General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

Measurement Assurance for End-Item Users

Speaker/Author:
Scott M. Mimbs
National Aeronautics and Space Administration (NASA)
John F. Kennedy Space Center
Kennedy Space Center, FL 32899
Phone: 321-861-5184
Fax: 321-867-1740
scott.m.mimbs@nasa.gov

Abstract

The goal of a Quality Management System (QMS) as specified in ISO 9001 and AS9100 is to assure the end product meets specifications and customer requirements. Measuring devices, often called measuring and test equipments (MTE), provide the evidence of product conformity to the prescribed requirements. Therefore the processes which employ MTE can become a weak link to the overall QMS if proper attention is not given to development and execution of these processes. Traditionally, calibration of MTE is given more focus in industry standards and process control efforts than the equally important proper usage of the same equipment. It is a common complaint of calibration laboratory personnel that MTE users are only interested in "a sticker."

If the QMS requires the MTE "to demonstrate conformity of the product," then the quality of the measurement process must be adequate for the task. This leads to an ad hoc definition; measurement assurance is a discipline that assures that all processes, activities, environments, standards, and procedures involved in making a measurement produce a result that can be rigorously evaluated for validity and accuracy. To evaluate that the existing measurement processes are providing an adequate level of quality to support the decisions based upon this measurement data, an understanding of measurement assurance basics is essential.

This topic is complimentary to the calibration standard, ANSI/NCSL Z540.3-2006, which targets the calibration of MTE at the organizational level. This paper will discuss general measurement assurance when MTE is used to provide evidence of product conformity, therefore the target audience of this paper is end item users of MTE. A central focus of the paper will be the verification of tolerances and the associated risks, so calibration professionals may find the paper useful in communication with their customers, MTE users.

Introduction

The new calibration program standard, ANSI/NCSL Z540.3-2006 [1], places its emphasis on organizations that use measuring and test equipment (MTE) and not solely on the calibration service provider (in traditional terms, the calibration lab). Combined with the Quality Management requirements of ISO 9001, end-item users and calibration providers will need to increase their communication. There are volumes of literature written to assist calibration providers, but with this subtle shift found in Z540.3, more literature for

the end-item user will now be required. More than in the past, MTE end-item users will want to understand the nuances of metrology as it relates to their specific circumstances. For the purpose of this paper, an MTE end-item user is anyone using calibrated MTE that is not a part of a calibration service.

In all cases regarding measurements, including calibration, two questions should be posed prior to making a measurement.

1. How good does the measurement need to be?
2. How good can the measurement be made?

Many factors can affect the answer to these questions. These factors can range from available funds to available technology. In most cases, the first question is answered in respect to a business case balancing between cost and quality, while the second question involves technology and process.

The idea to "begin with the end in mind" from Stephen R. Covey's [2] book on personal growth, is very applicable to the first question. The quality of a measurement depends on the reason for the measurement. In the simplest terms, measurements are made to either learn something or to make a decision. For example, measurements can be made in basic research to extend our knowledge, while they can also be used to check the conformance of a product prior to acceptance. By focusing first on the reason the measurement is required, the "how good is good enough" becomes much clearer.

The answer to the second question requires an understanding of how to evaluate the "goodness" of a measurement. There are many factors which influence a measurement result. These factors have to be understood, evaluated, and controlled to achieve measurement results that meet the requirements established in the first question. A "better" instrument does not always guarantee a better measurement.

This paper will provide a high-level overview of measurement assurance from the perspective of the MTE end-item user. The discussion will cover the fundamentals and associated risks of measurements. The paper also provides a discussion concerning methods for verifying conformance to specified tolerances. This paper provides an overview of the subject and not an in-depth coverage of measurement assurance. The intent is to provide an introduction for MTE end-item users (i.e., calibration customers) and hopefully open the communication with calibration providers which is necessary for the implementation of Z540.3

Metrology's Two Fundamental Tenets

Metrology's fundamental tenets are the first step in understanding measurement assurance. It is the rigorous application of these two tenets that characterizes the success of any good, cost-effective measurement assurance program.

- Traceability
- Measurement Uncertainty

Traceability

Traceability establishes the link for a given measurement to the national or international standard for that unit of measure. Prior to the establishment of international agreements on standardized units of measure, this was the largest hurdle to international commerce and scientific cooperation. Today, the Bureau International des Poids Mesures (BIPM) is the international organization charged with maintaining the International System of Units (SI) reference standards, which are the global references for measurement units.

Traceability is accomplished through an unbroken series of competent and documented calibrations. It relates the measurement result to the SI unit, which in turn provides the ability to compare measurements (and the related decisions) across an organization as well as international borders. Documented calibration is important to the quality of a measurement, especially in areas of research, testing and product acceptance. Figure 1, taken from the BIPM website, illustrates this "chain of traceability."

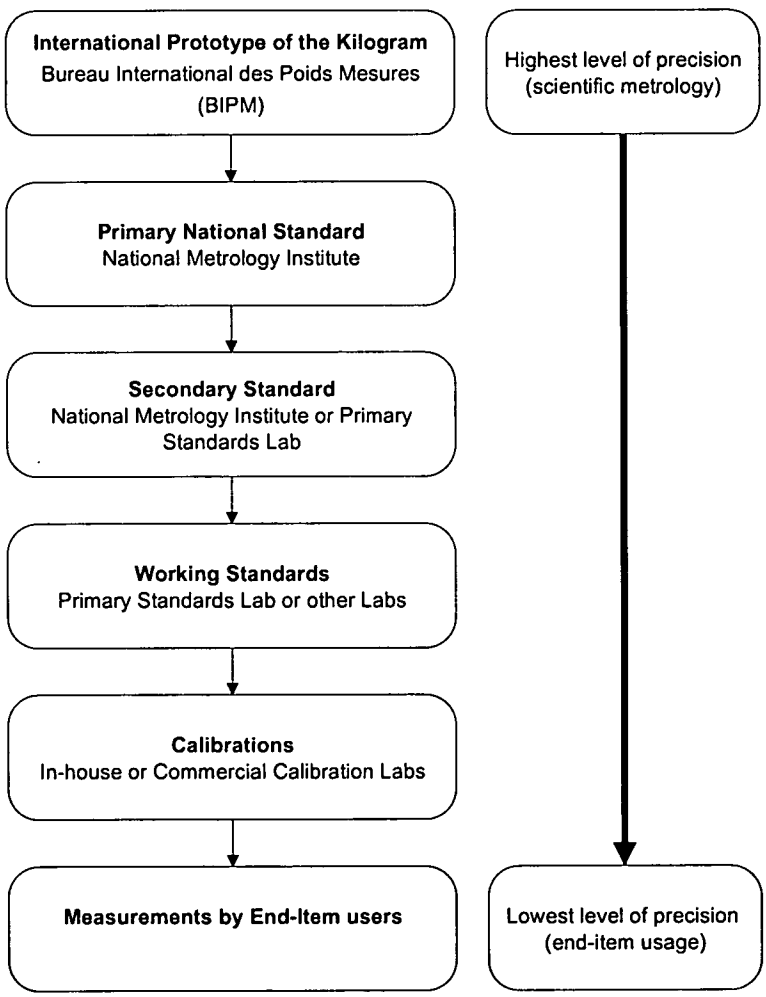

Figure 1: Traceability flow-down for a kilogram, based on the BIPM [3].

Measurement Uncertainty

Uncertainty in measurements is the second fundamental tenet of metrology. This is how the "goodness" of a measurement can be evaluated.

Measurement uncertainty is the estimate of the error in a measurement that results from the equipment, the processes, and other sources, including the originating international standard, as well as the quality of the traceability chain. Every element within the measurement process contributes errors to the measurement result. Evaluating the measurement uncertainty provides an estimate of the quality of the measurement [4]. Not all tasks require the same level of quality, thus measurement uncertainty provides the tool to adjust measurement processes according to technical as well as business needs. The typical components that make up measurement uncertainty are summarized below: [4]

> *Measurement process errors are the basic elements of uncertainty analysis. The errors most often encountered in making measurements include, but are not limited to the following:*
>
> - *Measurement Bias*
> - *Repeatability Error*
> - *Resolution Error*
> - *Digital Sampling Error*
> - *Computation Error*
> - *Operator Bias*
> - *Environmental Factors Error*
> - *Stress Response Error*

For many years, uncertainty estimation had the same problem as units of measure – a lack of standardization. The lack of standardization has caused (and still causes) disagreements and confusion. After almost two decades of work, an international consensus standard was developed. The International Organization for Standardization (ISO) *Guide to the Expression of Uncertainty in Measurement* (GUM) [5] provides a standardized approach to estimating uncertainty. ANSI/NCSL Z540.2-1997 (R2007), *U.S. Guide to the Expression of Uncertainty in Measurement* (U.S. Guide) [6], is the U.S. adoption of the ISO GUM. Additional guidance on estimating measurement uncertainty can be found in many discipline-centered voluntary consensus standards and complimentary documents. BUT, for consistent results within an organization, among organizations, or across national economies, it is imperative that the uncertainty analysis is based on the ISO GUM.

Evaluating Measurement Results

Evaluating measurement results combines the two questions proposed in the Introduction. In almost all cases, the actual value of the item being measured varies from the value observed on the instrument used in the measurement process. Understanding how this variance impacts the functionality of the end-item answers the question, "How good does the measurement need to be." Knowing the size of the variance answers the question, "How good can the measurement be made." Measurement uncertainty is the best way to

evaluate this variance. It provides an estimate of the range of possible values an observed measurement may be from the actual value. This determines the usefulness of the measurement process. The amount of effort given to the uncertainty evaluation should be determined by how important the measurement result is to the end-item's functionality. It should be noted that in most cases, there is an insignificant difference in the amount of effort required to perform a thorough versus a minimal measurement uncertainty evaluation.

Figure 2 illustrates three unique measurements processes with a different level of uncertainty for each as indicated by the error bars. When three different measurement processes are used to measure the same item, it is reasonable to expect three different answers. Although Figure 2 shows all three processes with the same nominal value, in an actual situation, this would be a rare occurrence. The intent of Figure 2 is to illustrate the large differences that are possible between measurement processes. An example would be the difference between using a steel machinist rule, dial caliper, and micrometer to make the same measurement.

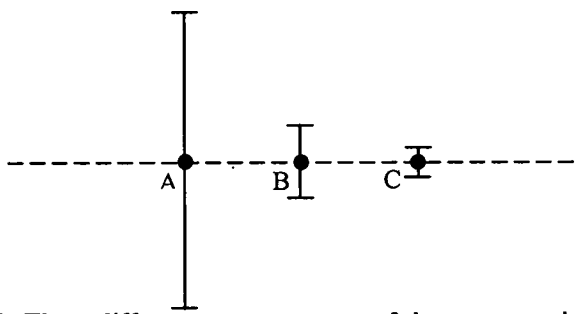

Figure 2: Three different measurements of the same nominal value with varying measurement uncertainty.

Figure 3 shows a more common occurrence when measuring the same item with three different measurement processes. Each of the three processes provides a different indication of the same value. The actual value is contained within each measurement's process uncertainty range. In the case of Figure 3, one may ask which is correct. The answer is all three are correct within the capability of the measurement process. A better question is which measurement process will support the objectives of the measurement requirement.

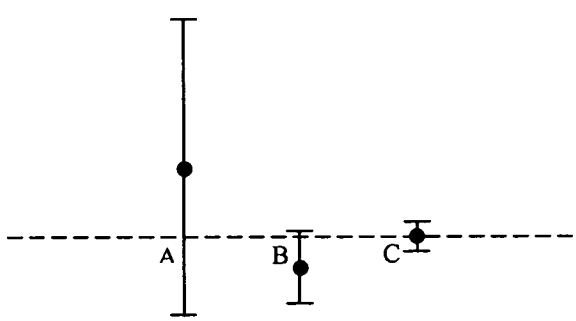

Figure 3: Three different measurement indications with the nominal value contained by each measurement process uncertainty range.

The key to evaluating any measurement result is an understanding of the relationship between the measurement and the end results. In other words knowing how "good is good enough."

Verifying Tolerances

A requirement to "demonstrate conformity of the product" is another way of saying "verify a specified tolerance." Although the topic dates back into the antiquity of human endeavors, there has been an increasing amount of discussion on this subject since the debut of the ISO 9000 series of quality standards in the early 1990's. More and more companies around the world are adopting these standards which call for a demonstration of conformity of product. Even NASA has adopted the AS9100 [8] for the quality management of its mission critical space launch systems. Striving for quality products is not new, but now the contractual requirement to "demonstrate conformity" meets head-on the technical problem of proving it. Demonstration of conformity is not always straightforward.

Figure 4 visually illustrates the problem with proving conformity to a specification. Figure 4 uses the same three measurement processes from Figures 2 and 3 and adds a tolerance limit of ± L to contain the specified nominal measurement. In Figure 4, all three measurements indicate the same value which is off-nominal. A portion of measurement A's uncertainty range extends beyond the + L limit, which means there is some probability that the true value as indicated by measurement A is in reality, outside of the limit.

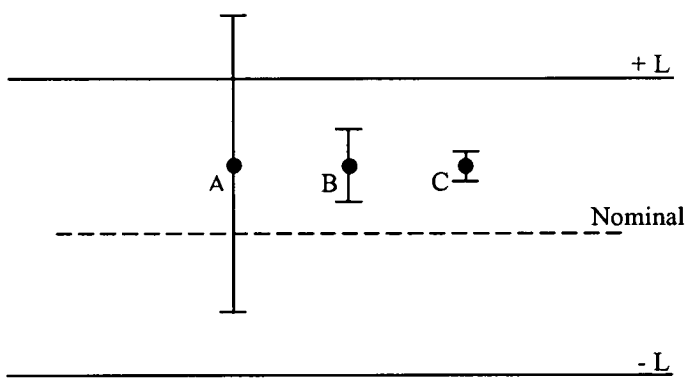

Figure 4: Three different measurement processes are used to verify a nominal value with a tolerance of ± L. Although all three have the same indication, measurement A has some probability of non-conformance.

Although no measurement system can provide 100% assurance of conformity, the probability of non-conformance can be greatly reduced using measurement assurance best practices. In the end, though, how much "proof" is required must be an agreement between the producer and the customer. There are many National and International standards, discipline-centered guidance documents, conference papers, and other literature that can help develop conformance strategies and agreements, but the underlying foundation for the best ones is found in the fundamental tenets of metrology. Without traceability and measurement uncertainty, assurance of conformity can become a risky gamble or impossible.

This section provides a discussion covering different aspects of using measurement results to demonstrate conformity of products. This will include a discussion on one the most common methods used for verifying a tolerance in the United States.

Risks associated with measurement

The best starting point for this discussion is to define the risks associated when measurements are used to verify a tolerance. The bottom-line is the decisions which are based on these measurements. Measurement assurance procedures, policies, and techniques are devised to control the risks to decisions which are associated with measurements. Thus, measurement decision risk is the probability of an incorrect decision based on a measurement. Measurement decision risk is affected by all parts of the metrology process from initial design parameters, to final measurement processes, to an instrument's calibration.

There are two basic types of risk when making a measurement;

1. **False Accept Risk** is the probability of an *out-of-tolerance* item or parameter being accepted by test or calibration. It is also known as a pass error, Consumer Risk, or Type 2 error.

   Depending on the criticality of the measurement, this type of error can lead to loss of mission or loss of life, reduced end-item function or capacity, damaged corporate reputation, warranty expenses, shipping and associated costs for returned items, loss of future sales, punitive damages, legal fees, etc [7].

2. **False Reject Risk** is the probability of an *in-tolerance* item or parameter being rejected by testing or calibration. It is also known as fail errors, Producer Risk, or Type 1 error.

   False rejects in test and calibration processes lead to unnecessary rework and handling. For test processes, higher rejection rates imply poorer production controls, thus false rejects also create an excessively pessimistic view of the quality of the end-item production process. This view may lead to more frequent disassembly and repair of end items than is necessary [7].

Both false accept risk and false reject risk are functions of the uncertainty in the measurement process. Jointly, they comprise measurement decision risk, which is a key element and metric of measurement assurance.

Ratio-Based Verification

One of the oldest and most common methods of tolerance testing involves a comparison of the tolerance range to either the instrument accuracy or the measurement process uncertainty. Originally called accuracy ratios, these methods were "rules of thumb" that were not always clearly linked to an engineering or scientific basis. The most common accuracy ratio was where the measuring instrument was required to be ten times more precise than the tolerance it was measuring. This was referred to a 10:1 rule. In the late 1940's and early 1950's, Alan Eagle, Frank Grubbs, and Helen Coon pioneered work on consumer and producer risk analysis [9, 10], which was a rigorous, statistical method of controlling measurement decision risk in manufacturing. Building upon Eagle's

published work, Jerry Hayes set out to establish a basis linking decision risks to accuracy ratios for application in the Navy's calibration program [11]. The target was to provide a reasonable rule without the statistical calculations which were very tedious with the limited computing power of the early 1950's [12]. Although developed for calibration, these "rules of thumb" were, and can still be, applicable to any areas where measurements are used to demonstrate conformity to a specified tolerance.

A ratio-based "rule of thumb" aims at providing some level of assurance that the measurement is capable of adequately verifying the tolerance. It may be perceived as a rudimentary rule that controls false accept and false reject risks, by attempting to assure the resulting measurement data will validate an in-tolerance condition of a specified tolerance. As a "rule of thumb," there are two major pitfalls which can undermine their usefulness.

The first major pitfall with using a "rule of thumb" is that it is a "one size fits all" approach. It may not be appropriate for all circumstances. In the case of life or mission critical measurements, detailed analysis of the measurement process should always be used. However, there is valid rationale for using the accuracy ratio methods in the majority of test process situations.

The second pitfall is a lack of standardization with using a "rule of thumb," which can lead to very different results and a false sense of security in some measurement processes. For the accuracy ratio rule, there are three prevalent, but very different methods for applying the rule:

1. **Resolution error method.** The lowest resolution (minor division) of the measuring instrument must be no greater than a given percent (such as 10%) of the overall tolerance of the article being measured.

   The first method is not very effective for measurement assurance. It accounts only for errors due to the resolution of the measuring instrument without consideration of other sources of uncertainty (or error) within the measurement process. Other sources such as readability, temperature, humidity, operator, and data collection methods may contribute more significantly to the process uncertainty than resolution alone

2. **Measuring instrument error method.** The tolerance (or stated accuracy) of the measuring instrument will be no greater than a given percent (such as 10%) of the tolerance of the article being measured.

   This approach is sometimes referred to as a Test Accuracy Ratio (TAR) method, although there is a lack of agreement as to how TAR should be defined. This application only looks at the accuracy specifications for the measuring instrument. Although better than the first approach, this method also lacks provision for the uncertainty due to measurement process errors, which can be more significant than the uncertainty in the error in the instrument.

3. **Measurement process uncertainty method.** The combination of the uncertainties (sometimes stated as "accuracies") of all the elements of the measurement process must be no greater than a given percent (such as 10%) of the overall tolerance of the article being measured.

This third method, which is sometimes referred to as a Test Uncertainty Ratio (TUR), does provide the means for effective measurement assurance. Until recently, there was no universal agreement in the metrology community on the definition of TUR. The Z540.3 [1] attempts to correct this deficiency by providing an explicit definition of TUR, which accounts for all relevant uncertainties:

3.11 Test uncertainty ratio
The ratio of the span of the tolerance of a measurement quantity subject to calibration, to twice the 95% expanded uncertainty of the measurement process used for calibration.

NOTE: This applies to two-sided tolerances.

The definition uses the expanded uncertainty as defined in Z540-2 where k is the coverage factor and u is the combined uncertainty of the measurement process. The definition in equation form:

$$\text{TUR} = \frac{\text{Upper} - \text{Lower}}{2 \cdot U} \qquad U = k \cdot u \qquad k = 2$$

Although the Z540.3 definition is found within a calibration standard, it represents a very useful tool for test processes and would provide the best and most realistic measurement assurance of the three examples discussed above.

Visualizing Accuracy-ratios

Figure 5, again, uses the same three measurement processes as in previous figures. This time the indicated measurements are positioned so that the extreme edge of the uncertainty range is approximately even with the tolerance limit. To put the figure in perspective, using the Z540.3 TUR definition, the approximate ratios are: measurement A = 1:1, measurement B = 4:1 and measurement C = 10:1. Figure 5 illustrates how close to the tolerance (± L) a measurement can be made and still have confidence the measurement is in-tolerance.

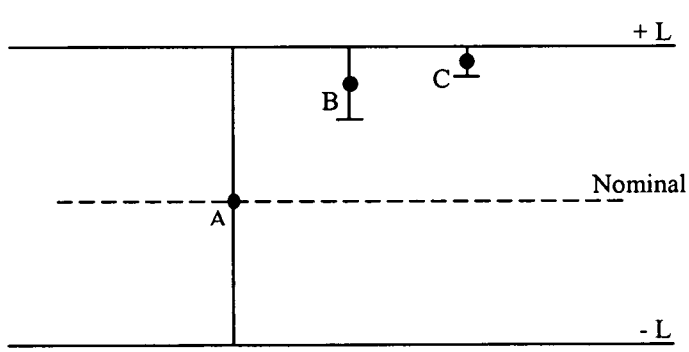

Figure 5: Three different measurement processes are used to verify a nominal value with a tolerance of ± L. Each measurement is shown at its capability limit prior to exceeding the tolerance limit.

It is interesting to note that measurement A, with its 1:1 TUR, only provides a minimal confidence at the nominal measurement, while the other two measurements can cover

much more of the tolerance range. In a conformance agreement, it would be difficult to defend using measurement A if other measurement processes were available.

Relationship of the TUR to False Accept Risk

Figure 5 showed that smaller uncertainty ratios (e.g., 1:1) are less effective on off-center measurements than larger ratios (e.g., 10:1). Figure 6 puts the same concept in terms of the out-of-tolerance risk incurred. Figure 6 illustrates that smaller ratios are less capable of demonstrating conformity over the full range of the specification. For example, a TUR of 1:1, measured at the nominal value will have approximately a 5% probability of being out-of-tolerance due to the large distribution of measurement uncertainty. For a TUR of 10:1, a majority of the tolerance can be utilized with a high level of confidence. Case in point is a measurement made at 86.5% of the specification range would only have a 0.3% probability of being out-of-tolerance. In other words, there is a 99.7% chance of being in-tolerance at that measurement point.

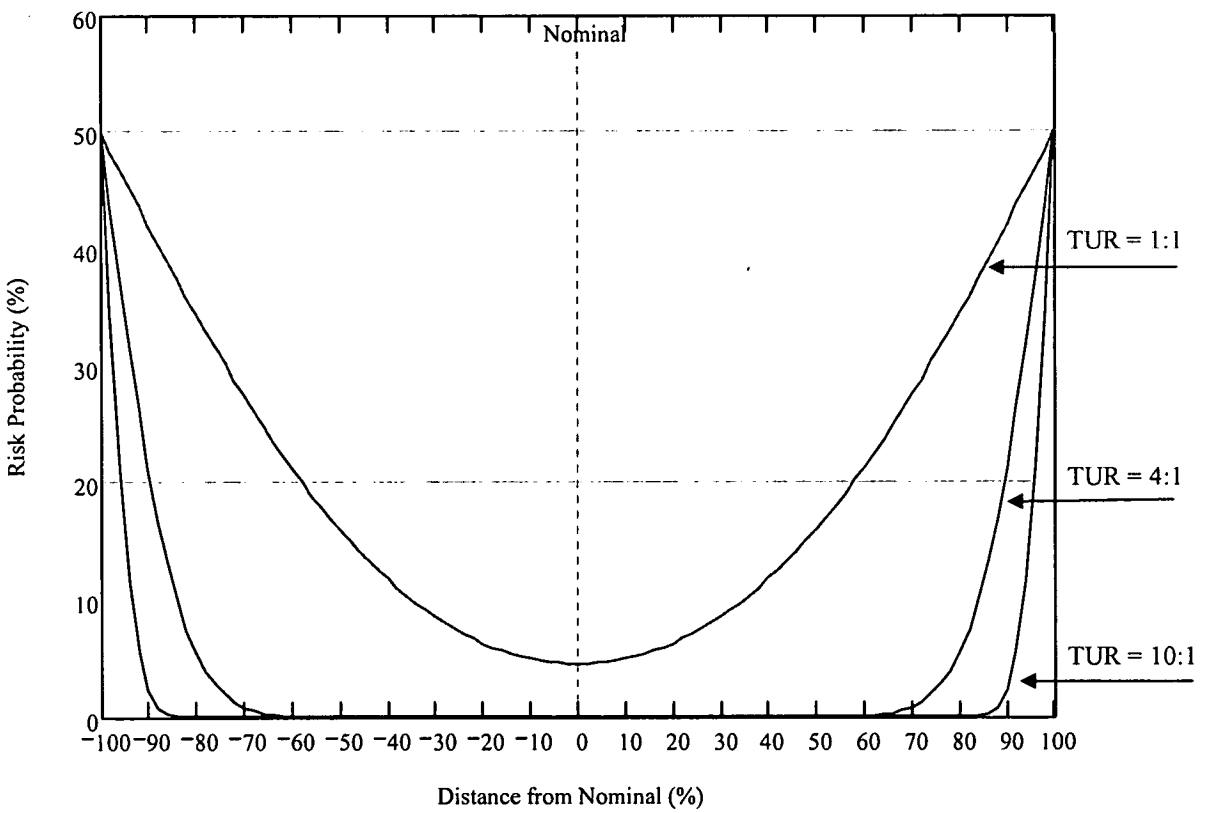

Figure 6: The risk probabilities for the three different measurement processes over the tolerance range of ± L. The probabilities are shown for each measurement process over the tolerance range as a percentage from the nominal. TURs are calculated per the Z540.3 definition.

One of the more important aspects of Figure 6 is that it shows the direct link of the measurement to the acceptance risk. In the case of end-item users that must develop agreements or conformity demonstration strategies, this can be an invaluable tool. Demonstration of conformity is much easier when it can be quantified.

Figure 6 links the false accept risk to the TUR as defined by the calibration standard Z540.3, but decision risk calculations can be made without the step of the TUR.

Considerations when applying measurement assurance rules

There are two important points that should be considered when developing or implementing measurement assurance rules. The first deals with the "rules of thumb," and the more important second deals with the relationship between the specification and its verification process.

A first consideration is the difference between the TAR and TUR. Although the Z540.3 TUR is defined in a calibration standard, it fits any situation where measurements are used for verifying a tolerance. The TUR differs from the TAR in the inclusion of all pertinent measurement process errors. Many times the process errors are larger than the instrument error by itself; therefore the TAR can only provide an "optimistic" view of the measurement quality.

The second consideration deals with how the specification is developed and the impact on its verification. Whereas it is fairly evident that an understanding of the measurement process is essential to estimating the uncertainty, it may not be as obvious that it may be just as important to understand how the specified tolerance was developed to provide proper measurement assurance. The basis for any ratio-based measurement assurance rule, whether it is 10% (10:1), 20% (5:1), or 25% (4:1), is the independence of the measurement process uncertainty and the specified tolerance. This is also applies to other verification methods, such as measurement decision risk analysis. Due to the relationship between the tolerance that is being verified and the measurement process providing the verification, an integrated approach to developing both provides the "best practice" to measurement assurance. This is true from a technical as well as cost effective perspective.

Design centers providing additional margin for the measurement process without proper documentation can lead to excessive costs to the program by "margin-stacking" if those responsible for verifying the tolerance are not aware of the additional margin and apply strict measurement assurance techniques such as the 10% rule. The converse can also be true when those responsible for verifying the tolerance believe that the measurement assurance margins are contained within the tolerance and in reality they are not. Without the application of appropriate measurement assurance techniques in this latter case, there is not only an increased measurement decision risk, but more importantly, there may also be increased safety or hazard risks if critical limits are unintentionally exceeded due to improper measurement assurance application.

Summary

A requirement to "demonstrate conformity of the product" may sound straightforward, but the implementation of this requirement can be challenging. The metrology concepts of traceability and measurement uncertainty (i.e., "the fundamental tenets") provide a solid foundation for developing methods that make "proving" conformity a lot easier. The goal of a QMS is to provide evidence of conformity, thus when measurement

assurance methods that are founded in measurement science are applied, costly disagreements between producers and customers can be avoided.

It is a key point to remember that measurements support decisions – accept, reject, rework, or scrap. The idea behind measurement assurance is to provide a level of confidence that the measurement will provide the proper data for the decision it supports. A measurement which reports that the tolerance is within specification when in reality it is not is called a False Accept, and when it indicates that it is outside the tolerance when in reality it is not is called a False Reject. The sum of these two establishes an indication of measurement decision risk. Ratio-based rules such as the 10% rule were derived from this concept and were intended to provide a simple to execute basis to control measurement decision risks and achieve appropriate measurement assurance in both testing and calibration processes. In-depth measurement decision risk analysis can even provide a higher level of confidence for more critical measurement applications.

Specification limits can complicate measurement assurance applications because not all specified tolerances are equal. Some may have additional margin, which means the actual risk could be lower than indicated by measurement assurance analysis or measurement assurance rules. Although this is true, it should not be considered in the measurement assurance planning unless there is engineering documentation stating otherwise.

Measurement assurance is not just about providing evidence of product conformity Decisions based on measurement data can be mundane or life critical. The criticality of the measurement carries the same criticality as the decision. Therefore, it is imperative that end-item uses understand and implement measurement assurance techniques that can be rigorously evaluated for validity and accuracy.

> *The more critical the decision, the more critical the data. The more critical the data, the more critical the measurement.*
>
> NASA Reference Publication 1342

References

1. *Requirements for the Calibration of Measuring and Test Equipment,* ANSI/NCSL Z540.3-2006, NCSL International, 2006
2. Covey, Stephen R., *The 7 Habits of Highly Effective People,* Simon & Schuster Inc., 1989
3. Bureau International des Poids Mesures, *Traceability*, February 17, 2008, http://www.bipm.org/en/convention/wmd/2004/figure.html
4. *Measurement Uncertainty Analysis: Principles and Methods*, KSC-UG-2809 Rev: Basic, NASA, November 2007
5. *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization (ISO), 1995
6. *American National Standard for Calibration – U.S. Guide to the Expression of Uncertainty in Measurement*, ANSI/NCSL Z540-2-1997 (R2007), NCSL International, 2002
7. *Metrology – Calibration and Measurement Processes Guidelines*, NASA Reference Publication 1342, NASA, 1994
8. *Quality Management Systems – Aerospace – Requirements*, AS9100, Rev. B, SAE Aerospace, 2004
9. Eagle, Alan R., "A Method for Handling Errors in Testing and Measuring," *Industrial Quality Control*, March 1954
10. Grubbs, Frank, and Coon, Helen, "On Setting Test Limits Relative to Specification Limits," *Industrial Quality Control*, March 1954
11. Hayes, Jerry L., *Factors Affecting Measuring Reliability,* U.S. Naval Ordnance Laboratory, TM No. 63-106, 24 March 1955
12. Mimbs, Scott, *Measurement Decision Risk – The Importance of Definitions,* Proc. NCSL International Workshop & Symposium, St. Paul, August 2007.

Additional related information

1. *Guidelines for Decision Rules: Considering Measurement Uncertainty in Determining Conformance to Specifications*, ASME B89.7.3.1, American Society of Mechanical Engineers,
2. *Dimensional Measurement Planning,* ASME B89.7.2-1999, American Society of Mechanical Engineers, March 2000
3. *Geometrical Product Specifications (GPS) — Inspection by measurement of workpieces and measuring equipment — Part 1: Decision rules for proving conformance or nonconformance with specifications,* ISO 14253-1:1998 (E), International Organization for Standardization (ISO), 1998

4. *Measurement management systems —Requirements for measurement processes and measuring equipment,* ISO 10012:2003 (E), International Organization for Standardization (ISO), 2003

5. *The Expression of Uncertainty and Confidence in Measurement,* M3003, United Kingdom Accreditation Service (UKAS), January 2007