A Pragmatic Method for Pass/Fail Conformance Reporting that Complies with ANSI/NCSL Z540.3, ISO/IEC 17025, and ILAC-G8

Michael C Dobbert[1], *Robert M Stern*[2], *Rui Villela Ferreira* [3]

[1] Agilent Technologies Inc, Santa Rosa/CA, USA, dobbert@agilent.com
[2] Agilent Technologies Inc, Santa Rosa/CA, USA, bob_stern@agilent.com
[3] Agillent Technologies Brasil Ltda, Barueri/SP, Brazil, rui_villela@agilent.com

Abstract: A practical criteria for stating Pass/Fail conformance when calibrating an instrument and comparing the measured results against manufacturer specifications, considering regional and regulatory requirements. A flexible calibration result reporting includes Pass/Fail conformance statements, which is especially true when serving a global market. This paper explores the different requirements or guidelines in standards documents, such as ANSI/NCSL Z540.3-2006, ISO/IEC 17025:2005, ILAC-G8:1996, and EURAMET/cg-15/v.01 and presents a non-obvious, yet simple method for expressing statements of Pass/Fail conformance. It employs flexible acceptance limits resulting in straightforward "Pass" and "Fail" conformance labels, with unobtrusive annotation to communicate additional information required by the standards documents. The result is a concise, uniform method flexible enough to satisfy all of the aforementioned standards, regardless of the chosen acceptance limits.

Key words: Statement of conformance, Acceptance and Tolerance limits, False Accept/Reject risks.

1. INTRODUCTION

When making a statement of conformance, we must acknowledge the risk that the statement may be incorrect. Of particular concern is how to report measurement results that fall outside the acceptance limit, yet are within the manufacturer's tolerance.

Various calibration standards each address risk management in a different way. The key differences are:

• ANSI/NCSL Z540-1 [1]: Pass/Fail criteria was a simple comparison to the instrument manufacturer's specified tolerance, so acceptance limits were equal to tolerance limits.

• ANSI/NCSL Z540.3 [2]: The probability of false acceptance (PFA) associated with any test point labeled "Pass" shall not exceed 2 %. (5.3 b)

• ISO/IEC 17025 [3]: States Pass/Fail criteria as, "When statements of compliance are made, the uncertainty of measurement shall be taken into account." (5.10.4.2) Accreditation bodies provide local regional interpretation of the international standard.

• ILAC-G8:1996 [4]: Pass/Fail criteria uses the 95%expanded uncertainty for making statements of conformance. For measured values where the specified tolerance is within the 95%expanded uncertainty interval, no declaration of conformance is made. Most European accreditation bodies require ILAC-G8 for statements of conformance for ISO/IEC 17025 calibrations.

• EURAMET/cg-15/v.01 [5]: Though targeted for digital multimeters, EURAMET/cg-15/v.01 can be applied to other instruments. No guard band is applied when assessing conformance during calibration. "Subsequent to calibration and under normal conditions of use, the uncertainty associated with the readings of a DMM will be the combination of the DMM's specification and the calibration uncertainty." (4.2).

2. UNDERSTANDING FALSE ACCEPT AND FALSE REJECT RISK

False accept risk depends on several factors, including the specified tolerance limits, the acceptance limits, calibration process uncertainty and the distribution of true values from a device under test population.

Visualizing risk is possible by looking at the relationship between the true values from a device under test population and the corresponding measured values observed during calibration. Because of measurement error, the measured values obtained during calibration only approximate the true values. Figure 1 illustrates this relationship graphically. The x-axis represents the true values of a population of devices and is described by a distribution.

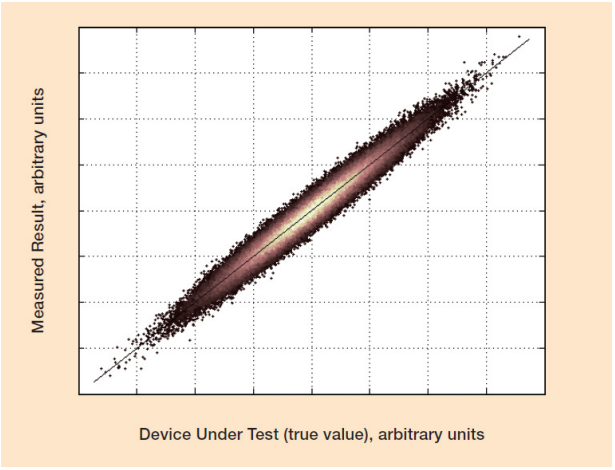

Fig. 1. Measured Results versus True Value

(For the graphic in Fig. 1, a Gaussian distribution represents both the device under test population and the measurement error. The ratio of the device population standard deviation and measurement error standard deviation is 4:1. For more information, see reference [8]).

The y-axis represents the measured values and includes measurement error. As long as the measurement error is not significant, measured values correlate very well with the true values, which is a desired attribute for a quality calibration. However, even with low measurement error, it is possible to make an incorrect in- or out-of-tolerance assessment.

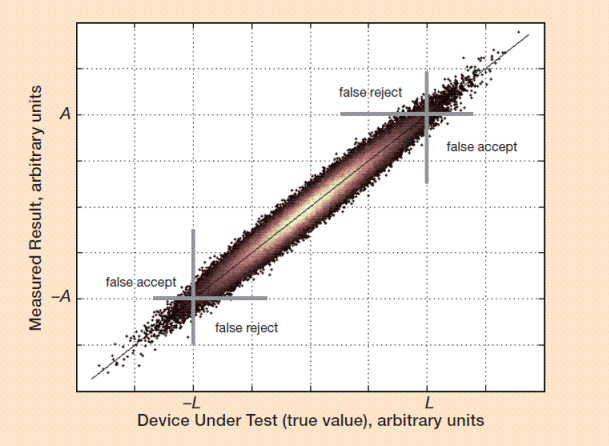

Fig. 2. False accept and false reject regions based on specified tolerance limits (-L, L) and acceptance limits (-A, A).

In Fig. 2, the tolerance limits (-L, L) represent two-sided symmetrical limits for the device under test. Devices with true values within the tolerance limits are in-tolerance. Devices with true values outside the tolerance limits are out-of-tolerance. However, because it is not possible to ever know the true value, to assess in- or out-of-tolerance status, the only recourse is to apply acceptance limits against the measured value. The acceptance limits (-A, A) represent two-sided symmetrical limits, in this case. As illustrated in Fig. 2, the tolerance limits and acceptance limits, together, define several regions. The regions labeled false accept include devices with measured values within the acceptance limits but with true values outside the tolerance limits. These devices appear to be in-tolerance as measured, but in reality, are out-of-tolerance. One strategy for reducing the number of false accept occurrences is to tighten the acceptance limits. Doing so, however, increases the frequency of false reject occurrences; that is, of devices observed to be out-of-tolerance that are actually in-tolerance. Both false accept occurrences and false reject occurrences have financial consequences, and it is worth noting that eliminating false accept occurrences at the expense of false reject occurrences does not always yield the best economic outcome.

The number of devices in the false accept regions relative to the number of devices in the entire population represents the risk of incorrectly stating in-tolerance status. This is unconditional false accept risk. Viewed in this way, unconditional false accept risk describes the likelihood of observing a device as in-tolerance when actually, it is out-of-tolerance. As a practical application, considering calibration as a process with selected acceptance limits and known measurement uncertainty and applying it to a specific population of devices produces a predictable number of false accept occurrences over time. In this case, acceptance limits are process control limits.

It is not possible to identify, with certainty, a device incorrectly deemed as intolerance. However, devices observed near the tolerance limit have a higher probability of being truly out-of-tolerance than devices observed well within the tolerance limits.

The likelihood that a specific device is truly out-of-tolerance, given a measured value, is conditional false accept risk. In this case, the attribute upon which risk is conditioned is the measured value. Conditional risk is a function of the tolerance limits, the calibration process uncertainty and the distribution representing the device population (See reference [8] for more information.). Figure 3 illustrates a set of devices having approximately the same measured value, m1. As shown above, it is possible to observe the same measured value for a set of devices with a range of true values. Assuming the measured value is within the acceptance limit, devices with true values outside the tolerance limit represent false accept occurrences. However, for a measured value further from the acceptance limit, m2, the likelihood of an out-of-tolerance true value is very small. If desired, it is possible to determine the risk of false accept for an individual device based on an observed measured value. It is also possible to set acceptance limits to contain the false accept risk for any given device within a desired level. For methods to determine conditional false accept risk and to set acceptance limits to limit conditional false accept risk, see reference [7], Appendix A, Method 4.

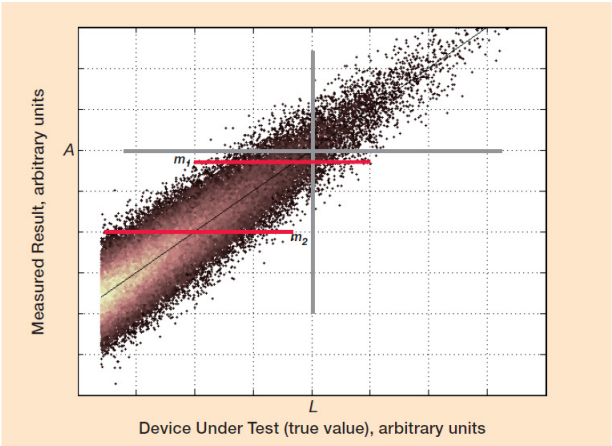

Fig. 3. Conditional risk for tolerance limit, L, and acceptance limit, A.

Managing false accept risk, either for calibration as a process for a population, or for individual devices, represents two distinct approaches to risk management. The acceptance limits for either approach can be significantly different. The choice of which approach to take may vary by application and is influenced by accreditation body requirements, quality management requirements, and historical tendencies. Either approach is viable when

considering ISO/IEC 17025 or ANSI/NCSL Z540.3 compliance. The approach to risk management also influences the language for statements of conformance.

With risk management for a population of devices, the acceptance limits represent process control limits. One purpose of the acceptance limits is to make possible statements of conformance. Performing a calibration results in a device declared either in-tolerance or out-of-tolerance. The declaration is within the context of a calibration process with either explicitly known, or maximum controlled false accept and false reject risk. Some compliance methods in reference [7] result in the determination of the actual PFA, while others provide a methodology to limit PFA to ≤ 2%. Of course, the level of risk is a function of the tolerance and acceptance limits, measurement error (uncertainty), and the device population.

With risk management for individual devices, it is common to state conformance for measured values extended by the uncertainty at a 95% level of confidence. While level of confidence is distinctively different6 from false accept or false reject risk, employing measurement uncertainty in this way is an effective approach to manage risk for individual devices. For a definition of "level of confidence" see reference [6], Section 6.2.2. For additional information related to "level of confidence" and false accept risk, see reference [9]. ILAC-G8 describes stating conformance considering measurement uncertainty and an associated 95% coverage probability. Conformance is stated as either in-tolerance or out-of-tolerance, but if the uncertainty interval about the measured value extends beyond the tolerance limit, a statement of conformance cannot be made at the 95% level of confidence. In that case, calibration produces a third outcome, which is neither in-tolerance, nor out-of-tolerance based on a 95% level of confidence.

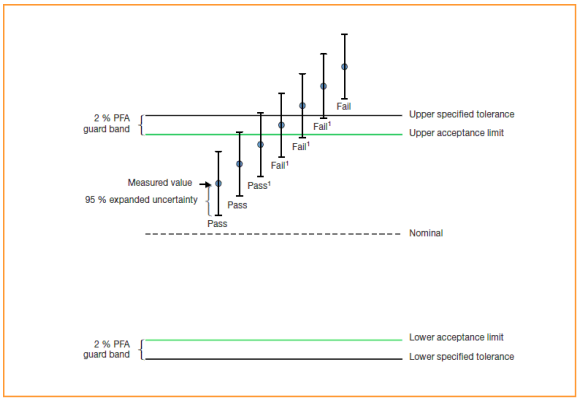

Fig. 4. Acceptance limits for ≤ 2% probability of false accept guard band (ANSI/NCSL Z540.3).

3. STATEMENTS OF PASS OR FAIL CONFORMANCE

Despite different risk management approaches, a simple method for expressing statements of Pass or Fail conformance, which also meets the ILAC-G8 reporting guidelines on assessment of conformance, considering that statements of Pass or Fail conformance accompany records of measured values, uncertainties, acceptance limits, and tolerance limits (See Appendix A for an example), follows:

1. Define the acceptance limit based on application requirements, accreditation body requirements, quality management requirements and/or other criteria.
2. Assign Pass or Fail status by comparing all measured points to the acceptance limits.
3. Note the 95% expanded uncertainty associated with the measured value.
   a. Annotate those points already assigned a Pass status (e.g. Pass[1]), where the 95% expanded measurement uncertainty extends outside the tolerance limit.
   b. Annotate those points already assigned a Fail status (e.g. Fail[1]), where the 95% expanded measurement uncertainty extends inside the tolerance limit.

The acceptance limits simply define the boundary for making pass or fail decisions. This allows for flexibility when setting the value for the acceptance limits. The Pass[1], Fail[1] annotation provides additional information helpful for managing conditional risk associated with a particular measured point.

Meeting the ≤ 2% probability of false accept (PFA) requirement of ANSI/NCSL Z540.3 may require the use of an acceptance limit different from the tolerance limit. The (guard band) difference between the tolerance limit and the acceptance limit is typically a fraction of the 95% expanded uncertainty (For more information on ANSI/NCSL Z540.3 compliance methods, see reference [7]). Figure 4 illustrates a typical scenario. Note the 3rd point from the left: it passes because the measured value is less than the upper acceptance limit. It is denoted "Pass[1]" because a portion of the 95% expanded uncertainty exceeds the upper specified tolerance. The 4th, 5th, and 6th points from the left exceed the acceptance limit, so they fail. However, a portion of the 95% expanded measurement uncertainty is within the specified tolerance, so each point is annotated "Fail[1]." Users need to perform an end-item impact analysis for any measured result denoted as Fail or Fail[1] in an "As received report." However, the likely negative impact and corresponding urgency is lower for a Fail[1] than a Fail.

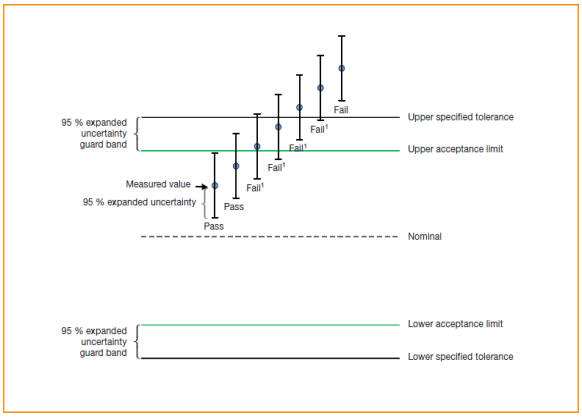

Fig. 5. Acceptance limits set using the 95 % expanded uncertainty.

ISO/IEC 17025 provides no specific guidance for taking the measurement uncertainty into account when assigning Pass/Fail status. The ≤ 2% PFA requirement represents a convenient unconditional risk threshold for managing a

population of instruments that meets not only ANSI/NCSL Z540.3, but also ISO/IEC 17025. Thus, many laboratories could use the limits employed in Figure 4 to comply with both standards simultaneously.

In Fig. 5 the guard band is set to the 95% expanded measurement uncertainty. The resulting conformance states are shown. Note that there is no Pass[1] state for this choice of guard band. Figure 6 shows acceptance limits equal to the tolerance limits. This choice of limits is appropriate for calibrations compliant with EURAMET/cg-15/v.01 ("Guidelines on the Calibration of Digital Multimeters"). From the standard: "Subsequent to calibration and under normal conditions of use, the uncertainty associated with the readings of a DMM will be the combination of the DMM's specification and the calibration uncertainty."

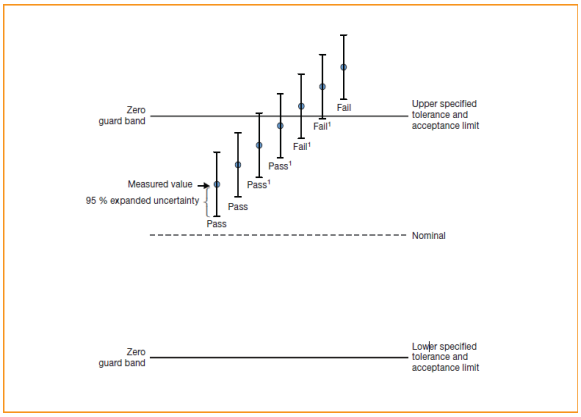

Fig. 6. Acceptance limits equal to the tolerance limits.

4. CONCLUSION

The simple method for expressing statements of Pass/Fail conformance provides key information for managing conditional and unconditional risk and meets the requirements of ANSI/NCSL Z540.3-2006, ISO/IEC 17025:2005, ILACG8:1996 and EURAMET/cg-15/v.01.

AKNOWLEDMENTS

The authors acknowledge the important contributions of Bruce Krueger, Ed Dempsey, Alan Dietrich, Jean-Claude Krynicki, and John Wilson (all of Agilent Technologies) to the work that is reported in this paper.

REFERENCES

[1] "Calibration Laboratories and Measuring and Test Equipment – General Requirements," ANSI/NCSL Z540-1-1994, National Conference of Standard Laboratories International, Boulder, CO, 1994.

[2] "Requirements for the Calibration of Measuring and Test Equipment," ANSI/NCSL Z540.3-2006, National Conference of Standard Laboratories International, Boulder, CO, 2006.

[3] "General requirements for the competence of testing and calibration laboratories," ISO/IEC 17025:2005, International Organization for Standardization, Geneva, Switzerland, 2005.

[4] "Guidelines of the Assessment and Reporting of Compliance with Specification," ILAC-G8:1996, International Laboratory Accreditation Cooperation, 1996.

[5] "Guidelines on the Calibration of Digital Multimeters," Calibration Guide EURAMET/cg-15/v.01, European Association of National Metrology Institutes, July 2007.

[6] "Guide to the Expression of Uncertainty in Measurement (GUM)," BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, International Standards Organization (ISO), Geneva, Switzerland, 1995, section 6.2.2.

[7] "Handbook for the Application of ANSI/NCSL Z540.3-2006 – Requirements for the Calibration of Measuring and Test Equipment," National Conference of Standard Laboratories International, Boulder, CO, USA, 2009.

[8] M. Dobbert, "Understanding Measurement Risk," Proceedings of NCSL International Workshop and Symposium, St. Paul, MN, August 2007.

[9] M. Kuster, "Tee Up with Confidence Revisiting t-Distribution-Based Confidence Intervals," Proceedings of NCSL International Workshop and Symposium, Orlando, FL, 2008.

APPENDIX A: ANSI/NCSL Z540.3 Example Measurement Results Table

In a table of measurement results, calibration customers have a clear preference for being able to view all relevant info rmation in one horizontal row. The common format Agilent Technologies has adopted includes the specification (tolerance), the measured result, the acceptance limits employed, the 95% expanded measurement uncertainty, and the Pass/Fail status.

Of course, even with a common format, a single table style does not fit all situations. A particular table style needs to address:

1. Is the specification expressed as the difference from an expected value or as a measured value?

2. Is the specification single sided (e.g., > 5 dBm) or double sided (e.g., 5 dBm ± 0.4 dB)?

3. Is the specification symmetrical or asymmetrical?

Table A-1 is an example of a symmetrical specification expressed as the difference from the expected value. This example is for a ANSI/NCSL Z540.3 measurement report where the managed guard band compliance method is employed (Method #6 in reference [7]).

Table A1. Sample ANSI/NCSL Z540.3 table for a symmetrical specification expressed as a measured difference from an expected value.

Power Level Accuracy (Software Revision A.1.14) Measured Diff = Measured – Expected						
Frequency	**Amplitude (Expected)**	**Specification**	**Measured Difference**	**Acceptance Limit**	**Measured Uncertainty**	**Status**
1 GHz	–60 dBm	±1.00 dB	+0.60 dB	±0.91 dB	0.40 dB	Pass
2 GHz	–60 dBm	±1.00 dB	+0.80 dB	±0.91 dB	0.40 dB	Pass[1]
3 GHz	–60 dBm	±1.00 dB	+1.00 dB	±0.91 dB	0.40 dB	Fail[1]
4 GHz	–60 dBm	±1.00 dB	+1.30 dB	±0.91 dB	0.40 dB	Fail[1]
5 GHz	–60 dBm	±1.00 dB	+1.50 dB	±0.91 dB	0.40 dB	Fail