Living with Uncertainty

By: Cameron Miller and Yoshi Ohno
National Institute of Standards and Technology

Understanding and Quantifying Uncertainty is Key to Accurate and Cost-effective Testing

Webster's Collegiate Dictionary defines uncertainty as "the quality or state of being indefinite, indeterminate, or not reliable: doubt." Can you imagine a customer asking your metrology or test staff if they doubt their measurements? Unfortunately, the answer must always be yes...to an extent. The extent is what we define as the measurement uncertainty.

A perfect measurement would obtain the true value of a quantity, which is the value consistent with the definition of a given quantity.[1] True values are, by nature, indeterminate because a perfect measurement cannot be performed. In fact, says the International Organization for Standardization (ISO), it is impossible to fully describe the measurand (the quantity to be measured) without an infinite amount of information. In other words, the final corrected result of a measurement is, at best, an estimate of the true value of the quantity that someone intended to measure. The measurement uncertainty is a parameter that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

Confusion exists, even among scientists and engineers, as to the exact meaning of the terms such as accuracy, precision, and error. The *accuracy* of measurement is the closeness of agreement between the result of a measurement and a true value of the measurand. The *error* of measurement is the result of a measurement minus a true value of the measurand. As stated above, the true value of a quantity cannot be known; therefore, accuracy and error cannot be quantified for a measurement result. *Precision* is the closeness of agreement for independent measurements obtained under stipulated conditions.[3] *Repeatability* is the closeness of the agreement between the results of successive measurements carried out under the same conditions of measurement over a short period of time. *Reproducibility* is the closeness of the agreement between the results of measurements carried out under changed conditions of measurement. Precision may be stated as "precision under repeatability conditions" or "precision under reproducibility conditions." These terms (accuracy, error, precision) should not be used to specify the uncertainty of measurements.

The cost of uncertainty

Uncertainty is significant because it can affect the cost of production. A particular quality of a product must fall within a given tolerance limit to meet a specification. If the product quality does not fall within the tolerance limit, it is rejected. We make this judgment by measurement, but to do so effectively, we must define how close the measured results can

be to the tolerance. The tolerance interval for the measured value must be smaller than the tolerance of the product (see figure). Here, the measurement uncertainty plays an important role. The smaller the uncertainty, the closer the measurement results can approach the tolerance limit without being rejected.

Properly understood, uncertainty can also provide a useful tool in the planning phase of production. By completing an uncertainty budget (a list of the uncertainty components and their respective uncertainty contributions), a manager can immediately determine which sources of uncertainty contribute most and reduce the largest sources of uncertainty by spending money wisely.

Uncertainty of measurement is becoming increasingly important as manufacturers implement quality systems on the production floor. Quality standards such as ISO 9000 and ISO/IEC 17025 require traceability of all measurements and calibrations performed to products. The ISO defines traceability as the property of the result of a measurement or the value of a standard that allows that standard to be related to stated references (generally national or international standards) through an unbroken chain of comparisons that possess stated uncertainties". This means that there should be a chain of calibration certificates between the instruments used by the manufacturers and the national standards, and all certificates involved must have uncertainty statements.

'NIST traceable' is a commonly used term, but it does not, in and of itself, guarantee the lowest uncertainty. If the calibration chain is too long or the intermediate calibrations are not of high quality, the uncertainty may become significantly large, much larger than that of the NIST scale. However, as long as an unbroken chain of comparisons with stated uncertainties is made, the measurement is considered traceable to NIST.
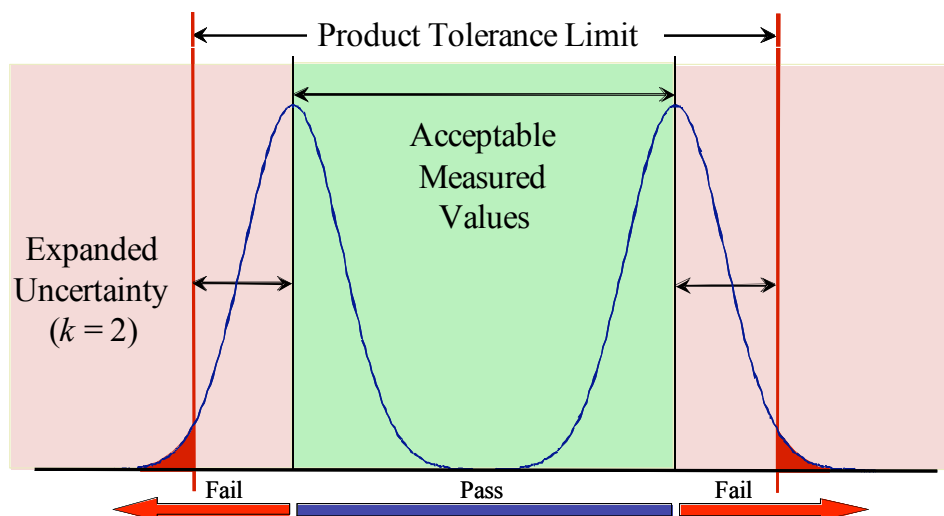
For a product to pass a test, the measured value must exceed the tolerance and the expanded uncertainty (green area). The expanded uncertainty defines a confidence interval. Due to the confidence interval, products that are not acceptable will be taken as acceptable (solid red area) or indeterminate (blue).

NIST is not a regulatory body. It is the customer's job to request and verify the calibration chain. A reputable laboratory will provide this information leading back to a NIST calibration report.

Determination of Uncertainty

To ensure that calibration certificates are properly interpreted by any calibration laboratory worldwide, it is critical that we state uncertainty values in a uniform manner. Over the years, the community has taken many different approaches to evaluating and expressing measurement uncertainty. In 1977, the International Committee for Weights and Measures (CIPM, Comité International des Poids et Measures), the world's highest authority in the field of metrology, asked the International Bureau of Weights and Measures (BIPM, Bureau International des Poids et Measures), to collaborate with national metrology institutes to propose a solution to this problem. The Working Group on the Statement of Uncertainties proposed Recommendation INC-1 (1980) Expression of Experimental Uncertainties. Recommendation INC-1 (1980) provided only a brief outline of requirements, so CIPM asked ISO to develop a comprehensive guide based on INC-1.

In 1993, the end result of the work was published as the Guide to the Expression of Uncertainty in Measurement (GUM).[2] Minor corrections were made and it was reprinted in 1995. Seven international organizations supported the development of this document, which has become the definitive authority on uncertainty evaluation worldwide. Recently, a new international organization, the Joint Committee for Guides in Metrology (JCGM), was formed to assume responsibility for the maintenance and revision of the GUM.[4]

The GUM describes eight steps for the determination of measurement uncertainty. *Step 1* is to write a measurement equation for the measurand. A measurement equation is expressed as a mathematical function $Y = f(X_1,…,X_n)$ that represents the process that is used to determine the result of a measurement and its associated standard uncertainty.[5] The quantities X_i are not only parameters for physical laws but can include corrections and other quantities that account for sources of variability. Often, the corrections have a magnitude of one and do not affect the value of the result but possess inherent uncertainty. For example, a test method may state the temperature must be 25.0ºC. If the temperature at the time of measurement is 25.0ºC with an uncertainty of 0.5ºC, the uncertainty of 0.5ºC contributes to the overall uncertainty of the measurement result. In many situations, defining all of the parameters that indirectly affect a measurement is the most difficult step. *Step 2* is to determine all of the input quantities x_i, the estimated values of X_i, based on statistical analysis of a series of measurements or by other means.

Step 3 is to determine the standard uncertainty, $u(x_i)$, associated with each estimate x_i. The GUM categorizes components of uncertainty by the methods used to evaluate them. Type A evaluation is based on statistical analysis of a series of observations. An uncertainty component obtained by a Type A evaluation is a statistically estimated standard deviation s_i. The standard uncertainty for a Type A evaluation is $u(x_i) = s_i$. Type B evaluation is based on means other than statistical analysis; that is, scientific judgment

using all relevant information, including previous measurement data, previous experience, manufacture's specifications, and data in calibration reports. Typically Type B evaluations are based on probability distributions such as a uniform distribution on some interval [a, b]. Specification of a uniform distribution, requires information on the upper and lower bounds only. We can determine the mean and standard deviation of a uniform distribution and other probability distributions using moment-generating functions. A uniform distribution, for example, has a mean μ and a standard deviation σ of

$$\mu = \frac{a+b}{2} \text{ and } \sigma = \frac{(b-a)}{2\sqrt{3}} \quad . \tag{1}$$

In this case, the estimate equals the mean, $x_i = \mu$ and the standard uncertainty equals the standard deviation, $u(x_i) = \sigma$. For a further discussion of common probability distributions for Type B evaluation, see the GUM.

Step 4 is to evaluate the correlation coefficient for all pairs of estimates. Two quantities are correlated when they are affected by a common quantity. For example, when luminous intensity of a lamp is measured with a photometer, the lamp current and color temperature are quantities of the measurement equation. The current and color temperature are correlated because as the current increases so does the color temperature. Another example of correlation is measuring a length standard with four NIST calibrated meter sticks. If correlation is ignored, the calibrated length standard would appear to have an uncertainty half that of the meter sticks. In reality, the uncertainty would barely decrease because the four calibrated instruments have a strong correlation because their calibration is based on the same unit realization. The correlation coefficient, $r(x_i,x_j)$, is the covariance divided by the standard deviation of the two quantities. The correlation coefficient has a range of -1 to +1, and can reduce or enlarge the combined standard uncertainty. The value of the correlation coefficient can be found by additional statistical analysis as presented in the GUM.

Step 5 is calculating the result of measurement y for Y from the measurement equation using the estimates x_i for the input quantities X_i. Thus, $y = f(x_1,\dots,x_n)$.

Step 6 is determining the combined standard uncertainty $u_c(y)$. The combined standard uncertainty is the estimated standard deviation for the measurand which is the positive square root of the estimated variance $u_c^2(y)$. The variance is approximated by a first order Taylor series of the measurement equation about y. The Taylor series approximation gives the law of propagation of uncertainty:

$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial Y}{\partial X_i} \bigg|_{x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial Y}{\partial X_i} \bigg|_{x_i} \frac{\partial Y}{\partial X_j} \bigg|_{x_j} u(x_i)u(x_j)r(x_i,x_j) \tag{2}$$

A partial derivative of Y with respect to X_i is called the sensitivity coefficient. The sensitivity coefficient can be determined experimentally if the exact functional relation cannot be expressed in a closed form.

Step 7 is to give an expanded uncertainty U, which defines an interval $[y \pm U]$ about the measurement result y that encompasses a large fraction of the distribution of values that could reasonably be attributed to the measurand. Standard uncertainty is converted to expanded uncertainty by multiplying a coverage factor, k. The coverage factor is selected on the basis of the level of confidence or coverage probability required. The level of confidence depends on the distribution for Y represented by y and $u_c(y)$ and the coverage factor k. If the distribution is normal and $k = 2$, the coverage probability is 95%. If the distribution is uniform and $k = 2$, the coverage probability is 100%. The minimum coverage probability with $k = 2$ for any distribution that has an expected value of y and standard deviation of $u_c(y)$ is 75%.[5] NIST policy on expression of uncertainty states the value of k to be used in the NIST calibration services for calculating U is, by convention, $k = 2$.

Step 8 is to report the result of measurement y along with its combined standard uncertainty $u_c(y)$ or expanded uncertainty U and describe how y, $u_c(y)$, and U were obtained.

Traceability to a standard is not sufficient to guarantee an effective measurement. To achieve meaningful results without incurring undue cost, it is important to understand uncertainty, know your calibration chain, and plan an efficient testing process.

References
1. International Vocabulary of Basic and General Terms in Metrology, International Organization for Standardization, 1993.
2. Guide to the Expression of Uncertainty in Measurement, International Organization for Standardization, 1995.
3. B.N. Taylor and C.E. Kuyatt, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Technical Note 1297, 1994. Available for download at http://physics.nist.gov/Document/tn1297.pdf
4. For details on the Joint Committee for Guides in Metrology, see http://www.bipm.org/.
5. R. Kacker and A. Jones, *Metrologia*, **40,** 235 – 248 (2003).

About the authors
Cameron Miller is Photometry Project Leader and Yoshi Ohno is Optical Sensors Group Leader in the Optical Technology Division at the National Institute of Standards and Technology, Gaithersburg, MD. For questions, contact Cameron Miller, Tel: 301-975-4713, Fax: 301-840-8551, or c.miller@nist.gov.