



INTRODUCTION

Calibration Interval Analysis

MudCats Engineering Tools focus on Measurement Uncertainty and analysis of parametric data. Both are essential elements of quality and statistical foundations.

Calculation and reporting of Measurement Uncertainty in compliance with ANSI/ISO/IEC 17025 is becoming mandatory for today's calibration laboratories. Section 5.4.6.2 Estimation of uncertainty of measurement states in part: "Testing laboratories shall have and apply procedures for estimating uncertainty of measurement. In certain cases the nature of the test method may preclude rigorous, metrologically and statistically valid, calculation of uncertainty of measurement. In these cases, the laboratory shall at least attempt to identify all the components of uncertainty and make a reasonable estimation. ..."

MudCats reduces the overwhelming task of complying with today's stringent requirements for evaluating and reporting of measurement uncertainty. MUDCATS automatically calculates each test point's Test Accuracy Ratio (TAR) and Estimated Measurement Uncertainty (EMU) when a calibration standard(s) is assigned to the datasheet. Utilizing the SpecTrack specification library, MUDCATS effortlessly interprets Units of Measure and synchronizes the UUT and Standards test parameters, recalculating them to account for what typically is the most prevalent source of error. Estimates of measurement uncertainty and test accuracy ratio can be quickly accessed using guidelines from ANSI/NCSL Z540-2-1997, U.S. Guide to the Expression of Uncertainty in Measurement (GUM), Type B methods, and ANSI/ISO/IEC 17025, Section 5.4.6 Estimation of Uncertainty of Measurement, and optionally reported to the customer.

Instrument failures can be analyzed utilizing the Instrument Range Utilization Report. This report enables the user to automatically evaluate out-of-tolerance conditions by tracking the function, range and test point information between each standard utilized and the UUT. Traceability is maintained from the specific standard's function and range through the UUT-specific test point. The report identifies each UUT and test point(s) affected by the standard's failed range.

Instrument trends can be analyzed using the historical data plotter. This invaluable tool enables Metrology Engineers to view parametric As Found and As Left calibration test point data by Asset ID or datasheet family within a given time frame. Out-of-Tolerance trends can be readily identified and adjustments to instrument specifications or intervals can be recommended with confidence.

Method A3 - Interval Test Method Module

Ultimately, drift analysis performed on historical parametric data in order to determine optimum calibration intervals is not only a highly desirable engineering tool but also one that provides potentially great cost savings to the organization.

While other interval methods (the Incremental Response Method in particular) allow an item to find its correct interval and remain there, it does not do so expeditiously. Often, periods as long as five to sixty



7300 Fenwick Lane, Westminster, CA 92683
Phone 714 895-0476 Fax 714 895-0781
Toll Free 1-866-723-2257
www.edisonMudCats.com

years are required to reach intervals commensurate with established reliability targets. In addition, although the correct interval is reached eventually, considerable flopping around is done in the process.

A reactive method that both attains correct intervals in reasonable periods of time and does not produce spasmodic interval fluctuations is the Interval Test Method. In this method, intervals are adjusted only if recent accumulated calibration results are not consistent with expectations. This consistency is evaluated statistically.

Method A3 - Interval Test Method

Method A3 employs accumulated calibration history for a given item to test whether the item's assigned interval is appropriate. The test is statistical in nature. The result of a test is whether to adjust or not based on whether or not calibration results are consistent with expectations. For instance, in-tolerance events are expected if an interval is believed to be associated with a high reliability target. In this case, an in-tolerance event is not likely to trigger an interval change. Method A3 bases adjustments on statistically significant results.

Interval Change Criteria

In Method A3, if the percentage of calibrations observed in-tolerance at a given interval proves to be significantly different from the desired reliability, an interval change is called for. Interval changes may require extrapolation or interpolation. Method A3 employs an exponential reliability model to extrapolate intervals when extrapolation is advisable and a numerical bisection method to interpolate intervals when interpolation is advisable. This method leads to intervals that are distribution-free and can be highly accurate under certain conditions.

Significant Differences

Since the occurrence of an in- or out-of-tolerance condition is a random event, it is not advisable to adjust calibration intervals in response to a single in- or out-of-tolerance conditions. Under certain circumstances, it may not even be advisable to adjust intervals in response to the occurrence of two or even three or more successive in- or out-of-tolerance conditions. Given the specific reliability target and the number of observed calibrations, it may be that such combinations of events are expected to occur fairly frequently at the correct interval. Whether to adjust a calibration interval or not depends on whether in- or out-of-tolerance events occur in a way that is highly unlikely, i.e., in a way that is not consistent with the assumption that the interval is correct.

Method A3 uses a statistical test to evaluate whether calibration results are consistent with a correct interval. If the test shows that the observed measurement reliability is significantly different from the target reliability, then an interval change is required. That is, if the observed measurement reliability is found to be significantly higher or lower than the reliability target, the interval is lengthened or shortened. What is meant by "significantly higher" or "significantly lower" is that the observed rate of occurrence of out-of-tolerance events causes a rejection of the notion that the calibration interval is correct. This rejection is made with a predetermined level of statistical significance. Hence, the use of the term "significant."

For example, suppose that all interested parties have agreed to reject a calibration interval if the observed out-of-tolerance behavior had less than a 30% chance of occurring if the interval were correct. Another way of saying this is that the calibration interval would not be adjusted (up or down) unless the out-of-tolerance rate observed at an interval fell outside statistical 70% confidence limits.

For example, if the reliability target were equal to 80%, interval increases would be recommended with 70% confidence if eight out of eight, nine out of nine, ten out of ten or ten out of eleven calibrations were



7300 Fenwick Lane, Westminster, CA 92683
Phone 714 895-0476 Fax 714 895-0781
Toll Free 1-866-723-2257
www.edisonMudCats.com

in-tolerance. From the foregoing, it can be appreciated that, with Method **A3**, a key objective is obtaining reliable initial interval estimates. With high reliability targets (e.g., 80% or higher), it takes a considerable number of calibrations before an interval can be justifiably extended on the grounds that it yields an in-tolerance percentage that is significantly different from the reliability target.

Combining calibration results from a grouping of individual items can reduce the time required to obtain sufficient numbers of calibrations for making interval increase or decrease decisions. In combining data in this way, it is important to bear in mind that what is being statistically tested is a particular calibration interval. *This means that, to apply Method A3 to a group of items, all items must be on the same calibration interval.* Another equally important requirement is that the grouping be homogeneous with respect to physical characteristics, usage, operating environment, tolerance limits, etc.

Considerations for Use

Since Method A3 concerns itself with testing a current assigned interval, all items contributing data to the statistical significance test described above must be on the same interval. This could restrict the use of the method in some organizations to application at the individual serial number level. If applied at the serial number level, it may take so long to accumulate enough data to justify an interval change that historically older data are not homogeneous with recent data. This would be the case if the stability of an item were to change as the item aged. If so, the older data should be excluded on the grounds that it is no longer relevant to the statistical test. The upshot of this is that, for items whose stability changes over a period of less than ten or twenty calibration intervals, there may *never* be enough representative data to justify an interval change.

Another consideration in the use of Method A3, though not unique to it, is that data taken prior to any change that bears on an item's in-tolerance probability cannot be used to evaluate the current interval. Such a change might be a calibration procedure revision or a modification of tolerance limits. Whatever the variable, the behavior of an item prior to the change may not be relevant to the item's current situation. For example, suppose that an item's tolerance limits are cut in half. Clearly, with half the original tolerances, it could require substantially less time for the item to drift out-of-tolerance than it did prior to the change. Thus, if the item's prior history consists of a string of in-tolerance observations, these observations cannot be taken to have any relevance to current tendencies for in- or out-of-tolerance. It may be that, with the new limits, a string of out-of-tolerances are on the horizon, even if the current interval is maintained. Under these circumstances, if the current interval is lengthened on the strength of past behavior, the likelihood for out-of-tolerances may increase dramatically.

Criteria for Use

Given the foregoing, Method A3 should be used only if the following criteria are met:

- Data used to test a given calibration interval consists of calibration results taken at the end of the current interval.
- Data used to test a given calibration interval of an item are homogenous with respect to the current use equal to (or nearly equal to) the interval in question, and stability of the item.
- Data used to test a given calibration interval are homogenous with respect to calibration procedure, tolerance limits, and other variables that impact measurement reliability over time.

Method A3 Pros and Cons

Pros

- Method A3 adjusts intervals to meet specified reliability targets.



7300 Fenwick Lane, Westminster, CA 92683
Phone 714 895-0476 Fax 714 895-0781
Toll Free 1-866-723-2257
www.edisonMudCats.com

- Method A3 is resistant to spurious interval changes. Intervals are adjusted only if a change is justified on the basis of statistical significance.
- Compared to statistical predictive methods, the design and implementation of Method A3 is inexpensive.
- Operating costs are low.

General Comment:

Method A3 provides most of the advantages of statistical predictive methods at a fraction of the development cost of other such methods.

Cons

Method A3 suffers from the following drawbacks:

- Except for interval extrapolation, method A3 makes no attempt to model underlying uncertainty growth mechanisms. Consequently, if an interval change is required, the appropriate magnitude of the change cannot be accurately determined.
- If initial intervals are grossly incorrect, Method A3 may require substantial time to arrive at correct intervals.



7300 Fenwick Lane, Westminster, CA 92683
Phone 714 895-0476 Fax 714 895-0781
Toll Free 1-866-723-2257
www.edisonMudCats.com