National Aeronautics and Space Administration

A Classical Repeatability Problem

Mihaela Fulop, SGT, Inc.

Problem Statement

The Guide to the Expression of Uncertainty in Measurement (GUM, Ref. 1) implies that repeatability is a key parameter in measurement uncertainty analysis (MUA), but NASA Glenn Research Center's Pressure Calibration Laboratory (Pressure Cal Lab) had not been conducting repeatability tests regularly as part of its calibration procedures for pressure transducers and other pressure instruments.

To address this contribution to uncertainty, Cal Lab staff leveraged the Lean Six Sigma certification process to formally evaluate the effect of adding repeatability to the Cal Lab MUA and to determine implementation solutions that would balance calibration quality and productivity.

Definitions From the "International Vocabulary of Metrology (VIM)" (Ref. 3)

Repeatability condition (para. 2.20)—"condition of measurement, out of a set of conditions that includes the same measurement procedure, same operators, same measuring system, same operating conditions and same location, and replicate measurements on the same or similar objects over a short period of time"

Measurement repeatability (para. 2.21)—"measurement precision under a set of repeatability conditions of measurement"

As defined in the VIM, repeatability is hard, or impossible, to achieve with reasonable productivity.

Potential Impacts of Implementation

Depending on the method chosen to implement repeatability, significant hours are needed for engineering analysis and checklist revision or for a combination of increased technician test time and engineering checklist modification.

Implementing repeatability across the board could reduce productivity by 20% and increase turnaround time. Additional technician (permanent) and engineering (temporary) staff would be needed if customers considered quick-turnaround calibrations critical. However, the failure to conduct a repeatability analysis for our MUA had left a hole in our due diligence documentation.

How should repeatability data be gathered?

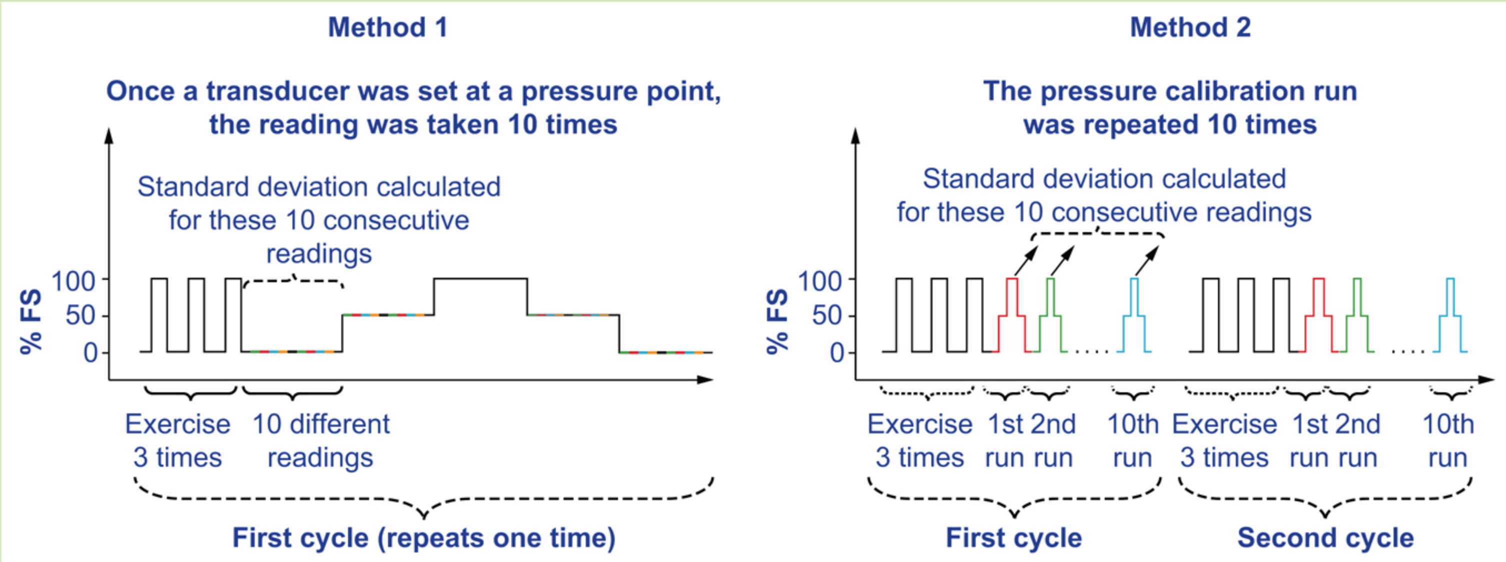
Approach Options

Data were gathered for 30 different transducers via two methods (see the following figures), and the results were compared statistically. For both methods, the pressure was set to 0%, 50%, and 100% of full scale (FS).

In its definition of repeatability, Ref. 2 refers implicitly to Method 2.

Method 2 is the more time consuming method, resulting in a significant loss of productivity.
Method 1 is already heavily implemented in the Pressure Cal Lab and is not resulting in loss of productivity.

If proven that the two methods are statistically equivalent, continuing to use Method 1 instead of switching to Method 2 would save time and money.



Evaluation of Test Results

The estimated standard deviation obtained by Method 1 (s1) was compared statistically with the standard deviation obtained by Method 2 (s2). An F-hypothesis test was performed to check the equality of the two variances:

$$H_0: \quad \sigma_1^2 = \sigma_2^2 \qquad F_{\text{test}} = \frac{s_1^2}{s_2^2}$$

$$H_a: \quad \sigma_1^2 \neq \sigma_2^2$$

The hypothesis that the two variances are equal is rejected if

$$F < F_{1-\frac{\alpha}{2},N_1-1,N_2-1} \quad \text{or} \quad F > F_{\frac{\alpha}{2},N_1-1,N_2-1}$$

where α is the significance level set at 0.05, N_1 is the numerator degrees of freedom, and N_2 is denominator degrees of freedom.

Data Analysis, Results, and Conclusions

Asset number	Repeatability,[a] % FSO		Variance		F-test	$F_{1-\frac{\alpha}{2},N_1-1,N_2-1}$	$F_{\frac{\alpha}{2},N_1-1,N_2-1}$	Significant?	Flag significant
	Method 1	Method 2	Method 1	Method 2					
M599766, Setra 204, 500 psig	0.002271	0.001054	1.960×10^{-8}	0.2809×10^{-8}	6.98	0.25	4.03	Variance between the two methods is ***not*** consistent	1
M606106, Setra 204, 500 psig	0.002667	0.003706	1.778	3.433	0.52	0.25	4.03	Variance between the two methods is consistent	
M600450, Setra 204, 500 psig	0.002667	0.001033	1.778	0.2667	6.67	0.25	4.03	Variance between the two methods is ***not*** consistent	1
M608479, Setra 204, 500 psig	0.001886	0.002633	0.889	1.733	0.51	0.25	4.03	Variance between the two methods is consistent	

[a]FSO, full-scale output.

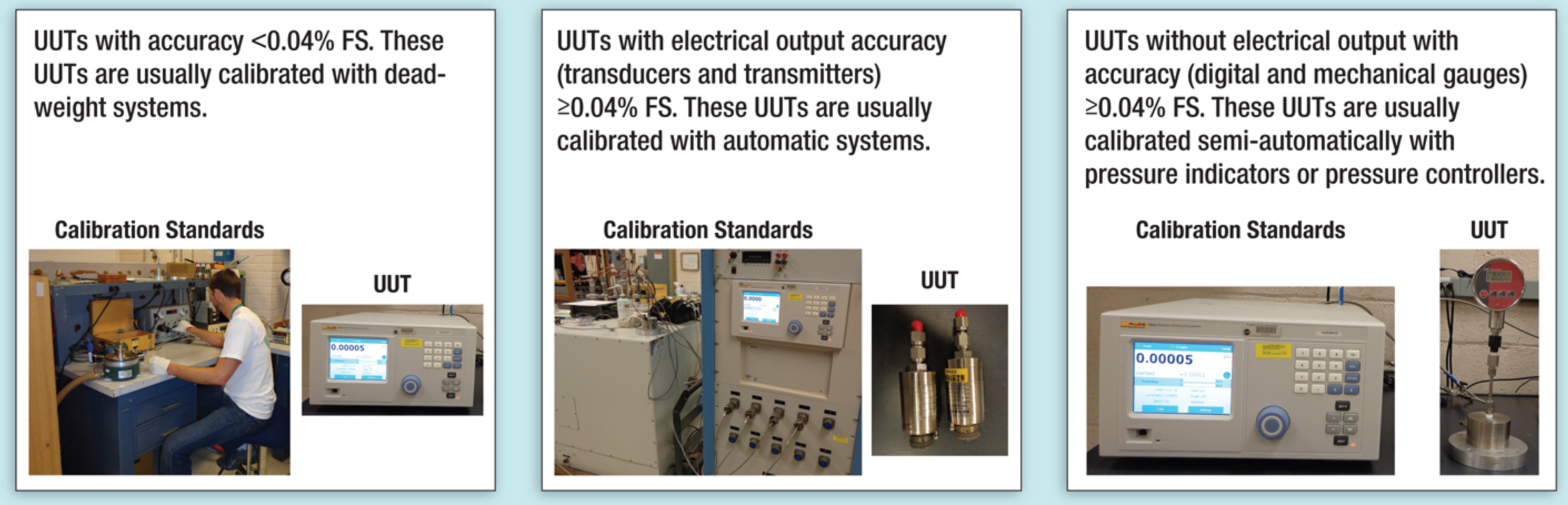
The F-test indicates that in 51.4% of units under test (UUTs) we have enough evidence to reject the null hypothesis that the two methods' variances are equal at the 0.05 significance level. For the rest of the tests, we used Method 2, as implied by Ref. 2.

What will our customers gain if repeatability is implemented in the Pressure Cal Lab?

Two test types are used in the Pressure Cal Lab: tolerance testing and characterization testing. Characterization always needs repeatability tests. We are testing to find out if repeatability is significant for the customer in tolerance testing.

Approach

- The UUTs in the Pressure Cal Lab were divided into three groups according to accuracy class and type (see the following figure).
- Up to this point, only pressure transducers and pressure transmitters with an accuracy of ≥0.4% FS have been tested (see the center of the figure). These units represent about 80% of the instruments in the Pressure Cal Lab.
- In 2013, a total of 2014 pressure transducers and pressure transmitters with an accuracy ≥0.04% FS were calibrated in the Pressure Cal Lab (see pie chart).

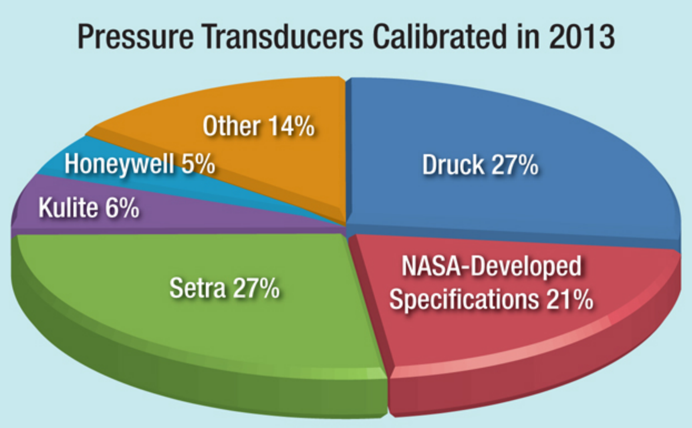


Test Description

A sample of 273 pressure transducers with different outputs were tested using automatic pressure systems. The UUTs

- Were from different manufacturers
- Had different accuracies
- Had pressures ranging from 15 to 3000 psi
- Had different modes: absolute, differential, and gauge

The number of units tested from each manufacturer was based on the percentage in the total population of pressure transducers calibrated in 2013 (see the pie chart). Each transducer was tested with 10 consecutive cycles of 0%, 50%, and 100% FS using Method 2. We assumed that customers who sent units to the Cal Lab for calibration planned to use the units for subsequent tests. For those tests, customers would have to perform their own MUA, and the main contributor to that MUA would be the unit accuracy.



Other is the total UUTs for which Pressure Cal Lab production was 2% or less for an individual manufacturer.

For each transducer, the customer's subsequent test MUA was calculated in two ways:

1. Including Cal Lab uncertainty ***but not*** Cal Lab repeatability.
2. Including Cal Lab uncertainty ***and*** Cal Lab repeatability.
   The results were compared, and the differences were calculated as a percentage of UUT tolerance (see the table).

- If the difference was ≥5%, we could conclude that the Cal Lab repeatability contribution was significant to the customer's subsequent test MUA.
- The 5% limit was based on Ref. 4, which defines a contributor as significant if its contribution increases the calibration measurement capability by ≥5%.

Data Analysis, Results, and Conclusions

An example for model, PMP 4000, accuracy, 0.08% of full-scale output (FSO); output, 0 to 5 V_{DC}

Customer MUA with Cal Lab MUA ***but no*** Cal Lab repeatability, V_{DC}	Customer MUA with Cal Lab MUA ***and*** Cal Lab repeatability, V_{DC}	Difference between columns 1 and 2 as percentage of UUT tolerance, %	Significant? (>5%)
0.00402	0.00410	1.97	No
0.00402	0.00408	1.60	No
0.00402	0.00406	0.90	No
0.00402	0.00410	1.97	No
0.00402	0.00420	4.52	No

Based on our sample, for 95.6% of the UUTs, with a margin of error of 2.3%, adding the Pressure Cal Lab's repeatability to the customer MUA made an insignificant difference (an increase of <5%).

The following formula was used to calculate the margin of error—how closely the sample result should match the population, where p is the sample proportion, n is the sample size, and Z is the Z-value corresponding to a desired confidence level. In our case $Z = 1.96$ (for a 95% confidence level): $Z\sqrt{p(1-p)/n}$

What value should be used in the MUA as a repeatability contribution?

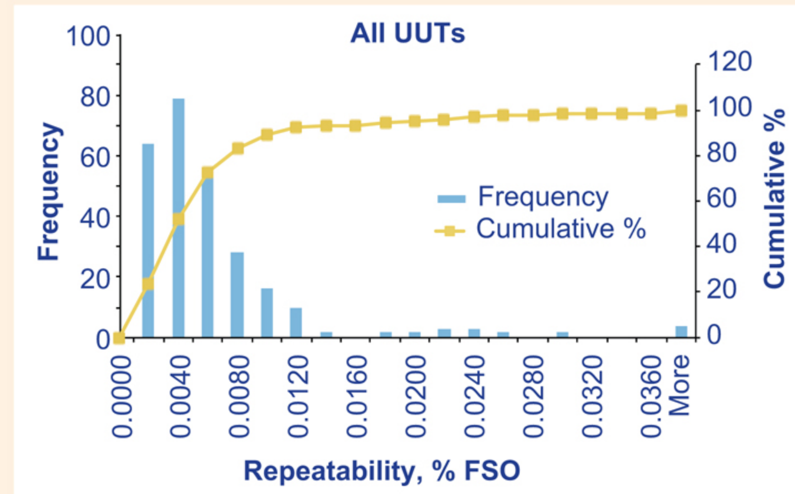
Test Description

A sample of 273 pressure transducers was selected as it was for the previous test (see the center column of this poster). Each transducer was tested for 10 consecutive cycles of 0%, 50%, and 100% of FS using Method 2.

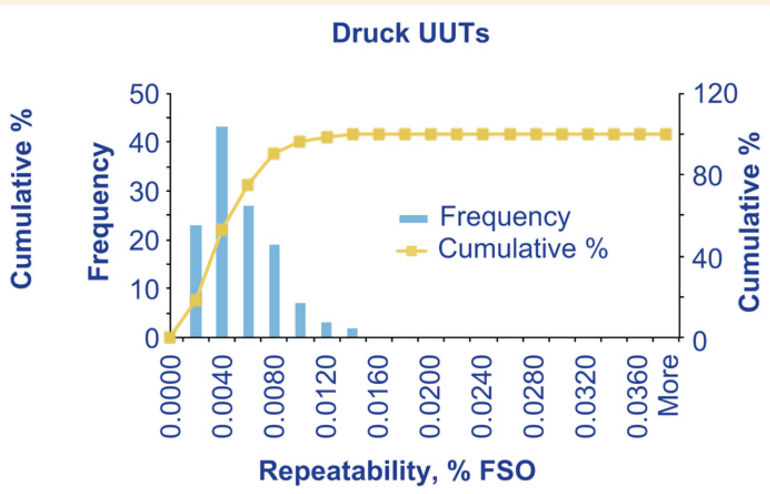
- This time, repeatability was calculated for each UUT as a standard deviation of the 10 readings and was expressed as % FSO.
- A histogram was plotted with the cumulative percentage for the repeatability for the sample of UUTs.

Data Analysis, Results, and Conclusions

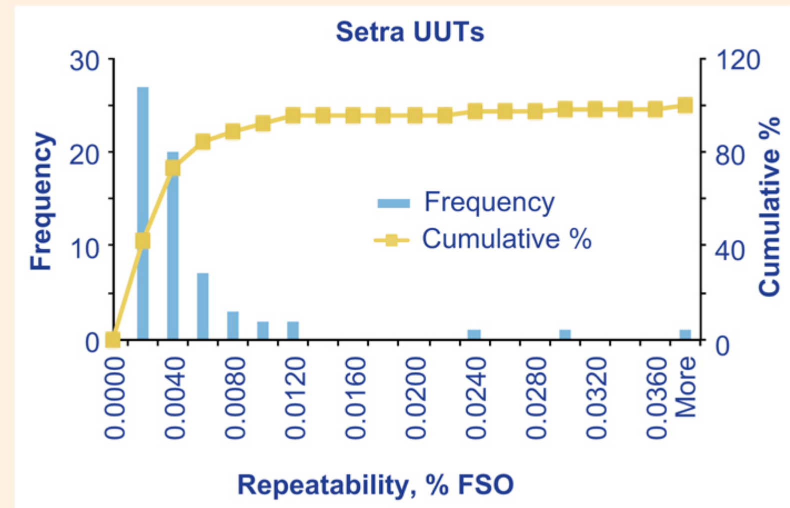
Based on our sample, 95.6% of all UUTs, with a margin of error of 2.5%, have a repeatability of <0.022% FS.



Based on our sample, 95.97% of all Druck UUTs, with a margin of error of 3.4%, have a repeatability of <0.01% FS.



Based on our sample, 95.31% of all Setra UUTs, with a margin of error of 5.1%, have a repeatability of <0.012% FS.



A Word of Caution: These results represent only a snapshot in time. If the process is not stable, we could get different results in the future.

If these values are used as a repeatability contribution to the MUA, it will be important to monitor the Pressure Cal Lab's variation through statistical process control. Consequently, a statistical process control was implemented in the Pressure Cal Lab for the automatic calibration of these types of transducers:

- Check standards are used for all 10-V excitation pressure transducers.
- Historical data from these check standards are analyzed with Shewhart control charts to determine if the process is statistically in control.

References

1. Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement, GUM 1995 With Minor Corrections. JCGM 100:2010, Corrected version, 2010.
2. Electrical Transducer Nomenclature & Terminology: ANSI–ISA Standard S37.1, ISA 37.1–1975 (R1982), The International Society of Automation, Apr. 1, 1975.
3. International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM). JCGM 200:2012, Corrected version, 2012.
4. A2LA Policy on Measurement Uncertainty in Calibration. A2LA–P110, May 2011.