What Is Measurement Risk?

Henry A. Zumbrun

Morehouse Instrument Company, Inc.

Introduction

Imagine that a satellite is launched into space and communications are intermittent. This happens because the satellite is wobbling, which causes connection problems in the receiver. The cause of the wobbling is identified: it is the result of not using a calibration provider with a low enough uncertainty. The load cells used to measure the amount of fuel stored in the satellite must be highly accurate with very low uncertainties. However, if a calibration provider does not have the right measurement capability, the load cells will not be accurate enough to make the measurement. In this case, the result is a wobbling satellite and significant resources to fix the problem.

If the problem is not using a calibration provider with an uncertainty adequate to perform the calibration, how does one figure out how low is good enough? This article answers this question by defining measurement risk, and the role Test Uncertainty Ratios (T.U.R) play in reducing measurement risk.

Understanding Measurement Risk

AS9100C defines risk as "[a]n undesirable situation or circumstance that has both a likelihood of occurring and a potentially negative consequence." It further states that "The focus of measurement quality assurance is to quantify and/or manage the 'likelihood' of incorrect measurement-based decisions. When doing so, there must be a balance between the level of effort involved in, and the risks resulting from, making an incorrect decision. In balancing the effort versus the risks, the decision (direct risk) and the consequences (indirect risk) of the measurement must be considered."

ANSI/NCSLI Z540.3-2006 defines Measurement decision risk as probability that an incorrect decision will result from a measurement.

What Does This Really Mean?

All measurements have a percentage 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; a product that doesn't meet quality standards passes undetected through a manufacturer's quality control system and enters the consumer market.

An example of this would be the batteries in the Samsung Note 7 phone. The batteries can potentially overheat, causing the phone to catch on fire. In this case, the faulty battery/charging system of the phone device was approved through the quality control process of the manufacturer, which was a 'false accept decision.' If you owned one of these phones, there was a risk of fire and potential damage and injury.

In metrological terms, consumer's risk is like the false accept risk, or Probability of False Accept (PFA). The biggest difference is that in the metrology field, the false accept risk is usually limited to a maximum of 2 percent. In cases where the estimation of this probability is not feasible, there is a requirement for a Test Uncertainty Ratio (TUR) to be 4:1 or greater to ensure lowering the PFA to a low risk level.
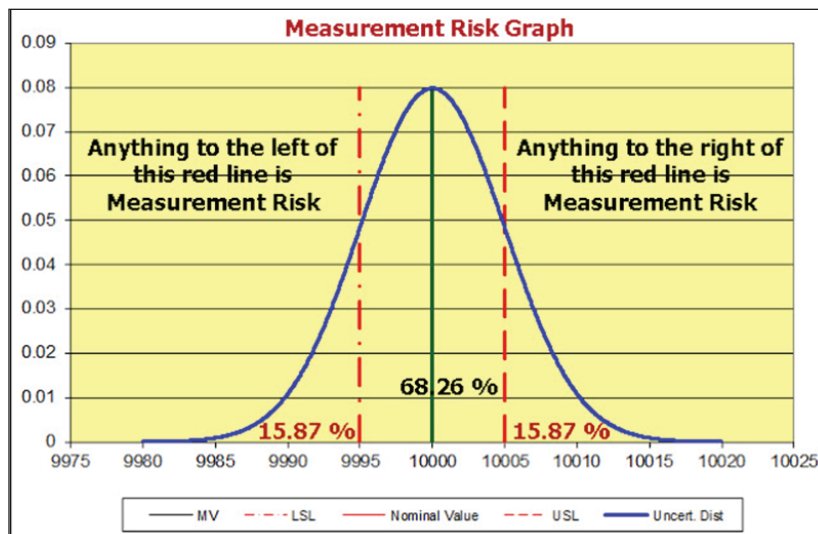

Figure 1. Graph showing the measurement risk which is the Probability of False Accept (PFA).

So, what does this mean for a metrology laboratory? It means that any lab making a statement of compliance, calling an instrument "in tolerance," must consider measurement uncertainty and properly calculate T.U.R. considering the location of the measurement. In simplistic terms, T.U.R. = Tolerance Required/Uncertainty of the Measurement (at a 95% confidence interval). If the Uncertainty of the Measurement is not less than the tolerance required, there will be a significant risk of false accept. In simplistic terms, a TUR that produces less than +/- 2 % upper and lower risk would be required to ensure the measurement is valid.

Keys to lowering measurement risk include having your calibration provider replicate how the instrument is used in the field, having competent technicians, using the right equipment, and lowering overall uncertainties by the calibration provider. There is quite a bit of difference between force measurement labs with CMCs of 0.1 percent, 0.05 percent, 0.02 percent, 0.01 percent, 0.005 percent and 0.002 percent of applied force. Not using the laboratory with the right capability to meet your requirements is like using a ruler to calibrate a gauge block.

Table 1 shows the Test Uncertainty Ratios (TUR) that force calibration labs with different calibration capabilities can provide for various levels of required tolerances. The far-left column represents the calibration standard required for force measurements. Deadweight primary standards are often required to achieve CMCs of better than 0.01 % of applied force. A high-end load cell calibrated by deadweights would be required to achieve CMCs of better than 0.05 %. This table indicates the best TUR that the labs can provide for the same load cell at similar conditions. Per this table, only calibration labs with CMCs around 0.02 % or better can calibrate devices with a tolerance of 0.1 %. They may still need to adjust the device to read closer to the nominal value. We will discuss guard banding later.

The table was derived from TUR and uncertainty formulas found in JCGM 100:2008 and ANSI/NCSLI Z540.3-2006. The formulas used to determine TUR and

Uncertainty are as follows:

$$TUR = \frac{Tolerance}{Expanded\ Uncertainty}$$

$$TUR = \frac{(USL - LSL)}{4 * u}$$

where:
TUR = Test Uncertainty Ratio,
USL = Upper Specification Limit,
LSL = Lower Specification Limit, and
u = standard uncertainty.

Note: We are using 4 assuming k=2, the proper formula would be 2 times the actual k value is for a 95 % confidence interval.

The Calculation of TUR for Tolerances

- ((Upper Specification Limit - Lower Specification Limit))/(4 * Standard Uncertainty)

- Combined Uncertainty (u) – The square root of the sum of the squares of all the input quantity uncertainty components.

$$u = \sqrt{\left(\frac{CMC}{k}\right)^2 + \left(\frac{Res}{3.464}\right)^2 + \left(\frac{Rep}{1}\right)^2}$$

- CMC = Calibration and Measurement Capability. This should be found on the calibration report.

- Res = This is the resolution of the Unit Under Test (UUT) The divisor for resolution will either be 3.464 or 1.732 (depending on how the UUT least significant digit resolves).

- Rep = Repeatability of the Unit Under Test (UUT). Repeatability of UUT must be used if repeatability studies were not previously accounted for in the CMC. If accounted for in the CMC, this would not be required.

How Good Does Your Calibration Provider Have to Be? (T.U.R. Table)

Calibration Standard Required			Tolerance Required					
			0.010%	0.020%	0.050%	0.100%	0.200%	0.500%
Deadweight	Calibration Lab capability (CMC)	0.002%	4.471	8.941	22.353	44.706	89.413	223.532
Deadweight		0.005%	1.961	3.922	9.805	19.610	39.221	98.052
Deadweight / Lever		0.010%	0.995	1.990	4.975	9.950	19.900	49.751
High End Load Cell		0.020%	0.499	0.999	2.497	4.994	9.987	24.969
High End Load Cell		0.050%	0.200	0.400	1.000	2.000	3.999	9.998
Good Load Cell		0.100%	0.100	0.200	0.500	1.000	2.000	5.000
This table is based on a Calibration Grade Load Cell with 0.01 lbf Resolution; 0.05 lbf Repeatability.								
Anything in Red would have too much measurement risk.								

Table 1. TUR Table

- Expanded Uncertainty - Typically 2 times the standard uncertainty. However, the appropriate k value should be used to ensure a coverage probability of 95 %, based on the effective degrees of freedom using the Welch Satterthwaite formula.

Is Your Calibration Provider Reporting Pass/Fail Criteria Properly?

If the calibration provider is accredited, it needs to follow the requirements per ISO/IEC 17025. ISO/IEC 17025:2005 Clause 5.10.4.2 states that "When statements of compliance are made, the uncertainty of measurement shall be taken into account."

This translates to minimizing the Probability of False Accept (PFA) by applying a guard banding method. ANSI/NCSLI Z540.3 -2006 Handbook discusses guard banding in section 3.3. Section 3.3 paragraph 2 states "As used in the National Standard, a guard band is used to change the criteria for making a measurement decision, such as pass or fail, from some tolerance or specification limits to achieve a defined objective, such as a 2 % probability of false accept. The offset may either be added to or subtracted from the decision value to achieve this objective."
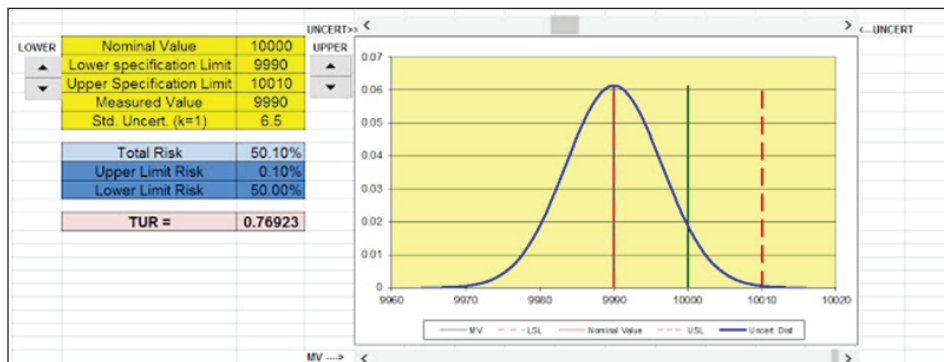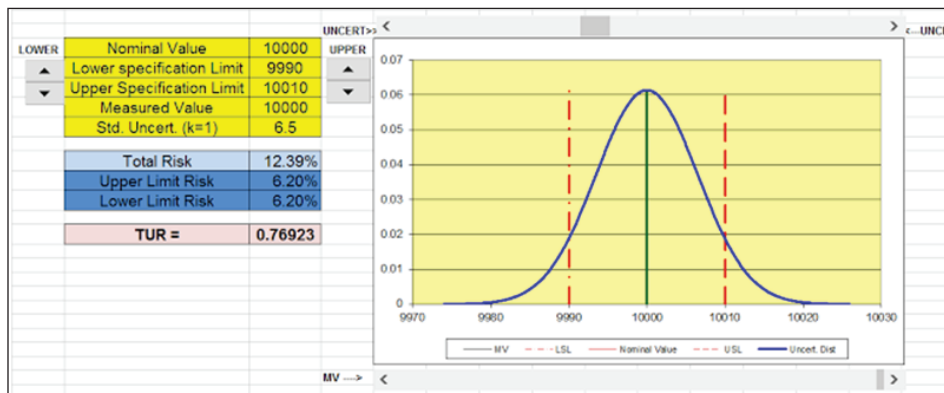
Examples of Calculating Measurement Risk with Guard Banding

Assume we are testing a load cell at 10,000 lbf force. The accuracy specification is 0.1 % of reading (or +/- 10 lbf at this force), and the measured value was 9990. Is the device in tolerance? After all, the calibration laboratory applied 10,000 lbf and the unit under test (UUT) read 9990. The bias is -10 lbf and the device meets its accuracy specification (accept decision without taking the uncertainty of measurement into account). The report is issued and the end user is happy. However, the problem is that the end user should not be happy. If the calibration and measurement capability (CMC) of the calibration laboratory using a specific reference standard was not considered, the end user will not know whether the device meets the accuracy specification required. Basically, this measurement was passed based on the assumption that the calibration providers reference was perfect and they applied exactly 10,000 lbf to the load cell. However, no measurements are perfect. That is why we estimate the uncertainty of measurement to quantify this "imperfection of the measurement." This is a false assumption which neglected the uncertainty in the calibration provider's measurement.

Let us assume that the standard uncertainty was calculated at 6.5 lbf for k=1. In Figure 2, the item being calibrated would normally be considered "in tolerance" by a large percentage of calibration laboratories since the accuracy specification is 0.1 % of reading or +/- 10 lbf and the measured value was within the accuracy specification at 9990 lbf. There is a 50.1 % chance of the calibration being accepted when it is not in tolerance.

Figure 2.

LOWER			UPPER
▲	Nominal Value	10000	▲
▼	Lower specification Limit	9990	▼
	Upper Specification Limit	10010	
	Measured Value	9990	
	Std. Uncert. (k=1)	6.5	
	Total Risk	50.10%	
	Upper Limit Risk	0.10%	
	Lower Limit Risk	50.00%	
	TUR =	0.76923	

Figure 3 shows the risk when the measured value of the UUT reads 10,000 lbf. In this scenario, the bias or measurement error is 0. However, there is still a 12.39 % chance that the UUT is not "in tolerance." Simply put, there is too much risk. We need to lower the standard uncertainty to reduce the risk. Note that the TUR remains the same since it is a ratio not dependent on the location of the measurement.
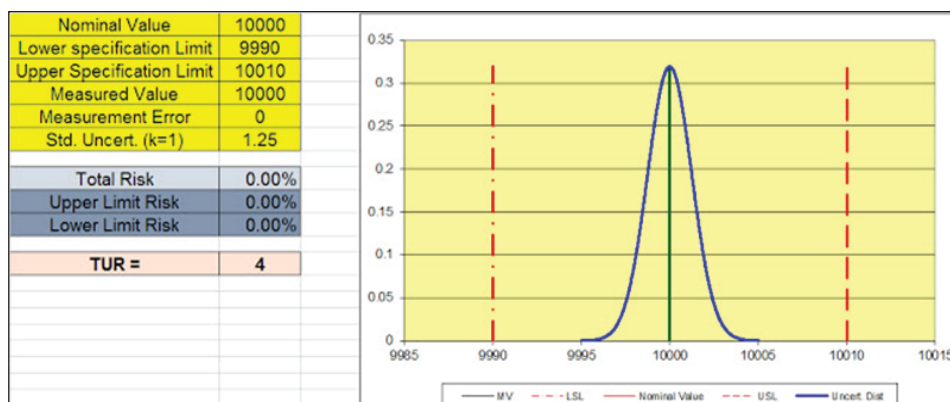

Figure 3.

LOWER			UPPER
▲	Nominal Value	10000	▲
▼	Lower specification Limit	9990	▼
	Upper Specification Limit	10010	
	Measured Value	10000	
	Std. Uncert. (k=1)	6.5	
	Total Risk	12.39%	
	Upper Limit Risk	6.20%	
	Lower Limit Risk	6.20%	
	TUR =	0.76923	

Figure 4.

How to Lower the Risk (PFA) By Lowering the Uncertainty

1. Use better equipment with a lower resolution and/ or better repeatability; e.g., higher quality load cell for force measurement.

2. Use a better calibration provider with a Calibration and Measurement Capability (CMC) low enough to reduce the measurement risk.

3. Pay attention to the uncertainty values listed in the calibration report issued by your calibration provider. Make sure to get proper T.U.R. values for every measurement point (but pay attention to the location of the measurement).

The last graph (Figure 4) shows the same test instrument with a lower Standard Uncertainty. This was a real scenario where an instrument was modified from a 10 lbf resolution to a 2 lbf resolution. The total risk is now 0 and the device will be "in tolerance" with less than 2 % total risk from reading of 9,996 through 10,006 lbf. There are several acceptable methods for applying a guard band to obtain what the measured value needs to be in order to maintain less than 2 % total risk.

These graphs comply with Method 5: Guard Bands Based on Expanded Uncertainty in the ANSI/NCSLI Z540.3 Handbook and is described in ISO 14253-1, and included in ILAC G8, and various other guidance documents.

After reading this paper, you may be standing at a crossroads and wondering if any of this extra work is necessary. To the left is the same rough path you've been travelling all along. This is the path that says, "If it's not broken, why fix it?" You might be thinking that measurement risk has not been an issue before, or you'll just wait until an auditor questions you about it (or there is a train wreck). Yet, to the right is the road that fewer people realize will help solve their measurement problems today.

This road is not more difficult; it's just different from the current way you may be doing things. Choosing to consider the impact of not doing things right—and making the decision to select the best calibration provider—will make all the difference. The rest is just putting formulas in place to report and know your measurement risk.

References

[1] JCGM 100:2008 *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*, BIPM.

[2] ILAC P14:01/2013 *Policy For Uncertainty in Calibration.*

[3] ANSI/NCSLI Z540.3-2006 *Requirements for the Calibration of Measuring and Test Equipment.*

[4] ISO/IEC 17025:2005: *General requirements for the competence of testing and calibration laboratories.*

[5] AS9100C *Quality Management Systems - Requirements for Aviation, Space, and Defense Organizations.*

Henry A. Zumbrun (hzumbrun@morehouse.com), Morehouse Instrument Company, Inc., York, Pennsylvania.