

Data Comparison 10360-2

Table of Contents

Data Comparison 10360-2.....	3
Introduction.....	3
Measurement 10360-2.....	3
Data Reporting.....	4
Data Comparison.....	5
Point Measurement Simulation.....	5
Length Measurement Simulation.....	7
Simulation With 1um Of Error.....	8
Simulation With 2um Of Error.....	9
Simulation With 3um Of Error.....	10
Simulation With 4um Of Error.....	11
Simulation With 5um Of Error.....	12
Comparison Summary.....	13
Comparison Using an En Score.....	13
Calculation of En Score.....	14
En Score Using Max Deviation.....	14
En Score Using Mean Deviation.....	15
General Comments.....	15
Compilation of Measurement10360 Utility.....	16
Compilation Example on MacOS.....	16
Revision History.....	18

Data Comparison 10360-2

Introduction

This article describes a method for comparing E_L measurement results from 10360-2 measurement tests on a CMM. The reason a comparison might be necessary can vary but a good example would be some form of proficiency test.

The 10360-2 standard allows the use of a variety of equipment that includes step gauges, ball plates, or multi-sphere ball bars among other artifacts. This article focus is on gauge measurements where the length is determined from a single point measured on both ends of an artifact which applies to Koba bars, step gauges, and gauge blocks.

The method used to demonstrate the preferred approach for comparison is by simulation. The utility created for this purpose is called *measurement10360* and provided as source code. Compilation of this utility is described at the end of this article.

Measurement 10360-2

The method of measurement for a 10360-2 performance test is outlined in the 10360-2 standard which is either AMSE B89.4.10360-2:2008 or ISO/IEC 10360-2:2009. Excerpts from the standard related to the measurement are in the following sections.

Requirements from 10360-2 section 6.2 Principle:

The principle of the assessment method is to use a calibrated test length, traceable to the metre, to establish whether the CMM is capable of measuring within the stated maximum permissible error of length measurement for a CMM with a specified ram axis stylus tip offset (both 0 and 150), E_0 , MPE and E_{150} , MPE, and within the stated maximum permissible limit for the repeatability range, R_0 , MPL.

The assessment shall be performed by comparison of the indicated values of five different calibrated test lengths, each measured three times, relative to their calibrated values. The indicated values are calculated by point-to-point length measurements projected onto the alignment direction (see also Annex C).

Each of the three repeated measurements is to be arranged in the following manner: if one end of the calibrated test length is labelled "A" and the other end "B", then the measurement sequence is either A1 B1, A2 B2, A3 B3 or A1 B1, B2 A2, A3 B3. Other sequences such as A1 A2 A3, B1 B2 B3 are not permitted. Each of the three repeated measurements shall have its own unique measured points. That is, in general, B1, B2 and B3 shall be different actual points of the same target point B. Once the measurement sequence for a test length has begun, no additional probing points shall be measured other than those required to measure its length; for instance, no alignment points are permitted between the measurement of A1 and B3.

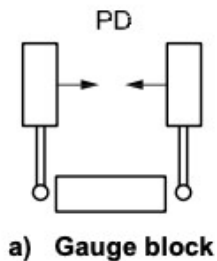
Requirements from 10360-2 section 6.3.3.2 Measurement Procedure:

Data Comparison 10360-2

For each of the five calibrated test lengths, obtain three measurement results. See Annex B for details regarding the measurement procedure for specific types of test lengths. Repeat for all seven measurement positions for a total of 105 measurement results from the calibrated test lengths.

Annex B section B.2.1 General:

Bidirectional measurements of calibrated gauges represent a calibrated test length. A bidirectional measurement involves probing a single point on each end of the gauge, and approaching these probing points from diametrically opposite directions (see Figure B.1).



Data Reporting

The method used to report the deviation from any 10360-2 measurement line varies based on interpretation of the standard. Some 10360-2 calibration reports were found that don't generate a measured value but instead results are shown graphically with respect to the tolerance limits. Although this is fine it makes comparison next to impossible without access to the values.

The method used by SCI is to report the largest absolute deviation from any individual 10360-2 measurement line as shown in the following samples. Each of the 15 individual length measurement are treated uniquely and all length measurements are expected to fall within the specification limits taking into account the measurement uncertainty. The reported value is the worst absolute deviation found.

Table 1: 10360-2 Measurement Sample 1. Units are mm and um for deviations.

Nominal	Length 1	Dev 1	Length 2	Dev 2	Length 3	Dev 3	Uc (k=2)
440.0044	440.0057	1.3	440.0041	-0.4	440.0040	-0.4	0.0012
880.0076	880.0052	-2.5	880.0049	-2.8	880.0042	-3.5	0.0016
1320.0105	1320.0052	-5.3	1320.0048	-5.7	1320.0042	-6.2	0.0021
1760.0075	1760.0060	-1.5	1760.0046	-2.9	1760.0053	-2.2	0.0026
2200.0100	2200.0057	-4.3	2200.0051	-4.8	2200.0049	-5.1	0.0032
						E0:	0.0062

Data Comparison 10360-2

Table 2: 10360-2 Measurement Sample 2. Units are mm and um for deviations.

Nominal	Length 1	Dev 1	Length 2	Dev 2	Length 3	Dev 3	Uc (k=2)	
439.9972	440.0000	2.8	440.0007	3.4	440.0006	3.3	0.0012	
879.9968	880.0013	4.5	880.0013	4.5	880.0014	4.6	0.0016	
1319.9994	1320.0008	1.4	1320.0017	2.3	1320.0019	2.5	0.0021	
1760.0043	1760.0007	-3.5	1760.0022	-2.0	1760.0014	-2.9	0.0027	
2200.0058	2200.0005	-5.3	2200.0010	-4.8	2200.0018	-4.0	0.0033	
							E0:	0.0053

Table 3: 10360-2 Measurement Sample 3. Units are mm and um for deviations.

Nominal	Length 1	Dev 1	Length 2	Dev 2	Length 3	Dev 3	Uc (k=2)	
230.0004	230.0008	0.5	230.0005	0.2	230.0000	-0.4	0.0011	
470.0010	470.0002	-0.8	470.0000	-1.0	469.9992	-1.8	0.0012	
710.0010	709.9993	-1.7	710.0004	-0.7	709.9996	-1.5	0.0013	
950.0029	950.0020	-0.9	950.0011	-1.8	949.9990	-3.9	0.0015	
1190.0037	1190.0011	-2.6	1190.0010	-2.8	1189.9990	-4.8	0.0018	
							E0:	0.0048

For the purposes of comparison it would be necessary to treat the max deviations as a signed result. From a reporting point of view showing the maximum absolute deviation makes sense but from a comparison point of view that does not work if the max deviations have opposite signs.

Data Comparison

Two methods were considered for comparison of results:

- Individual deviation of lengths.
- Mean deviation of the three measurement lengths for each of the five lengths.

Regardless of the comparison method it is necessary to recreate measurement data that is unbiased and excludes influence from the environment or other aspects of a real CMM measurement. This is the purpose of the *measurement10360* utility as data is created with only a varying amount of randomness to the individual points used for any specific length measurement.

Point Measurement Simulation

The point data used for the length measurement simulation is from a normal distribution random number generator. The important part is that the data is generated with a normal distribution as

Data Comparison 10360-2

this is what would be expected when performing an actual measurement on a CMM.

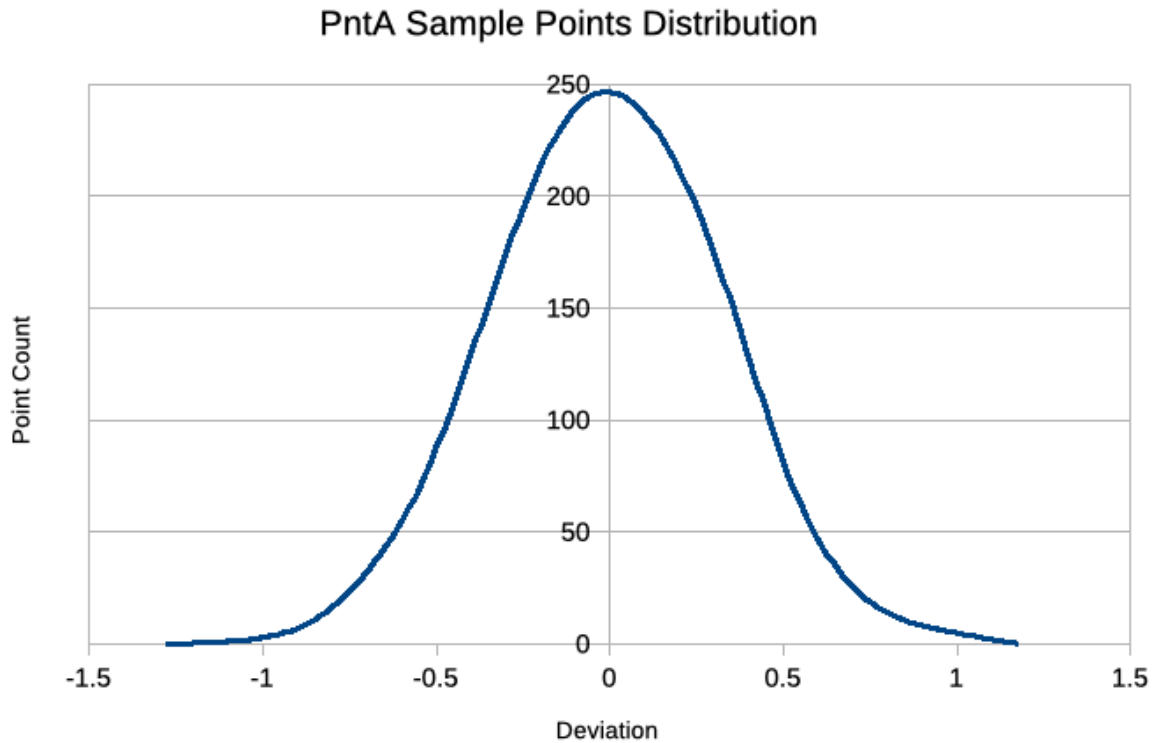


Illustration 1: Example of point distribution from 1000 samples at one end of a simulated measurement artifact.

The data shown in illustration 1 shows the distribution of the points measured on one end of a measurement artifact. The point data was created with the following command line options from the *measurement10360* utility:

```
./measurement10360 -gauge_length 1000 -max_error 1.0 -meas_count 1000 -output_name output.txt
```

Where:

-gauge_length: Longest measurement length.
-max_error: Maximum error.
-meas_count: Total number of length measurements.
-output_name: Name of output file to create.

When processed the following results for the example data shown in illustration 1 for one of the length measurement points was found:

Std. Dev.	0.33204
Std. Dev. X2	0.66409
Std. Dev. X3	0.99613
Mean:	-0.00681

Data Comparison 10360-2

With an input error of 1.0 mm it was expected that three standard deviations of the data would be just under 1.0 mm (technically 99% of the input range).

Length Measurement Simulation

The length measurement data was created with the following command line options from the *measurement10360* utility:

```
./measurement10360 -gauge_length 1000 -max_error 0.001 -meas_count 3 -output_name output1.txt
```

Where:

-gauge_length: Longest measurement length.
-max_error: Maximum error.
-meas_count: Total number of length measurements.
-output_name: Name of output file to create.

Output File Example:

```
# Input Parameters
# Gauge Length:      1000.000000000 mm
# Deviation Slope:   0.000000000 mm/m
# Maximum Error:     0.005000000 mm

Nominal PtA: 0.000000000, Nominal PtB: 200.000000000
PtA:  -0.000660432, PtB:  199.996803571, Len:  199.997464003
PtA:  -0.000114047, PtB:  199.998144546, Len:  199.998258593
PtA:  -0.000019148, PtB:  200.000874950, Len:  200.000894098

Nominal PtA: 0.000000000, Nominal PtB: 400.000000000
PtA:  -0.001610229, PtB:  400.000889032, Len:  400.002499260
PtA:   0.001059191, PtB:  400.001581556, Len:  400.000522365
PtA:  -0.001431095, PtB:  399.996997707, Len:  399.998428802

Nominal PtA: 0.000000000, Nominal PtB: 600.000000000
PtA:   0.002463755, PtB:  600.001515855, Len:  599.999052100
PtA:  -0.000344578, PtB:  600.000843152, Len:  600.001187730
PtA:  -0.000925294, PtB:  600.002048265, Len:  600.002973558

Nominal PtA: 0.000000000, Nominal PtB: 800.000000000
PtA:  -0.001576710, PtB:  799.997778338, Len:  799.999355048
PtA:   0.001265917, PtB:  799.999269568, Len:  799.998003651
PtA:   0.002634509, PtB:  800.002425810, Len:  799.999791302

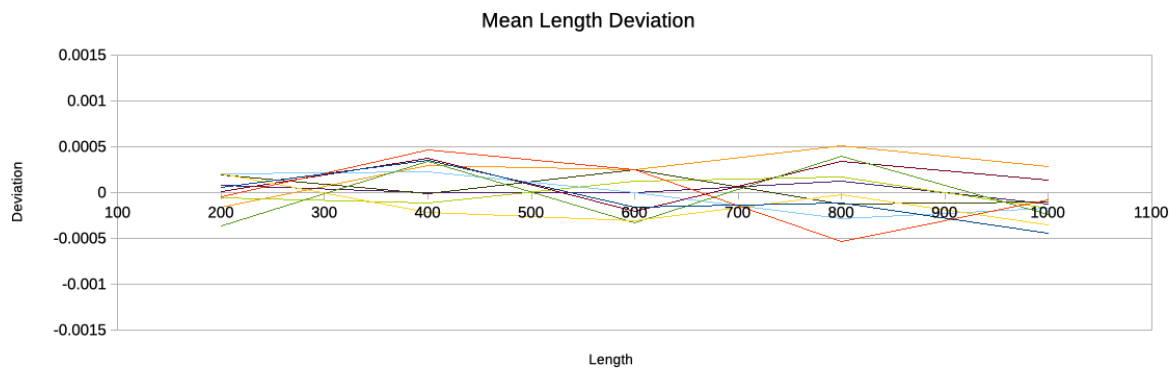
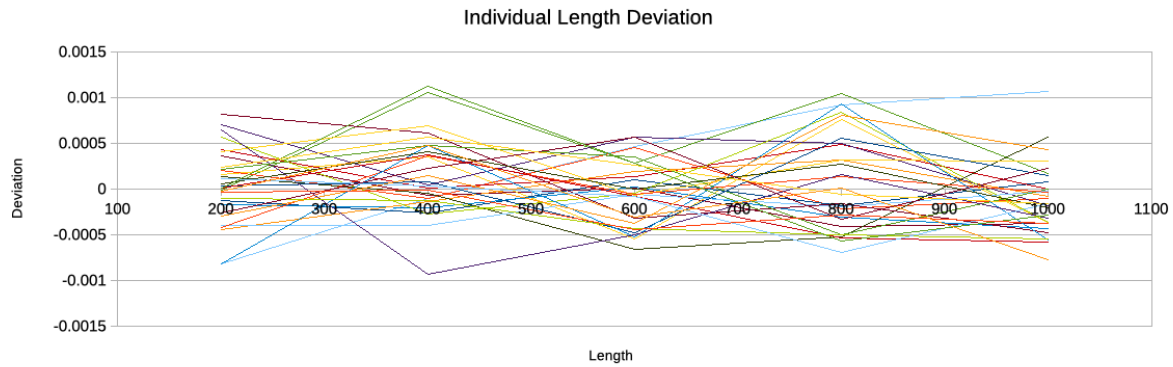
Nominal PtA: 0.000000000, Nominal PtB: 1000.000000000
PtA:  -0.002621254, PtB:  999.999308172, Len:  1000.001929426
PtA:  -0.000210498, PtB:  999.999674280, Len:  999.999884778
PtA:  -0.002014852, PtB:  999.999554536, Len:  1000.001569388
```

The *measurement10360* utility was run ten times for each error level (0.001 to 0.005 mm) with output files named output1.txt to output10.txt for a total of 50 results.

Data Comparison 10360-2

Simulation With 1um Of Error

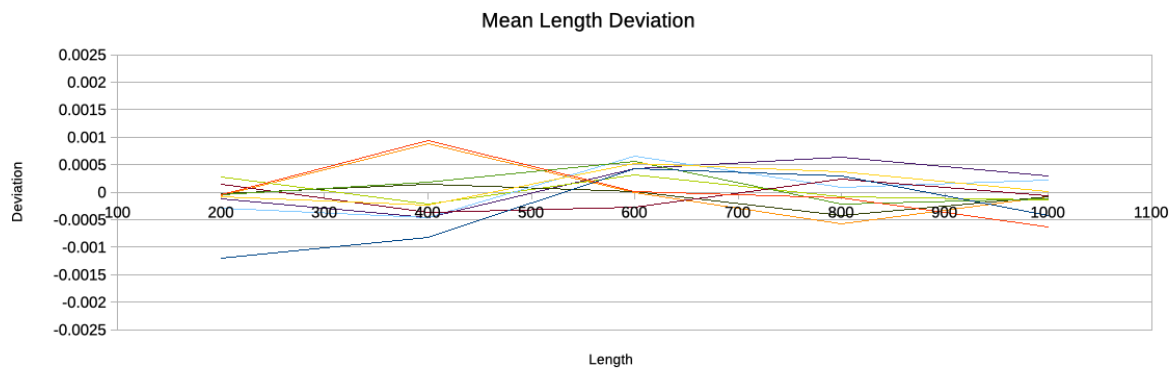
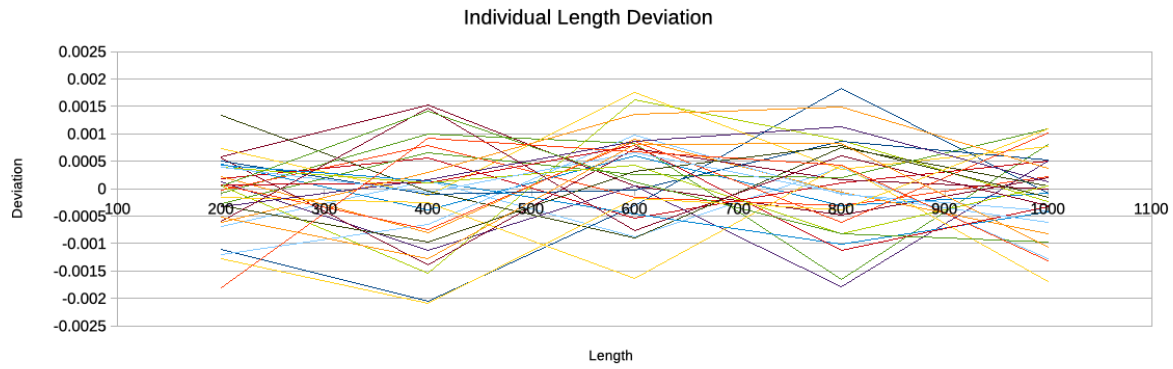
The following graphs show results of 10 simulated measurement samples with a random error of 0.001 mm on each end of the artifact. The vertical resolution of the two graphs is set to the same value in order to show the difference in method.



Data Comparison 10360-2

Simulation With 2um Of Error

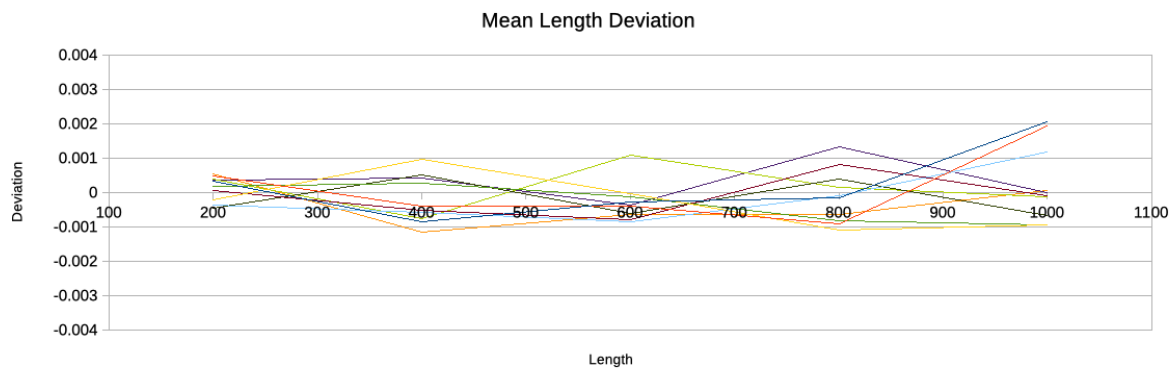
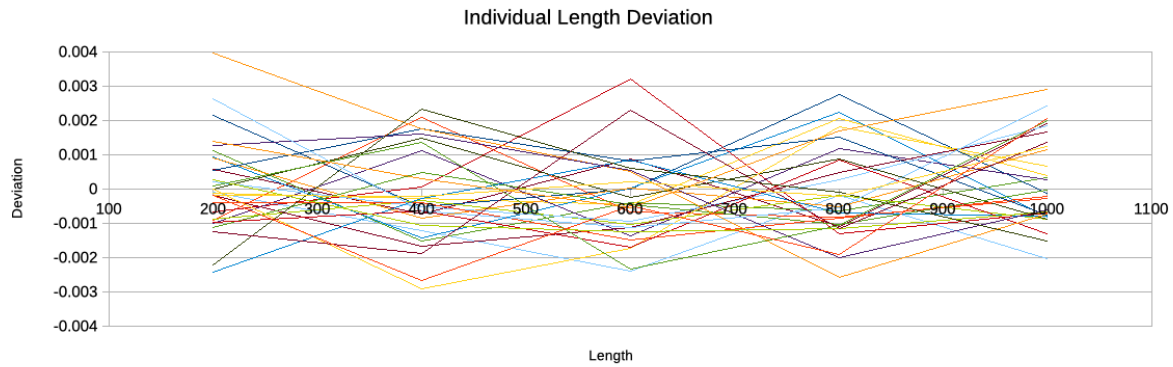
The following graphs show results of 10 simulated measurement samples with a random error of 0.002 mm on each end of the artifact. The vertical resolution of the two graphs is set to the same value in order to show the difference in method.



Data Comparison 10360-2

Simulation With 3um Of Error

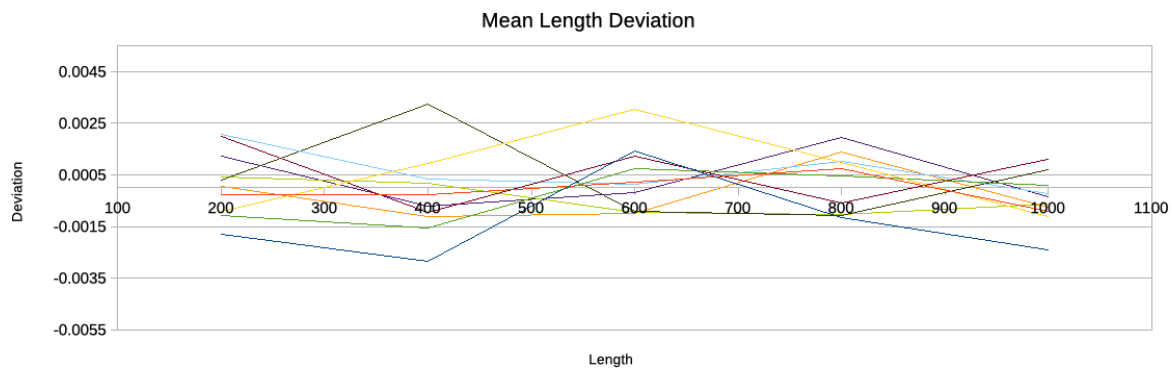
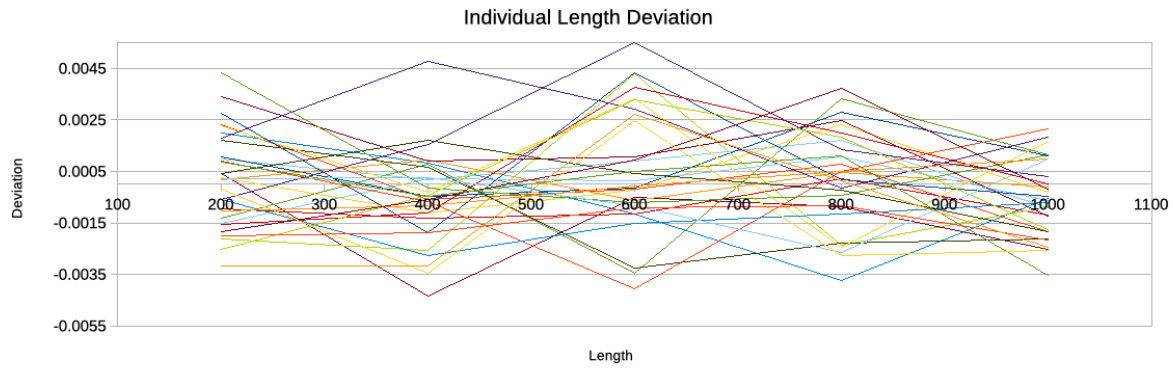
The following graphs show results of 10 simulated measurement samples with a random error of 0.003 mm on each end of the artifact. The vertical resolution of the two graphs is set to the same value in order to show the difference in method.



Data Comparison 10360-2

Simulation With 4um Of Error

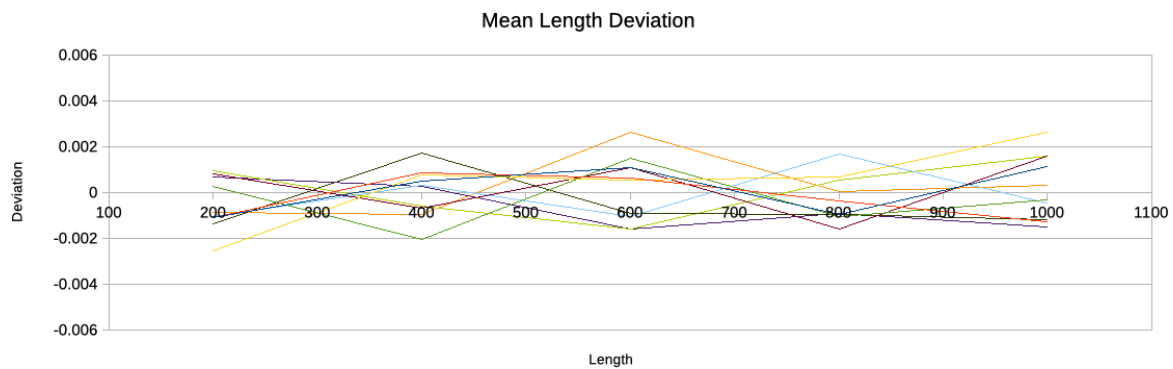
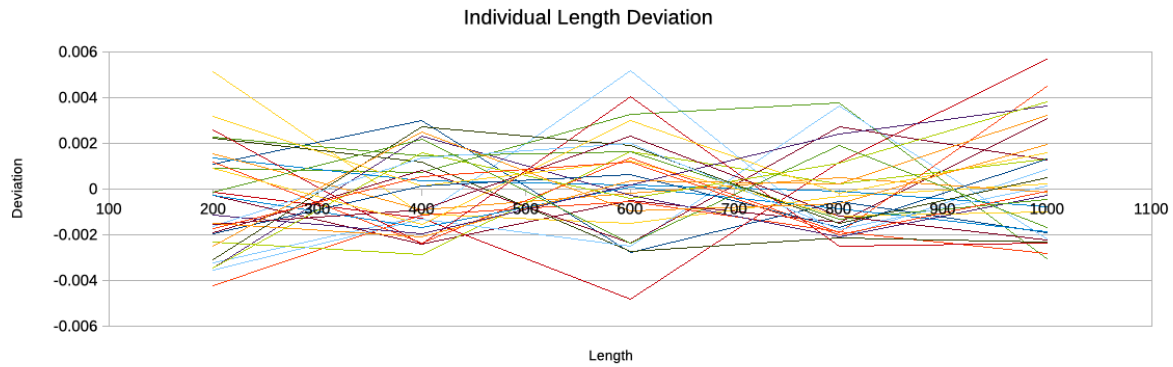
The following graphs show results of 10 simulated measurement samples with a random error of 0.004 mm on each end of the artifact. The vertical resolution of the two graphs is set to the same value in order to show the difference in method.



Data Comparison 10360-2

Simulation With 5um Of Error

The following graphs show results of 10 simulated measurement samples with a random error of 0.005 mm on each end of the artifact. The vertical resolution of the two graphs is set to the same value in order to show the difference in method.



Data Comparison 10360-2

Comparison Summary

The following table summarizes the results from the individual length deviations:

1 um	Max Dev:	0.00112	Max R0:	0.00149
	Min Dev:	-0.00094	Stdev R0:	0.00072
	Range:	0.00206	Mean R0:	0.00038
2 um	Max Dev:	0.00184	Max R0:	0.00300
	Min Dev:	-0.00208	Stdev R0:	0.00149
	Range:	0.00392	Mean R0:	0.00077
3 um	Max Dev:	0.00396	Max R0:	0.00640
	Min Dev:	-0.00290	Stdev R0:	0.00223
	Range:	0.00686	Mean R0:	0.00128
4 um	Max Dev:	0.00548	Max R0:	0.00840
	Min Dev:	-0.00432	Stdev R0:	0.00292
	Range:	0.00981	Mean R0:	0.00173
5 um	Max Dev:	0.00566	Max R0:	0.00938
	Min Dev:	-0.00482	Stdev R0:	0.00353
	Range:	0.01047	Mean R0:	0.00178

The following table summarizes the results from the mean length deviations:

1 um	Max Dev:	0.00051
	Min Dev:	-0.00053
	Range:	0.00104
2 um	Max Dev:	0.00094
	Min Dev:	-0.00120
	Range:	0.00214
3 um	Max Dev:	0.00206
	Min Dev:	-0.00117
	Range:	0.00322
4 um	Max Dev:	0.00325
	Min Dev:	-0.00285
	Range:	0.00610
5 um	Max Dev:	0.00264
	Min Dev:	-0.00257
	Range:	0.00521

Based on the simulation it appears that the range of lengths (R0) is roughly 2:1 of the individual point range. When using the mean length the range is roughly 1:1 as compared to the individual point range.

Comparison Using an En Score

The typical method for comparison is using the En score. The 10360-2 standard treats each length measurement uniquely which must fall within the acceptance limits therefore the estimated

Data Comparison 10360-2

uncertainty would be to the level of each individual measurement length without any consideration of the length repeatability (a separate evaluation of the data).

Consider this measurement from the 5um sample data:

Table 4: Sample data for 5um error.

<i>Nominal</i>	<i>Actual</i>	<i>Deviation</i>
600.0000	599.9991	-0.0009
600.0000	600.0012	0.0012
600.0000	600.0030	0.0030

The reported length deviation from the 10360-2 report data would be 0.0030 mm as this is the largest deviation with a length repeatability R0 of 0.0039 mm at this length. The odds of a repeated measurement showing the exact same result is very small as demonstrated by the measurement simulations. Typically the uncertainty would not include anything related to the repeatability as each individual length is considered separately. This highlights the problem when trying to compare results using the maximum error. There may be a way to do this but it seems problematic particularly since the number of samples is limited.

Calculation of En Score

The standard formula to calculate the En value is:

$$En = \frac{(x - X)}{\sqrt{(Ux^2 + UX^2)}}$$

Where:

En: Result
x: Max length deviation from source 1
X: Max length deviation from source 2
Ux: Estimated expanded measurement uncertainty from source 1
UX: Estimated expanded measurement uncertainty from source 2

En Score Using Max Deviation

For this to work it would be necessary to have more than 3 length measurement samples to start with. When using the worst deviation value from a 10360-2 test for comparison purposes there is a random element to this as it is using the extreme result. Assuming a sufficient number of samples existed where the max deviation was stable then the En score could be calculated normally. Running simulations to find the minimum number of samples is possible but also outside of the requirements of 10360-2 which makes this pointless as the number of length measurement samples will always be 3. If a method exists to handle this better then this article will be updated but, at the time of writing, no practical way is known to do this reliably.

Data Comparison 10360-2

En Score Using Mean Deviation

Using the mean value of a limited set of length measurements is a more stable approach. It is assumed the estimated uncertainty from either source does not include any contribution from measurement length repeatability as the 10360-2 standard treats each length uniquely so this would need to be included. The formula to calculate a comparison En score using the mean result and combining the length repeatability would therefore be:

$$En = \frac{(x - X)}{\sqrt{(Ux^2 + U(xR_0)^2 + UX^2 + U(XR_0)^2)}}$$

Where:

En: Result
x: Mean length deviation from source 1
X: Mean length deviation from source 2
Ux: Estimated expanded measurement uncertainty from source 1
U(xR₀): Estimated expanded measurement uncertainty based on R0 data from source 1.
UX: Estimated expanded measurement uncertainty from source 2
U(XR₀): Estimated expanded measurement uncertainty based on R0 data from source 2.

Where:

$$UR_0 = 2x \frac{\text{stdev}(R_0)}{\sqrt{\text{count}}}$$

UR₀: Estimated expanded measurement uncertainty for length repeatability.
R₀: Length repeatability range.
count: Number of samples.

Since the comparison is done against the mean value the uncertainty contribution from the range of the repeatability data needs to be adjusted to take into account the number of length measurements. When using the standard 3 length measurements the contribution will be larger than if there was 1000 length measurements as the mean value is expected to be central with respect to the range of data samples.

It is not clear if this is the proper way to compare measurements but what is clear is that, when the mean and effect of repeatability is taken into account, the results are more realistic. The 10360-2 standard treats the length measurements more like a go-nogo gauge which is not suitable for comparison purposes.

General Comments

When doing any kind of comparison between 10360-2 results that are based on single point length measurements some form of altered comparison method is needed as repeatability are not taken into account. Standards such as ASME B89.4.1:1997 ballbar didn't show the same kinds of problems as each sphere was measured with a number of points that essentially filter out the single point errors. As shown in the simulation data the variations in length are roughly 2x worse

Data Comparison 10360-2

for 10360-2 tests than what would be expected for simulated ballbar measurements (it depends on the number of points but 0.5:1 is reasonable for an 8 point sphere measurement of a ballbar).

In addition it makes more sense to compare a mean measurement value than an extreme value particularly when more than one measurement exists. Measurements that are intended to be compared are always thorough and repeated in order to produce the best result. For example, if measuring the diameter of a ring gauge on a CMM it is unlikely that a user will rely on a three point circle but instead measure with many points and repeat a number of times under different conditions. Results submitted for comparison will always be what is believed to be the best mean result.

Compilation of Measurement10360 Utility

The *measurement10360* utility used to generate the test data is provided as source code. The *measurement10360* utility make file is created using the cmake utility. In order to compile this utility on any target operating system it is necessary to have the following:

- C++ compiler with a minimum C++11 support.
- CMake

To compile the *measurement10360* utility the following steps are needed:

- Open a terminal window (console window on Windows).
- Navigate to the root directory of the source code.
- Run the following commands:

```
cmake .  
make
```

NOTE: It may be necessary to specify the cmake generator depending on the operating system and preferred compiler. For example to use MinGW on Windows instead of the default MSVC the MinGW Makefiles must be specified at the command line:

```
cmake . -G "MinGW Makefiles"  
make
```

Compilation Example on MacOS

The following is an example showing the compilation of the *measurement10360* utility on MacOS Catalina. Software such as Xcode along with the command line utilities and cmake are installed.

Data Comparison 10360-2

```
Rons-iMac:~ ron$ cd MyProjects/measurement10360
```

```
Rons-iMac:measurement10360 ron$ cmake .
-- The C compiler identification is AppleClang 12.0.0.12000032
-- The CXX compiler identification is AppleClang 12.0.0.12000032
-- Detecting C compiler ABI info
-- Detecting C compiler ABI info - done
-- Check for working C compiler:
/Applications/Xcode.app/Contents/Developer/Toolchains/XcodeDefault.xctoolchain/usr/bin/cc - skipped
-- Detecting C compile features
-- Detecting C compile features - done
-- Detecting CXX compiler ABI info
-- Detecting CXX compiler ABI info - done
-- Check for working CXX compiler:
/Applications/Xcode.app/Contents/Developer/Toolchains/XcodeDefault.xctoolchain/usr/bin/c++ - skipped
-- Detecting CXX compile features
-- Detecting CXX compile features - done
-- Configuring done (2.0s)
-- Generating done (0.0s)
-- Build files have been written to: /Users/ron/MyProjects/measurement10360
```

```
Rons-iMac:measurement10360 ron$ make
[ 25%] Building CXX object CMakeFiles/measurement10360.dir/main.cpp.o
[ 50%] Building CXX object CMakeFiles/measurement10360.dir/cmdline.cpp.o
[ 75%] Building CXX object CMakeFiles/measurement10360.dir/measurement10360.cpp.o
[100%] Linking CXX executable bin/measurement10360
[100%] Built target measurement10360
```

```
Rons-iMac:measurement10360 ron$ bin/measurement10360 -h
Command line options:
  -dev_slope Deviation slope in mm/m
  -gauge_length Stepgauge length (50 - 5000 mm)
  -h Display argument help and exit
  -max_error Maximum error in mm
  -meas_count Measurement count of lengths
  -output_name Name of output file
  -quiet Suppress output
  -v Show version and exit
```

Data Comparison 10360-2

Revision History

<i>Date</i>	<i>Version</i>	<i>Changes</i>
Jun 4, 2026	1.0	Initial release