Uncertainty Analysis for Risk Management*

Howard Castrup, Ph.D.
Integrated Sciences Group
Bakersfield, California

Abstract

Measurement errors and error models are reviewed and measurement process error components are described. The computation of measurement uncertainty due to error sources and due to the process error components of each error source is discussed. Measurement decision risks are estimated based on the results of an uncertainty analysis example and risk management considerations are outlined. Classical measurement decision risk is also discussed, with special emphasis on the impact of process uncertainty on false accept and false reject risks. A new method for computing risks is given in Appendix B.

Introduction

In recent years, ISO and NIST guidelines [1, 2] and recommendations have been developed that provide a framework for analyzing and communicating measurement uncertainties. These guidelines constitute a major step in building a common analytical language for both domestic and international trade. This language is important for ensuring that tolerance limits and other measures of uncertainty have the same meaning for both buyer and seller.

Measurement Decision Risk

The motivation for developing such a common understanding derives primarily from the need to control *false accept risk*. False accept risk is the probability that out-of-tolerance product or other parameters are perceived to be in-tolerance. False accept risk constitutes a measure of the quality of a measurement process as viewed by individuals external to the measuring organization. The higher the false accept risk, the greater the chances for returned goods, loss of reputation, litigation and other undesirable outcomes. In a commercial context, individuals external to a measuring organization are labeled "consumers." For this reason, false accept risk has traditionally been called *consumer's risk*.

A counterpart to false accept risk is false reject risk. False reject risk is the probability that in-tolerance parameters are perceived to be out-of-tolerance. False reject risk is a measure of the quality of a measuring process as viewed by individuals within the measuring organization. The higher the false reject risk, the greater the chances for unnecessary re-work and re-test. In a commercial context, a measuring organization is labeled the "producer," and false reject risk is called *producer's risk*.

False accept risk and false reject risk, taken together, are referred to as *measurement decision risk*.

Managing Risks

It used to be that testing a product or calibrating a test system involved very little false accept and false reject risk. This situation has changed. Over the last two decades, market competition and other forces have been relentlessly pushing performance objectives for technology intensive products to levels approaching state-of-the-art envelopes. This means tighter tolerances for product parameters — tolerances so tight that they sometime equal or even exceed the best available measurement system tolerances. With product tolerances on a par with measurement system tolerances, false accept and false reject risks can become serious.

*Presented at the 1995 Measurement Science Conference, Anaheim, California, January 27, 1995.

This is compounded by the existence of measurement process uncertainties. As the tolerances of product parameters shrink, the sizes of such uncertainties become relatively larger. This means that measurement decision risk analyses that fail to account for measurement process uncertainties may produce misleading and sometimes dangerous results.

This paper examines the impact of measurement uncertainty on measurement decision risk. The discussion begins with a review of uncertainty analysis concepts. Following this, an example is presented that shows how an uncertainty analysis can be equated to a false accept risk analysis. Next, another example is presented that illustrates the impact of incorporating measurement uncertainties in traditional consumer/producer risk analyses. Detailed mathematical methods are discussed in Appendices A and B.

Error and Uncertainty

When we measure a physical parameter by any means (e.g., eyeballing, using off-the-shelf instruments, employing precise standards, etc.) we are making an estimate of the value of the quantity being measured. Two features of such estimates are measurement *error* and *measurement uncertainty*.

Measurement Error

The difference between the value of a measured quantity and a measurement estimate of its value is referred to as *measurement error*. Measurement error may be systematic or random.

Systematic errors are classified as those whose sign and magnitude remain fixed over a specified period of time or whose values change in a predictable way under specified conditions.

Random errors are those whose sign and/or magnitude may change randomly over a specified period of time or whose values are unpredictable, given randomly changing conditions. More will be said about random and systematic errors later.

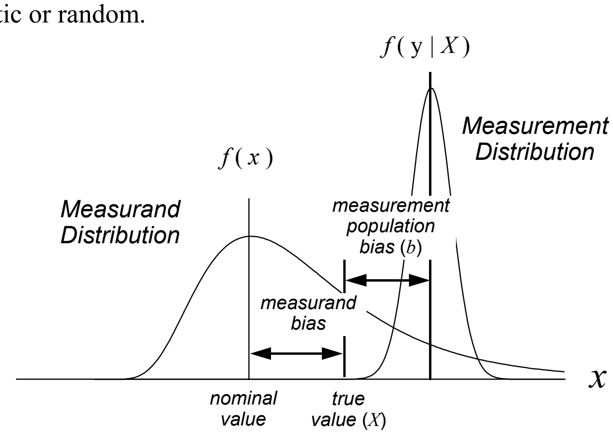


Figure 1 - Bias and Uncertainty. The relationship between the bias in a quantity and the bias and uncertainty in the measurement of the quantity.

Measurement Uncertainty

Measurement errors are never known exactly. In some instances they may be estimated and tolerated or corrected for; or they may be simply acknowledged as being present. Whether an error is estimated or acknowledged, its existence introduces a certain amount of measurement uncertainty.

This brings us to an operational definition of measurement uncertainty. Since measurement error is the discrepancy between the value of a parameter and a perceived or measured parameter value, we can think of measurement uncertainty as either a lack of knowledge concerning the value of a measured parameter or as a lack of knowledge concerning the *error* in the parameter's measurement. The latter view provides a workable framework for analyzing measurement uncertainty.

Error Sources

Before embarking on an analysis of measurement uncertainty, it is usually prudent to develop an error model. In developing an error model, we acknowledge that each factor that influences a measurement result constitutes a potential source of error. Given the view of measurement uncertainty stated above, this means that each factor that influences a measurement is likewise a source of measurement uncertainty. The contribution of each source to the total uncertainty is governed by a coefficient that weights the contribution of the source to the total. Coefficients are determined from a measurement error model.

Error Models

In developing a model, there are a few guidelines that should be followed:

Error Modeling Guidelines

1. Be aware of what is physically involved in making the measurement of interest and what the sources of error are.
2. Be conversant enough with the mechanics of your measurement to know what the measurement data are telling you. This is necessary for evaluating samples of data and for calling attention to special circumstances. This is especially so when sources of error are dependent on one another.
3. Whenever possible, take samples of measurements of variables that may influence your measurement result.

Error Model Example

To illustrate the development of an error model, consider the measurement of the velocity *v* of a moving body. The measurement is decomposed into measurements of distance *d* and time *t*. The system equation is

$$v = \frac{d}{t} .$$

The error $\varepsilon(v)$ in the velocity can be determined using a little high school math:†

$$\begin{aligned} v + \varepsilon(v) &= \frac{d + \varepsilon(d)}{t + \varepsilon(t)} = \frac{d[1 + \varepsilon(d)/d]}{t[1 + \varepsilon(t)/t]} \\ &\cong \frac{d}{t}[1 + \varepsilon(d)/d][1 - \varepsilon(t)/t] \\ &\cong \frac{d}{t}\left[1 + \frac{\varepsilon(d)}{d} - \frac{\varepsilon(t)}{t}\right] \\ &= v + \left(\frac{v}{d}\right)\varepsilon(d) - \left(\frac{v}{t}\right)\varepsilon(t) . \end{aligned}$$

so that

$$\varepsilon(v) = \left(\frac{v}{d}\right)\varepsilon(d) - \left(\frac{v}{t}\right)\varepsilon(t) \ .$$

This equation is the error model for the velocity measurement.

Error Model Uncertainty Analysis

Statistical Variance

A statistic that is useful in analyzing measurement uncertainty is statistical variance. The variance in a measurement may be thought of as the mean square error for the measurement.

† If a variable x is small, relative to unity, then $1/(1 + x) \cong 1 - x$.

Combining Variances

There is a simple rule that governs the variance in the sum of two quantities *x* and *y*. This rule states that, if a and b are constants (or "coefficients"), then

$$\mathrm{var}(ax+by) = a^2\,\mathrm{var}(x) + b^2\,\mathrm{var}(y) + 2ab\,\mathrm{cov}(x,y)\,. \tag{1}$$

The term cov(x,y) is the "covariance" of *x* and *y*.

Statistical Independence

In cases where uncertainties arise from distinct process error sources, the covariance term is usually zero. This is because many, if not most, error sources exhibit a property called "statistical independence." If *x* and *y* are statistically independent variables, the covariance term in Eq. (1) is zero and

$$\mathrm{var}(ax+by) = a^2\,\mathrm{var}(x) + b^2\,\mathrm{var}(y)\ \ . \tag{2}$$

Working from the Error Model

In our velocity measurement example, errors in distance and time are statistically independent — measurements made with a tape measure, say, have no influence on measurements made with a stopwatch. Accordingly, Eq. (2) applies to the velocity measurement error model, and

$$\mathrm{var}[\varepsilon(v)] = \left(\frac{v}{d}\right)^2 \mathrm{var}[\varepsilon(d)] + \left(\frac{v}{t}\right)^2 \mathrm{var}[\varepsilon(t)]\ \ .$$

Using the view of measurement uncertainty expressed earlier, we can see that the variance in the error of a measured quantity is just the variance in the measurement itself. Thus, for a source of error *x*,

$$\begin{aligned}\mathrm{var}(\varepsilon(x)) &= \mathrm{var}(x)\\ &\equiv \sigma_x^2\,.\end{aligned} \tag{3}$$

With this in mind, the variance of the velocity measurement is written

$$\begin{aligned}\sigma_v^2 &= \left(\frac{v}{d}\right)^2 \sigma_d^2 + \left(\frac{v}{t}\right)^2 \sigma_t^2\\ &= \frac{\sigma_d^2}{t^2} + \frac{v^2\sigma_t^2}{t^2}.\end{aligned}$$

Statistical Uncertainty - Standard Deviations

When we sample a value of a random variable, such as distance or time, we obtain a number that may take on a range of values. In many cases, the range of values accessible to a variable is infinite. This doesn't mean, however, that all values are equally likely. For the most part, sampled values of a random variable tend to be distributed about some mean or mode value.

If we look at a measurement as something that attempts to estimate a mean or mode value, then deviations in measurement from the mean or mode comprise measurement errors. The extent to which measurement errors are spread out from the mean or mode value of a variable is equated with the uncertainty in the variable. The statistic that quantifies this spread is the **standard deviation**. The standard deviation is just the square root of the variance. In the velocity measurement example, the measurement uncertainty is, therefore,

$$\sigma_v = \frac{1}{t}\sqrt{\sigma_d^2 + v^2\sigma_t^2}\ .$$

In general, the greater the spread, the larger the standard deviation. This means that, with large standard deviations, errors tend not to be "localized," i.e., the confidence with which they are known tends to be low. Equating the word "confidence" with the less precise but more comfortable word "certainty," we see why the standard deviation for a quantity is equated with its *un*certainty. To make the equivalence clearer, we will henceforth denote standard deviations and other uncertainties by the letter u, rather than the symbol σ:

$$u_v = \frac{1}{t}\sqrt{u_d^2 + v^2 u_t^2} \ .$$

Component Uncertainties

Process Error Components

The perceived values of distance d and time t are each obtained in a measurement process. If each process involves a definable set of error sources, then the errors in d and t can be expressed in terms of the component errors of their respective measurement processes. It has been found useful to break process errors down as follows:

Error Component	**Description**
Subject Parameter Bias	Systematic discrepancy between the "true" value and the nominal or reading value of a parameter being measured.
Measuring Parameter Bias	Systematic discrepancy between the "true" value and the nominal or reading value of a parameter performing a measurement.
Subject Parameter Random	Random fluctuations in the value of a parameter being measured.
Measuring Parameter Random	Random fluctuations in the value of a parameter performing a measurement.
Subject Parameter Resolution	Error due to the finite precision with which values of a parameter being measured can be perceived.
Measuring Parameter Resolution	Error due to the precision with which values of a parameter performing a measurement can be perceived.
Data Acquisition	Error due to acquiring data from measurements. Includes data sampling error, computation or "round off" error and operator bias.
Stress Response	Error due to stresses of shipping and handling of an item following measurement. Stress response error is important in cases where a measured parameter's value is reported externally and the measured item is physically moved from the measurement environment to another location.
Environment/Ancillary Equipment	Error due to environmental factors or to ancillary equipment, such as temperature monitoring devices.
Other	Error due to sources peculiar to a given measurement scenario.

With the above breakdown, we can write the process error experienced in measuring a quantity x as

$$\varepsilon(x) = \varepsilon_{sb}(x) + \varepsilon_{mb}(x) + \varepsilon_{sran}(x) + \varepsilon_{mran}(x) + \varepsilon_{sres}(x) + \varepsilon_{mres}(x) + \varepsilon_{da}(x) + \varepsilon_{stress}(x) + \varepsilon_{env}(x) + \varepsilon_{other}(x) \ . \quad (4)$$

The terms $\varepsilon_{da}(x)$, $\varepsilon_{stress}(x)$, $\varepsilon_{env}(x)$ and $\varepsilon_{other}(x)$ can each be further decomposed into subcomponent errors [3].

Process Uncertainty

The variance addition rule in Eq. (1) can be applied to the process error components of each error source. To illustrate, consider the distance measurement. Suppose that, after careful consideration, it is decided that the error in the distance measurement is expressed by

$$\varepsilon(d) = \varepsilon_b(d) + \varepsilon_{ran}(d) + \varepsilon_{res}(d) + \varepsilon_{da}(d) + \varepsilon_{env}(d) \; .$$

Invoking Eq. (3) and applying Eq. (1) gives

$$\begin{aligned} u_d^2(d) = {} & u_b^2(d) + u_{ran}^2(d) + u_{res}^2(d) + u_{da}^2(d) + u_{env}^2(d) \\ & + 2\,\mathrm{cov}[\varepsilon_b(d), \varepsilon_{ran}(d)] + 2\,\mathrm{cov}[\varepsilon_b(d), \varepsilon_{res}(d)] + 2\,\mathrm{cov}[\varepsilon_b(d), \varepsilon_{da}(d)] + 2\,\mathrm{cov}[\varepsilon_b(d), \varepsilon_{env}(d)] \\ & + 2\,\mathrm{cov}[\varepsilon_{ran}(d), \varepsilon_{res}(d)] + 2\,\mathrm{cov}[\varepsilon_{ran}(d), \varepsilon_{da}(d)] + 2\,\mathrm{cov}[\varepsilon_{ran}(d), \varepsilon_{env}(d)] \\ & + 2\,\mathrm{cov}[\varepsilon_{res}(d), \varepsilon_{da}(d)] + 2\,\mathrm{cov}[\varepsilon_{res}(d), \varepsilon_{env(d)}] \; . \end{aligned}$$

Correlated Process Errors

If two error sources or process error components are not statistically independent, then we say that they are correlated. This correlation is quantified by a *correlation coefficient*. A correlation coefficient of +1 means that two quantities vary in perfect step with one another. If one goes up or down by a certain amount, the other goes up or down by a proportional amount. A correlation coefficient of zero means that two quantities are statistically independent. A correlation coefficient of -1 means that two quantities vary in perfect step with one another but in opposite directions. If one goes up or down by a certain amount, the other goes down or up by a proportional amount.

The correlation coefficient can be either estimated heuristically or computed from statistical samples. As an example of a heuristic determination, consider a case where the same measuring parameter is used to perform two separate measurements. Since the same parameter is used for both measurements, the error due to parameter bias will be the same for both measurements. If the parameter is used to transfer one value to the other, the errors offset and cancel. If the two measurements combine in a total, the errors add together. In the former case, the correlation coefficient is -1, in the latter it is +1.

Statistical estimation of correlation coefficients is described in References 1 and 2.

Correlations and Covariances

If ρ_{ij} is the correlation coefficient for two error sources or error components x_i and x_j, then the covariance between x_i and x_j is given by

$$\mathrm{cov}(x_i, x_j) = \rho_{ij} u_i u_j \; , \tag{5}$$

where u_i and u_j are the uncertainties in x_i and x_j. In general, it is simpler to estimate ρ_{ij} than $\mathrm{cov}(x_i, x_j)$. The procedure is found in standard upper division textbooks on statistics. Using Eq. (5) in the expression for the variance in the distance measurement of the velocity example gives

$$\begin{aligned} u_d^2(d) = {} & u_b^2(d) + u_{ran}^2(d) + u_{res}^2(d) + u_{da}^2(d) + u_{env}^2(d) \\ & + 2\rho_{b,ran} u_b(d) u_{ran}(d) + 2\rho_{b,res} u_b(d) u_{res}(d) + 2\rho_{b,da} u_b(d) u_{da}(d) + 2\rho_{b,env} u_b(d) u_{env}(d) \\ & + 2\rho_{ran,res} u_{ran}(d) u_{res}(d) + 2\rho_{ran,da} u_{ran}(d) u_{da}(d) + 2\rho_{ran,env} u_{ran}(d) u_{env}(d) \\ & + 2\rho_{res,da} u_{res}(d) u_{da}(d) + 2\rho_{res,env} u_{res}(d) u_{env}(d) \; . \end{aligned}$$

(6)

Uncorrelated Process Errors

Suppose that, we can assert that all process error components are independent. Then the correlation coefficients in Eq. (6) are all set to zero, and the uncertainty in the measurement of d becomes

$$u_d = \sqrt{u_b^2(d) + u_{ran}^2(d) + u_{res}^2(d) + u_{da}^2(d) + u_{env}^2(d)} \; .$$

Estimating Process Uncertainties

Statistical (Category A) Estimates

An uncertainty estimate may be made on the basis of a statistical sample of measurement data. If so, the uncertainty estimate is just the sample standard deviation. A standard deviation obtained by statistical sampling is called a "Category A estimate" [1, 2]. If confidence limits are desired for a Category A uncertainty estimate, a "*t*-statistic" is calculated and applied as described in Appendix A.

Heuristic (Category B) Estimates

Estimating an uncertainty for errors from a given source where statistical samples are not available is very simple and straightforward. Such an uncertainty can be obtained from a heuristic estimate of limits that serve as bounds for errors from the source together with an estimate for the probability that these limits are expected to contain the errors. This level of probability reflects the confidence that the limits will contain the errors.

For this reason, it is tempting to refer to the heuristically estimated limits as "confidence limits." However, confidence limits are quantities that are usually determined statistically. To avoid confusing heuristic limits with statistical confidence limits, we will instead call them "error limits" or "containment limits." The probability that they contain errors from the source of interest will be called the "containment probability."

For normally distributed errors, with an estimated containment probability P and symmetric containment limits $\pm L_P$, the uncertainty in the errors is obtained from

$$u = \frac{L}{\Phi^{-1}\left(\frac{1+P}{2}\right)}, \tag{7}$$

where the function $\Phi^{-1}(\cdot)$ is the inverse normal distribution function. For uniformly distributed errors, contained within limits $\pm L$, the uncertainty estimate is

$$u = \frac{L}{\sqrt{3}}. \tag{8}$$

There are other ways to determine heuristic uncertainty estimates. For instance, an uncertainty may be estimated directly without recourse to containment limits or containment probabilities. However it is arrived at, a heuristic uncertainty estimate is termed a **Category B estimate**.

Occasionally, Category B estimates are just as good as statistical estimates. Estimating the uncertainty due to parameter resolution in a digital readout is a case in point. With a digital readout, the information we have is a displayed number and a position for the least significant digit.

The readout provides an indication of the analog value sensed by the measuring parameter. All we know about this sensed value is that it lies somewhere between ± 0.5 times the magnitude of the value represented by the least significant digit of the indicated value. For example, suppose we obtain a display of 10.00156 for a measured quantity. We know only that the sensed value lies somewhere between 10.001555 and 10.001565.[‡] We don't know anything about how sensed values are distributed between these points.

In the absence of such information, we have to confine any statements we might make regarding the resolution error of the measurement to saying that the sensed value lies between the two numbers with equal or uniform probability. In this case, we can estimate the uncertainty due to parameter resolution as

$$u_{res} = \frac{0.000005}{\sqrt{3}}.$$

[‡]Of course, we can't say the same thing about where the *actual* value is.

In obtaining the Category B estimate u_{res}, there was no need to estimate approximate error limits and a containment probability. The error limits were set by the resolution of a digital display, and the containment probability was 100%. Such circumstances are atypical. Usually, the error limit/containment probability approach is advisable in estimating Category B process uncertainty components.

Bias Uncertainty

When systematic errors are of known sign and magnitude, they should be extracted from measurement results. When this knowledge is unavailable, the uncertainty surrounding a systematic error is treated statistically. An example of such a systematic error is the bias of a measuring parameter of an instrument drawn randomly from a pool or “population” of like instruments. The bias of the parameter is the difference between the parameter’s actual value and its nominal, reading or expected value.

Bias Distributions

Although a parameter’s bias may be systematic over the period during which the parameter is used, we usually don’t know its sign and magnitude. At best, all we can expect to know is that the bias we get when we select an item from a population or pool of like items may differ from item to item to a definable extent. Another way of saying this is that, the population of biases in a parameter's value may be distributed in a way that can be defined.

Estimating Bias Uncertainty

Knowledge of just a few details about a population of parameters can often help in forming an estimate of the population's distribution of bias values. We can usually specify the bias distribution once we know the mode or mean value for the population of biases and can compute the distribution’s standard deviation. This standard deviation represents the uncertainty in the value of parameters drawn from the population. It is the statistic to be used for the bias uncertainty component of the measurement process.

There are two principal methods for estimating the standard deviation of a bias distribution. One involves the use of test or calibration history, while the other involves making a heuristic estimate.

Using Test / Calibration History

The standard deviation for a parameter's bias distribution can be computed from a random sample of parameter values taken by a higher accuracy device. Such a sample could be obtained from periodic test or calibration results. This method is simple and straightforward, but rarely used.

This is because the economics of equipment management usually lead to data collection and storage practices in which calibrated or tested values are not kept or are not readily available. Normally, all that is kept are in- or out-of-tolerance results at the instrument or unit level.

Although sampled parameter values would be nice to have in estimating bias standard deviation, we can often get reasonably good estimates from the unit level test or calibration results. For normally distributed or uniformly distributed biases, this involves a two-step process:

1. Obtain an in-tolerance probability (observed percent in-tolerance) estimate for the parameter of interest at the time of test or calibration.
2. Use this in-tolerance probability, along with the parameter tolerances in Eq. (7) or (8) to compute the bias uncertainty.

For biases that are not normally distributed or uniformly distributed, the procedure is similar, but Eqs. (7) and (8) do not apply. Other, somewhat more complicated means must be applied [4].

Making a Heuristic Estimate

For many applications, test or calibration history data may not be available. This is the case, for instance, when the parameter of interest is not periodically tested or calibrated. If not, then we need to make some kind of engineering or heuristic estimate for the parameter's in-tolerance probability.

Heuristic estimates are based on considerations of parameter stability and "inherent" accuracy. The latter is often no more than a reflection of our expectations for the parameter's ability to yield accurate measurements, considering its design and fabrication. For instance, we tend to put more faith in measurements from a precision scale than from a bathroom scale.

When possible, heuristic estimates can be made on the basis of detailed engineering knowledge about the parameter of interest and the conditions under which it's used. In these cases, heuristic estimates can be as good as or better than (i.e., more representative) estimates made using test or calibration history.

However it is arrived at, a heuristic estimate for a parameter's in-tolerance probability is used (together with parameter tolerance information) in the same way that an estimate based on history data is used.

Random Uncertainty

Random error can appear as random differences in sampled values observed during measurement. Calculating the uncertainty due to random error usually involves taking a sample of measurements and computing a sample mean and standard deviation. Expressing confidence limits for random error often involves multiplying the standard deviation by a "*t*-statistic." The method is described in Appendix A.

Resolution Error

Measurement uncertainty can result from the fact that values sensed by a measuring parameter or generated by a source may fall between indicated meter points or between least significant digits on decimal readout displays. Resolution uncertainty can be computed from estimated limits of resolution error, together with estimated confidence limits, or from the available number of significant digits.[§]

Data Acquisition Error

Data acquisition uncertainty may arise from a variety of sources including quantization error, aperture time error, hysteresis error, aliasing error, operator bias, finite sampling rates and so on [3, 7]. Operator bias will be discussed later on.

Stress Response

If a parameter being measured drifts or fluctuates in value in response to shipping, handling or usage stress, it may be necessary to include the uncertainty in these changes in the parameter's uncertainty analysis. This is done by multiplying assumed limits of stress by a stress response coefficient and estimating the probability that stresses will be confined within the estimated limits.

Alternatively, if statistical stress response data are available for the parameter of interest, then it may be possible to obtain a statistical estimate of stress response uncertainty.

Environment / Ancillary Error

Errors arise from environmental factors or can be due to ancillary equipment such as power supplies, secondary monitoring devices, etc. For instance, if ambient temperature corrections are applied to measured values, then the

[§]See the earlier discussion on Category B estimates.

error in a given ambient temperature measurement constitutes an ancillary error. Environment/ancillary uncertainty is obtained by estimating a standard deviation for the environmental or ancillary parameter and multiplying it by an interaction coefficient. For instance, if the coefficient of thermal expansion for a gage block is κ, and σ_T is the uncertainty in the ambient temperature, then the process uncertainty in gage block length due to errors or deviations in ambient temperature is given by $\kappa\sigma_T$.[**]

Uncertainty Analysis and Risk Management

We will explore the rudiments of using uncertainty analysis results to manage decision risks by working through a simple example. The example addresses the problem of managing risks involved in matching parts together. The problem consists of cutting carpet to fit in a room.

Example - Fitting Parts Together

In the example, a length of carpeting is to be measured with a steel tape measure to fit in a room whose nominal dimensions are known from a set of blueprints. The carpet will be cut in the shop and delivered to the job site. The objective is to avoid coming up short at the job site while at the same time not wasting an excessive amount of carpet.

Preparation for Analysis

To restate the analysis objective, we want to know the minimum amount over the room's nominal length we need to cut the carpet. Of course, we could easily use a large safety margin and cut the carpet to something like a half meter or so over the nominal room length, but we want to avoid waste to improve our profit margin. So, we want to be "fairly certain" that we have cut enough carpet, but realize that we can't be "absolutely certain" without losing money.

For the sake of discussion, suppose that we equate "fairly certain" with a 99% probability that our carpet length measurement will contain the room length's true value. This is equivalent to setting a maximum acceptable false accept risk criterion of 1%.

Estimating Subject Parameter Bias Uncertainty

Often, a critical step in estimating measurement uncertainty is getting some idea of the expected uncertainty in the parameter being measured. In the present example, this means estimating the uncertainty in the nominal value of the room.

We define this uncertainty by specifying the nominal value for the room and developing a set of tolerances or uncertainty limits that bound the nominal value. One way to do this would involve going to the job site and measuring the room. In this example, however, we will not do this. Instead, we will attempt to determine these values from the builder's plans.

Suppose that, according to these plans, the room has an indicated nominal length of five meters. Imagine that experience has shown us that roughly ninety percent of room lengths fall within about one percent of nominal.[††] From this, we set the room's specification at five meters $\pm$1% with an error containment probability of 90%.

Assuming that room length biases are normally distributed around the five meter design value, we use Eq. (7) to get

---

[**] It is assumed that any calculated thermal expansion length corrections have been implemented.

[††] This includes effects due to shifting strata, temperature variations, etc.

$$
\begin{aligned}
u_{sb} &= \frac{0.01 \times 5.0 \text{ meters}}{\Phi^{-1}\left(\frac{1+0.90}{2}\right)} \\
&= 0.030398 \text{ m} \\
&= 3.0398 \text{ cm},
\end{aligned}
$$

where *sb* refers to "subject parameter bias."

Estimating Measuring Parameter Bias Uncertainty

In the present example, the measuring parameter is the reading from a steel tape measure. Suppose we have recently purchased the tape measure and that it comes with a tolerance specification of $\pm$ 0.1% of reading. We call the manufacturer and discover that; given the tape measure's design, fabrication and testing; approximately 95% of the units sold measure to within $\pm$ 0.1% or less of the true value of whatever is being measured. In other words, the manufacturer's tolerance limits of $\pm$ 0.1% are error limits that are expected to bound tape measure biases with a 95% containment probability.

Assuming normally distributed measuring parameter biases, we use Eq. (7) to get

$$
\begin{aligned}
u_{mb} &= \frac{0.001 \times 5.0 \text{ meters}}{\Phi^{-1}\left(\frac{1+0.95}{2}\right)} \\
&= 0.2551 \text{ cm}.
\end{aligned}
$$

The *mb* refers to "measuring parameter bias."

Random Uncertainty

Before proceeding, we need to briefly review what we have at this point. We have successfully accounted for uncertainty due to the length of the room and due to measurement bias in the tape measure. We realize that we will be using the tape measure to measure carpet in the shop and that we will cut the carpet based on our measurement results.

From experience, we know that we may make repeated measurements of the same quantity and get slightly different results from measurement to measurement. These differences are due to random error. To account for random error, we go to the shop and take a sample of measurements using the measuring parameter (the tape measure). Suppose that we obtain the sample shown below

measurement	**value (x)**
1	5.005
2	5.004
3	5.006
4	5.006
5	5.008
6	5.004
7	5.006
8	5.003

We next use Eqs. (A-1) and (A-2) of Appendix A to compute random uncertainty

$$
\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 5.00525 \text{ m},
$$

and

$$u_{ran} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$
$$= 0.1581 \text{ cm},$$

where the *ran* subscript denotes "random" uncertainty.

The Sampling Distribution

The computed random uncertainty of 0.1581 cm applies to single measurements. That is, if we make a measurement of a carpet length, the estimate of the random uncertainty in our measurement is 0.1581 cm.

If, instead of a single measurement, we base cutting the carpet on a mean value taken from a sample of measurements, then the magnitude of the random uncertainty component is somewhat reduced. This reduced magnitude turns out to be equal to the random uncertainty for a single measurement divided by the square root of the number of measurements in the sample:

$$u_{ran} \rightarrow \frac{u_{ran}}{\sqrt{n}} .$$

This result is based on what is called the *sampling distribution*. Whereas the standard deviation for a sample applies to single measurements of a quantity, the standard deviation for the sampling distribution applies to the mean or average value of measurements of the quantity.

In many applications, where an action is taken or a value is published based on a mean of measurements, the sampling distribution applies. In the present example, use of the sampling distribution reduces the estimated random uncertainty from 0.1581 cm to 0.0559 cm

$$u_{ran} = 0.0559 \text{ cm} .$$

Resolution Uncertainty

In bias uncertainty and random uncertainty, we have encountered two of the major aspects of measurements, namely, **accuracy** and **repeatability**. We now come to grips with **precision**. "Precision" corresponds to how many places past the decimal point we can express a measurement result. For example, a measuring device may yield the number 5.1 when measuring a subject parameter. Another measuring device may yield the number 5.1022 for the same parameter. The second measurement is said to be more precise than the first.

Although higher precision does not necessarily mean higher accuracy, the lack of precision in a measurement is a source of measurement error. The error due to the finiteness of the precision of a measurement is referred to as **resolution error** and the uncertainty in this error is called **resolution uncertainty**.

It may seem at first that the two statements "higher precision does not necessarily mean higher accuracy" and "the lack of precision is a source of measurement error" are contradictory. To see that this is not so, consider the following cases.

The statement "higher precision does not necessarily mean higher accuracy" is exemplified by a meter stick whose nominal markings are all short by 10%, but whose measurement resolution is 10 significant digits. If we measure a ten cm long artifact, we get 9.0000000000 cm. We have abundant precision, but are still off by a full centimeter.

To illustrate the statement "the lack of precision is a source of measurement error," consider a voltmeter with a digital display with one significant digit. Assuming perfect accuracy and repeatability for the device, a 12.35v subject parameter value would be displayed as 12.4v. Despite its perfect accuracy, the voltmeter provides a measurement result that is high by 0.05 volts.

Back to our example. Suppose that the tick marks of the tape measure used to make the sample of measurements are one millimeter apart. This means that measurements of carpet can be resolved to within one-half millimeter or ± 0.05 cm. To compute the uncertainty due errors arising from this resolution, we use Eq. (8) to get

$$u_{res} = \frac{0.05 \text{ cm}}{\sqrt{3}} = 0.0289 \text{ cm}.$$

Data Acquisition Uncertainty

Data acquisition involves acquiring, processing and perceiving physical stimuli from a measuring parameter. Data acquisition uncertainty is influenced by error from three primary sources:

- Data Sampling
- Data Reduction (Computation Error)
- Operator Bias

In the present example, we are concerned with data acquisition uncertainty because error may arise from a discrepancy between the readings provided by the measuring parameter and our perception of these readings. Depending on our perspective relative to the subject parameter and the measurement readings, the measurements we perceive may differ from what the tape measure is trying to tell us.

We may shift this perspective in random ways during the taking of the measurement sample. Variations due to these random shifts contribute to random error, discussed earlier.

In addition to random shifts in perspective, we may also be prone to a *systematic* bias due to perspective. This bias is referred to as **operator bias**. We do not know the sign or magnitude of operator bias, so it cannot be factored out and we must treat it as an unknown measurement error. The uncertainty in this error is determined as a Category B estimate.

Suppose that we have good reason to believe that 95% of errors due to operator bias will not exceed 25% of the measuring parameter resolution. Since the tape measure resolution is 1 mm, the operator bias error limits are ± 0.25 mm (or ± 0.025 cm), and, using Eq. (7), we get

$$\begin{aligned} u_{op} &= \frac{0.025 \text{ cm}}{\Phi^{-1}\left(\frac{1+0.95}{2}\right)} \\ &= 0.0128 \text{ cm}, \end{aligned}$$

where the *op* subscript denotes "operator" bias.

Uncertainty due to Environment

The length of the steel tape measure varies with temperature. We suspect that contributions to uncertainty due to this variation will not be significant for carpet cutting purposes. But we are "in the hunt" so to speak and need to satisfy our growing lust for knowledge about measurement uncertainty.

In this example, we examine the uncertainty due to variations in measuring parameter length by doing the following:

1. Determine how tape measure length varies with temperature (i.e., the tape measure's thermal expansion).
2. Estimate the uncertainty in the temperature of the measuring environment.
3. Estimate the thermal expansion uncertainty in the tape measure readings.

Tape Measure Thermal Expansion

Suppose that the manufacturer of the tape measure has informed us that the markings on the tape were made against a reference standard at a nominal temperature of 25 degrees C. If the tape measure is used at any temperature other than 25 degrees C, its readings will be off by some amount due to thermal expansion or contraction. For steel, the thermal expansion coefficient is around 5.30×10^{-6} per degree C.

Measuring Environment Temperature Uncertainty

We could easily measure the temperature in the shop and correct for it by an appropriate amount. For now, however, we want to see what happens if we don't apply this correction — that is, if we just use shop measurements "as is" without correction. Let us say that, from experience, we know that, 95% of the time, the temperature in the shop is contained within ± 5 degrees of 25 degrees C. This means that we can state the temperature as 25 deg C ± 5 deg C with a 95% error containment probability. We now proceed to find out what the uncertainty in the shop temperature does to our measurement uncertainty. The shop temperature uncertainty is computed using Eq. (7) as

$$
\begin{aligned}
u_T &= \frac{5 \text{ deg C}}{\Phi^{-1}\left(\dfrac{1+0.95}{2}\right)} \\
&= 2.5511 \text{ deg C.}
\end{aligned}
$$

Thermal Expansion Uncertainty

Given this temperature uncertainty, the uncertainty in the thermal expansion of a five-meter long steel tape measurement is

$$
\begin{aligned}
u_{env} &= (5.30\times10^{-6} / \text{deg C})(5.0 \text{ m})(2.5511 \text{ deg C}) \\
&= 0.0068 \text{ cm,}
\end{aligned}
$$

where *env* refers to "environmental" uncertainty. From this, we conclude that tape measure variations due to temperature can be ignored in future carpet measurements — as we suspected earlier.

Uncertainty Combination

In the present example, all process error sources are statistically independent.[‡‡] This means that the total uncertainty is given by Eq. (A-11) with all correlation coefficients set to zero:

$$
\begin{aligned}
u_{total}^2 &= u_{sb}^2 + u_{mb}^2 + u_{ran}^2 + u_{res}^2 + u_{op}^2 + u_{env}^2 \\
&= [(3.0398)^2 + (0.2551)^2 + (0.0559)^2 + (0.0289)^2 + (0.0128)^2 + (0.0068)^2] \text{ cm}^2 \\
&= 9.3098 \text{ cm}^2 ,
\end{aligned}
$$

and

$$
u_{total} = 3.0512 \text{ cm} .
$$

Risk Analysis

We are now ready to use this uncertainty estimate to evaluate the measurement decision risks associated with fitting carpet into five meter long rooms. To do this, we review our risk management objectives.

We wanted to be able to cut a length of carpet in the shop that would be neither too short for the room at the job site nor so long that we incur unnecessary waste. We know that we have to cut the carpet longer than the nominal room dimension, but by how much? We have an uncertainty estimate that includes uncertainty due to the room deviating from its nominal value and due to the carpet cutting measurement process. What do we do with this uncertainty to determine how much the carpet length needs to exceed the five meter nominal value?

[‡‡] Although operator bias is based on tape measure resolution, the latter is not a variable quantity. For this reason, there is no covariance between operator bias and resolution error and the two can be taken to be statistically independent in that sense.

At the outset of the analysis, we decided that we wanted to be 99% confident that the carpet would not be cut too short. We can use our uncertainty estimate to obtain an error limit that allows us to achieve this goal. This limit will be unlike other limits we have encountered so far. These were all "±" two-sided limits. We now want a one-sided limit. For normally distributed total uncertainty estimates,[§§] a one-sided limit L_P, corresponding to a containment probability P, is given by

$$L_P = \Phi^{-1}(P) u_{total} .$$

For this example, $P = 0.99$, $u_{total} = 3.0512$ cm, and

$$\begin{aligned} L_{0.99} &= \Phi^{-1}(0.99)(3.0512) \text{ cm} \\ &= (2.3635)(3.0512) \text{ cm} \cong 7.2 \text{ cm.} \end{aligned}$$

What this means is that a carpet length cut to 5 m + 7.2 cm in the shop has a 99% chance of being long enough to fit the room at the job site. Another way of looking at this is to say that, by cutting carpet to 5.072 m, we incur a 1% risk of cutting carpet too short. Recalling earlier discussions on measurement decision risk, this statement is equivalent to saying that cutting carpet 7.2 cm longer than nominal incurs a 1% false accept risk.

Caution

In this example, we did not make a preliminary measurement of the room into which carpet was to be fit. Instead, we used a nominal estimate of five meters and allowed for uncertainty in this estimate. In order to work from nominal estimates instead of preliminary measurements, we need to make sure that the error limits we set for nominal values are valid. In this example, we felt that 90% of actual room lengths would be within ±1 percent of the nominal value of five meters.

If this is an unjustified assumption, basing decisions on the uncertainty analysis we just did would be risky. In other words, we could easily show up at the job site with carpet that was either too short to fit or excessively long. For instance, suppose that only 50% of room lengths fell within ±1 cm of nominal. If so, then we would find that we should be cutting about 17 cm of extra carpet instead of 7.2 cm.

Consumer/Producer Risk Analysis

We now examine how measurement process uncertainty can influence false accept and false reject risk in a calibration or testing application. The risk assessment methodology is described in References 6 and 8 -19 (see also Appendix B). We discuss the contribution of process uncertainty to false accept and false reject risk within the context of a hypothetical example.

The example involves the calibration of a test system parameter (the subject parameter) by a calibration system parameter (the measuring parameter). The relative nominal accuracy ratio between measuring and subject parameters is 10:1. The in-tolerance probabilities for both parameters is 95% at the time of measurement. The "nominal" risk situation is shown in Figure 2.

[§§]The combined uncertainty estimate is really *t*-distributed rather than normally distributed. However, the computed degrees of freedom for this example turn out to be a little over sixty-two million. For such a large number, the *t*-distribution and the normal distribution are the same.

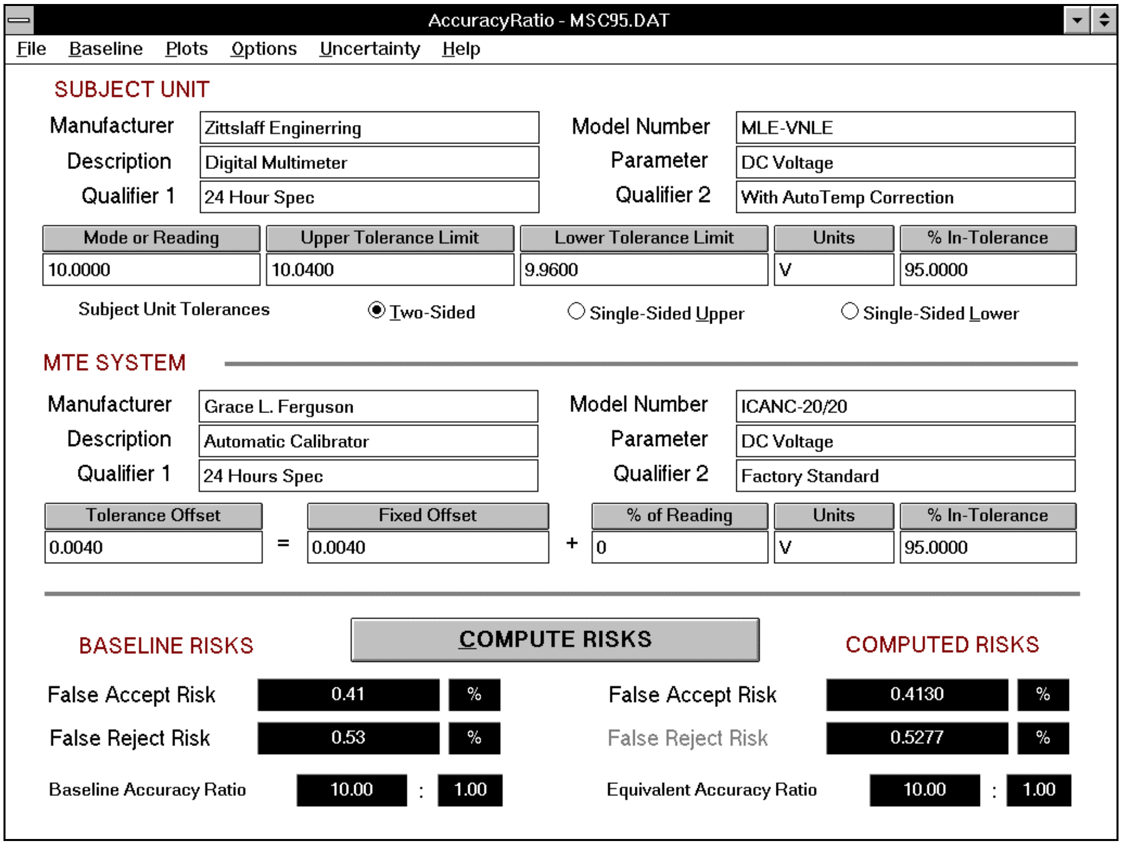


Figure 2 - Nominal Risks. False accept and false reject risks with process uncertainty excluded.

The nominal risks are those that would exist if there were no measurement process uncertainty other than the bias uncertainties in the measuring and subject parameters. We now introduce the process uncertainties

Process Error Component	**Uncertainty Estimate (mV)**
measuring parameter random	0.212
subject parameter random	1.455
measuring parameter resolution	0.0255
subject parameter resolution	0.0255
operator bias	0.912
stress response	0.892
environmental	1.114
Combined Uncertainty:	**2.2431**

The impact of these uncertainties on measurement decision risk is shown in Figure 3. By comparing the computed risks of Figure 3 with those of Figure 2, we see that process uncertainty can have a significant impact on false accept and false reject risks. We conclude that, in evaluating measurement decision risk, it is important to perform a thorough measurement uncertainty analysis and to incorporate the results of this analysis in risk calculations.

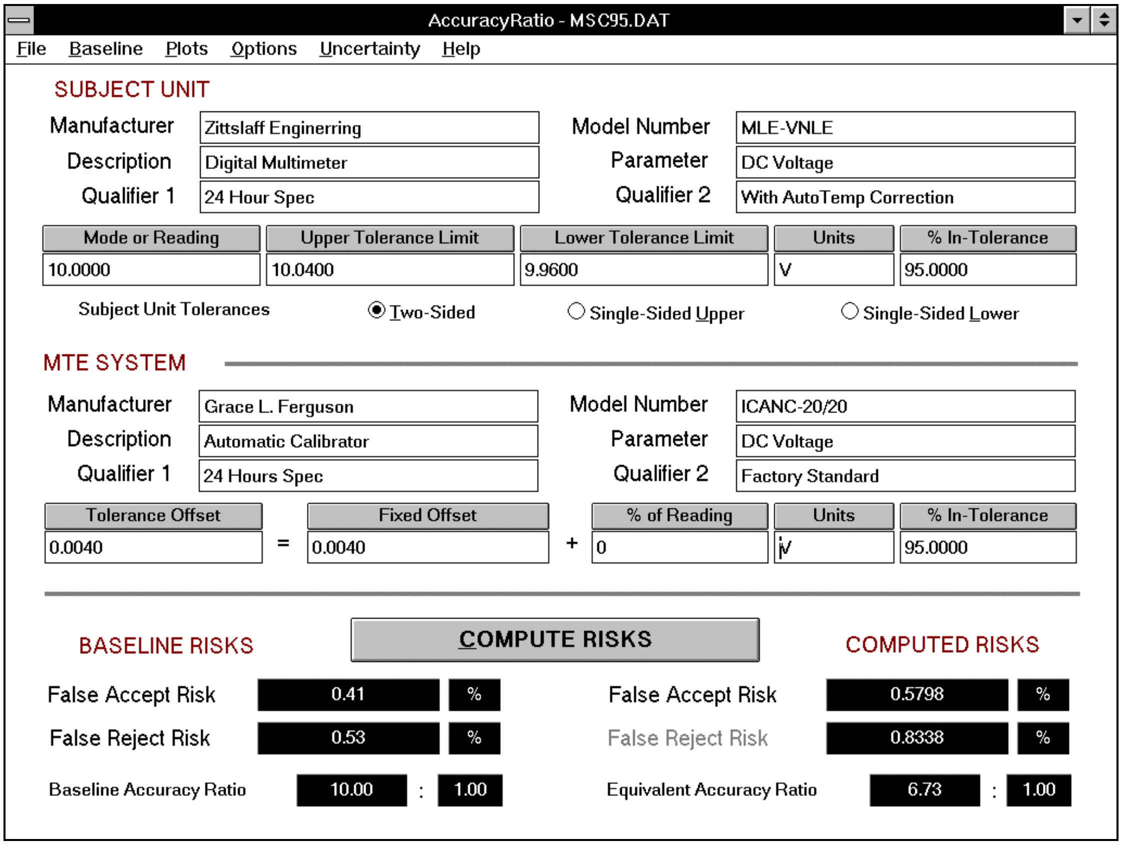


Figure 3 - Actual Risks. The false accept and false reject risks reflect the inclusion of process uncertainty.

Conclusion

In employing uncertainty estimates in risk analyses, it is important to identify all relevant sources of measurement error. This identification is facilitated by first developing an error model.

Following the identification of error sources, each source is decomposed into its constituent process error components, and the uncertainty in each component is then estimated. The component uncertainties for each error source are combined to yield an estimate of the uncertainty in the source. Source uncertainties are then combined in accordance with the error model to yield a total measurement uncertainty.

In this paper, we have endeavored to place uncertainty analysis in a utilitarian context. Rather than simply produce uncertainty estimates as quantities that have value in themselves, we have attempted to employ these estimates in a way that has relevance for making decisions based on measurement results.

In doing so, we have seen that measurement uncertainty can be a significant factor in evaluating measurement decision risk. In one example, we developed an estimate of false accept risk directly from an uncertainty estimate. In another, we incorporated uncertainty estimates in a consumer/producer risk analysis.

Appendix A - Uncertainty Analysis Methodology

Compliance with National and International Guidelines

National and international guidelines for computing and communicating uncertainties in measurement have been developed to ensure that tolerances and other variables that serve to document uncertainty have the same meaning to both buyer and seller. To comply with this objective, it is recommended that Category A (statistical) and Category B (heuristic) uncertainties be computed using the methods documented in References 1 and 2. In addition, bias estimates and bias uncertainties should be computed using the methods documented in Appendix D of Reference 6.

Category A Estimates

Random uncertainties are due to random fluctuations in measurements made with the measuring parameter or the subject parameter. If a sample of *n* measurements yields the values $x_1, x_2, \cdots, x_n$, then the random uncertainty in a measurement is estimated by the sample standard deviation

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} \ , \qquad \text{(A-1)}$$

where

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ . \qquad \text{(A-2)}$$

In cases where the estimate is said to represent the uncertainty in the mean value $\overline{x}$, then the applicable expression is

$$s = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \overline{x})^2} \ . \qquad \text{(A-3)}$$

Degrees of Freedom

The degrees of freedom ν for a random uncertainty estimate, based on a sample of size *n* is given by

$$\nu = n - 1.$$

Category B Estimates

The Category B estimate for an isolated error source is obtained from a heuristically derived probability that errors from the source will be contained within stated limits (also heuristically derived). Errors for a given source are assumed to be normally distributed except in cases where the containment probability is 100%. In these cases, the uniform distribution is assumed over the range of the error limits. For normally distributed errors, contained within symmetric two-sided tolerance offsets $\pm L$ with probability *P*, the relevant equation is

$$u = \frac{L}{\Phi^{-1}\left(\frac{1+P}{2}\right)}, \qquad \text{(A-4)}$$

where Φ^{-1} is the inverse normal distribution function. For uniformly distributed errors, the relevant equation is

$$u = \frac{L}{\sqrt{3}} \ . \qquad \text{(A-5)}$$

Degrees of Freedom

Degrees of freedom for Category B estimates may be obtained using a method documented in Reference 1. The use of this method requires estimating what are essentially confidence limits for the estimated confidence limits. This is a refinement that may be beyond the scope of what can be done in practical situations. Moreover, since Category B estimates are heuristic in nature and, therefore, of a somewhat subjective nature anyway, embellishing a Category B estimate with such a refinement is not overly productive in the first place. Consequently, the degrees of freedom for Category B estimates may be regarded as infinite in most cases without incurring any loss of credibility.

Bias Uncertainty Estimates

Subject parameter and measuring parameter biases are estimated using the method described in Reference 20. The method encompasses cases where a measurement sample is taken by either the subject parameter, the measuring parameter or both. The variables are

$\overline{y}_s =$ sample mean of subject parameter measurements
$s_s =$ subject parameter sample standard deviation
$u_s =$ subject parameter bias uncertainty
$\overline{y}_m =$ sample mean of measuring parameter measurements
$s_m =$ measuring parameter sample standard deviation
$u_m =$ measuring parameter bias uncertainty
$u_{process} =$ process uncertainty .

Subject parameter and measuring parameter biases are estimated according to

$$\text{Subject Parameter Bias} = \frac{(\overline{y}_s - \overline{y}_m)(u_s / u_m)^2}{1 + (u_s / u_m)^2}$$

$$\text{Measuring Parameter Bias} = \frac{(\overline{y}_m - \overline{y}_s)(u_m / u_s)^2}{1 + (u_m / u_s)^2},$$

where

$$u_s = \sqrt{u_{s-bias}^2 + \frac{s_s^2}{n_s} + u_{process}^2}$$

and

$$u_m = \sqrt{u_{m-bias}^2 + \frac{s_m^2}{n_m} + u_{process}^2} \quad .$$

The variable $u_{process}$ is a combination of uncertainties due to process error sources, excluding the random components and the bias uncertainties. The quantities u_{sb} and u_{mb} are *a priori* estimates for subject parameter and measuring parameter bias uncertainties. They are computed from tolerance limit and percent in-tolerance information on the populations of items from which the subject unit and measuring unit are drawn.

For instance, if the percent in-tolerance for the subject parameter population is $100 \times p$ %, the tolerance limits are L_1 and L_2, and the population's probability density function is $f(x,u_{sb})$, then $f(x,u_{sb})$ is determined by inverting the equation

$$p = \int_{L_1}^{L_2} f(x, u_{sb}) dx .$$

Uncertainties in the parameter bias estimates are computed from

$$\text{Subject Parameter Bias Uncertainty} = \frac{(u_s / u_m)^2 u_m}{1 + (u_s / u_m)^2}$$

and

$$\text{Measuring Parameter Bias Uncertainty} = \frac{(u_m / u_s)^2 u_s}{1 + (u_m / u_s)^2} \quad .$$

Uncertainty Combination

Error Sources

Suppose that a quantity y is a function of n measured variables $x_1, x_2, \cdots, x_n$:

$$y = y(x_1, x_2, \cdots, x_n) \ . \tag{A-6}$$

The measured value of each variable x_i is a source of error in the determination of y. The total error in y due to measurements of its constituent variables x_i is written

$$\begin{aligned} \varepsilon(y) &= \sum_{i=1}^{n} \left(\frac{\partial y}{\partial x_i} \right) \varepsilon(x_i) \\ &= \sum_{i=1}^{n} c_i \varepsilon(x_i) \ , \end{aligned} \tag{A-7}$$

where

$$c_i \equiv \left(\frac{\partial y}{\partial x_i} \right) . \tag{A-8}$$

Using Eqs. (1), (3) and (5) of the text, we obtain the variance in y as

$$\sigma_y^2 = \sum_{i=1}^{n} \left(\frac{\partial y}{\partial x_i} \right)^2 \sigma_{x_i}^2 + 2 \sum_{i=1}^{n} \sum_{j>i} c_i c_j \rho_{ij} \sigma_{x_i} \sigma_{x_j} \ . \tag{A-9}$$

Although the variances are ordinarily thought of as statistical operators, their use here is extended to represent whatever processes (statistical or heuristic) are employed to estimate the uncertainties in the variables x_i. With this in mind, we rewrite Eq. (A-9) as

$$u_y^2 = \sum_{i=1}^{n} \left(\frac{\partial y}{\partial x_i} \right)^2 u_{x_i}^2 + 2 \sum_{i=1}^{n} \sum_{j>i} c_i c_j \rho_{ij} u_{x_i} u_{x_j} \ . \tag{A-10}$$

Process Error Components

The measurement of each variable x_i is subject to uncertainty due to the measurement process. Basically, if the error of the *ith* error source is comprised of process error components u_{ir}, $r = 1, 2, \cdots, n_i$, then

$$u_{x_i}^2 = \sum_{r=1}^{n_i} u_{ir}^2 + 2 \sum_{r=1}^{n_i} \sum_{q>r} \rho_{irq} u_{ir} u_{iq} \ . \tag{A-11}$$

In this expression, the correlation coefficients ρ_{irq} describe the correlations between process uncertainty components r and q for the *ith* error source. These components include

- subject parameter bias
- measuring parameter bias

- random error
- resolution error
- data acquisition error
- error due to shipping and handling stress
- error due to environmental factors or ancillary equipment
- other process error.

Expanded Uncertainty

Throughout this paper, the term "uncertainty" has been used interchangeably with the term "standard deviation." Whether obtained statistically or heuristically, uncertainty estimates are variables that can be used as statistics in describing error distributions.

However, many, if not most, statements of uncertainty do not refer to such variables but, instead, refer to limits based on them. These limits serve as boundaries for error containment. In statements like " the value of the parameter is $a \pm b$," the range $\pm b$ is offered as something that is expected to contain errors around the quantity a. The probability of this expectation is called a "confidence level" or "containment probability." The upper and lower values of the range $\pm b$ are called "confidence limits" or "error limits." The range itself is sometimes called a "confidence interval."

Since uncertainty statements often refer to confidence limits rather than uncertainties *per se*, it has been suggested that the terms "confidence limit" and "error limits" be replaced by the term "expanded uncertainty" [1, 2].

Coverage Factors

In recent years, the recognition of the importance of Category B estimates in developing uncertainty statements has prompted some investigators to speak in terms of "coverage" rather than "confidence" [1, 2]. The outcome of this is to calculate expanded uncertainties by multiplying an uncertainty estimate by some arbitrary number that is loosely related to a confidence multiplier. Other than this loose relationship, the number, referred to as a "coverage factor," is devoid of statistical content. What value is has for communicating error containment limits is not completely understood.

Category A Confidence Limits

Determining confidence limits for an uncertainty estimate for errors from a given source is a straightforward and simple exercise. Confidence limits for normally distributed errors are determined under the assumption that the uncertainty estimate is t-distributed with ν degrees of freedom. The degrees of freedom can be obtained from a sample size n according to $\nu = n - 1$.

When an uncertainty estimate is a combination of uncertainty estimates for p sources of error, each with k_r process error components, $r = 1,2, \cdots, p$, the degrees of freedom is determined from the Welch-Satterthwaite formula

$$\nu_r = \frac{u_r^4}{\sum_{i=1}^{k_r} \frac{u_{ri}^4}{\nu_{ri}}},$$

and

$$\nu_t = \frac{u_t^4}{\sum_{r=1}^{p} \frac{u_r^4}{\nu_r}},$$

where u_t is the combined uncertainty estimate and u_r and ν_r, $r = 1,2, \cdots, p$, are the uncertainty estimates and degrees of freedom, respectively, for each of p error sources.

Once the degrees of freedom has been established for an uncertainty estimate u, two-sided confidence limits $\pm b_P$ for an error containment probability P are determined according to

$$b_P = t_{\alpha,\nu} u \ ,$$

where the variable α is given by

$$\alpha = (1+P)/2 \ .$$

Note:

The assumption of a *t*-distributed total uncertainty breaks down if any component of the total process uncertainty is not normally distributed or approximately normally distributed. In this event, the validity of a confidence limit for the total uncertainty is compromised (although the estimate of the uncertainty itself remains valid). The extent of this compromise depends on the number of non-normal components and the departure from normality in each component.

It should be mentioned that, in the overwhelming majority of cases, the confidence limit estimate can be taken to be valid. Nevertheless, general computing methods are currently being developed to enable convolving the distributions of process uncertainty components into a total uncertainty distribution from which generally valid confidence limits can be determined [20].

Category B Estimates

For Category B estimates, the effective degrees of freedom can be set to infinity.[***] For normally distributed errors and a Category B uncertainty estimate u, the confidence limits $\pm b_P$ are determined from

$$b_P = z_P u \ ,$$

where

$$z_P = \Phi^{-1}(1 - P/2) \ .$$

For uniformly distributed errors that are symmetrically distributed around a quantity a and bounded by limits $a \pm L$, the confidence limits are $b_P = PL$.

Appendix B - Risk Analysis and Uncertainty Estimates

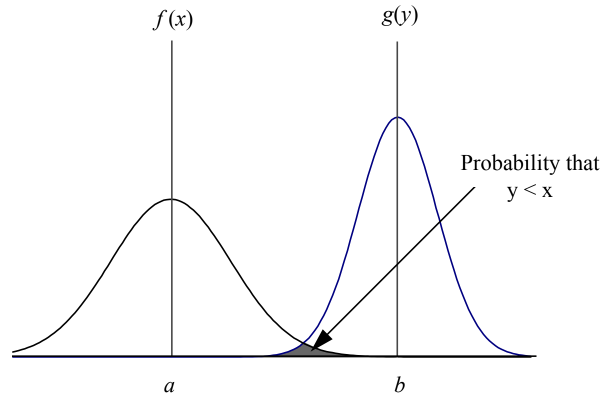
Introduction

In consumer/producer risk analysis, expressions for false accept and false reject risk for normally distributed attributes are computed from joint probabilities functions [6, 8-19]. Until now, this has necessitated numerically integrating cumulative normal distribution functions over ranges of subject attribute or measurement attribute values.

[***] A method for refining estimates of Category B degrees of freedom and employing the *t*-distribution is described in References 1 and 2.

This appendix describes how these numerical integrals can be avoided under certain conditions by substituting cumulative distribution functions developed from a simple convolution of subject parameter and measurement parameter attributes.

Probability Density



Matching Attributes Risk Analysis

Single-Sided Risk

Suppose we have two variables x and y distributed according to

$$f(x) = \frac{1}{\sqrt{2\pi}u_x} e^{-(x-a)^2/2u_x^2} \qquad \text{(B-1a)}$$

and

$$g(y) = \frac{1}{\sqrt{2\pi}u_y} e^{-(y-b)^2/2u_y^2} \ . \qquad \text{(B-1b)}$$

The quantities u_x and u_y are the uncertainty estimates for the values of x and y. These estimates are based on error source uncertainties for x and y that are, in turn, composed of appropriate combinations of process uncertainty estimates.

Imagine that the variable y represents an artifact dimension and that the variable x represents a companion artifact dimension. For instance, x could be the outside diameter of a bolt and y the inside diameter of a nut. The variable (b - a) represents the nominal "play" between nut and bolt. Obviously, the nut will not fit the bolt if $y < x$.

The risk that this will happen is given by

$$\begin{aligned} P(y < x) &= \int_{-\infty}^{\infty} dx f(x) \int_{-\infty}^{x} dy g(y) \\ &= \frac{1}{\sqrt{2\pi}\sigma_x} \int_{-\infty}^{\infty} \Phi\left(\frac{x-b}{u_y}\right) e^{-(x-a)^2/2u_y^2} dx \\ &= 1 - \frac{1}{\sqrt{2\pi}\sigma_x} \int_{-\infty}^{\infty} \Phi\left(\frac{b-x}{u_y}\right) e^{-(x-a)^2/2u_y^2} dx \ . \end{aligned} \qquad \text{(B-2)}$$

We now express the same risk through the use of the distribution of the variable $z = y - x$. By convolving x and y, we obtain this distribution as

$$h(z) = \frac{1}{\sqrt{2\pi}u} e^{-[z-(b-a)]^2/2u^2} \ , \qquad \text{(B-3)}$$

where

$$u^2 = u_x^2 + u_y^2 \ . \qquad \text{(B-4)}$$

Since $P(y < x) = P(z < 0)$, we have

$$\begin{aligned} P(y < x) = P(z < 0) &= \int_{-\infty}^{0} h(z) dz \\ &= 1 - \Phi\left(\frac{b-a}{\sigma}\right). \end{aligned} \qquad \text{(B-5)}$$

Equating (B-5) with (B-2) gives the result

$$\frac{1}{\sqrt{2\pi}u_x}\int_{-\infty}^{\infty}\Phi\left(\frac{b-x}{u_y}\right)e^{-(x-a)^2/2u_x^2}dx=\Phi\left(\frac{b-a}{u}\right). \tag{B-6}$$

Equivalence with Uncertainty Analysis

From Eq. (B-6), it is apparent that setting a confidence level for the risk management of two normally distributed variables with means a and b and uncertainties u_x and u_y is equivalent to setting a confidence level for a normally distributed variable z with mean b - a and variance $u^2 = u_x^2 + u_y^2$.

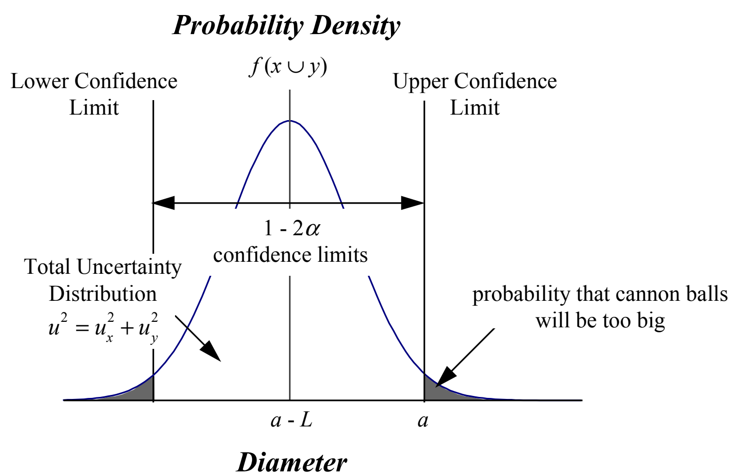
To see how this works, suppose that a manufacturer of ordnance claims that the inside diameter of his cannons follows a $N(a, u_x^2)$ distribution. If we want to make cannon balls for these cannons, we need to be confident that our cannon balls will not get stuck in his barrels. In other words, we need to control to some amount α the probability that the diameter of one of our cannon balls will be larger than one of his inside cannon diameters. Suppose that we settle on $\alpha = 0.01$, i.e., a 99% chance that this won't happen.

To control risks to α, we need to know the uncertainty in our manufacturing and testing process. Assuming that our combined manufacturing and testing error is a $N(0, u_y^2)$ variable, we develop a 1 - α confidence limit according to:

$$L = \Phi^{-1}(1-\alpha)u\,, \tag{B-7}$$

where $u^2 = u_x^2 + u_y^2$.

We next move the mean diameter of our cannon ball manufacturing process to the point $b = a - L$ and get a 1 - α probability that cannon balls will not be too big.



We verify this as follows. Starting with a risk analysis for the variables x and y we write

$$P(y<x)=\int_{-\infty}^{\infty}dxf(x)\int_{-\infty}^{x}dyg(y)$$

$$=1-\frac{1}{\sqrt{2\pi}u_x}\int_{-\infty}^{\infty}\Phi\left(\frac{x-b}{u_y}\right)e^{-(x-a)^2/2u_x^2}dx\,.$$

We now substitute from Eq. (B-6) to get

$$P(y<x)=1-\Phi\left(\frac{b-a}{u}\right)=\Phi\left(\frac{L}{u}\right).$$

Substituting for L from Eq. (B-7) gives

$$\begin{aligned}P(y<x)&=\Phi\left[\frac{\Phi^{-1}(1-\alpha)u}{u}\right]\\&=\Phi\left[\Phi^{-1}(1-\alpha)\right]\\&=1-\alpha \quad \text{(QED)}.\end{aligned}$$

To illustrate, suppose that we want 99% of our cannon balls to fit within cannons whose nominal inside diameter is a. We have an estimate of u_x for the bias uncertainty in cannon diameters and an estimate of u_y for our manufacturing and testing process. We combine these uncertainties in rss to get

$$u = \sqrt{u_x^2 + u_y^2} \quad .$$

Since we want 1 - α = 0.99, we set

$$\begin{aligned} L &= \Phi^{-1}(0.99)u \\ &= 2.32635u \,. \end{aligned}$$

We now move the bias b of our manufacturing process to a - L to get a 99% chance that $y < x$. We verify this from

$$\begin{aligned} P(y < x) &= P(z < 0) \\ &= \frac{1}{\sqrt{2\pi}u} \int_{-\infty}^{0} e^{-[z-(b-a)]^2/2u^2} dz \\ &= \frac{1}{\sqrt{2\pi}u} \int_{-\infty}^{0} e^{-[z-(-L)]^2/2u^2} dz \\ &= 1 - \Phi\left(\frac{-L}{u}\right) = \Phi\left(\frac{L}{u}\right) \\ &= \Phi(2.32635) \\ &= 0.99 \,. \end{aligned}$$

Two-Sided Risk

The foregoing dealt with analyzing the risk of whether a given attribute was smaller (or larger) than another attribute. We now look at the problem of analyzing the probability of whether a given attribute is simultaneously neither smaller nor larger than another.

Let x represent the value of the second attribute and let y represent the value of the first. If the limits L_1 and L_2 bound the acceptance region for the first attribute then the second attribute is acceptable if

$$x - L_1 \le y \le x + L_2 \quad .$$

Suppose that we want to implement the following probabilistic constraints:

$$P(y < x - L_1) = \alpha_L$$

and

$$P(y > x + L_2) = \alpha_U \quad .$$

If x and y are $N(a, u_x^2)$ and $N(b, u_y^2)$ variables, respectively, then

$$\begin{aligned} P(y < x - L_1) &= \int_{-\infty}^{\infty} dx\, f(x) \int_{-\infty}^{x-L_1} g(y) dy \\ &= 1 - \frac{1}{\sqrt{2\pi}u_x} \int_{-\infty}^{\infty} \Phi\left(\frac{x - L_1 - b}{u_y}\right) e^{-(x-a)^2/2u_x^2} dx \qquad \text{(B-8)} \\ &= \alpha_L \,, \end{aligned}$$

and

$$\begin{aligned} P(y > x + L_2) &= \int_{-\infty}^{\infty} dx\, f(x) \int_{x+L_2}^{\infty} g(y) dy \\ &= 1 - \frac{1}{\sqrt{2\pi}u_x} \int_{-\infty}^{\infty} \Phi\left(\frac{x + L_2 - b}{u_y}\right) e^{-(x-a)^2/2u_x^2} dx \qquad \text{(B-9)} \\ &= \alpha_U \,, \end{aligned}$$

We return to Eq. (B-6) and rewrite it as

$$\frac{1}{\sqrt{2\pi}u_x}\int_{-\infty}^{\infty}\Phi\left(\frac{x-b}{u_y}\right)e^{-(x-a)^2/2u_x^2}dx = 1-\Phi\left(\frac{b-a}{u}\right). \tag{B-10}$$

Substituting Eq. (B-10) in Eq. (B-8) gives

$$\alpha_L = \Phi\left(\frac{L_1+(b-a)}{u}\right). \tag{B-11}$$

Likewise, substituting Eq. (B-10) in Eq. (B-9) gives

$$\alpha_U = \Phi\left(\frac{L_2-(b-a)}{u}\right). \tag{B-12}$$

Eqs. (B-11) and (B-12) can be used to solve for tolerance limits for the bias in y:

$$\begin{aligned} b_L &= a - L_1 + \Phi^{-1}(\alpha_L)u \\ b_U &= a + L_2 - \Phi^{-1}(\alpha_U)u\,, \end{aligned} \tag{B-13}$$

where, as before,

$$u^2 = u_x^2 + u_y^2\,.$$

In these expressions, the parameters b_L and b_U are the tolerance limits of the manufacturing process for y. We next need to set a nominal bias b for the process so that the probability $P(b_L \le y \le b_U)$ is maximized:

$$\begin{aligned} \frac{d}{db}P(b_L \le y \le b_U) &= \frac{d}{db}\frac{1}{\sqrt{2\pi}u_y}\int_{b_L}^{b_U} e^{-(y-b)^2/2u_y^2}dy \\ &= \frac{1}{\sqrt{2\pi}u_y}\int_{b_L}^{b_U}\left(\frac{y-b}{u_y}\right)e^{-(y-b)^2/2u_y^2}dy \\ &= \frac{1}{2\sqrt{\pi}}\left[e^{-(b_L-b)^2/2u_y^2} - e^{-(b_U-b)^2/2u_y^2}\right] = 0\,, \end{aligned}$$

which yields, to nobody's surprise,

$$b = \frac{b_L + b_U}{2}\,.$$

To illustrate the use of these expressions, we continue with the cannon/cannon ball problem. In this particular application, we don't want cannon balls to be too small ($y < x - L_1$), because this reduces their range. On the other hand, we don't want cannon balls too large ($y > x + L_2$), because this gets them stuck in the cannon. The first alternative reduces cannon performance, while the second eliminates it. Accordingly, we may want to assign different levels of risk to each.

Suppose that, because of engineering considerations L_1 = 0.100 cm and L_2 = 0.000 cm for a cannon of diameter a = 10 cm. Suppose further that we settle on risks of α_L = 0.05 and α_U = 0.01. Then, since $\Phi^{-1}(0.05) = -1.64485$, and $\Phi^{-1}(0.01) = -2.32635$, we have

$$\begin{aligned} b_L &= [9.900 \text{ cm} - (1.64485)u] \\ b_U &= [10.000 \text{ cm} + (2.32635)u]\,, \end{aligned}$$

and

$$\begin{aligned} b &= \frac{[9.900 \text{ cm} - (1.64485)u] + [10.000 \text{ cm} + (2.32635)u]}{2} \\ &= 9.950 \text{ cm} + 0.3408u\,, \end{aligned}$$

where $u^2 = u_x^2 + u_y^2$.

References

[1] ISO/TAG 4/WG 3, *Guide to the Expression of Uncertainty in Measurement*, June 1992.

[2] Taylor, B. and Kuyatt, C., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Tech. Note 1297, January 1993.

[3] UncertaintyAnalyzer 1.0, Integrated Sciences Group, 1994.

[4] AccuracyRatio 1.1, Integrated Sciences Group, 1992-1994.

[5] *Establishment and Adjustment of Calibration Intervals*, National Conference of Standards Laboratories, Recommended Practice RP-1, 15 November 1989.

[6] *Metrology — Calibration and Measurement Processes Guidelines*, NASA Reference Publication 1342, Jet Propulsion Laboratory, June 1994.

[7] Himelblau, et al., *Guidelines for Dynamic Data Acquisition and Analysis*, Jet Propulsion Laboratory, Pasadena, California, 26 April 1993.

[8] Hayes, J., "Factors Affecting Measuring Reliability," U.S. Naval Ordnance Laboratory, TM No. 63-106, 24 October 1955.

[9] Castrup, H., *Evaluation of Customer and Manufacturer Risk vs. Acceptance Test In-Tolerance Level*, TRW Technical Report No. 99900-7871-RU-00, April 1978.

[10] Ferling, J., *Calibration Interval Analysis Model*, SAI Comsystems Technical Report, Prepared for the U.S. Navy Metrology Engineering Center, Contract N00123-87-C-0589, April 1979.

[11] Kuskey, K., *Cost-Benefit Model for Analysis of Calibration-System Designs in the Case of Random-Walk Equipment Behavior*, SAI Comsystems Technical Report, Prepared for the U.S. Navy Metrology Engineering Center, Contract N00123-76-C-0589, January 1979.

[12] Kuskey, K. and Kraft, R., *An Analysis and Evaluation of Management and Information Systems for the Calibration and Repair of the Navy's Electronic Test Equipment*, DDI Technical Report TR 81-9-177, December 1981.

[13] Ferling, J., "The Role of Accuracy Ratios in Test and Measurement Processes," *Proc. 1984 Measurement Science Conference*, Long Beach, January 19-20.

[14] Weber, S. and Hillstrom, A., *Economic Model of Calibration Improvements for Automatic Test Equipment*, NBS Special Publication 673, April 1984.

[15] Castrup, H., "Navy Analytical Metrology RD&E," *Navy Metrology Research & Development Program Conference Report*, Dept. of the Navy, Metrology Engineering Center, NWS, Seal Beach, Corona Annex, March 1988.

[16] Castrup, H., "Calibration Requirements Analysis System," *Proc. 1989 NCSL Workshop and Symposium*, Denver, July 1989.

[17] Ferling, J. and Caldwell, D., "Analytic Modeling for Electronic Test Equipment Adjustment Policies," *Proc. 1989 NCSL Workshop and Symposium*, Denver, July 1989.

[18] Deaver, D., "How to Maintain Your Confidence," *Proc. 1993 NCSL Workshop and Symposium*, Albuquerque, July 1993.

[19] Deaver, D., "Guardbanding with Confidence," *Proc. 1994 NCSL Workshop and Symposium*, Chicago, July - August 1993.

[20] Integrated Sciences Group R&D efforts in-progress, Integrated Sciences Group, Bakersfield, CA.