Calibration Requirements Analysis System[1]

Howard T. Castrup, Ph. D.
Manager, Advanced Measurement Science Department
Science Applications International Corporation

Abstract
Metrology management and technical personnel have long puzzled over determining appropriate quality assurance standards and practices for effective management of calibration and test equipment. Such standards and practices include accuracy ratio criteria, measurement reliability (percent in-tolerance) targets, test tolerance limits vs. performance tolerance limits, "significant" out-of-tolerance points, and equipment adjustment or renewal policy. This paper reports on recent developments which promise to yield a user capability for establishing these standards and practices in a rigorous and cost effective manner. The analytical methodology is structured in such a way that quality, reliability and cost requirements and parameters at each level in the test and calibration support hierarchy are linked to their counterparts at every other level in the hierarchy by an integrated infrastructure model. This relates requirements and capabilities at any given level in the hierarchy to the performance objectives of the end items which the hierarchy is established to ultimately support. Through use of the integrated model, the impact of decisions at any level on test decision risks, calibration intervals and support costs at other levels in the hierarchy is computed. Included in the model is the capability to quantify the effect of calibration/test support quality on costs resulting from the risk of degraded end item performance, either in terms of loses suffered through poor performance or expenses incurred from returned products, warranty rework or reimbursement, legal damages, or retrofit of product improvements.

This work was accomplished under contract to the Naval Weapons Station, Seal Beach, Metrology Engineering Center.

Introduction

Test and calibration infrastructures are characterized by a number of technical and management parameters. These parameters include calibration system, test system and end item performance tolerances; calibration system and test system calibration intervals; test intervals for fielded end items; accuracy ratios between calibration systems and test systems and between test systems and end items; measurement reliability (percent in-tolerance) targets; acceptable false alarm (reporting in-tolerance equipment as out-of-tolerance) rates and missed fault (reporting out-of-tolerance equipment as in-tolerance) rates; equipment renewal policies; and the like.

Over the past fifteen years, the problem of analyzing and evaluating such technical parameters has received considerable study by several investigators [1-9]. The results of this work have been expressed in various papers and technical reports in the form of tabulated or plotted comparisons and trends. While useful in the specific applications treated by each author, these tables and plots are usually of little value in general application. This is because the specifics of a given test and calibration support scenario typically take on a multitude of possible combinations which require the development of tables and/or plots tailored to the combination or combinations of interest.

What is needed are tools by which a non-specialist can enter information defining his or her specific infrastructure, provide relevant technical and cost data, punch a button and receive technical and management parameter values and other decision support outputs by which management of the infrastructure can be evaluated and optimized.

Such a tool has been in development and is now available in a user-interactive PC based prototype decision support system. This system incorporates the results of prior work as well as several new concepts and developments which have emerged from recent research. It incorporates numerous suggestions and various key developments by Navy Metrology Engineering Center technical personnel and by SAIC personnel, subcontractors and consultants under

---

[1] Presented at the 1989 NCSL Workshop & Symposium, Denver, July 9-13, 1989.

contract to the Metrology Engineering Center. Because the system can be used as an analytical tool for defining and adjusting calibration, test and prime system tolerances, among other things, it has been named the Equipment Tolerancing System (ETS). The concepts, methodologies and assumptions on which the ETS is built are described in this paper.

The development of the theory and practice of test and calibration has been traditionally based on the now classical concept of straightforward comparisons between simple devices, each characterized by a single measurement attribute or parameter. Prior to the late `60s, the accuracy of higher-level devices or standards was generally significantly greater than the devices being tested or calibrated. Under such idyllic circumstances, all that was required for ensuring adequacy of testing or calibration was to legislate that only standards whose accuracy exceeded some easily attainable minimum level relative to the device under test could be employed.

In the modern technological environment of optical data encryption and retrieval, Mach 3+ missiles, and nuclear power plants, the classical concept of test and calibration is no longer generally applicable. The pressures of the competitive international marketplace and of national aerospace, energy, environmental and defense performance capability and reliability requirements have led to a situation in which end item performance tolerances rival and in some cases exceed the best accuracies attainable at the primary standards level. In such cases, the interpretation and management of test or calibration data and systems requires that the subtleties of the test/calibration process be accounted for and their impact on end item performance and support costs be quantified.

These subtleties are manifested in a number of statistics, the study of which is referred to as *analytical metrology*. The statistics which deal with analyzing and evaluating the impact of the test and calibration hierarchy on end item quality, performance capability and support costs are described in this paper. The terms which are relevant to this description are defined below in alphabetical order.

accuracy ratio - the ratio of the performance tolerance limits of a UUT measurement attribute to the performance tolerance limits of a corresponding TS measurement attribute.

adjustment limits - specified upper and lower limits outside of which a measurement attribute is considered to require adjustment.

average over period (aop) - refers to the average value of a quantity, where the average is computed between bop and eop reference points.

beginning of period (bop) - the beginning of a calibration or test interval, referenced to the time of return to the user of a UUT following test or calibration.

center spec - see *nominal value*.

end item - a consumer product, military system, or other item which fulfills a requirement other than the test or calibration of another system within the test and calibration hierarchy.

end of period (eop) - the end of a calibration or test interval, referenced to the time of shipping of a UUT for test or calibration to the TS facility.

in-tolerance - a condition in which a measurement attribute value lies within the specified tolerance limits for the attribute.

measurement attribute - a measurable individual equipment parameter characterized by performance specifications.

measurement reliability - the probability that a measurement attribute of an item of equipment is in conformance with performance specifications.

measurement reliability target - a specified measurement reliability objective commensurate with quality, cost and logistic objectives.

nominal value - the value of a measurement attribute corresponding to optimal performance.

performance tolerance limits - specified upper and lower limits for a measurement attribute which bound attribute values corresponding to acceptable performance.

rebound - the response of a measurement attribute to renewal.

renewal - the process of adjustment of a measurement attribute to its nominal value.

shipping - handling and transportation of a UUT either from the user to the TS facility or from the TS facility to the user.

test accuracy ratio (TAR) - the ratio of the test tolerance limits of a UUT measurement attribute to the performance tolerance limits of a corresponding TS measurement attribute.

test limits - specified upper and lower limits for a measurement attribute which bound acceptable measured values of the attribute obtained during test or calibration.

test process - the procedure, environment and application of the test or calibration activity, including the test/calibrating personnel, the physical ancillary equipment, the test/calibration environment temperature and humidity, the mechanical and other stresses experienced by the UUT during test/calibration, etc.

test system (TS) - an item of test or calibrating equipment used to test or calibrate an end item or another item of test or calibrating equipment.

unit under test (UUT) - an end item or test system submitted for test or calibration.

Analytical metrology links each level of the test and calibration support hierarchy in an integrated model by describing each level of the hierarchy in terms of the support it gives to the next highest level and the support it receives from the next lowest level (see Fig. 1). For any given level, the support given to the next highest level is measured in terms of several parameters. These are:

- bop measurement reliability
- length of UUT test or calibration interval
- probability of incorrectly reporting out-of-tolerance UUT as in-tolerance
- probability of incorrectly reporting in-tolerance UUT as out-of-tolerance
- UUT aop measurement reliability
- UUT availability
- cost of test, calibration and repair
- cost of rejection (with consequent adjustment, repair or rework and downtime) of in-tolerance UUT
- cost of acceptance of tested/calibrated UUT.

Of these, the cost of acceptance of tested/calibrated UUT is a new concept developed in the course of recent RD&E efforts [9,10]. This and related concepts will be discussed in detail under cost modeling later.

The support received from the next lowest level is measured in terms of the following parameters:

- TS bop measurement reliability

- length of TS test or calibration interval
- TS aop measurement reliability
- TS availability
- cost of test, calibration and repair of the TS

These parameters connect from one level of the hierarchy to the next in a contiguous sequence. Hence, any change in one or more of these parameters at any given level affects the parameters at other levels within the hierarchy. This fact makes possible the development of methods and techniques which enable the analysis of costs and benefits in such a way that both summary results for the entire hierarchy and detailed visibility at each level are provided.

The Test and Calibration Support Hierarchy

A simplified diagram of the test and calibration support hierarchy is shown in Figure 1. In the hierarchy, the end item is placed at the top of the chain. Below the end item is the test system and below the test system is a series of calibration systems, culminating in a primary calibration system (e.g., NIST), labeled calibration system 1.

END ITEM

TS workload ↓ ↑ test results/adjusted end items

TEST SYSTEM

CS workload ↓ ↑ calibration results/adjusted TS items

CALIBRATION SYSTEM *n*

CS workload ↓ ↑ calibration results/adjusted CS items

CALIBRATION SYSTEM *n*-1

CS workload ↓ ↑ calibration results/adjusted CS items

.
.
.

CALIBRATION SYSTEM 2

CS workload ↓ ↑ calibration results/adjusted CS items

CALIBRATION SYSTEM 1

Figure 1. The test and calibration hierarchy. With the exception of the test system to end item level, all levels can be considered equivalent with respect to mode of operation. The test system to end item level differs from the other levels in that additional cost factors arise in the treatment of end items. These are discussed under Cost Modeling later in the paper. At lower levels, calibration systems which are classified as reference or transfer (i.e., "fixed-value") standards can be considered as calibration equipment which are set to center spec at calibration via the reported attribute values that accompany reports of their calibration. This is because reported values for fixed-value standards function equivalently to center spec adjustments of calibration or test systems.

Testing of a given end item measurement attribute by the test system yields a reported in-or out-of-tolerance indication (referenced to the end item test tolerance limits), an adjustment (referenced to the end item adjustment limits) and a bop measurement reliability (referenced to the end item performance tolerance limits). Similarly, the results of calibration of the test system attribute are a reported in- or out-of-tolerance indication (referenced to the test system test limits), an attribute adjustment (referenced to the test system adjustment limits) and a bop measurement reliability (referenced to the test system performance limits). The same sort of data result from calibration of the calibration system and accompany calibrations down through the hierarchy until a point is reached where the UUT of interest is a calibration standard.

It should be noted that in many applications, end items are not tested at designated periodic intervals. In military weapon system applications, for example, end item testing often arises in response to detected operational failure or may be performed prior to use. In such cases, the end item test interval may be thought of as the *average* time elapsed between tests. In commercial applications, end item testing may take the form of receiving inspection of purchased equipment. In these cases, the end item test interval can be regarded as the duration between factory testing and customer testing.

Ordinarily, calibration standards are not managed to specified performance or test tolerances and reported as in- or out-of-tolerance per se, but instead receive a reported measured value, accompanied by confidence limits. Since such standards are not managed to specified tolerances, a statement of bop measurement reliability is apparently not applicable. In addition, the treatment of calibration standards differs from that of calibration or test systems in that calibration standards' measurement attribute values are reported rather than adjusted.

These observations appear to set calibration of standards apart from other test or calibration scenarios. With regard to reported attribute values in place of adjustments, however, such reports can be considered to be completely equivalent to non-intrusive adjustments to nominal in that reported values are used as nominal values until the next calibration. Additionally, the lack of specified tolerances for calibration standards will likely be eliminated in future calibration standard management systems. This is due to the fact that such standards are assigned calibration intervals, which can be optimized only if specified tolerances accompany reports of calibration. Specifically, the calibration standard attribute's reported measured value needs to be accompanied by both a set of limits (i.e., performance specifications) which are expected to contain the attribute value over the course of the calibration interval and an estimate of the probability that this expectation will be realized (i.e., a measurement reliability target). The methodology presented herein assumes that this practice will be followed.

From the foregoing, with regard to the applicability of the reliability analysis methodology described in this paper, test or calibration at any pair of successive levels in the hierarchy is equivalent to test or calibration at any other pair of successive levels. This is not true, however, for cost modeling and analysis. For cost modeling end item testing is treated somewhat differently than other tests or calibrations in that the cost consequences of □□accepting □□end items as having passed testing can be evaluated in intrinsic terms. This will be elaborated on later.

BOP Measurement Reliability - The Test Process

At any two consecutive levels of the test/calibration hierarchy, both the unit under test or calibration (UUT) and the test system (TS) are assumed to be drawn randomly from their respective populations. The UUT and TS attribute values of interest are assumed to be normally distributed with zero population means and with variances (uncertainties) which grow with time elapsed since prior testing and/or adjustment. UUT attribute adjustments are assumed to be made using TS attributes as reference values.

Attribute values are taken to be toleranced with two-sided performance specifications and to be assigned associated two-sided test tolerance and adjustment limit specifications; if the "true" value of the UUT attribute is represented by x, and its value as reported by measurement with the TS is represented by y, then the performance, test and adjustment specifications are defined as follows:

$-L_{per} \le x \le L_{per}$	:	UUT attribute is in-tolerance
$-L_{test} \le y \le L_{test}$	:	UUT attribute is observed (reported) in-tolerance
$y \le -L_{adj}$ or $L_{adj} \le y$	:	observed value of the UUT attribute is adjusted to center spec using the TS attribute as a reference.

UUT items are assumed to be tested at periodic intervals, referred to as test or calibration intervals. The start of each interval is termed the "beginning of period" (bop), and the end of each interval is called the "end of period" (eop). The beginning of period starts upon receipt of the UUT by its user, and the end of period is marked at the point where the UUT is sent for test by the user facility. Hence, the random selection of items from the UUT population is referenced to eop. This is in contrast with the random selection of TS items used to test UUT items. The former are assumed to be drawn from their populations at random times within their TS calibration interval. Consequently, the distribution of TS attributes is referenced to average-over-period (aop) values.

At this point, it is worthwhile to note that the test or calibration interval is a quantity which can adopt three identities. From the standpoint of UUT availability to the user, it is the elapsed time between a given bop date and the successive eop date. From the standpoint of recall of the UUT for test or calibration, it is the time elapsed between successive bop dates. From the standpoint of the TS facility, it is the time elapsed between successive test or calibration dates. In the bulk of the analytical treatment presented in this paper, the interval will be taken to be synonymous with the period of time the UUT is available for use. Effects of shipping and storage are accounted for by the introduction of a variable labeled σ_s, defined below. For most discussions, this constitutes a good approximation. An exception is the analysis of equipment availability, discussed later. In availability analysis, all factors contributing to downtime are examined.

The test process is characterized by several sources of uncertainty, which are quantified by the following set of standard deviations:

σ_{eop} = the true standard deviation of UUT attribute values prior to shipping to the TS facility (i.e., at end of period).

σ_s = the contribution to the UUT "as received" standard deviation due to shipping stresses (set to zero if the UUT is not shipped to the TS facility).

σ_{ts} = the true standard deviation of TS attribute values at time of test or calibration. If random demand of TS items is assumed this is set equal to the aop value of the TS attribute standard deviation. Determination of aop values is discussed later.

σ_{tp} = the standard deviation of the test process.

As a result of UUT testing, we "observe" an eop measurement reliability given by

$$R_{obs} = 2F\left(\frac{L_{test}}{\sigma_{obs}}\right) - 1 \qquad (1)$$

where F(·) is the cumulative distribution for the normal distribution, and where

$$\sigma_{obs}^2 = \sigma_{eop}^2 + \sigma_s^2 + \sigma_t^2 \ , \qquad (2)$$

and

$$\sigma_t^2 = \sigma_{ts}^2 + \sigma_{tp}^2 \ . \qquad (3)$$

The reliability at eop is given by

$$R_{eop} = 2F\left(\frac{L_{per}}{\sigma_{eop}}\right) - 1 \ , \qquad (4)$$

where

$$\sigma_{eop}^2 = \sigma_{obs}^2 + \sigma_t^2 + \sigma_s^2 \qquad (5)$$

The true measurement reliability at time of test or calibration is given by

$$R_{true} = 2F\left(\frac{L_{per}}{\sigma_{true}}\right) - 1 \ , \qquad (6)$$

where

$$\sigma_{true}^2 = \sigma_{eop}^2 + \sigma_s^2 = \sigma_{obs}^2 - \sigma_t^2 \qquad (7)$$

UUT items are tested to test tolerance limits and adjusted to adjustment limits. Adjustment limits are set in accordance with the renewal policy of the TS facility. There are three main renewal policy categories:

$$L_{adj} = L_{test} \quad : \quad \text{Renew if "failed" only}$$

$$L_{adj} = 0 \quad : \quad \text{Renew always}$$

$$0 < L_{adj} < L_{test} \quad : \quad \text{Renew "as needed."}$$

In the foregoing, the term "renew" refers to adjustment of the UUT attribute value to nominal or center spec. UUT attribute adjustment may consist of a physical adjustment or may take the form of a correction factor. In some instances, UUT attribute adjustment may result in resetting the attribute value to a quasi-stable point, subject to gradual drift, random fluctuation or response to external stress. In these cases, a renew always or renew as-needed policy is often preferred. Alternatively, for certain types of UUT attributes, adjustment may result in resetting the attribute value to an unstable point from which the UUT will attempt to spontaneously revert or "rebound." The latter behavior contributes an additional source of uncertainty characterized by

σ_{rb} = the standard deviation due to reversion or rebound of UUT attributes away from values set as a result of adjustment.

In these cases, a renew if failed only policy is often the best choice.

Regardless of renewal policy, UUT items are received by the TS facility with attributes distributed according to the pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_{true}} e^{-x^2/2\sigma_{true}^2} \ , \tag{8}$$

and are tested with TS items yielding observed attribute values distributed according to

$$f(y\,|\,x) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-(y-x)^2/2\sigma_t^2} \ . \tag{9}$$

As a result of the test process, UUT items are delivered to the user with a measurement reliability reflecting the quality of the test process. In general, the higher the bop measurement reliability, the longer a UUT item can remain in use before subsequent testing is required. Consequently, determination of this quantity is an important aspect of the test quality management process. Accordingly, we seek to determine the distribution of UUT attribute values following test and adjustment. This "post test" distribution is given by

$$f_{pt}(x) = f(x\,|\,\text{not adjust})P(\text{not adjust}) + f(x\,|\,\text{adjust})P(\text{adjust}) \ , \tag{10}$$

where $f(x|E)$ is the pdf for x given the event E has taken place, and $P(E)$ is the probability that E has occurred. The first component of the RHS of (10) is obtained using the Bayes' relation

$$f(x\,|\,\text{not adjust})P(\text{not adjust}) = f(\text{not adjust}\,|\,x) f(x) \ . \tag{11}$$

The pdf $f(x)$ is given in Eq. (8). The pdf $f(\text{not adjust}|x)$ is readily obtained from Eq. (9), using the definition of adjustment limits:

$$\begin{aligned} f(\text{not adjust}\,|\,x) &= \int_{-L_{adj}}^{L_{adj}} f(y\,|\,x)dy \\ &= F\left(\frac{L_{adj}+x}{\sigma_t}\right) + F\left(\frac{L_{adj}-x}{\sigma_t}\right) - 1 \, . \end{aligned} \tag{12}$$

The pdf $f(x|\text{adjust})$ is given by

$$f(x\,|\,\text{adjust}) = \frac{1}{\sqrt{2\pi(\sigma_t^2+\sigma_{rb}^2)}} e^{-x^2/2(\sigma_t^2+\sigma_{rb}^2)} \ , \tag{13}$$

where rebound from adjustment has been included. The probability $P(\text{adjust})$ is equal to 1 - $P(\text{not adjust})$:

$$
\begin{aligned}
P(\text{adjust}) &= 1 - \int_{-\infty}^{\infty} dx\, f(x) \int_{-L_{adj}}^{L_{adj}} dy\, f(y\,|\,x) \\
&= 2\left[1 - F\left(\frac{L_{adj}}{\sqrt{\sigma_{true}^2 + \sigma_t^2}}\right)\right] - 1\,,
\end{aligned} \tag{14}
$$

Combining Eqs. (11) - (14) in Eq. (10) gives

$$
f_{pt}(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi(\sigma_t^2 + \sigma_{rb}^2)}} e^{-x^2/2(\sigma_t^2+\sigma_{rb}^2)}\,, & \text{renew always} \\[2ex] \phi(x)\dfrac{e^{-x^2/2\sigma_{true}^2}}{\sqrt{2\pi}\sigma_{true}} + K\dfrac{e^{-x^2/2(\sigma_t^2+\sigma_{rb}^2)}}{\sqrt{2\pi(\sigma_t^2+\sigma_{rb}^2)}}\,, & \text{otherwise}\,, \end{cases} \tag{15}
$$

where

$$
\phi(x) \equiv F\left(\frac{L_{adj} + x}{\sigma_t}\right) + F\left(\frac{L_{adj} - x}{\sigma_t}\right) - 1\ , \tag{16}
$$

and

$$
K \equiv 2\left[1 - F\left(\frac{L_{adj}}{\sqrt{\sigma_{true}^2 + \sigma_t^2}}\right)\right]\,. \tag{17}
$$

Since the beginning of period reliability is referenced to the point of return of the UUT to the user, the effects of shipping need to be considered. This is done in accordance with the following assumptions:

1) Stresses due to shipping occur in such a way that responses of measurement attributes toward increasing values occur with equal probability to responses toward decreasing values.

2) Stresses due to shipping occur in such a way that measurement attribute responses are random in magnitude.

3) Stresses due to shipping occur at some average rate r.

4) Shipping requires some average duration of time τ.

Given these assumptions, responses due to shipping are seen to follow the classic random walk behavior. Letting the variable ζ represent the value of the measurement attribute following shipping, the pdf for ζ can be expressed as

$$
q(\varsigma\,|\,x) \equiv \frac{e^{-(\varsigma - x)^2/2\sigma_s^2}}{\sqrt{2\pi}\sigma_s}\ , \tag{18}
$$

where $\sigma_s^2 = \langle \varsigma^2 \rangle r\tau$. The bop measurement reliability is given by

$$
R_{bop} = \int_{-\infty}^{\infty} dx\, f_{pt}(x) \int_{-L_{per}}^{L_{per}} d\varsigma\, q(\varsigma\,|\,x)\ , \tag{19}
$$

With renew if failed only and renew as-needed policies, Eq. (19) is solved numerically. For the renew always adjustment policy, Eq. (19) can be solved in closed form:

$$
R_{bop} = 2F\left(\frac{L_{per}}{\sqrt{\sigma_t^2 + \sigma_{rb}^2 + \sigma_s^2}}\right) - 1 \quad \text{(renew always)}\ . \tag{20}
$$

Interval Adjustment

It is assumed that a goal of effective management of UUT equipment is the attainment of a measurement reliability objective or set of objectives consistent with end item quality or performance goals. Such measurement reliability objectives, expressed in terms of the probability that a UUT item is performing within its performance tolerance limits over its test or calibration interval, are typically met by setting test or calibration intervals in such a way that a minimum percentage of UUT are received in-tolerance for calibration at eop. These minimum percentages are referred to as *measurement reliability targets*.

For purposes of discussion, it will be assumed that either some level of observed measurement reliability, R_{obs}, or some measurement reliability target R^* is known or projected which corresponds to a test or calibration interval I and a set of tolerance limits, $\pm L_{per}$, $\pm L_{test}$ and $\pm L_{adj}$.

Immediately following test or calibration, the value of a UUT attribute is localized to a neighborhood of values defined by the accuracy of the associated TS and the uncertainty of the test or calibration process. As time elapses from the point of test or calibration, the UUT experiences various stresses due to transportation, storage, use, etc. These stresses contribute to a growing lack of confidence that the neighborhood of values contains the true value of the UUT attribute. This uncertainty growth is depicted in Figure 2.

Let the UUT measurement reliability at some time t be denoted $R(t)$. Since test or calibration intervals are set to achieve $R(I) = R_{obs} = R^*$, any change which brings about either a change in R^* or in R_{obs} will require a change in I as follows:

$$R^* \to R^{*\prime} \Rightarrow I \to I' : R(I') = R^{*\prime} \ ,$$

or

$$R_{obs} \to R'_{obs} \Rightarrow I \to I' : R'_{obs} = R^* \ .$$

From this simple scheme, it can be seen that an interval change is in order if either the UUT measurement reliability target is changed or if the observed UUT measurement reliability is altered. In general, if the interval I is held constant, the observed measurement reliability of an item of equipment is altered if either the item is altered in some physical way or if its in-tolerance criteria are altered. Physical equipment alterations are handled by a redefinition of the various parameters which govern measurement uncertainty growth over time. Such alterations are not explicitly covered in this paper. Alteration of in-tolerance criteria are manifested in changes of L_{per}, L_{test} and L_{adj}.

Interval changes in response to measurement reliability target changes and alterations in tolerance limits are discussed below.

Adjustment to Reliability Target Changes

Two alternative mathematical functions are used in this paper to model $R(t)$. One function assumes a constant out-of-tolerance rate and the other assumes that measurement attribute values fluctuate randomly with respect to direction and magnitude and with equal probability for positive and negative fluctuations. The former is labeled the *exponential model* and the latter is called the *random walk model*.

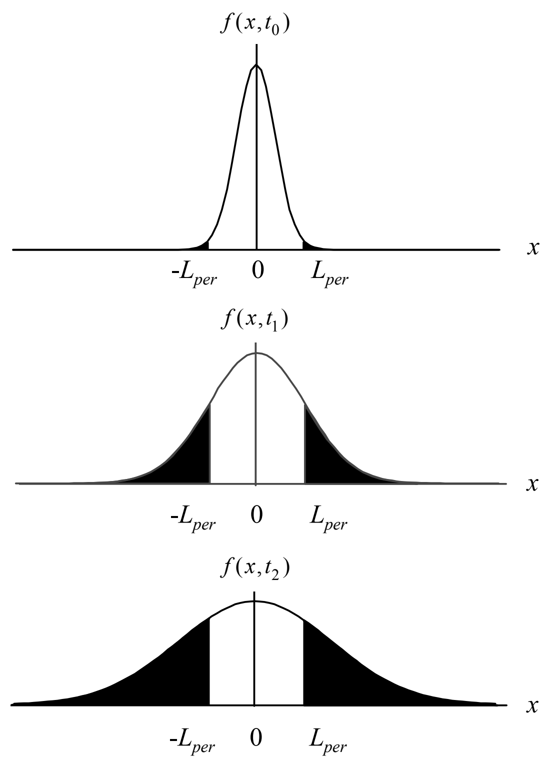

Figure 2. Uncertainty growth of a measurement attribute value over time elapsed since test or calibration. The shaded areas represent the probability that the attribute is out-of-tolerance.

Exponential Model

If the measurement reliability of an item is characterized by a constant out-of-tolerance rate, λ, the measurement reliability in effect after an interval I is given by

$$R_{eop} = R_{bop} e^{-\lambda t} . \tag{21}$$

From which

$$\lambda = -\frac{1}{I} \ln\left(\frac{R_{eop}}{R_{bop}} \right) . \tag{22}$$

Using Eq. (4) in (22) gives

$$\lambda = -\frac{1}{I} \ln\left[\frac{2F(L_{per} / \sigma_{eop}) - 1}{R_{bop}} \right] , \tag{23}$$

where R_{bop} is obtained using Eq. (19) or (20), and σ_{eop} is given in Eq. (5). The quantity σ_{obs} is obtained from Eq. (1):

$$\sigma_{obs} = \frac{L_{test}}{F^{-1}\left[(1 + R_{obs}) / 2 \right]} . \tag{24}$$

Now suppose the reliability target is changed to $R^{*\prime}$. A new interval I' is set as follows. As before,

$$\sigma'_{obs} = \frac{L_{test}}{F^{-1}\left[(1 + R^{*\prime}) / 2 \right]} , \tag{25}$$

and, from Eqs. (21) and (22),

$$\begin{aligned} R'_{eop} &= R'_{bop} \exp\left[\frac{I'}{I} \ln\left(\frac{R_{eop}}{R_{bop}} \right) \right] \\ &= 2F\left(\frac{L_{per}}{\sigma'_{eop}} \right) - 1 , \end{aligned} \tag{26}$$

where

$$(\sigma'_{eop})^2 = (\sigma'_{obs})^2 - \sigma_t^2 - \sigma_s^2 , \tag{27}$$

and R'_{bop} is as given in Eq. (19) or (20) with σ'_{true} in place of σ_{true} in Eqs. (15) - (17). The quantity σ'_{true} is obtained as in Eq. (7):

$$(\sigma'_{true})^2 = (\sigma'_{eop})^2 + \sigma_s^2 . \tag{28}$$

Solving for I' in Eq. (26) gives

$$I' = I \frac{\ln\left\{ \left[2F(L_{per} / \sigma'_{eop}) - 1 \right] \big/ R'_{bop} \right\}}{\ln\left(R_{eop} / R_{bop} \right)} , \tag{29}$$

with R_{eop} given by Eq. (4) and R_{bop} given in Eq. (19) or (20).

Random Walk Model

With the random walk model, the variance in the attribute value of interest (prior to shipping) is a linear function of the elapsed interval I:

$$\begin{aligned} \sigma_{eop}^2 &= \sigma_{bop}^2 + \alpha I \\ &= \sigma_{true}^2 - \sigma_s^2 , \end{aligned} \tag{30}$$

where the coefficient α is a constant dependent only on the measurement attribute's inherent stability. Eq. (30) will be used to determine a new interval I' in response to a reliability target change from R to R'.

The first step is to compute a new value for σ'_{obs} using Eq. (25), and σ'_{eop} using Eq. (27). Next, R'_{bop} is calculated using Eq. (19) or (20) with σ'_{true} in Eqs. (15) - (17). From this, σ'_{bop} is computed according to

$$\sigma'_{bop} = \frac{L_{per}}{F^{-1}\left[(1+R'_{bop})/2\right]} \ . \tag{31}$$

Finally, I' is calculated using Eq. (30):

$$I' = \frac{1}{\alpha}\left[(\sigma'_{true})^2 - (\sigma'_{bop})^2 - \sigma_s^2\right] \ . \tag{32}$$

Adjustment to Tolerance Limits Changes

An alteration of an item's UUT performance limits results in a redefinition of the standard by which the item is judged in- or out-of-tolerance. This is shown in Figure 3.

Such a redefinition results in changes in R_{bop}, $R(t)$ and $R(I)=R_{eop}$. In addition, performance tolerance limit changes are normally accompanied by test tolerance limit changes and adjustment limit changes. The former impacts R_{obs} and the latter affects R_{bop} and support costs in terms of increased or decreased numbers of equipment adjustments performed.

In order to maintain measurement reliability objectives, such changes necessitate an alteration in $I \to I'$ such that

$$R(I') = R(I) = R^* \ , \tag{33}$$

where R is referenced to L_{per}, L_{test} and L_{adj}, and R' is referenced to L'_{per}, L'_{test} and L'_{adj}. This alteration is discussed for UUT governed by the exponential and random walk measurement reliability models.

Exponential Model

If the UUT attribute performance tolerance is modified, the standard by which its measurement reliability is defined is also modified. This results in a change in the attribute out-of-tolerance rate. For the exponential model, the out-of-tolerance rate is given by the parameter λ. Hence, performance tolerance limit changes result in changes in λ and corresponding changes in calibration interval.

To compute a new λ and a new interval I', given a performance limit change from L_{per} to L'_{per}, we imagine the following sequence of events to occur:

1) The UUT is received for calibration at the end of an interval I. The UUT is tested/adjusted using L'_{test}/L'_{adj}.
2) The performance tolerance limits are changed from $\pm L_{per}$ to $\pm L'_{per}$.
3) The UUT is delivered to the user with the new performance limits.
4) The UUT is again recalled at the end of I, at which point the measurement reliability is equal to R' (before shipping).
5) A new λ is calculated.

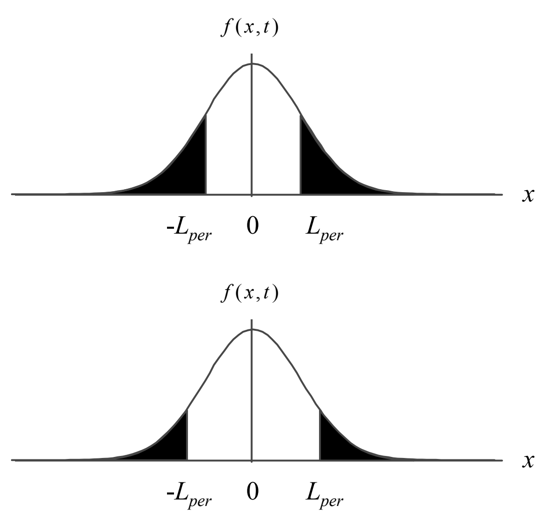

Figure 3. Change in out-of-tolerance criteria. The sum of the shaded areas in the upper figure represents the out-of-tolerance probability for a given distribution of attribute values under the original performance specifications. The sum of the shaded areas in the lower figure represents the out-of-tolerance probability for the same distribution under modified performance specifications.

6) The test tolerance limits are changed from $\pm L_{test}$ to $\pm L'_{test}$, and the adjustment limits are changed from $\pm L_{adj}$ to $\pm L'_{adj}$ (these changes are optional but normally accompany a performance tolerance change).

7) A new interval, I', is calculated.

At step 1 above, the observed measurement reliability is given by Eq. (1), from which σ_{obs} is computed using Eq. (24). Using Eqs. (2) and (24), a value for σ_{eop} is calculated. In this calculation, the quantities σ_t and σ_s are known.

At steps 2 and 3, the beginning of period measurement reliability is given as in Eq. (19) or (20)

$$R'_{bop} = \int_{-\infty}^{\infty} dx\, f_{eop}(x) \int_{-L'_{per}}^{L'_{per}} d\varsigma\, q(\varsigma \mid x) \ , \tag{34}$$

where the pdfs are as defined in (15) - (18).

At step 4, the measurement reliability is obtained with the aid of Eqs. (21) and (4):

$$R'_{eop} = R'_{bop} e^{-\lambda' I} = 2F\left(\frac{L'_{per}}{\sigma_{eop}}\right) - 1 \ , \tag{35}$$

from which the new out-of-tolerance rate is given by

$$\lambda' = -\frac{1}{I} \ln\left\{\left[2F(L'_{per} / \sigma_{eop}) - 1\right] \Big/ R'_{bop}\right\} \ . \tag{36}$$

At step 6, new test tolerance limits and adjustment limits are determined. These changes necessitate calculation of a new beginning of period measurement reliability R''_{bop}. This is accomplished by employing Eq. (1) and (15) - (19) or (20) with L'_{per}, L'_{test}, L'_{adj} and R'_{bop} in place of their unprimed counterparts.

At step 7, a new interval is calculated:

$$I' = -\frac{1}{\lambda'} \ln\left\{\left[2F(L'_{per} / \sigma'_{eop}) - 1\right] \Big/ R''_{bop}\right\} \ , \tag{37}$$

where σ'_{eop} is given in Eq. (27).

Since the calibration interval I was presumably managed to achieve a value of R_{obs} equal to the desired target measurement reliability, it is assumed that the observed measurement reliability will be unchanged from its original value. Given this assumption, we have from Eq. (24),

$$\sigma'_{obs} = \frac{L'_{test}}{F^{-1}\left[(1 + R_{obs})/2\right]} \ . \tag{38}$$

Random Walk Model

Unlike the previous calculations for the exponential model, the adjusted new interval I' can be determined for the random walk model by converting to L'_{per}, L'_{test} and L'_{adj} up front. In computing the new interval, σ'_{obs} and σ'_{eop} are computed using Eqs. (37) and (27), respectively. Next, R'_{bop} is obtained using Eqs. (15) - (19) or (20) with L'_{per}, σ'_{true} and L'_{adj} in place of L_{per}, σ_{true} and L_{adj}. The beginning of period standard deviation is next calculated using

$$\sigma'_{bop} = \frac{L'_{per}}{F^{-1}\left[(1 + R'_{bop})/2\right]} \ , \tag{39}$$

and I' is obtained using this result with Eqs. (27) and (37) in Eq. (32).

Test Decision Risk

Implied in the foregoing treatment is the recognition that, given that test or calibration systems and processes are imperfect, the *true* condition of a UUT may not necessarily match its *apparent* condition observed and recorded as a result of test or calibration. The discrepancy between true condition and observed/reported condition is referred to as *test decision risk*. We discuss this risk in terms of *true vs reported measurement reliability* and in terms of the probability for *false alarms* (in-tolerance items reported out-of-tolerance) and *missed faults* (out-of-tolerance items reported in-tolerance).

True vs Reported Measurement Reliability

The discrepancy between the true eop measurement reliability of an item of UUT and its observed/reported measurement reliability is measured in terms of the discrepancy between the probability that the item is truly in-tolerance at the time of test/calibration and the probability that it is *observed* within its test tolerance limits during test/calibration, i.e., the probability that it "passes" test or calibration [5]. These quantities are, respectively, given by

$$\begin{aligned} P(\text{in-tolerance}) &= \int_{-L_{per}}^{L_{per}} f(x)dx \\ &= 2F\left(\frac{L_{per}}{\sigma_{true}}\right) - 1\,, \end{aligned} \tag{38}$$

and

$$\begin{aligned} P(\text{pass test}) &= \int_{-\infty}^{\infty} f(x)dx \int_{-L_{test}}^{L_{test}} f(y\,|\,x)dy \\ &= 2F\left(\frac{L_{test}}{\sqrt{\sigma_{true}^2 + \sigma_t^2}}\right) - 1\,, \end{aligned} \tag{39}$$

where $f(x)$ and $f(y|x)$ are given in Eqs. (8) and (9). From these expressions, it can be readily appreciated that, in most cases, a discrepancy exists between the true and observed/reported in-tolerance levels. This discrepancy can be eliminated, however, by adjusting L_{test} according to

$$L_{test} = L_{per}\sqrt{1 + (\sigma_t / \sigma_{true})^2}\ . \tag{40}$$

As this expression indicates, since uncertainties are present in the test or calibration process (i.e., $\sigma_t > 0$), the test limits should be placed outside the performance limits if reported in-tolerance levels are to match true measurement reliabilities.

False Alarms/Missed Faults

A false alarm is defined as a case in which an in-tolerance UUT item is falsely reported as out-of-tolerance. This can constitute a costly error in that, in some instances, such a report may lead to unnecessary rework and/or repair. Moreover, false out-of-tolerances can have a significant impact on calibration or test intervals, particularly if intervals are adjusted to meet high (over 50%) measurement reliability targets. This is because, in these cases, intervals are shortened in response to a reported out-of-tolerance to a greater extent than they are lengthened in response to a reported in-tolerance test or calibration result.

The probability of a false alarm is given by

$$
\begin{aligned}
P(\text{false alarm}) &= P(|x| \le L_{per}, |y| \le L_{test}) \\
&= \int_{-L_{per}}^{L_{per}} f(x)\,dx \int_{L_{test}}^{\infty} f(y\,|\,x)\,dy + \int_{-L_{per}}^{L_{per}} f(x)\,dx \int_{-\infty}^{-L_{test}} f(y\,|\,x)\,dy \\
&= 2 - \frac{1}{\sqrt{2\pi}} \int_{-L_{per}/\sigma_{true}}^{L_{per}/\sigma_{true}} \left[F\left(\frac{L_{test} + \sigma_{true}\xi}{\sigma_t} \right) + F\left(\frac{L_{test} - \sigma_{true}\xi}{\sigma_t} \right) \right] e^{-\xi^2/2}\, d\xi \,.
\end{aligned} \tag{41}
$$

Corresponding to the probability of a false alarm is the probability of a missed fault. From the viewpoint of the UUT user, a missed fault is an item returned to the user facility from test or calibration in an out-of-tolerance state. Recalling the earlier discussion on bop reliability, the probability of this occurrence is given by

$$
P(\text{missed fault}) = 1 - R_{bop} \ ,
$$

where R_{bop} is given in Eq. (19) or (20).

Average Over Period Reliability

From Eq. (42), it can be seen that a viable measure of the quality of the test or calibration process is the UUT bop reliability. Likewise, from Eq. (41), since the probability of a false alarm is a function of σ_{true}, the unnecessary rework cost is seen to be controlled to some extent by the eop reliability. While these quantities are of interest, the UUT user is generally more concerned about the measurement reliability of the UUT over the period of use, i.e., over the test or calibration interval. To put this in a somewhat more quantifiable framework, the user is interested in the probability that the UUT is in-tolerance under the conditions of the demand for its use experienced during the test or calibration interval. If the usage demand is random, i.e., if the likelihood for use is uniform over the interval, then the appropriate measure of this in-tolerance probability is the average over period or aop measurement reliability.

The aop measurement reliability is the mathematical average of the reliability from time $t = 0$ to time $t = I$, where the zero point corresponds to R_{bop} and $t = I$ corresponds to R_{eop}:

$$
R_{aop} = \frac{1}{I} \int_0^I R(t)dt \ . \tag{43}
$$

For the exponential model, this is given by

$$
\begin{aligned}
R_{aop} &= \frac{R_{bop}}{I} \int_0^I e^{-\lambda t} dt \\
&= \frac{R_{bop}}{\lambda I} \left(1 - e^{-\lambda I} \right) \quad \text{(exponential model)} \ .
\end{aligned} \tag{44}
$$

For the random walk model, there are two possibilities. The first covers cases governed by the renew always policy ($L_{adj} = 0$) and the second applies to other policies. For renew always cases,

$$
\begin{aligned}
R_{aop} &= \frac{1}{\sqrt{2\pi} I} \int_0^I \frac{dt}{\sigma(t)} \int_{-L_{per}}^{L_{per}} dx\, e^{-x^2/2\sigma^2(t)} \\
&= \frac{1}{I} \int_0^I \left\{ 2F\left(L_{per} / \sigma(t) \right) - 1 \right\} dt \quad \text{(random walk/renew always)} \ ,
\end{aligned} \tag{45}
$$

where, from Eqs (30) and (20),

$$
\sigma^2(t) = \sigma_t^2 + \sigma_{rb}^2 + \sigma_s^2 + \alpha t \ . \tag{46}
$$

For cases where $L_{adj} \ne 0$, setting $I = 0$ in Eq. (45) will not return R_{bop} as expressed in Eq. (19). This is because, if only a portion of the UUT population is adjusted using the test system, the resulting distribution of UUT attribute values is not strictly Gaussian. For these cases, numerical Monte Carlo or Markov process techniques are needed to evaluate R_{aop} precisely. Unfortunately, use of these methods is somewhat unwieldy. From experience with several simulated examples, however, a simplification is possible. This simplification consists of obtaining an approximate aop value for $\sigma(t)$, referred to as σ_{aop}, and plugging this quantity into the appropriate expression for $R(t)$ to get R_{aop}. Not only is this approximation useful for the $L_{adj} \neq 0$ case, but also works well for the renew always case.

Determination of σ_{aop} begins with obtaining an approximate value for σ_{bop}. This is given by

$$\sigma_{bop} \cong \frac{L_{per}}{F^{-1}\left[(1+R_{bop})/2\right]} \ , \tag{47}$$

where R_{bop} is given in Eq. (19) for $L_{adj} \neq 0$ renewal policies, and in Eq. (20) for the $L_{adj} = 0$ renewal policy (for which Eq. (47) is an exact expression). Working from Eq. (30), σ_{aop} can be expressed as

$$\begin{aligned} \sigma_{aop} &= \sqrt{\frac{1}{I}\int_0^I \left(\sigma_{bop}^2 + \alpha t\right) dt} \\ &= \sqrt{\sigma_{bop}^2 + \tfrac{1}{2}\alpha t} \quad \text{(random walk model)} \ . \end{aligned} \tag{48}$$

Note that if the UUT is used as the TS for the next highest level in the test and calibration hierarchy, σ_{aop} is the value used for σ_{ts} in Eq. (3). This is because TS items are assumed to be selected and used for UUT test/calibration at random times over their calibration intervals.

For the exponential model, use of Eq. (44) gives

$$\begin{aligned} \sigma_{aop} &= \frac{L_{per}}{F^{-1}\left[(1+R_{aop})/2\right]} \\ &= \frac{L_{per}}{F^{-1}\left\{\frac{1}{2}\left[1 + \frac{R_{bop}}{\lambda I}\left(1 - e^{-\lambda I}\right)\right]\right\}} \quad \text{(exponential model)} \ , \end{aligned} \tag{49}$$

with R_{bop} as given in Eq. (19) or (20).

Availability

The cost of operating a test and calibration program, as well as the cost of maintaining a functioning field capability is impacted by the need for equipment spares. Spares costs are minimized by maximizing equipment availability. The availability of an item of UUT is defined as the probability that the item will be available for use over the period of its administrative test or calibration interval. If this interval is thought of as the time elapsed between successive bop dates, then the availability of an item is given by

$$\text{availability} = \frac{I}{\text{administrative interval}} \ , \tag{50}$$

where I is the "active" portion of the test or calibration interval as defined in Eqs. (21) and (30). The difference between the administrative interval and the variable I is the downtime:

$$T_d = \text{administrative interval} - I \ . \tag{51}$$

For our purposes, the composition of T_d is assumed to be described according to

$$\begin{aligned} T_d &= \text{calibration downtime} + \text{adjustment downtime} \times P(\text{adjust}) \\ &\quad + \text{repair downtime} \times P(\text{repair}) \ . \end{aligned} \tag{52}$$

P(adjust) is given in Eq. (14). The probability for repair is defined as the probability that UUT items, submitted for test or calibration, will require repair action in addition to the various adjustments and corrections that normally accompany test or calibration. As the reader will note, this is a subset of the total repair downtime, which includes downtime due to user-detectable functional failures encountered during use. Since this paper is concerned primarily with cost and performance as impacted by test and calibration, only this subset is of interest in the present context. The quantity P(repair) is thus given by

$$P(\text{repair}) = 2\left[1 - F\left(\frac{L_{rep}}{\sigma_{true}}\right)\right] , \tag{53}$$

where L_{rep} is a parameter which marks a limiting measurement attribute value, beyond which repair actions are normally required to restore a UUT item to its nominal performance value.

The remaining quantities in Eq. (52) will now be considered. First we define the following variables

- T_{cal} - mean time required to perform a test or calibration action.
- T_{css} - mean shipping and storage time experienced between eop and bop dates.
- T_{rep} - mean time required to perform a repair action.
- T_{rss} - mean shipping and storage time experienced between submittal and return of an item of UUT submitted for repair.
- T_{adj} - mean time required to perform a routine adjustment of a UUT measurement attribute.

Given these definitions, we have

$$\begin{aligned} \text{calibration downtime} &= T_{cal} + T_{css} \\ \text{adjustment downtime} &= T_{adj} \\ \text{repair downtime} &= T_{rep} + T_{rss} , \end{aligned}$$

It is assumed that, under ordinary circumstances, these quantities are known. Substituting these variables in Eq. (50) and using Eqs. (51) and (52) gives

$$\begin{aligned} P(\text{available}) &= \frac{I}{I + T_{cal} + T_{css} + T_{adj}P(\text{adjust}) + (T_{rep} + T_{rss})P(\text{repair})} \\ &= \frac{1}{1 + \dfrac{T_{cal} + T_{css} + T_{adj}P(\text{adjust}) + (T_{rep} + T_{rss})P(\text{repair})}{I}} \\ &= \frac{1}{1 + T_d / I} . \end{aligned} \tag{54}$$

From Eq. (54), it is evident that availability approaches unity as $I \to \infty$ and/or as $T_d \to 0$. Eq. (54) also shows that availability improves as P(adjust) and P(repair) are minimized.

Cost Modeling

Calibration intervals, test decision risks, and availability are parameters which have a direct bearing on the costs associated with operating and maintaining a test and calibration support hierarchy. These parameters also impact indirect costs associated with end item quality and/or performance capability.

End item quality and/or performance capability is measured in terms of the extent to which an end item achieves a desired effect or avoids and undesired effect. These effects can be referenced to the Program Management level, in the case of a military system, to an end item producer in the case of a commercial product or to any category of end item disposition for which the end item quality and/or performance capability can be quantified in economic terms. Examples of desired effects may include successful strike of an offensive weapon, follow-on reorders of a product item, creation of a desirable corporate image, etc. Examples of undesired effects may include unsuccessful response to a military threat, warranty expenses associated with poor product performance, return of products rejected by customer receiving inspection, etc. In each case, the end item experiences an "encounter" (approach of an intended target, approach of an incoming missile, appearance of an unexpected highway obstruction, etc.) which results in a perceived "outcome" (successful missile strike, missile interception, obstruction avoidance, etc.). The effect is determined by the "response" of the end item to the encounter (timely sighting and ranging, early detection and warning, nominal braking and maneuvering, etc.). The cost of a given outcome is a variable which is assumed to be known. If an outcome is associated with a benefit, the cost is expressed in negative dollars.

Table 1. Cost Modeling Variables

Variable Description	Variable Name
End item attribute value corresponding to the onset of degraded performance	x_d
End item attribute value corresponding to loss of function	x_f
Cost of a given outcome	C_f
Quantity of end items sold or in inventory	N_{UUT}
Acquisition cost of an end item unit	C_{UUT}
End item spare coverage desired (in percent)[2]	S_{UUT}
Probability of a successful outcome, given successful end item performance	P_{sr}
Probability of an encounter	P_e
Hours to calibrate/test	H_c
Additional hours required for adjustments	H_a
Cost per hour for test/calibration and/or adjustment	C_{hr}
Cost per repair action	C_r

The analytical methodology developed herein provides a means for determining the probability for a successful or unsuccessful outcome as a function of various technical parameters which characterize the test and calibration support hierarchy.

The hierarchy impacts costs associated with fielding, selling or otherwise dispatching the supported end item. An end item which has been dispatched is referred to as one which has been "accepted" by the end item test system. Accordingly, the costs which derive from a dispatched end item are termed "acceptance costs" in this paper. The variables employed in modeling acceptance cost are shown in Table 1. The variables resulting from cost modeling and analysis are shown in Table 2. In this table, total annual calibration, adjustment, repair, and support costs relate to costs incurred from support of a UUT of interest (calibration system, test system or end item). Annual acceptance cost applies only if the UUT of interest is an end item.

Key to the cost modeling discussed in this paper is the assumption that the quality and/or performance capability of an end item is related to the value of the measurement attribute supported by test and calibration. In other words, the attributes which are tested prior to end item dispatch can be out-of-tolerance to a degree that end item performance is negatively impacted. The variables x_d and x_f mark the onset of degraded performance and the complete loss of performance points, respectively, for the relevant attribute. To relate end item quality or capability to values between these points, the following model is adopted:

$$P(\text{success}) = \begin{cases} 1, & |x| \le x_d \\ 1 - \sin^2\left[\dfrac{(|x| - x_d)\pi}{2(x_f - x_d)}\right], & x_d \le |x| \le x_f \\ 0, & x_f \le |x|, \end{cases} \qquad (55)$$

[2] This variable controls the number of spares maintained to cover the end item inventory or the population of end items sold to customers.

where $P(\text{success}|x)$ is the probability for successful performance of the end item, given that its attribute value is equal to x. The probability of a successful outcome is given by

$$P(\text{success}) = P_{sr} \int_{-\infty}^{\infty} f_{aop}(x) P(\text{success} \mid x)\, dx \ , \tag{56}$$

where $f_{aop}(x)$ is obtained from Eq. (8) with "aop" in place of "true" to indicate that the end item is employed in accordance with the random demand assumption throughout its test interval:

$$f_{aop}(x) = \frac{1}{\sqrt{2\pi}\sigma_{aop}} e^{-x^2/2\sigma_{aop}^2} \ . \tag{57}$$

As Eqs. (48) and (49) show, σ_{aop} is a function of σ_{bop} or, equivalently, R_{bop}. These quantities are, in turn, determined by the accuracy of the test system and the quality of the test and calibration support hierarchy.

Table 2. Acceptance Costs Modeling Variables

Variable Description	Variable Name
Total annual cost	C_{tot}
Annual acceptance cost	C_{acc}
Total annual support cost	C_{ts}
Annual calibration cost	C_{cal}
Annual adjustment cost	C_{adj}
Annual repair cost	C_{rep}
Total spares acquisition cost	C_{sa}

The acceptance cost for dispatched end items is the product of the cost of a given outcome, the number of end items dispatched, the probability of an encounter occurring and the probability of unsuccessful end item performance:

$$C_{acc} = C_f N_{UUT} P_e [1 - P(\text{success})] \ , \tag{58a}$$

where $P(\text{success})$ is given in Eq. (56). If C_{acc} represents a benefit, the appropriate expression is

$$C_{acc} = C_f N_{UUT} P_e P(\text{success}) \ , \tag{58b}$$

where C_f would be given in terms of payoff rather than cost. The quantity C_{acc} can be "annualized" by expressing P_e in terms of the probability of encounter per end item unit per year. In some cases, it may be desirable to set P_e equal to the probable *number* of encounters experienced per end item unit per year. (The reader may note that this quantity may be a function of N_{UUT}.)

As stated earlier, acceptance cost applies only to the end item. The quantities which follow, however, apply to any UUT encountered at any level of the test and calibration support chain. Of these, we first consider costs associated with UUT downtime. This downtime results in a requirement for replacement spares to have on hand to cover items submitted for test or calibration.

The number of available UUT spares required is equal to the number needed to cover the unavailable UUT items multiplied by the extent of coverage desired (the number of spares desired on hand to cover an out of use UUT):

$$N_s P(\text{available}) = N_{UUT} [1 - P(\text{available})] S_{UUT}$$

or

$$N_s = \frac{[1 - P(\text{available})]}{P(\text{available})} N_{UUT} \, S_{UUT} \ ,$$

which becomes, with the aid of Eq. (54),

$$N_s = (T_d / I) N_{UUT} \, S_{UUT} \ . \tag{59}$$

The cost to purchase these spares is given by

$$C_{sa} = N_s C_{UUT} \ , \tag{60}$$

and the annual cost resulting from the requirement for these spares is given by

$$C_s^{year} = C_d C_{sa} \ , \tag{61}$$

where C_d is either the annual depreciation cost per UUT item, for private sector applications, or is the unit rate at which UUT items expire from use and require replacement, in the case of government applications.

The annual cost due to calibration or test is given by

$$C_{cal} = H_c C_{hr} (N_{UUT} + N_s) / I \quad , \tag{62}$$

where I is expressed in years. The annual cost of UUT adjustments is given by

$$C_{adj} = \frac{(N_{UUT} + N_s)}{I} H_a C_{hr} P(\text{adjust}) \quad , \tag{63}$$

and the annual cost of UUT repair is

$$C_{rep} = \frac{(N_{UUT} + N_s)}{I} C_r P(\text{repair}) \quad , \tag{64}$$

where P(adjust) is given in Eq. (14), P(repair) is given in Eq. (53) and, again, I is expressed in years.

The total annual support cost is the sum of Eqs. (61), (62), (63) and (64):

$$C_{ts} = C_s^{year} + C_{cal} + C_{adj} + C_{rep} \quad . \tag{65}$$

The total annual cost, including support and acceptance costs, is given by the sum of Eq. (65) and Eq. (58):

$$C_{tot} = C_{acc} + C_{ts} \quad . \tag{66}$$

Acknowledgment

The author is indebted to Mr. Pete Strucker of the U.S. Navy Metrology Engineering Center (MEC) for his continued enthusiasm and encouragement; to Mr. Del Caldwell of MEC for ideas, critiques and unfailing support of this and earlier work; to Mr. Jerry Hayes of Hayes Technology for introducing calibration/risk assessment concepts into the field of metrology [10] and for his many valuable contributions to the present work; to Mr. Dale Rockwell of MEC and Dr. Robert Munk of SAIC for key questions and suggestions; to Dr. John Ferling of Claremont McKenna College and Dr. Ken Kuskey of DDI for years of invaluable collaboration; and to Mr. Ken Drew and Mr. Marshall Holstrom of SAIC for innumerable contributions without which the work would not have been possible.

References

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

[2] Kuskey, K., *A 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.

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

[4] 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.

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

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

[7] Kuskey, K., *An Analysis of Single-Sided Tolerance Testing*, U.S. Navy Metrology Engineering Center, Analytical Metrology Note AMN 85-2, April 1985.

[8] Larsen J., *A Handy Approach to Examine and Analyze Calibration Decision Risks and Accuracy Ratios*, U.S. Navy Metrology Engineering Department, Analytical Metrology Note AMN 87-2, August 1987.

[9] 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.

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

Appendix - Measurement Uncertainty Analysis

A prescription is offered in this appendix to assist in the determination of the various standard deviations employed in the methodology described in this paper. These standard deviations are constructed from several uncertainty components listed below. In the listing, the performance limit of the UUT is labeled L_{per}^{UUT}, and the performance limit of the TS is labeled L_{per}^{TS}.

Uncertainty Component	**Definition**	**Standard Deviation**
UUT resolution	$= \rho_1 + \rho_2 L_{per}^{UUT}$	σ_R^{UUT}
TS resolution	$= \rho_3 + \rho_4 L_{per}^{TS}$	σ_R^{TS}
process error	$= \rho_5 + \rho_6 L_{per}^{UUT} + \rho_7 L_{per}^{TS}$	σ_P
technician error	$= \rho_8 + \rho_9 L_{per}^{UUT} + \rho_{10} L_{per}^{TS}$	σ_{tech}
rebound error	$= \rho_{11} + \rho_{12} L_{per}^{UUT}$	σ_{reb}
shipping error	$= \rho_{13} + \rho_{14} L_{per}^{UUT}$	σ_s

The coefficients ρ_i, $i = 1,2, \cdots, 14$, are provided as estimates by persons knowledgeable of the test or calibration process and associated equipment. Of the uncertainty components, *UUT resolution* and *TS resolution* refer to the coarseness of respective UUT or TS attribute readings. *Process error* refers to uncertainties introduced into the test or calibration process by fluctuations in ancillary equipment, shifts in environmental factors, etc. *Technician error* arises from the fact that different technicians may, under identical circumstances report different measured values for a given UUT attribute. *Rebound error* was defined earlier. *Shipping error* is an estimate of the upper limits to which the UUT attribute can be displaced as a result of shipping and storage.

In the absence of more specific information, we assume that each uncertainty component provided constitutes an upper limit estimate outside of which no values are expected to lie. Although we make no claim to privileged expertise regarding the cerebral mechanisms by which human minds develop such estimates, we feel it is safe to regard these components as approximate 3σ limits. Accordingly, the standard deviation corresponding to each uncertainty component is obtained by dividing the magnitude of each estimated component by 3. Thus, for example, $\sigma_R^{UUT} = (\text{UUT resolution})/3$.

The component standard deviations σ_{rb} and σ_s have already been encountered. The other components can be used to determine the test process standard deviation σ_{tp}:

$$\sigma_{tp}^2 = (\sigma_R^{UUT})^2 + (\sigma_R^{TS})^2 + \sigma_P^2 + \sigma_{tech}^2 \ .$$