Procedure for Validating Automatic Consistence Measurement Systems

Rev	Date	Author	Reviewer	Comments
01	01/12/2023	Ed Waugh, Emeso Ojo	Mark Steadman	
02	17/12/2025	Ed Waugh, Emeso Ojo	Mark Steadman	

Contents

Executive Summary

This document describes the procedure for validating that the performance of an automatic consistence measurement system is equivalent to or better than the manual test. Using the method described in the Cloud Cycle white paper *Validating Automatic Consistence Measurement Systems*[1], each step of the procedure is described in detail, allowing technology providers, Ready-Mix Concrete producers, and construction firms to perform the validation. The parameters of the procedure are mandated to guarantee correctness when used with the analysis provided, they have been chosen pragmatically to balance the need for accuracy with practical limits when working on-site.

1. Collect 20 measurements of 3 testers each completing 2 tests and corresponding automatic measurements.
2. Calculate the upper and lower limits of agreement and compare them to the thresholds in Table 1 to determine if the validation is passed
3. Prepare the data in a standard format for dissemination

1. Introduction

This document describes the procedure for validating that the performance of an automatic consistence measurement system is equivalent to the manual test. Its purpose is to provide a systematic approach that can be used by interested parties to independently assess the performance of a system and validate technology provider claims.

This document does not explore the mathematics behind the procedure. For an in-depth explanation and examples please see the Cloud Cycle white paper: Validating Automatic Consistence Measurement Systems [1].

The document is split into two sections. The first describes the procedure for collecting reference data to ensure its suitability for analysis and the second describes the statistical analysis and reporting.

1.1. Scope

This document describes a method for evaluating the performance of automatic consistence reporting systems. It is intended to be technology-agnostic and applies regardless of the measurement principle employed or whether the system is installed on fixed or mobile equipment. The method is applicable to automatic systems reporting concrete consistence, including, but not limited to, slump, slump-flow, and flow measurements.

This document does not address the calibration of digital consistence monitoring systems. It is assumed that calibration has been completed by the technology provider prior to system validation in accordance with their internal procedures and any applicable standards.

1.2. Background

Consistence (also referred to as workability) is typically assessed in terms of the outcome of several physical tests including slump, flow table or slump flow tests. Each test type has an associated precision, which determines the range of results obtained when a sample of concrete is subject to repeated tests. Averaging the results of these tests tends towards the notional "true" value [1] that is an intrinsic property of the material, which is approximated by either the corresponding manual test or an automatic consistence measurement system.

1.3. Definition of Terms

Calibration
A process by which a measurement system is adjusted to remove systematic errors. This normally involves making a measurement with the system alongside a reference measurement and calculating an error correcting function. That function is then embedded in the system and used in future calculations.

Validation
A process that ensures that the measurement system meets the needs of the user. This shall include comparing the system output to a reference value and assessing it against a set of criteria. This process may be repeated periodically.

Verification
A process that ensures that the measurement system meets a set of design requirements. This is normally an internal process where non-customer facing parameters may be tested.

Consistence
Property of freshly mixed concrete that determines the ease at which it can be mixed, placed, consolidated, and finished to a homogeneous condition.

Consistence test
An assessment of the consistence of fresh concrete. A single test performed by one tester.

Consistence measurement
Multiple consistence tests performed simultaneously by multiple testers on concrete from the same sample.

Consistence Class
A range of consistence measurements grouped to provide a meaningful short-hand description.

Technology Provider
Any provider of an automatic system that can be used to monitor or control the consistence of fresh concrete.

Accuracy
Both Trueness and Precision

Precision
The spread of the results of repeated measurements determined by random errors.

Trueness
The closeness of the result to the true value, determined by systematic errors.

Standard Error
A measure of how much a sample statistic, such as the mean, is likely to differ from the true population value.

Variance
A measure of how widely spread a set of values are from their mean. Defined as the average squared difference from the mean.

Standard Deviation
A measure of how widely spread a set of values are from their mean. Defined as the square root of the variance, it is in the same units as the underlying value.

Normal Distribution
Also known as a Gaussian distribution or bell curve, a continuous probability distribution defined by a mean and variance.

Tester
A person trained in performing the manual test method.

True Consistence
A notional value that is estimated by different measurement methods.

Reproducibility
Reproducibility is the spread of the differences between two Testers, with different sets of test equipment, performing the same test.

Repeatability
Repeatability is the spread of the differences between successive measurements performed by the same Tester with the same equipment.

Bias
A constant difference between the reference value and the measured value caused by systematic errors.

Concrete Mix
A combination of a set of constituents designed to achieve concrete for use in a specific application

Concrete Family
A grouping of similar concretes that have the same behaviour when measured automatically. These are technology dependent and determined by the technology provider.

True Positive
In this context, a true positive is the case where the measurement system is genuinely at least as precise as the manual test and passes validation criteria.

False Positive
In this context, as false positive is the case where a measurement system is less precise than the manual test but is erroneously determined to be valid.

2. Preparation

2.1. Overview

To perform validation, 20 measurements are required that cover the consistence range being validated [1]. This can normally be performed over two days at a busy batch plant, although other locations can be considered. Each delivery of a Concrete that is part of the Concrete Family under validation will be paused while any technology provider requirements are met and a sample is taken. The truck can then continue its delivery as normal.

2.2. Concrete Selection

The performance of a consistence monitoring system will be assessed according to Concrete Family and the mixes used for testing should be representative of that family. The consistence of the selected mixes shall fall within the range supported by the automatic consistence measurement, and the validation shall be performed across the specified consistence range for the system. The validation applies to the Concrete Family that has been tested as agreed between the customer and system provider.

2.3. Data Collection from the Automatic System

The consistence value reported by the automatic system shall be recorded in accordance with the technology providers instructions at a time as close as possible to the manual measurement. The exact time of the measurement shall be recorded.

2.4. Location

The validation can be performed at the batch plant, a delivery site or another location, the requirements for the site are:

- The truck can stop there for a short time while performing a normal delivery
- Sufficient space for the truck to park and be operated safely
- A flat and level surface for the testers to work
- A structure that provides sun or rain protection if required
- Access to a water supply for cleaning equipment
- A way to use the tested concrete or recycling/disposal facilities
- The weather conditions at the time of testing are suitable

2.5. Recommended Roles

It is helpful to assign roles to the different members of the validation team to ensure that all steps are performed correctly and safely.

2.5.1. Tester

Three testers are required to make the measurements and should be appropriately trained in advance of the validation.

2.5.2. Test Supervisor

The Test Supervisor is responsible for the quality of the data collected.

- Record the information for each tester including their full name, experience and qualifications and company that supplied them
- Record details of each test location and photograph the site

- Brief the testing team at the start of day on what is being done
- Review the procedure for each test method being used with the Testers
- Ensure everyone has the correct equipment for each of the tests and that equipment is in a good condition
- Monitor the testers during the testing process and ensure they are following it correctly
    - Correct any errors in technique and record in the day notes
    - Ensure equipment is cleaned correctly between each test
    - Ensure they do not know the results from other testers
    - Ensure they do not know the expected slump of the load they are testing
- Record the results of all the tests

2.6. Equipment

- Three identical sets of consistence measuring equipment, in good condition and made according to the appropriate standard
- Three wheelbarrows with a minimum capacity of 100 Kg
- Buckets for water, brushes and squeegees for cleaning equipment between tests
- An accurate clock used for recording all times
- A method for recording the test results as they are made

3. Procedure

The validation procedure requires 20 measurements of the Concrete Family being validated.

3.1. Setup

Before the first truck is brought for testing the site and team shall:

- Perform all required safety procedures
- All equipment is laid out, cleaned and checked by the Test Supervisor
- Testers are briefed on the test plan and the importance of measuring independently

3.2. Procedure for a Single Measurement

(1) A truck fitted with the automatic consistence measurement system containing a load of Concrete is moved into position
(2) The batch record shall be obtained and recorded
(3) Fully mix the concrete according to the design
(4) Perform any technology provider required steps e.g. drum rotations
(5) Discharge enough concrete into wheelbarrows to complete the required number of consistence tests:
    - For standard consistence tests one wheelbarrow is sufficient to perform six tests
    - The first wheelbarrow is discarded in all standard methods
    - A third wheelbarrow is helpful to capture excess discharge and collect waste
(6) Thoroughly remix the sample prior to testing to ensure homogeneity
(7) Each tester shall perform two consecutive consistence tests in accordance with the relevant regional standard, without prior knowledge of:
    - The value reported by the automatic consistence measuring system
    - The specified consistence class or target of the concrete
(8) Measurements shall be recorded by the Test Supervisor. To ensure the independence of each test the Tester must report their test result without informing the other Testers. They shall also record:
    - The time of the first and last tests to the minute
    - The automatic consistence measurement system result, and the time it was taken
    - Any observations of the concrete or slump behaviour
(9) Following testing, the used sample is discarded and not returned to the test barrow
(10) All testing equipment shall be cleaned thoroughly before further use
(11) The Test Supervisor assesses the measurement against the Data Exclusion Criteria (section 3.3) and determines if it can be used in the subsequent analysis

3.3. Data Exclusion Criteria

Data collected during validation shall be excluded from performance analysis if any of the following conditions apply:

- The Tester has not complied with the specified sampling or testing procedures, including deviations from the relevant consistence test standard
- The concrete exhibits visible signs of segregation or excessive bleeding
- Measurements shall be excluded where the average of all test results falls outside the supported measurement range set out by the technology provider.

All excluded data shall be clearly identified, with the reason for exclusion recorded. The number and proportion of excluded measurements shall be reported as part of the validation results.

4. Performance Evaluation

4.1. Data Preparation

Prior to analysis, the dataset shall be reviewed to confirm that:

- A manual and automatic test result were recorded
- All measurements comply with the data exclusion criteria defined in section 3.3
- Each digital measurement is within 10 minutes of its corresponding manual reference test.

4.2. Analysis

To aid in the analysis an accompanying spreadsheet is supplied that calculates the results of the validation automatically[2]. For a detailed explanation of the mathematics please refer to the supporting document *Validating Automatic Consistence Measurement Systems*[1].

First, calculate the mean manual result for each measurement as shown in Equation 4.1, where Test X, Y is the first test X made by Tester Y e.g. Test 2.3 is the second test made by Tester 3.

$$\overline{Manual} = \frac{Test1.1 + Test1.2 + Test1.3 + Test2.1 + Test2.2 + Test2.3}{6} \tag{4.1}$$

Then calculate the difference between each of the mean manual results and the automatic measurement taken alongside it as shown in Equation 4.2.

$$\bar{d} = \frac{\sum_{i=1}^{N}\left(Automatic_i - \overline{Manual_i}\right)}{n} \tag{4.2}$$

To calculate the Limits of Agreement the unbiased standard deviation of the differences is also required and is calculated as shown in Equation 4.3.

$$s = \sqrt{\frac{\sum_{i=1}^{N}\left(d_i - \bar{d}\right)^2}{n-1}} \tag{4.3}$$

The Limits of Agreement are a pair of values calculated as shown in Equation 4.4 and rounded to mm.

$$LoA = round\left(\bar{d} \pm 2s\right) \tag{4.4}$$

Once these values are obtained, they can be compared to the appropriate threshold taken from Table 1. If both Limits of Agreement are less than or equal to the ± threshold, $\boldsymbol{\Delta_{adj}}$, then the system is determined to be a valid replacement for the manual test. See [1] for details on determination of the threshold values.

Test type	Test range	R (mm)	Threshold Δ_{adj} (mm)
Slump	S1	20	21
	S2	25	27
	S3 - S4	37	39
Flow	F2 – F5	91	96
Slump-flow	SF1 – SF2	43	44
	SF3	28	29

Table 1 - Reproducibility taken from the standards [3][4][5][6]. Δ_{adj} is the threshold criterion for calculated limits of agreement including the tolerance of 90%

5. Reporting validation outcomes

Validation outcomes shall be reported in a standardised way for each campaign.

The required components of a validation report are:

1. A scatter plot of the manual vs automatic measurements and a line of perfect agreement
2. A histogram of the differences showing the upper and lower limits of agreement and the target (adjusted) limits of agreement.
3. A statement of validation including the average manual consistence and its standard deviation, the average difference (bias), and the upper and lower Limits of Agreement.

Below is an example of a report based on validation testing of the Cloud Cycle flow measurement system.

"For the validation test carried out on 19th and 20th March, 2026 the Cloud Cycle system was found to be a valid replacement for the manual flow table test across the F2-F5 range. The mean manual flow across all tests was 586 mm with a standard deviation of 67 mm. The mean difference (bias) between automatic and manual measurements was 5 mm. The limits of agreement were -46 mm to 56 mm against a target of ±96 mm."

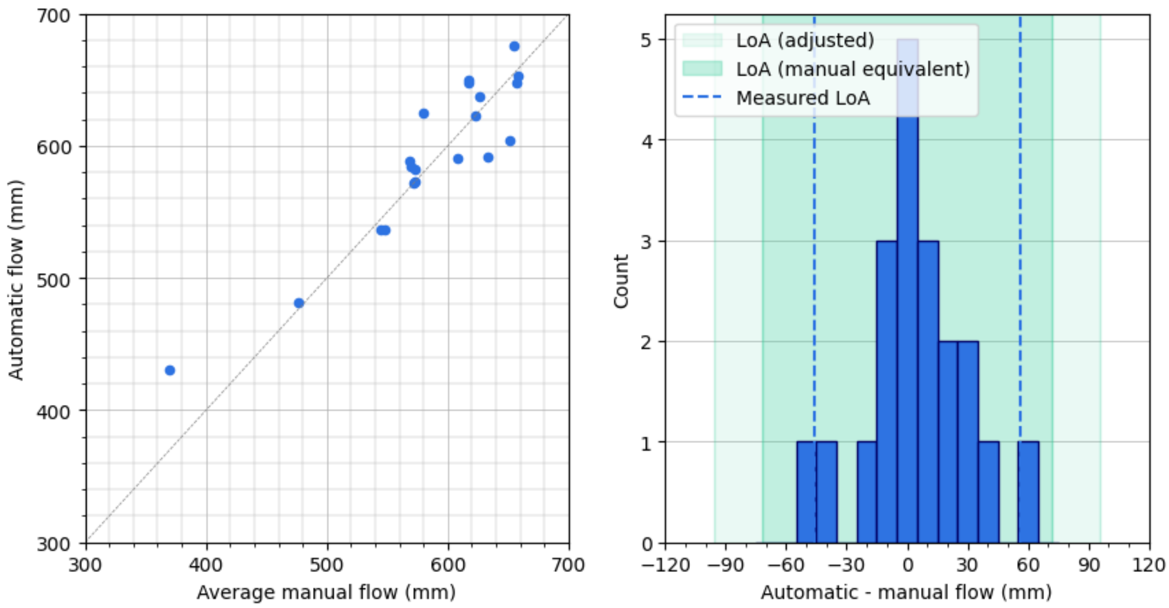

Figure 1 - Simulated validation of System 3, (A) example scatter plot (B) example histogram

6. References

[1] Validating Automatic Consistence Measurement Systems, Cloud Cycle, 2026

[2] Consistence Validation Record, Cloud Cycle, 2026

[3] BS EN 12350-2:2019 Testing Fresh Concrete Part 2: Slump Test

[4] BS EN 12350-5:2019 Testing Fresh Concrete Part 5: Flow Table Test

[5] BS EN 12350-8:2019 Testing Fresh Concrete Part 8: Self-compacting Concrete, Slump Flow Test

[6] ASTM C143/C143M Standard Test Method for Slump of Hydraulic-Cement Concrete

7. FAQ

Can an automatic consistence measurement system be more accurate than the manual test?

Yes. There is no reason why an automatic system cannot be more accurate than a manual test. An automatic system with a low-bias and a low standard deviation can outperform a single tester and reach the performance of multiple averaged tests.

Would more data points improve the accuracy of the validation test outcome?

Yes, they would. It is possible to design the validation test procedure based on different requirements such as an acceptable limit of false positives and false negatives under certain conditions. The validation test procedure has been designed pragmatically such that it can be carried out within a reasonable time at a reasonable cost.

Can I carry out more than the specified number of tests, if this gives a more accurate outcome?

No. The tolerance applied to the limits of agreement is based on a defined number of data points. Adding more without adjusting the tolerance would increase the risk of a false positive.

Would it be better for the outcome to be the result of a statistical test?

Yes, this would provide information about how confident we are in the outcome based on the data. A common approach is to estimate confidence intervals for the limits of agreement and carry out a “two one-sided tests” analysis to determine whether both limits are within the target range with sufficient degree of confidence. Bayesian approaches could enable statements like “based on the data, there is a 70% chance that the system is valid”. These are not described here as they require expert interpretation and therefore a higher barrier for adoption.

The automatic system reports consistence to the nearest 10 mm. Do I need to adjust the analysis?

No. It is not necessary to adjust the analysis. If the system only reports results to the nearest 10 mm this is a design choice that introduces rounding error to that system’s measurements. This will reduce the accuracy of the system, but the benchmark against which to validate it is based only on the properties of the manual test and remains the same.

For compliance, the standards allow greater tolerance for spot tests vs. composite tests. Why are they treated the same here?

There is no reason the method of sampling should affect the precision of the manual test. Correspondingly, the precision tables in the standards do not distinguish between the two sampling methods. However, if the load is not well mixed then it is reasonable to expect that the composite test could be more representative of the average for a batch of concrete, which is likely the justification for allowing different tolerances for consistence tests carried out on spot vs composite samples. When validating automatic consistence measurement, it is important that the samples used in both the automatic and manual tests are well matched, and the best way to do this is to ensure that the load is well mixed prior to carrying out any testing.

Why are some consistence classes missing from the table of limit values?

Annex L of (British Standards Institution, 2013) states:

Due to a lack of sensitivity of the test methods beyond certain values of consistence it is recommended to use the indicated tests for:

Slump ≥ 10 mm and ≤ 210 mm

Flow diameter > 340 mm and ≤ 620 mm

Slump flow diameter > 550 mm and ≤ 850 mm

Measurement comparison should only be made where the manual test results are within the sensitive range. This means that tests within the F1, F6 and S5 ranges are not included. These consistence classes are defined by having a consistence lower or higher than the sensitive ranges, this does not matter when using the tests to determine compliance. However, the precision metrics are not defined so no comparison can be made.

Can I use the limits of agreement to compare performance of multiple automatic measurement systems?

No. In a validation study comprising 20 tests, the calculated limits of agreement could have a high level of uncertainty associated with them. It would be misleading to conclude that one system is more accurate than another based on the limits of agreement determined from this procedure.

8. Appendix – Common Testing Problems

8.1. Sample preparation

The goal is to acquire a well-mixed sample that has not been subjected to vibration.

- Watch for testers lifting and dropping the barrows to level the concrete before moving it from the truck to the test site
- Make sure the sample is thoroughly remixed, including material at the bottom of the barrow and up against the sides (more heavy aggregates tend to end up at the back and bottom of the barrow)
- Ensure testers turn the material over during mixing without lifting and slapping concrete down.

8.2. Slump Test Procedure

- Testers will often step off the tabs to reach a rod or scoop causing the cone to lift early
- When tamping, check they are not striking the base when tamping the first layer
- Ensure the height of each layer is a third of the height of the cone
- Check the rod comes out of the sample each time (i.e. they are tamping, not stirring)
    - Check tamping is evenly distributed across the surface of each layer - testers often forget about spiralling in towards the middle
    - It's common for the tamping rod to always lean away from the tester, so the material at the far side is not tamped.
- It's common for testers to lift the side of the cone closest to them first, putting their weight the other side of the cone to help start the lift. This can cause loose / wet material to start pouring out before the cone is fully off the base. The solution is to stand / squat more over the top of the cone, not far away from it.
- Testers often skip the step of ensuring that the concrete surface is level after the final round of tamping