Morehouse Instrument Company
1742 Sixth Ave.
York, PA 17403-2675
(717) 843-0081

Why a 4:1 TUR is not Enough: The Importance of Analyzing the Probability of False Accept Risk

Introduction

Several organizations and publications reference the use of a 4:1 Test Uncertainty Ratio (TUR). Some standards even reference a TUR requirement greater than or equal to 4:1.

The question to ask is if they know why they may need a 4:1 TUR and if they may understand the rationale for requiring a 4:1. The thought here is that a 4:1 ratio is based on specific false accept and false reject risk, and a 4:1 ratio is a simple way of achieving it if certain conditions can be met.

That thought process alone is dangerous if one does not have enough historical data to use a joint probability density function associated with many TUR-based methods.

If one does the math, a 4:1 TUR with a coverage probability of k = 2 for the measurement uncertainty and a 95 % End of Period reliability can equate to less than 1 % false accept and slightly over 1.5 % false reject (these terms are covered later).

In simplistic terms, End of Period Reliability is defined as the number of calibrations resulting in acceptance criteria being met divided by the total number of calibrations. The formula to determine the required sample size from "In-Tolerance" Reliability from historical data is easy to replicate in Excel. The formula is Sample Size = ln(1-Confidence)/ln(Target Reliability)

If we use the formula for Sample Size above, we would need over 59 (58.4) samples to use a joint probability distribution associated with many TUR-based methods.

4:1 may sound good on paper, though many laboratories might use the boilerplate language on a purchase order asking for a 4:1 TUR, likely without the appropriate sample size.

And then, there are different disciplines that, like equipment, cannot easily be grouped into the calculation based on a global risk approach, equipment that might have different usages, fixturing, wear patterns, lots with sub-par quality control, different calibration intervals, and more. These different usages and conditions can lead to statistical independence from the population of like instruments.

When dealing with physical changes to the instrumentation, like material deformation, as found in many force and torque measuring equipment, it isn't the same as if we were measuring the voltage of batteries from a large production lot.

Thus, we must understand the limitations when we analyze requiring a 4:1 TUR as a risk mitigation strategy to control our probability of false accept risk.

Many labs may use a 4:1 TUR properly and understand what decision rule is best to use to manage their application's false accept and reject risks.

The paper makes several assumptions about standards, assuming the end-user might request a 4:1 TUR based on insufficient sample size. These assumptions are based on the author's perception of what might be happening in the industry.

We will discuss TUR, why the location of the measurement matters, PFA, and some common guard banding methods used to limit the PFA risk and stay within the lines.

4:1 TUR Requirements

ANSI/NCSL Z540.3 – Requirements for Calibration of Measuring and Test Equipment in section 5.3 b) allows for use of a test uncertainty ratio (TUR) equal to or greater than 4:1 when it is not practical to estimate the false accept risk of less than 2 %. Then goes on to say objective evidence of nonpracticability of this determination is expected as in an agreement with the customer TUR use. [1]

The assumption is that the higher the TUR, the higher the probability that the measuring equipment will have a Probability of False Accept (PFA) of less than 2 %.

Using TUR to control false accept and false reject risk typically requires the end-user to know their End of Period Reliability (EOPR). A 2 % PFA requirement can be achieved with either measurement process uncertainty or observed EOPR assuming a single variable is dominant.

NCSLI RP-18 in section 3.5.2 A Critique of the 4:1 Requirement, discusses some Z540.3 TUR requirements that deserve mention. These are:

- The requirement is merely a ratio of UUT tolerance limits relative to the expanded uncertainty of the measurement process. It is, at best, a crude risk control tool, i.e., one that does not control risks to any specified level. Moreover, in some cases, it may be superfluous. For instance, what if all UUT attributes of a given manufacturer/model are in-tolerance prior to test or calibration? In this case, the false accept risk is zero regardless of the TUR.
- The requirement is not applicable when UUT tolerances are single-sided.
- The requirement is only approximately applicable when tolerances are two-sided but asymmetric and the UUT bias is distributed such that its mode value is zero [2]

In addition, many fail to realize what is described in "Introduction to Statistics in Metrology," section 5.2.1.5 states, "While the 4:1 TUR requirement is commonly used to ensure a measurement is adequate for making an accept/reject determination, this metric assumes that the process distribution is centered between the specification limit. If this is not the case, TUR cannot be reliably used as an indicator of risk" [3]

All of this matters because it is a requirement of ISO/IEC 17025:2017. Section 7.8.6.1 states, "When a statement of conformity to a specification or standard for test or calibration is provided, the laboratory shall document the decision rule employed, taking into account the level of risk (such as false accept and false reject and statistical assumptions) associated with the decision rule employed and apply the decision rule." [4]

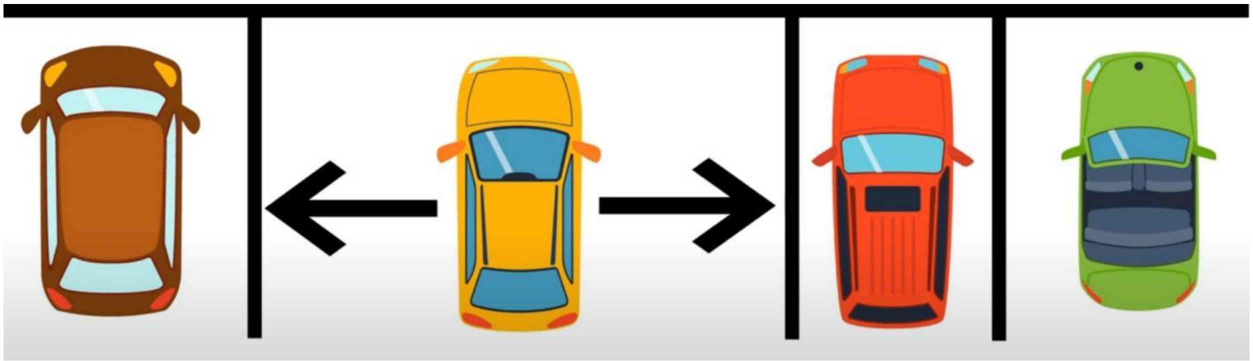

Figure 1 Parking Lot Example with Small Calibration Process Uncertainty

We can think about the measurement risk this way. We have a car, and we need to park it between two lines. The lines represent the upper and lower specification limit of our device. The width of our car is the Calibration Process Uncertainty, and parking lines are our tolerance specification limits.

The probability of us getting a ding or denting another vehicle is our PFA, depending on how centered we are within the parking lines. If we try to park too close to one side, we may risk not being able to open the door, or if we misjudge entirely,, we may run right into the car in the other lane and cause substantial damage. If we park centered on the line, 50 % of our car will be in the next lane no matter what size our car is.

Many examples cited are assumed to be based on discrete measurements at the bench level (Specific Risk).

Specific risk is that we are testing an instrument when we do not have a high enough sample size or information other than where the result is located in relation to the tolerance requested and calculate our uncertainty correctly to calculate the false accept risk. Figure 1 below shows this concept.

Nominal Value	10000
Lower specification Limit	9995
Upper Specification Limit	10005
Measured Value	10005
Measurement Error	5
Std. Uncert. (k=1)	0.08
Total Risk	50.00%
Upper Limit Risk	50.00%
Lower Limit Risk	0.00%
TUR =	**31.2296736**
Guard Band LSL	9995.164407
Guard Band USL	10004.83559

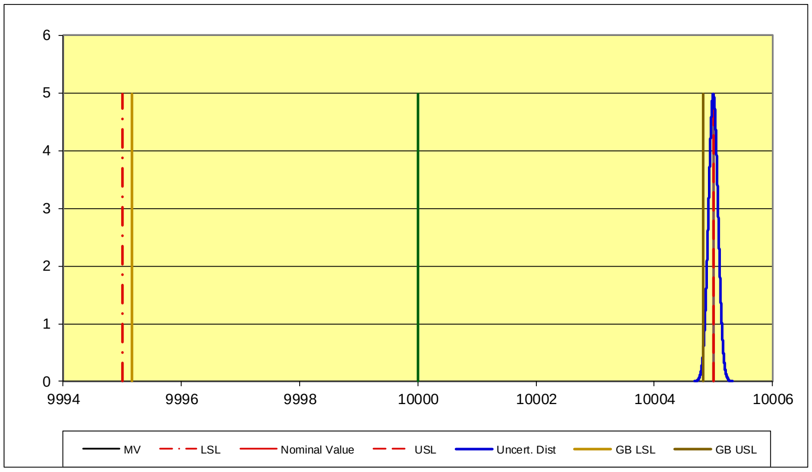

Figure 2 31.23:1 TUR with a 50 % PFA at the Upper Specification Limit

TUR

TUR or Test Uncertainty Ratio, is defined in Section 3.11 of ANSI/NCSL Z540.3 as, "The ratio of the span of the tolerance of a measurement quantity subject to calibration, to twice the 95% expanded uncertainty of the measurement process used for calibration." The TUR tells us how much space between the lines we must be "in-tolerance." [5]

$$\mathbf{TUR} = \frac{\textbf{Span of the } \pm \textbf{ UUT Tolerance}}{\mathbf{2\ x\ k_{95\%}(\textbf{Calibration Process Uncertainty})}}$$

Figure 3 TUR Formula found in ANSI/NCSLI Z540.3 Handbook

TUR is commonly used as a simplified approach of evaluating global risk. When we know the tolerance we are working to, we have a high enough sample size to know the shape and the distribution of the calibration results and our end-of-period reliability. We can calculate the appropriate uncertainty that corresponds to the maximum amount of false accept risk we are okay with.

Per the ANSI/NCSL Z540.3 Handbook, "For the denominator, the 95 % expanded uncertainty of the measurement process used for calibration following the calibration procedure is to be used to calculate TUR. The value of this uncertainty estimate should reflect the results that are reasonably expected from the use of the approved procedure to calibrate the M&TE. Therefore, the estimate includes all components of error that influence the calibration measurement results, which would also include the influences of the item being calibrated except for the bias of the M&TE. The calibration process error, therefore, includes temporary and non-correctable influences incurred during the calibration such as repeatability, resolution, error in the measurement source, operator error, error in correction factors, environmental influences, etc." [5]

PFA Risk

Nominal Value	10000
Lower specification Limit	9995
Upper Specification Limit	10005
Measured Value	10004.83559
Measurement Error	4.83559
Std. Uncert. (k=1)	0.08
Total Risk	2.00%
Upper Limit Risk	2.00%
Lower Limit Risk	0.00%
TUR =	**31.2296736**
Guard Band LSL	9995.164407
Guard Band USL	10004.83559

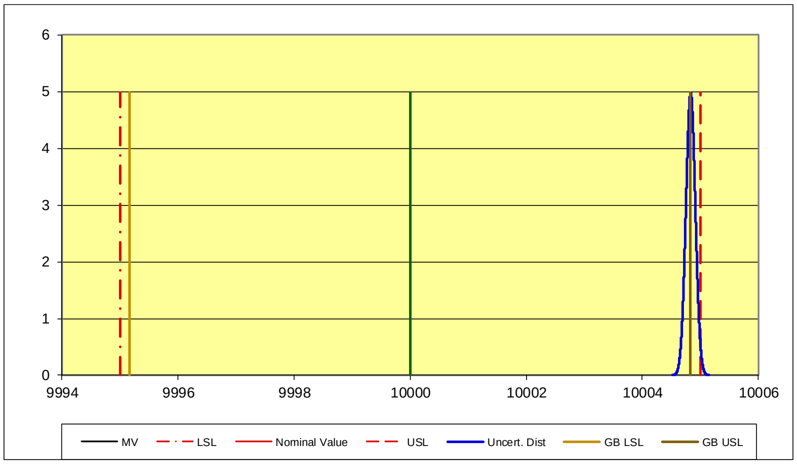

Figure 4: Guard band USL showing a 2 % PFA when Measured Value is at the GB USL

All measurements have a percentage of likelihood of calling something good when it is bad, and something bad when it is good. You might be familiar with the terms consumer's risk and producer's risk. Consumer's risk refers to the possibility of a problem occurring in a consumer-oriented product; occasionally, a product not meeting quality standards passes undetected through a manufacturer's quality control system and enters the consumer market. The Probability of False Accept (PFA) is similar to consumer's risk. It is the likelihood of calling a measurement "good" or stating something is "In Tolerance" when there is a percentage that the measurement is "bad" or "Out of Tolerance".

ANSI/NCSLI sub-clause 5.3 is the tolerance-type test requirement that "the probability that incorrect acceptance decisions (false accept) will result from calibration tests shall not exceed 2 %." With the preponderance of calibrations being of this type, the resources and conditions described by the calibration procedure will require careful evaluation and determination to achieve the measurement uncertainty needed for the calibration process to achieve this allowable probability of false accept." The measurement uncertainty must be accounted for, and the acceptance limits must be calculated to ensure the likelihood of the measurement being "Out of Tolerance" does not exceed 2 %.

The entire purpose of analyzing the PFA is to ensure your measurements are "In Tolerance" with risk that does not exceed 2 %. And why just knowing you have a 4:1 TUR without analyzing the PFA regarding the location of the measurement is not enough to minimize measurement risk as shown in Figure 3. Figure 3 shows the upper and lower guard banded limits to ensure a PFA of 2 % or less. If the measured value is not within the guard band limits, the PFA will be higher than 2 %.

Location of the Measurement

Nominal Value	10000
Lower specification Limit	9995
Upper Specification Limit	10005
Measured Value	10004
Measurement Error	4
Std. Uncert. (k=1)	0.08
Total Risk	0.00%
Upper Limit Risk	0.00%
Lower Limit Risk	0.00%
TUR =	**31.2296736**
Guard Band LSL	9995.164407
Guard Band USL	10004.83559

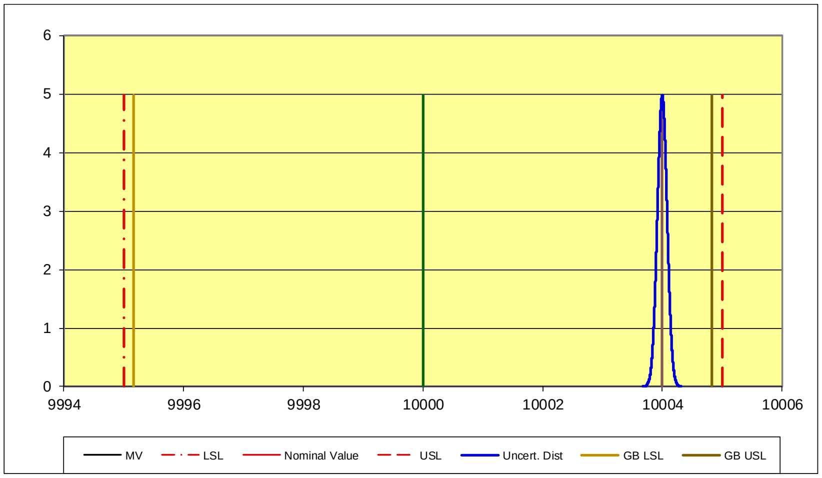

Figure 5: Graph Showing 10,004 as the measured value with a 31.23:1 TUR which is achieved by using a lab with low uncertainties

Calling an instrument "In Tolerance" is all about location, location, location. It's also about the uncertainty of the measurement, but a bad location will raise the Probability of False Accept (PFA) significantly.

The probability of false acceptance is the likelihood of a lab calling a measurement "In Tolerance" when it is not. The location we are referring to is how close the measurement is to the nominal value. If the nominal value is 10,000 lbf and the instrument reads 10,004 lbf, the instrument bias is 4 lbf, as shown in Figure 4.

The larger the bias, the worse the location of the measurement. If we go back to our parking scenario, the worse the bias from nominal, the more likely one side of our automobile will be damaged, or maybe we are still "in tolerance" but have to exit the vehicle from the other side.

Higher TURs help control PFA. If the End of Period Reliability (EOPR) is fixed, the TUR will decrease as the measurement process uncertainty increases.

Figure 5 shows this concept as risk level increases as we have switched calibration providers, and the new provider has a higher CMC uncertainty component of 0.025% than shown in Figure 4 where the calibration provider had a 0.0016% CMC uncertainty component; everything else has remained the same.

Nominal Value	10000
Lower specification Limit	9995
Upper Specification Limit	10005
Measured Value	10004
Measurement Error	4
Std. Uncert. (k=1)	1.25
Total Risk	21.19%
Upper Limit Risk	21.19%
Lower Limit Risk	0.00%
TUR =	**1.999994666**
Guard Band LSL	9997.567193
Guard Band USL	10002.43281

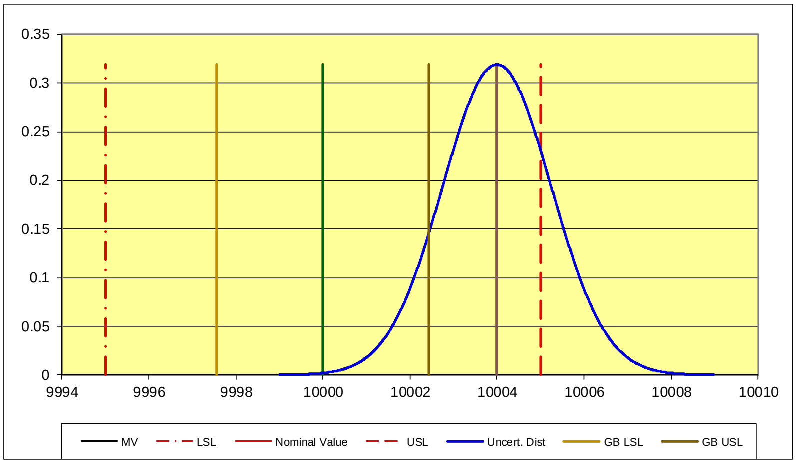

Figure 6: Graph Showing 10,004 as the measured value with a 1.99:1 TUR as the labs Expanded uncertainty is higher than in figure 4

Why do we care about the location of the measurement if the device is within tolerance? If a device has a specification of 0.1 % of full scale and the calibrating laboratory reports a value within 0.1%, the device is “In Tolerance,” right?

The answer is and always will be it depends on the measurement is uncertainty and if the lab performing the calibration has adequately calculated their Calibration Process Uncertainty correctly and followed the proper guidelines in determining the uncertainty of measurement when making the statement of compliance.

If the uncertainty of the measurement is significant, the lab performing the calibration will have to be very concerned with the location of the measurement.

If their uncertainty of measurement is too high, they may not even be able to perform the calibration at all, and if the measured value falls right on the specified tolerance line, the PFA can be 50 % or higher.

There are several methods to ensure a 2 % PFA requirement can be met. These TUR-based decision rules typically set acceptance limits to ensure the PFA is less than 2 %.

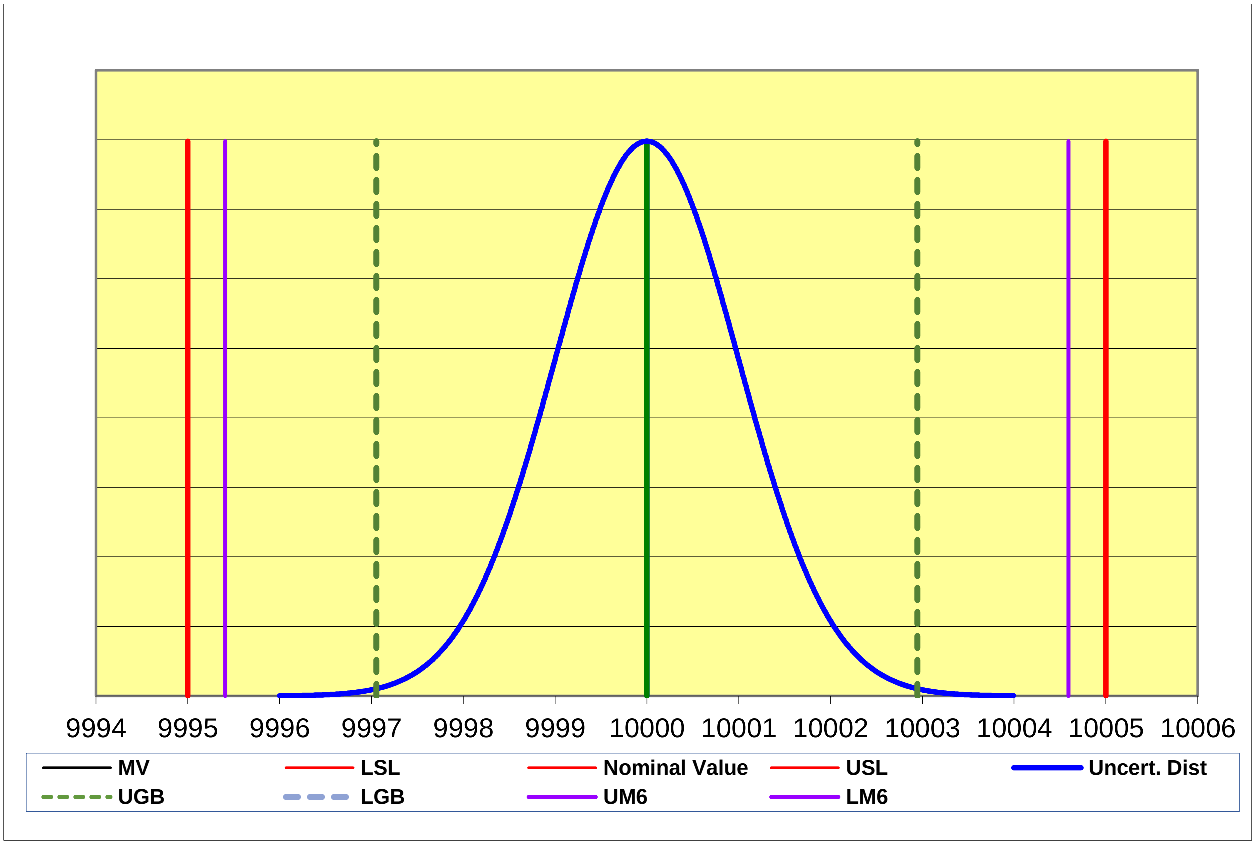

Figure 7: Graph Showing Specification Limits and Acceptance Limits for Both Method 5 and Method 6

Two Managed Risk Guard Banding Methods to Ensure the PFA is Less Than 2 %

ISO/IEC 17025:2017 section 7.8.6.2 states "The laboratory shall report on the statement of conformity such that the statement clearly identifies –a) to which results the statement applies; and –b) which specifications, standard or parts thereof are met or not met; –c) the decision rule applied (unless it is inherent in the requested specification or standard)".

In this paper, we are going to discuss three decision rules. Two of these rules, known as Method 5 and Method 6 are documented in ANSI/NCLI Z540.3 Handbook, and a third rule is something a lab may think about using to meet the criteria. The standard does not dictate what rules can or cannot be applied. It just requires that the calibration laboratory list the decision rule applied and that the laboratory discusses its decision criteria with the customer. There are several other guard banding and risk-based approaches than what is presented here. ILAC G8:09/2019 Guidelines on Decision

Rules and Statements of Conformity has several examples of decision rules and the corresponding risk associated with those rules.

Guard Band Method 5, Based on the Expanded Calibration Process Uncertainty. This method is simple as one subtracts the 95 % expanded process uncertainty from the tolerance limits. The above graphs in figures 1 through 4 use Method 5. It is very similar to the ILAC G8: 2009 rule.

ILAC-G8:03/2009 states if the specification limit is not breached by the measurement result plus the expanded uncertainty with a 95 % coverage probability, then compliance with the specification can be stated. The ILAC rule allows for a PFA of < 2.5 %.

Simply put, if one subtracts the expanded calibration process uncertainty from the upper limit of the specified tolerance, then the new acceptance limits will assure a PFA of less than 2.275 %. It is an interesting point as the ANSI/Z540.3 Handbook mentions 2 %, though the calculation gives a PFA of 2.275 %. One must assume some rounding took place. The only information needed to use Method 5 is the tolerance and the calibration process uncertainty formula in the figure below.

$$\mathbf{2\ x\ k_{95\%}}\left(\sqrt[2]{\left(\frac{\mathbf{CMC}}{\mathbf{k_{CMC}}}\right)^2+\left(\frac{\mathbf{Resolution_{UUT}}}{\sqrt[2]{\mathbf{12}}}\right)^2+\left(\frac{\mathbf{Repeatability_{UUT}}}{\mathbf{1}}\right)^2+\cdots(\mathbf{u}_{Other})^2}\right)$$

Figure 8: Calibration Process Uncertainty assuming 95 % confidence interval

Note: See ILAC-P14:09/2020 section 5.4 for requirements for calculating CMC and reporting measurement uncertainty. The requirements align very closely with the formula in Figure 7.

The downside of Method 5 is that the test limit is based on the worst-case PFA, which means they may be too aggressive, resulting in more false rejects from the reference laboratory. Being overly aggressive and needing to adjust more equipment lends one to look for an alternative method.

Guard Band Method 6, Based on Test Uncertainty Ratio:

This method is also simple as it depends only on the measurement uncertainty when compared with the specification limits of what is being calibrated. Per ANSI/NCSLI Z540.3 Handbook, "It makes use of an observation that for a given Test Uncertainty Ratio (TUR), there is a maximum PFA value for all values of the M&TE test point in-tolerance probability. Applying a guard band based on this maximum PFA value, and the corresponding TUR ensures that the PFA is 2 % or less regardless of the in-tolerance probability." It also results in guard bands with acceptance limits much larger than that of method 5.

The acceptance limit in Method 6 is given by:

$$A_{2\%} = L - U_{95\%} \times M_{2\%}$$

where:

$A_{2\%}$	=	Acceptance limit used to achieve a maximum 2 % PFA
L	=	Tolerance limit (generally the manufacturer's specification)
$U_{95\%}$	=	Calibration process 95 % expanded uncertainty
$M_{2\%}$	=	The multiplier of the 95 % expanded calibration process uncertainty that provides a guard band to ensure 2 % PFA

NOTE: In other methods described in this appendix, the symbol A is used to describe the acceptance limits for any specific application. In this method, $A_{2\%}$ is being used to be consistent with the paper cited.

Combining the expressions for $M_{2\%}$ and $A_{2\%}$ results in:

$$A_{2\%} = L - U_{95\%} \times \left[1.04 - e^{(0.38\log(TUR) - 0.54)}\right]$$

The downside of Method 6 is it only works with TUR ratios of 0.76:1 through 4.6:1. Any ratio higher or lower will cause errors with not calculating the acceptance limits properly.

Comparing Method 5 versus Method 6

Below is a table using the same 10,000 lbf device, using the same variables as shown in figures 2-4, which are a 0.01 resolution and a CMC uncertainty component of 0.0016 % from Morehouse, who was used as the reference laboratory resulting in a 0.08 lbf calibration process at the 10,000 lbf pt.

DIFFERENCE IN ACCEPTANCE LIMITS METHOD 5 VS METHOD 6

Force or Torque Applied	Avg Instrument Reading	Method 6 AL	PASS/FAIL	Method 5 AL	PASS/FAIL	% Diff in AL
1000	1000.00	5.00	PASS	4.98	PASS	0.33%
2000	2000.00	5.00	PASS	4.97	PASS	0.63%
3000	3000.00	5.00	PASS	4.95	PASS	0.94%
4000	4000.00	5.00	PASS	4.93	PASS	1.25%
5000	5000.00	5.00	PASS	4.92	PASS	1.57%
6000	6000.00	4.99	PASS	4.90	PASS	1.88%
7000	7000.00	4.99	PASS	4.88	PASS	2.19%
8000	8000.00	4.99	PASS	4.87	PASS	2.50%
9000	9000.00	4.99	PASS	4.85	PASS	2.81%
10000	10000.00	4.99	PASS	4.84	PASS	3.13%

Figure 9: Difference in Acceptance Limits Method 5 versus Method 6 using a Reference Standard with an Expanded Uncertainty of 0.0016 %

DIFFERENCE IN ACCEPTANCE LIMITS METHOD 5 VS METHOD 6

Force or Torque Applied	Avg Instrument Reading	Method 6 AL	PASS/FAIL	Method 5 AL	PASS/FAIL	% Diff in AL
1000	1000.00	4.99	PASS	4.74	PASS	4.89%
2000	2000.00	4.97	PASS	4.49	PASS	9.80%
3000	3000.00	4.96	PASS	4.23	PASS	14.74%
4000	4000.00	4.95	PASS	3.97	PASS	19.70%
5000	5000.00	4.93	PASS	3.72	PASS	24.69%
6000	6000.00	4.84	PASS	3.46	PASS	28.45%
7000	7000.00	4.72	PASS	3.20	PASS	32.10%
8000	8000.00	4.59	PASS	2.95	PASS	35.84%
9000	9000.00	4.46	PASS	2.69	PASS	39.71%
10000	10000.00	4.32	PASS	2.43	PASS	43.74%

Figure 10: Difference in Acceptance Limits Method 5 versus Method 6 using a Reference Standard with an Expanded Uncertainty of 0.025 %

DIFFERENCE IN ACCEPTANCE LIMITS METHOD 5 VS METHOD 6

Force or Torque Applied	Avg Instrument Reading	Method 6 AL	PASS/FAIL	Method 5 AL	PASS/FAIL	% Diff in AL
1000	1000.00	4.97	PASS	4.49	PASS	9.80%
2000	2000.00	4.95	PASS	3.97	PASS	19.70%
3000	3000.00	4.84	PASS	3.46	PASS	28.45%
4000	4000.00	4.59	PASS	2.95	PASS	35.84%
5000	5000.00	4.32	PASS	2.43	PASS	43.74%
6000	6000.00	4.04	PASS	1.92	PASS	52.48%
7000	7000.00	3.74	PASS	1.41	PASS	62.40%
8000	8000.00	3.43	PASS	0.89	PASS	73.97%
9000	9000.00	3.11	PASS	0.38	PASS	87.81%
10000	10000.00	2.78	PASS	-0.13	FAIL	104.83%

Figure 11: Difference in Acceptance Limits Method 5 versus Method 6 using a Reference Standard with an Expanded Uncertainty of 0.05 %

When we analyze the data in Figures 8 through 10, it becomes apparent that the differences between Method 5 and Method 6 start to become quite drastic as the calibration process uncertainty increases. The CMC uncertainty component of the reference laboratory impacts the calibration process uncertainty, the resolution of the Test Instrument, and possibly the repeatability of the Test Instrument, which may or may not have been included in the calibration process uncertainty.

The laboratory with the low CMC uncertainty component in Figure 8, shows the least amount of % difference from using Method 5. However, the formulas are based on the measurement process uncertainty, which includes the UUT's resolution and repeatability.

If the UUT's resolution and repeatability were to increase, the % difference would increase. Method 5 is the most affected as we subtract the measurement process uncertainty from the upper specification limits and add it to the lower specification limit to obtain our acceptance limits.

Figure 10 shows that the calibration laboratory would not make a conformity assessment using Method 5 at the last calibrated test point. However, using Method 6 allows that same laboratory to make a statement of conformity, assuming the measured value falls within the specified tolerance limits.

Conclusion

Any method used for calculating PFA will have both positive and negatives associated with implementation. The ISO/IEC 17025:2017 standard better addresses measurement risk by requiring the laboratory to report which specifications are not met and the decision rule applied. The decision rule applied needs to consider the location of the measurement for reporting False Accept Risk (PFA).

Throughout this paper, the author has demonstrated that TUR only shows the ratio of the specified tolerance compared to the calibration process uncertainty. If the ratio is too large, a laboratory may not be able to make a statement of conformance with complying with ISO/IEC 17025:2017.

Furthermore, the author shows why a 4:1 or better TUR might not be enough to control risk without several other conditions being met.

It is important to analyze the measurement's location to ensure the measured value falls within the acceptance limits calculated by the accepted guard banding method used.

The best chance of continually meeting tolerance requirements is to use a reference lab (Calibration vendor) with the lowest CMC uncertainty component that replicates how the instrument is used. Also, the end-user must purchase the right equipment capable of continually achieving the desired result or adjust the tolerance appropriately.

Annex (Sample Calculation of TUR)

Example: A customer sends a **10,000 lbf** load cell for calibration with an accuracy specification of **± 0.05 % of full scale**. The calibration provider uses a Universal Calibrating Machine to perform the calibration. When **10,000 lbf** is applied, the unit reads **10,001 lbf**. The display resolution is **1 lbf**.

Step 1: Calculate the numerator

$$\text{TUR} = \frac{\text{Span of the} \pm \text{Tolerance}}{2 \times k_{95\%}\left(\sqrt[2]{\left(\frac{\text{CMC}}{k_{\text{CMC}}}\right)^2 + \left(\frac{\text{Resolution}_{\text{UUT}}}{\sqrt[2]{12}}\right)^2 + \left(\frac{\text{Repeatability}_{\text{UUT}}}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 11: TUR Formula Nominator

The device is a **10,000 lbf** load cell with an accuracy specification of **± 0.05 %**

10,000 * 0.0005 = ± 5 lbf

The upper specification limit is 10,000 + 5 = 10,005 lbf

The lower specification limit is 10,000 – 5 = 9,995 lbf

Therefore, the Span of the ±Tolerance is 10,005 – 9,995 = 10 lbf

$$TUR = \frac{10\ lbf}{2\ x\ k_{95\%}\left(\sqrt[2]{\left(\frac{CMC}{k_{CMC}}\right)^2 + \left(\frac{Resolution_{UUT}}{\sqrt[2]{12}}\right)^2 + \left(\frac{Repeatability_{UUT}}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 12: TUR Formula with the Numerator added

Step 2: Calculate the denominator

Everything is calculated to **1 standard deviation (Standard Uncertainty)** for this calculation.
Calibration and Measurement Capability (CMC)

$$\text{TUR} = \frac{\text{Span of the} \pm \text{Tolerance}}{2 \times k_{95\%}\left(\sqrt[2]{\left(\frac{\text{CMC}}{k_{\text{CMC}}}\right)^2 + \left(\frac{\text{Resolution}_{\text{UUT}}}{\sqrt[2]{12}}\right)^2 + \left(\frac{\text{Repeatability}_{\text{UUT}}}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 13: CMC portion of the denominator

CMC is the uncertainty at the calibrated force. The Universal Calibrating Machine has an uncertainty of **0.02 % at 10,000 lbf**.

The CMC is **10,000 * 0.0002 = 2 lbf**

k_{CMC} **is 2**, which was listed on the calibration provider's certificate.

Dividing the CMC by 2, the standard uncertainty is reported at **one standard deviation**. In most cases, the **CMC uncertainty component is reported at approximately 95 %**, and **a coverage factor of $k = 2$** is used.

$$TUR = \frac{10\ lbf}{2\ x\ k_{95\%}\left(\sqrt[2]{\left(\frac{2\ lbf}{2}\right)^2 + \left(\frac{Resolution_{UUT}}{\sqrt[2]{12}}\right)^2 + \left(\frac{Repeatability_{UUT}}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 14: TUR Formula with CMC added

UUT Resolution

$$\text{TUR} = \frac{\text{Span of the} \pm \text{Tolerance}}{2\,x\,k_{95\%}\left(\sqrt[2]{\left(\frac{CMC}{k_{CMC}}\right)^2 + \left(\frac{Resolution_{UUT}}{\sqrt[2]{12}}\right)^2 + \left(\frac{Repeatability_{UUT}}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 15: Resolution portion of the denominator

Resolution$_{UUT}$ for force instrument is calculated by dividing the force applied by the output at applied force and then multiplying this by the instrument's readability.

The **Resolution**$_{UUT}$ is **(10,000 lbf / 10,000 lbf) * 1 = 1 lbf**

To convert **1 lbf** resolution to standard uncertainty, it is either divided by the **square root of 12**, or the square root of 3 depending on the Type of resolution.

$$TUR = \frac{10\ lbf}{2\ x\ k_{95\%}\left(\sqrt[2]{\left(\frac{2\ lbf}{2}\right)^2 + \left(\frac{1\ lbf}{\sqrt[2]{12}}\right)^2 + \left(\frac{Repeatability_{UUT}}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 16: TUR Formula with Resolution added

Repeatability

$$TUR = \frac{\text{Span of the} \pm \text{Tolerance}}{2 \times k_{95\%}\left(\sqrt[2]{\left(\frac{CMC}{k_{CMC}}\right)^2 + \left(\frac{Resolution_{UUT}}{\sqrt[2]{12}}\right)^2 + \left(\frac{Repeatability_{UUT}}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 17: Repeatability portion of the denominator

For this example, **five replicate readings** are taken.

Repeatability is obtained by applying a force of **10,000 lbf** to the **Unit Under Test (UUT)** five times, and the sample standard deviation of five replicated measurements is calculated.

Repeatability of sample size five: **(10,000, 10,001, 10,000, 10,001, 10,001) = 0.54772**
Since the repeatability is already expressed as one standard deviation, the divisor is 1.

$$TUR = \frac{10\ lbf}{2\ x\ k_{95\%}\left(\sqrt[2]{\left(\frac{2\ lbf}{2}\right)^2 + \left(\frac{1\ lbf}{\sqrt[2]{12}}\right)^2 + \left(\frac{0.54772}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 18: TUR Formula with Repeatability added

Other Error Sources

$$TUR = \frac{\text{Span of the} \pm \text{Tolerance}}{2 \times k_{95\%}\left(\sqrt[2]{\left(\frac{CMC}{k_{CMC}}\right)^2 + \left(\frac{Resolution_{UUT}}{\sqrt[2]{12}}\right)^2 + \left(\frac{Repeatability_{UUT}}{1}\right)^2 + \cdots (u_{Other})^2}\right)}$$

Figure 19: Other error sources in the denominator

Other error sources attributed to the **CPU** can be considered for the **UUT**. Some examples are environmental influences, error in correction factors, etc. For this example, other error sources are inherent in repeatability and **CMC**.

$$TUR = \frac{10\ lbf}{2\ x\ k_{95\%}\left(\sqrt[2]{\left(\frac{2\ lbf}{2}\right)^2 + \left(\frac{1\ lbf}{\sqrt[2]{12}}\right)^2 + \left(\frac{0.54772}{1}\right)^2}\right)}$$

Figure 20: TUR Formula with all error sources added

Calculate the Denominator

Sum of all the contributors = **SQRT((2/2)^2+(1/3.464)^2+(0.54772/1)^2) = 1.1762**

$$\textbf{TUR} = \frac{10\ lbf}{2\ x\ k_{95\%}\ (1.1762)}$$

Figure 21: TUR Calculated

The specification of **10 lbf** is divided by: **2 * k at 95 %** Calibration Process Uncertainty (**k = 2** for this example)

$$\textbf{TUR} = \frac{10\ lbf}{2 \times 2.35231} \qquad \textbf{TUR} = \frac{10\ lbf}{4.70462}$$

Figure22: TUR Calculated

TUR = 2.1256

Want to learn more?

Henry Zumbrun and Dilip Shah teach classes together at Morehouse Instrument Company about twice a year where the participants can learn more about the proper practices to ensure measurements are compliant to the ISO/IEC 17025:2017 and provide the tools to help minimize measurement errors.

References

[1] Requirements for the Calibration of Measuring and Test Equipment, 2006, ANSI/NCSL Z540.3-2006

[2] NCSLI RP – 18 Estimation and Evaluation of Measurement Decision Risk

[3]Crowder, Stephen; Delker, Collin; Forrest, Eric; Martin, Nevin. 2020. Introduction to Statistics in Metrology. Springer Nature Switzerland AG.

[4] ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories

[5] Handbook for the ANSI/NCSL Z540.3-2006, 2009, ANSI/NCSL International

[6] NCSLI RP-12 Determining and Reporting Measurement Uncertainty

[7] ILAC-P14:09/2020 ILAC Policy for Uncertainty in Calibration