i, μ is a measurement averaged over all possible operators and all possible parts, α_i represents the random effects of different parts, β_j represents the random effects of different operators, $\alpha\beta_{ij}$ represents the random joint effects specific to combinations of particular parts and operators, and ε_{ijk} is random measurement error.

The corresponding variances (σ_α^2, σ_β^2, $\sigma_{\alpha\beta}^2$, and σ^2) of the parameters in the model, called "variance components," govern the variability of the measurements.

According to the random effects model, the only difference between different measurements for a specific combination of part and operator is the measurement error (ε), so its standard deviation (σ) is a measure of the repeatability in the model:

$$\sigma_{\text{repeatability}} = \sigma. \tag{2}$$

For a fixed part "i," the value $\mu + \alpha_i$ is constant for different measurements, so the measure of operator bias for a fixed part, i.e., $\sqrt{\sigma_\beta^2 + \sigma_{\alpha\beta}^2}$, is an appropriate measure of reproducibility, which can be expressed as

$$\sigma_{\text{reproducibility}} = \sqrt{\sigma_\beta^2 + \sigma_{\alpha\beta}^2}. \tag{3}$$

Therefore, the overall variation due to repeatability and reproducibility ($\sigma_{\text{R\&R}}$) can be calculated as

$$\sigma_{\text{R\&R}} = \sqrt{\sigma_{\text{repeatability}}^2 + \sigma_{\text{repeatability}}^2}. \tag{4}$$

To obtain the parameters used in this model, a two-way ANOVA table such as Table 5 can be determined based on the test data using a statistical software package.

The number of parts (I) and number of operators (J) should be set as nominal variables for the two-way ANOVA analysis. The three standard deviations can then be calculated as follows:

$$\begin{aligned}
\sigma_{\text{repeatability}} &= \sigma = \sqrt{\text{MSE}}, \\
\sigma_{\text{reproducibility}} &= \sqrt{\max\left(0, \frac{\text{MSB}}{mI} + \frac{(I-1)\text{MSAB}}{mI} - \frac{\text{MSE}}{m}\right)}, \\
\sigma_{\text{R\&R}} &= \sqrt{\frac{\text{MSB}}{mI} + \frac{(I-1)\text{MSAB}}{mI} - \frac{(m-1)\text{MSE}}{m}}.
\end{aligned} \tag{5}$$

The degrees of freedom of the three quantities can be approximately determined using the Satterthwaite method [32] as

$$\begin{aligned}
\nu_{\text{repeatability}} &= IJ(m-1), \\
\nu_{\text{reproducibility}} &= \frac{\sigma_{\text{reproducibility}}^4}{(1/m^2)\left(\left(\text{MSB}^2/I^2(J-1)\right) + \left((I-1)\text{MSAB}^2/I^2(J-1)\right) + \left(\text{MSE}^2/IJ(m-1)\right)\right)}, \\
\nu_{\text{R\&R}} &= \frac{\sigma_{\text{R\&R}}^4}{(1/m^2)\left(\left(\text{MSB}^2/I^2(J-1)\right) + \left((I-1)\text{MSAB}^2/I^2(J-1)\right) + \left((m-1)\text{MSE}^2/IJ\right)\right)}.
\end{aligned} \tag{6}$$

Table 5: Typical two-way ANOVA table for the R&R study.

Source	Sum of squares, SS	Degrees of freedom, df	Mean square, MS
Part (I)	SSA	$I-1$	$\mathrm{MSA}=\mathrm{SSA}/(I-1)$
Operator (J)	SSB	$J-1$	$\mathrm{MSB}=\mathrm{SSB}/(J-1)$
Part × operator ($I\times J$)	SSAB	$(I-1)(J-1)$	$\mathrm{MSAB}=\mathrm{SSAB}/(I-1)(J-1)$
Error	SSE	$IJ(m-1)$	$\mathrm{MSE}=\mathrm{SSE}/IJ\ (m-1)$
Total	SSTot	$IJm-1$	—

The corresponding confidence limits for each of the quantities can be calculated based on the Chi-squared distribution (χ^2_v) using

$$\sigma\sqrt{\frac{v}{\chi^2_{\mathrm{upper}}}},\qquad \sigma\sqrt{\frac{v}{\chi^2_{\mathrm{lower}}}}. \tag{7}$$

The contributions of $\sigma_{\mathrm{repeatability}}$ and $\sigma_{\mathrm{reproducibility}}$ to $\sigma_{\mathrm{R\&R}}$ are quantified using

$$\text{Fraction of } \sigma_{\mathrm{R\&R}} \text{ due to } \sigma_{\mathrm{repeatability}} = \frac{\sigma^2_{\mathrm{repeatability}}}{\sigma^2_{\mathrm{R\&R}}},$$
$$\text{Fraction of } \sigma_{\mathrm{R\&R}} \text{ due to } \sigma_{\mathrm{repeatability}} = \frac{\sigma^2_{\mathrm{reproducibility}}}{\sigma^2_{\mathrm{R\&R}}}. \tag{8}$$

5.2. Results of Repeatability and Reproducibility Analysis. The two-way ANOVA-based R&R analysis was conducted on the results of the various laboratory plasticity tests detailed in the preceding sections. The testing matrix used for the R&R analysis is shown in Table 6. For each of the test methods, each operator conducted three replicate tests on the five samples. Hence, the data collected for analysis of each test method are from three different operators (J) measuring each of the five different parts (I) a total of three times (m).

The results of the analyses are summarized and compared to the R&R reported in ASTM D4318 in Table 7. For the Casagrande cup LL test, the R&R determined using the two-way ANOVA-based method presented herein (0.6% and 1.7%) are close to those reported in ASTM D4318 (0.5% and 1.3%).

The two-way ANOVA-based analysis results show that the overall variation ($\sigma_{\mathrm{R\&R}}$) of the fall cone LL tests is 0.7%, which is less than half that of the Casagrande cup test (1.8%). The analysis results also can identify the sources of error inherent in the test methods. For the fall cone test, the fraction of the $\sigma_{\mathrm{R\&R}}$ due to the $\sigma_{\mathrm{reproducibility}}$ (i.e., between-operator error) is 50%. However, for the Casagrande cup test, 89% of the overall $\sigma_{\mathrm{R\&R}}$ is contributed by the between-operator errors, even though all three operators were trained at the same time. Based on the two-way ANOVA analysis results, it can be concluded that the fall cone test used for measuring the LL is more consistent and less operator dependent than the Casagrande cup test. For the PL test conducted using the ASTM rolling device, the $\sigma_{\mathrm{R\&R}}$ determined using the two-way ANOVA-based method is 0.7%, which is close to the multilaboratory value reported in ASTM D4318. The between-operator error of the PL test is still the main source of the overall variation (73%), which is expected because the testing method is somewhat subjective. The use of the ASTM PL rolling device produces more consistent 3.2 mm-diameter threads compared to rolling by hand, which will improve both the repeatability and reproducibility of the PL test results. However, the R&R analysis was not specifically performed on the hand rolling method in this study.

For the BLS test, $\sigma_{\mathrm{R\&R}}$ is 1%, and the between-operator error accounts for 57% of the overall variation. As discussed in the previous section on correlations, however, the conventional PL and BLS test results were not sensitive to the bentonite content of the mixtures. Therefore, the five different samples (parts) prepared in this study could be regarded as nearly the same for these two tests, which may result in favorably smaller $\sigma_{\mathrm{reproducibility}}$ values.

5.3. Measurement Capacity Ratio. The ASTM standards typically use the d2s limit (i.e., $1.960\sqrt{2}\cdot 1\mathrm{s}$) to determine the acceptable range of two test results (ASTM E691 and E177), which are calculated based on either the single-operator or multilaboratory standard deviations, s. Based on the two-way ANOVA-based R&R analysis results, the measurement capacity ratio (MCR), which is the precision-to-tolerance ratio of a measurement, can be used to quantify the errors from both the testing device and multiple operators.

The MCR can be used to determine whether a measurement is suitable for verifying the conformance of a measured characteristic to engineering specifications. The MCR can also be considered when setting specification ranges based on measurements. For example, if the lower (L) and upper (U) boundaries of a specification for the LL of a material are 30% and 45%, and the $\sigma_{\mathrm{R\&R}}$ of the fall cone LL device is 0.7%, the MCR of the device can be calculated using Equation (9), which gives a MCR value of 0.28.

$$\mathrm{MCR}=\frac{6\sigma_{\mathrm{R\&R}}}{U-L}. \tag{9}$$

According to Vardeman and Jobe [31], "It is common to treat some multiple of $\sigma_{\mathrm{R\&R}}$ (often the multiplier is six, but sometimes 5.15 is used) as a kind of uncertainty associated with a measurement made using the gauge or measurement system in question," and "*The hope is that measurement uncertainty is at least an order of magnitude smaller than the*

Table 6: Laboratory testing matrix used in this study.

Test method	No. of soil samples (I)	No. of operators (J)	No. of replicate tests per operator (m)
Casagrande cup LL test Fall cone LL test ASTM PL test Bar linear shrinkage test	5[a]	3	3

[a]Minus No. 40 sieved granular-road surface material with 0%, 3%, 6%, 9%, and 12% added bentonite.

Table 7: Repeatability and reproducibility results reported in ASTM D4318 and determined by two-way ANOVA-based analysis for the various laboratory tests.

Parameters	Liquid limit		Plastic limit		Bar linear shrinkage
	Casagrande cup	Fall cone	ASTM roller[b]	Fall cone	
Repeatability and reproducibility (R&R) reported in ASTM D4318[a]					
Single-operator standard deviation (%) (within-laboratory repeatability)	0.5	NA	0.3	NA	NA
Multilaboratory standard deviation (%) (between-laboratory reproducibility)	1.3		0.9		
Two-way ANOVA-based R&R Analysis Results					
$\sigma_{\text{repeatability}}$ (%)	0.6	0.5	0.4	NA	0.6
$\nu_{\text{repeatability}}$	30	30	24		30
95% confidence interval (%)	0.5–0.8	0.4–0.6	0.3–0.5		0.5–0.9
$\sigma_{\text{reproducibility}}$ (%)	1.7	0.5	0.6		0.7
$\nu_{\text{reproducibility}}$	6	3	5		3
95% confidence interval (%)	1.1–3.8	0.3–1.7	0.4–1.5		0.4–2.7
$\sigma_{\text{R\&R}}$ (%)	1.8	0.7	0.7		1.0
$\nu_{\text{R\&R}}$	8	10	11		10
95% confidence interval (%)	1.3–3.5	0.5–1.1	0.5–1.2		0.7–1.7
Fraction of $\sigma_{\text{R\&R}}$ due to $\sigma_{\text{repeatability}(\%)}$	11	50	27		43
Fraction of $\sigma_{\text{R\&R}}$ due to $\sigma_{\text{reproducibility}(\%)}$	89	50	73		57

[a]The R&R analysis was conducted on a USCS: ML soil. [b]The ASTM PL test was conducted on four samples, because the 0% bentonite sample was nonplastic.

spread in specifications," which requires that the MCR should be no larger than 0.1 in order to use the measurements to check conformance to such specifications. However, this target MCR value of 0.1 may be too strict for geotechnical applications and needs to be reevaluated for different materials and testing methods.

6. Conclusions and Recommendations

In this study, several laboratory soil plasticity tests were selected for demonstration of the use of a two-way ANOVA-based R&R analysis to evaluate the repeatability, reproducibility, overall variation, source of error, and capacity of a given measurement. Such an analysis can provide useful suggestions for improvement of a testing method. The measurement capacity ratio (MCR) was also demonstrated, which considers errors from both the device and the interoperator variability, and it should therefore be considered when selecting QC/QA testing methods. Based on the findings of this study, the authors suggest using the two-way ANOVA-based analysis presented herein to determine the R&R and identify the sources of measurement error, and considering the MCR of a measurement when setting specifications or selecting QA/QC testing methods. Based on the laboratory testing results, some other key findings about correlations between the various testing methods are listed below:

(i) Correlations between the fall cone and Casagrande cup tests determined in the present and previous studies demonstrated that the fall cone test can be used to determine LL of a material with reduced variability between repeated tests. The two-way ANOVA-based repeatability and reproducibility analysis also revealed that the fall cone test can result in smaller overall variation than the Casagrande cup test, which is more prone to between-operator errors.

(ii) For measuring the PL and PI, the fall cone test and conventional test method using the ASTM plastic roller yielded significant discrepancies for the abraded crushed limestone granular materials with small percentages of bentonite incorporated. The fall cone test showed a dependence of PL on the bentonite content, whereas the conventional method was practically insensitive to the bentonite content. Further studies need to be conducted to evaluate the influence of the different testing mechanisms and whether PL is governed by the dominant minerals of a soil mixture.

(iii) The bar linear shrinkage results exhibited a linear correlation with the PI determined by conventional ASTM testing methods. However, as the PI increased significantly from 0 to 28% by incorporating

bentonite, the corresponding BLS values were much less sensitive, exhibiting a change of only 6%. Moreover, the ranges of measured BLS values for the different bentonite contents overlapped, prohibiting a reasonably accurate correlation between BLS and PI.

Nomenclature

LL_{cone}: Liquid limit by fall cone tests
LL_{cup}: Liquid limit by ASTM standardized Casagrande cup tests
MCR: Measurement capacity ratio
PIASTM: Plasticity index by ASTM standardized tests
PI_{cone}: Plasticity index by fall cone tests
PL_{cone}: Plastic limit calculated based on fall cone test results
s: Standard deviation
$\sigma_{repeatability}$: Repeatability standard deviation
$\sigma_{reproducibility}$: Reproducibility (between-operators) standard deviation
$\sigma_{R\&R}$: Combined R&R standard deviation.

Data Availability

The data used to support the findings of this study are available from the first author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank the Iowa Department of Transportation for sponsoring this study. The financial supports provided by Natural Science Foundation of Shaanxi Province in China (grant no. 2019JQ-498), the Science and Technology Projects of Gansu Transportation Department (grant nos. 2019-16 and 2019-17), and Opening Foundation of Research and Development Center of Transport Industry of Technologies, Materials and Equipments of Highway Construction and Maintenance (Gansu Road and Bridge Construction Group) (grant nos. GLKF201804 and GLKF201807), are greatly appreciated.

References

[1] ASTM E691-15, *Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method, Annual Book of ASTM Standards*, ASTM International, West Conshohocken, PA, USA, 2015.

[2] ASTM E177-13, *Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods, Annual Book of ASTM Standards*, ASTM International, West Conshohocken, PA, USA, 2013.

[3] Y. Li, Y. Yang, H.-S. Yu, and G. Roberts, "Principal stress rotation under bidirectional simple shear loadings," *KSCE Journal of Civil Engineering*, vol. 22, no. 5, pp. 1651–1660, 2018.

[4] J. Li, D. J. White, W. R. Stephenson, and C. Li, "Considerations for laboratory resilient modulus testing of unbound pavement base materials," *Construction and Building Materials*, vol. 195, pp. 515–523, 2019.

[5] C. Li, J. C. Ashlock, S. Lin, and P. K. R. Vennapusa, "In situ modulus reduction characteristics of stabilized pavement foundations by multichannel analysis of surface waves and falling weight deflectometer tests," *Construction and Building Materials*, vol. 188, pp. 809–819, 2018.

[6] ASTM D4318-10, *Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils*, Annual Book of ASTM Standards, ASTM International, West Conshohocken, PA, USA, 2013.

[7] BS 1377-2, *Methods of Test for Soils for Civil Engineering Purposes, Part 2*, BSI, London, UK, 1990.

[8] P. Paige-Green and D. Ventura, "The bar linear shrinkage test–more useful than we think," in *Proceedings of the 12th Regional Conference for Africa on Soil Mechanics and Geotechnical Engineering: Geotechnics for Developing Africa*, G. R. Wardle, G. E. Blight, and A. B. Fourie, Eds., pp. 379–388, A.A. Balkema, Durban, South Africa, October 1999.

[9] A. Atterberg, "Uber Die Physikalische Bodenuntei Suchung and Uber Die Plastizitat Der Tone," *Internationale Mitteilungenfiir Bodenkunde*, vol. 1, 1911.

[10] E. Bauer, "History and development of the Atterberg limits tests," in *Papers on Soils 1959 Meetings*, pp. 160–167, ASTM International, West Conshohocken, PA, USA, 1959.

[11] C. P. Wroth and D. M. Wood, "The correlation of index properties with some basic engineering properties of soils," *Canadian Geotechnical Journal*, vol. 15, no. 2, pp. 137–145, 1978.

[12] B. C. O'Kelly, P. J. Vardanega, and S. K. Haigh, "Use of fall cones to determine Atterberg limits: a review," *Géotechnique*, vol. 68, no. 10, pp. 843–856, 2018.

[13] C. Li, J. C. Ashlock, B. Cetin, C. T. Jahren, and V. Goetz, "Performance-based design method for gradation and plasticity of granular road surface materials," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 52, pp. 216–225, 2018.

[14] C. Li, J. C. Ashlock, D. J. White, and P. K. R. Vennapusa, "Mechanistic-based comparisons of stabilised base and granular surface layers of low-volume roads," *International Journal of Pavement Engineering*, vol. 20, no. 1, pp. 112–124, 2019.

[15] L. Di Matteo, "Liquid limit of low- to medium-plasticity soils: comparison between casagrande cup and cone penetrometer test," *Bulletin of Engineering Geology and the Environment*, vol. 71, no. 1, pp. 79–85, 2012.

[16] S. K. Haigh, "Mechanics of the casagrande liquid limit test," *Canadian Geotechnical Journal*, vol. 49, no. 9, pp. 1015–1023, 2012.

[17] P. T. Sherwood and M. D. Ryley, "An investigation of a cone-penetrometer method for the determination of the liquid limit," *Géotechnique*, vol. 20, no. 2, pp. 203–208, 1970.

[18] G. F. Sowers, A. Vesić, and M. Grandolfi, "Penetration tests for liquid limit," in *Papers on Soils 1959 Meetings*, pp. 216–226, ASTM International, West Conshohocken, PA, USA, 1960.

[19] A. Casagrande, "Notes on the design of the liquid limit device," *Géotechnique*, vol. 8, no. 2, pp. 84–91, 1958.

[20] G. Spagnoli, "Comparison between casagrande and drop-cone methods to calculate liquid limit for pure clay," *Canadian Journal of Soil Science*, vol. 92, no. 6, pp. 859–864, 2012.

[21] R. Belviso, S. Ciampoli, V. Cotecchia, and A. Federico, "Use of the cone penetrometer to determine consistency limits," *Ground Engineering*, vol. 18, no. 5, pp. 21-22, 1985.

[22] Y. Wasti and M. H. Bezirci, "Determination of the consistency limits of soils by the fall cone test," *Canadian Geotechnical Journal*, vol. 23, no. 2, pp. 241–246, 1986.

[23] L. Fojtová, M. Marschalko, R. Franeková, and L. Kovář, "Study of compatibility of methods for liquid limit measurement according to Czech state standard and newly adopted European standard," *GeoScience Engineering*, vol. 55, no. 1, pp. 55–68, 2009.

[24] V. U. Dragoni, N. Prosperini, and G. Vinti, "Some observations on the procedures for the determination of the liquid limit: an application on plio-pleistocenic clayey soils from Umbria region (Italy)," *Italian Journal of Engineering Geology and Environment*, vol. 1, pp. 185–197, 2008.

[25] M. Özer, "Comparison of liquid limit values determined using the hard and soft base casagrande apparatus and the cone penetrometer," *Bulletin of Engineering Geology and the Environment*, vol. 68, no. 3, pp. 289–296, 2009.

[26] J. Kodikara, H. N. Seneviratne, and C. V. Wijayakulasooryia, "Discussion of "using a small ring and a fall-cone to determine the plastic limit" by Tao-Wei Feng," *Journal of Geotechnical and Geoenvironmental engineering*, vol. 132, no. 2, pp. 276–278, 2006.

[27] A. W. Skempton and R. D. Northey, "The sensitivity of clays," *Géotechnique*, vol. 3, no. 1, pp. 30–53, 1952.

[28] S. Hansbo, *A New Approach to the Determination of the Shear Strength of Clay by the Fall-Cone Test*, Royal Swedish Geotechnical Institute, Stockholm, Sweden, 1957.

[29] L. R. Sampson, D. F. C. Ventura, and D. K. Kalombo, *Linear Shrinkage Test: Justification for Its Reintroduction as a Standard South African Test Method*, South African Roads Board, Pretoria, South Africa, 1992.

[30] K. L. Bergeson and A. M. Wahbeh, *Development of an Economic Dust Palliative for Limestone Surfaced Secondary Roads, Project HR-297*, Iowa Department of Transportation, Ames, Iowa, 1990.

[31] S. B. Vardeman and J. M. Jobe, *Statistical Quality Assurance Methods for Engineers*, John Wiley, New York, NY, USA, 1998.

[32] F. E. Satterthwaite, "An approximate distribution of estimates of variance components," *Biometrics Bulletin*, vol. 2, no. 6, pp. 110–114, 1946.