Table of Contents

Uncertainty Analysis: Instrument Calibration

1. PURPOSE OF PROCEDURE

The purpose of this procedure is to provide methods for the evaluation of instrument calibration uncertainty at the 95 % confidence level. ITTC (1999) describes calibration methods and documentation.

2. SCOPE

Contemporary laboratories acquire data with digital data acquisition systems. For conversion to engineering or physical units, instrumentation connected to these systems must be calibrated. This procedure describes methods for applying uncertainty estimates to these calibrations. Most instrumentation is highly linear; consequently, the calibration includes a linear fit to the data. Usually, most of the uncertainty is associated with the data scatter in the regression fit. This procedure also describes uncertainty estimates for non-linear curve fits, which are usually associated with an analytical model of the data such as the relation between thrust coefficient versus advance ratio in propeller performance.

Torque transducers, load cells, and block gages are typically calibrated in a calibration stand by mass. The uncertainty analysis procedure for force and torque calibration by mass is discussed.

Finally in some cases, signals are inherently digital such as pulse devices for rotating machinery such as propellers and carriage calibration wheels. A method for estimation of the uncertainty in rotational rates and carriage speed is described.

3. GENERAL

Since the laboratories within ITTC are considered to be world-class, all measurements should be traceable with the appropriate reference standard to the respective National Metrology Institutes (NMI) within each country. Usually, the uncertainty in the reference standard should be small relative to the data scatter in the calibration. All calibrations should be through system calibrations or end-to-end calibrations with the same data acquisition system and software as applied during the test. If the calibration is not end-to-end, the AD (analogue to digital converter) should be calibrated per the procedures outlined here. The uncertainty in the AD should be included with the uncertainty in the instrument calibration from another system such as a voltmeter or another computer system.

A schematic of the end-to-end calibration process is shown in Figure 1. A known measured physical input is applied to the instrumentation system such as roll angle, for example. The physical input is then measured by an NMI traceable measurement. The physical input is converted to a voltage by an electronic instrument. Amplification is then applied to the signal so that the expected voltage range matches the range of the AD converter. The output from the amplifier is then processed by a low-pass filter or anti-aliasing filter, which matches the frequency range of the electronic instrument. The filtered signal is digitized by the AD converter at a data rate, which is consistent with the Nyquist sampling theorem (Otnes and Enochson, 1972, and Bendat and Piersol, 2000). Finally, the data is processed by software and output as data in voltage units of the AD converter.

In principle, the physical input may also be accomplished with computer control. In that case, the total process may be automated under computer control. The details of implementation may vary from Figure 1.

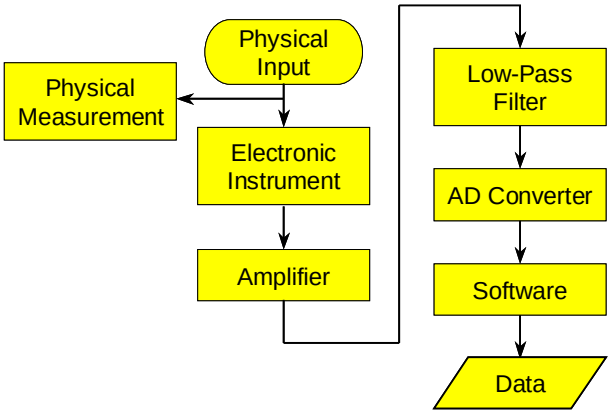
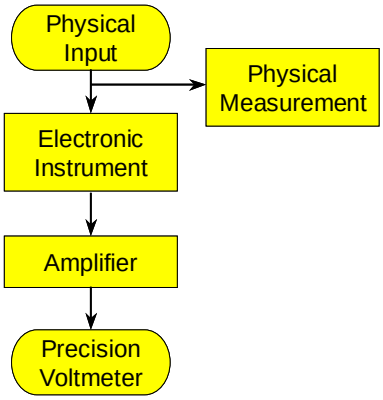


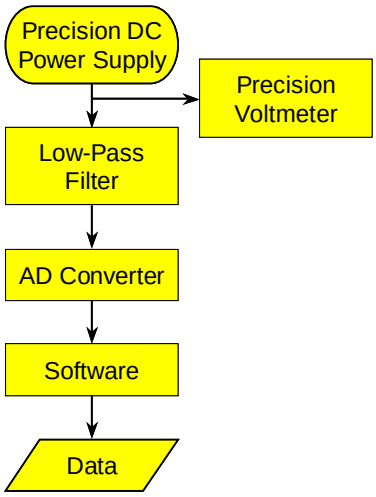
Figure 1: End-to-end calibration schematic

A typical alternative to this calibration is illustrated in Figure 2. In this case, the calibration process is split into 2 parts. In the first part, the instrument is benched calibrated with a voltmeter or another computer system as a voltmeter. The computer system is calibrated separately with a precision DC power supply and precision voltmeter. Typically, an AD converter is 16 bits. In this case, the contribution to the combined uncertainty from the computer system is likely small in comparison to the instrument calibration.

An instrument should be calibrated over approximately ten (10) equal increments over the range of the instrument in engineering or physical units. The highest and lowest values of the calibration should not over-range the AD. Nominally, the uncertainty will be constant in physical units over the range of calibration.



a. Instrument calibration



b. Computer system calibration

Figure 2: Data acquisition system calibration in parts

For most instrumentation, the uncertainty is stated as percent full-scale. After an instrument is calibrated, the constants in physical units are entered into the software for the test. As verification that the constants are correctly entered, the calibration should be checked through the system at 3 points: high, low, and mid-range.

At each data point for the calibration, approximately 100 to 1000 samples should be collected via the AD. The signal should be filtered

at a value consistent with the frequency response of the transducer. The sample rate should be at least 2.5 times the filter cut-off frequency of the low-pass anti-aliasing filter. The mean, standard deviation, and number of samples should be recorded. This information is necessary in the computation of the uncertainty by the Type A method per the JCGM Uncertainty Guide (2008) or GUM.

4. LINEAR REGRESSION ANALYSIS

4.1 Fundamentals of linear regression analysis

4.2 Fundamentals of linear regression analysis

The fundamentals of linear regression analysis may be found in a number of texts on statistical theory, such as Ross (2004) and Devore (2008). Many commercial codes are available. Some of the fundamentals quantities may be calculated from a spreadsheet. The fundamental equation is given by

$$y = a + bx \tag{1}$$

where y is the dependent value in either Volts from a voltmeter or digital Volts from an AD, x is the independent variable in physical units from the reference standard, a is the intercept and b is the slope. The intercept and slope from linear regression theory are then:

$$b = s_{xy} / s_{xx} \tag{1a}$$

$$a = \bar{y} - b\bar{x} \tag{1b}$$

$$\bar{x} = (1/n)\sum_{i=1}^{n} x_i \tag{1c}$$

$$\begin{aligned} s_{xy} &= \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\ &= \sum_{i=1}^{n} x_i y_i - (1/n)\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \end{aligned} \tag{1d}$$

Typically, calibration data have been plotted with equation (1). For highly linear calibration data, all data will lie on a straight line. A better representation of the statistical character of calibration data is the residual plot. A residual is defined as follows:

$$Resid = y_i - a - bx_i \tag{2}$$

A residual is the difference between the measured data and the curve fit. Example residual plots will be presented later in this procedure.

The sum of the square of the residuals is then

$$SS_R = \sum_{i=1}^{N}(y_i - a - bx_i)^2 \tag{3}$$

A measure of the standard deviation for the linear regression analysis is the standard error of estimate.

$$SEE = \sqrt{SS_R/(n-2)} \tag{4}$$

The standardized residual is equation (2) divided by *SEE* per Ross (2004). A plot of standardized residuals is useful in the identification of outliers.

The uncertainty in the slope and offset may be useful. The standard deviation or the standard uncertainty in the intercept and slope are, respectively, from Ross (2004)

$$u_a = s_a = SEE\sqrt{(1/n)\sum_{i=1}^{n} x_i^2 / s_{xx}} \tag{5a}$$

$$u_b = s_b = SEE / \sqrt{s_{xx}} \tag{5b}$$

For application of regression analysis in the conversion of digital data to physical units, the uncertainty for the curve fit is determined by the prediction limit.

$$\text{Re}sid / SEE = \pm t_{\alpha/2,n-2}\sqrt{(n+1)/n + (x-\bar{x})^2 / s_{xx}} \quad (6)$$

where t is the Student t inverse probability density function at $\alpha/2$ confidence limit from JCGM (2008). Care must be taken determining this value from a table. Tables can be based on either a 1-tail or 2-tail distribution. Usually at the 95 % confidence limit, α = 0.05 or $\alpha/2$ = 0.025 for a 2-tailed distribution. As a check, the Student t is Gaussian for an infinite number of samples so that t = 1.96 at the 95 % confidence level.

For implementation in a data processing code, the equation for conversion to physical units is given by

$$y' = A + Bx' \quad (7)$$

where y' is the dependent variable in physical units, x' the independent variable in digital Volts, and the slope and intercept from equations (1) are

$$\begin{aligned} A &= -a / b \\ B &= 1 / b \\ SEE' &= SEE / b \end{aligned} \quad (7a)$$

4.3 Calibration theory.

Scheffe (1973) has developed a statistical theory of calibration. A simplified method with detailed examples has been proposed by Carroll, et al. (1988). The prediction limit at the 95 % confidence limit in this case is given by

$$\begin{aligned} f(x) - SEE(c_1 + c_2 s_{xx}) \le y \le f(x) \\ + SEE(c_1 + c_2 s_{xx}) \end{aligned} \quad (8)$$

where

$$c_1 = t_{n-2}, c_2 = \sqrt{2F_{2,n-2}}$$

and F is the inverse Fisher probability density function.

From equation (7), the uncertainty in x for a linear equation in physical units is then from Scheffe (1973)

$$\begin{aligned} x_h = \bar{x} + [bD - (-1)^h (SEE \times c_2) \times \\ (C/N + D^2 s_{xx})^{1/2}] / C \end{aligned} \quad (9a)$$

$$C = b^2 - (c_2 SEE)^2 / s_{xx} \quad 9b)$$

$$D = b(x - \bar{x}) - (-1)^h SEE \times c_1 \quad (9c)$$

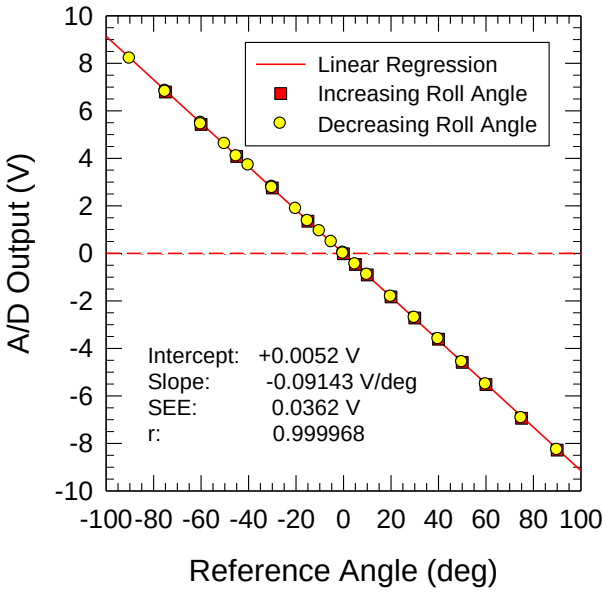
where h = 1 and 2 are the upper and lower bounds, respectively.

The inverse Student t and Fisher pdfs may be found in tables in standard statistical references and mathematical handbooks such as Ross (2004) and Devore (2008). These functions are also available in many computational codes and spreadsheets such as Microsoft Excel. Other statistical functions such as the slope (a), intercept (b), average ($\bar{x}$) and standard error of estimate (SEE) are available in these codes.

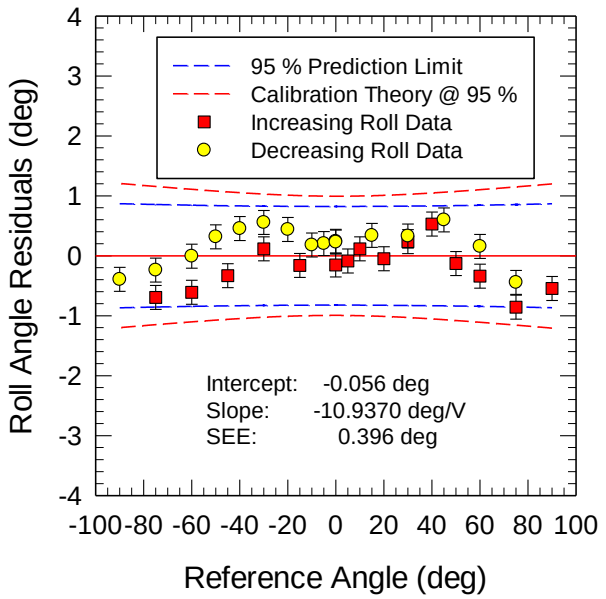
An example plot is shown in Figure 3 for calibration of a commercial vertical gyroscope in roll from Chirozzi and Park (2005). The reference angle was an electronic protractor with a measurement uncertainty of ±0.2° at the 95% confidence limit. The manufacturer rates the gyroscope with an uncertainty of ±1.0°.

From Figure 3a, all data points lie on a straight line. The error bars in such a plot are smaller than the symbols. The residual plot in Figure 3b yields significantly more information about the statistical character of the data. As the plot indicates, the data for increasing angle are

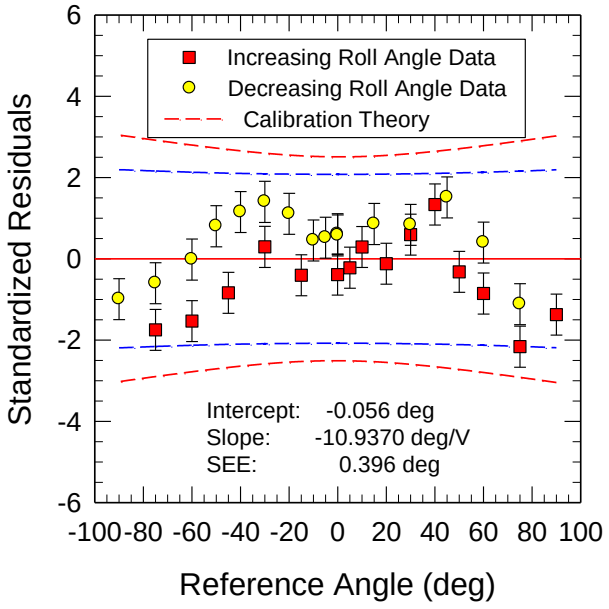
systematically different from the decreasing angle. The plot, thus, indicates a slight hysteresis in the data not evident in the linear plot.



a. Linear Plot



b. Residuals Plot



c. Standardized Residuals

Figure 3: Calibration Data for Vertical Gyroscope in Roll

Also, the error bars are readily apparent in the residuals plot. The error bars in this case are the uncertainty in the reference measurement standard at the 95 % confidence limit (±0.2°).

The dashed lines in the residual plot (Figure 3b) are the prediction limits from the curve fit at the 95 % confidence level. The blue dashed line is the conventional prediction limit from equation (5) while the red dashed line is from calibration theory, equation (8). As the figure shows, the uncertainty from calibration theory is similar to the value claimed by the manufacturer. In subsequent figures, the red dashed lines from calibration theory are indicated as the 95 % prediction limit.

Finally, the calibration data are presented as standardized residuals in Figure 3c. In this case, the *SEE* is computed from the linear regression of equation (6) so that *SEE* has units of degrees.

As the figure indicates, the conventional predication limit at the 95 % confidence level is near $2\times SEE$ while calibration theory yields $3\times SEE$. If *SEE* is applied as the uncertainty estimate, $3\times$ *SEE* is recommended as the uncertainty at the 95 % confidence level. The statistical quantities described in this section may be computed in a standard commercial spreadsheet or statistical program.

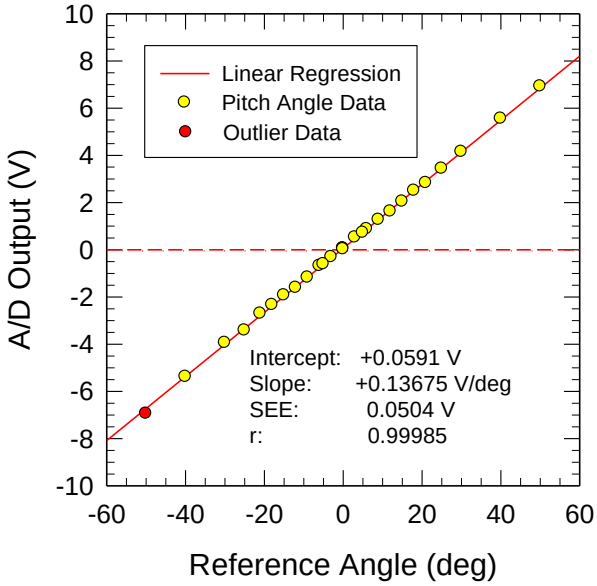
4.4 Hypothesis tests and outliers.

The standardized residual plot is useful in identification of outliers. The coverage factor or threshold for exclusion of an outlier may be determined by either Chauvenet's criterion in Coleman and Steele (1999) or the Student *t*, $t_{a/2,n-2}$ via a hypothesis test.

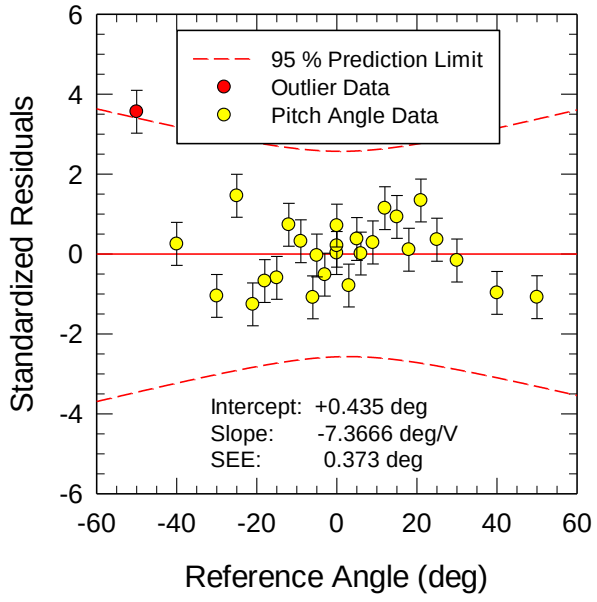
An example is presented in Figure 4 for a vertical gyroscope in pitch from Strano and Park (2005). Again all of the data lie on a straight line, including an outlier, as shown in Figure 4a as the red symbol. In the standardized residual plot, this data point as an outlier is more evident. The statistics for this figure are based on the exclusion of the outlier.

Before the data point was removed, the *SEE* was 0.465°. The standardized residual for this point was 2.76. From Chauvenet's criterion for 27 points, the threshold value was 2.36, and from the Student *t* the value was 2.06. By both criteria, this data point is an outlier. After exclusion of the outlier, the standard error of estimate is 0.373°, and the standardized residual for the outlier becomes 3.56.

If a data point were an outlier, a physical cause should be determined and corrected. In this case, none was evident. However, the angle is –50°, and encountering a pitch angle of this value during a test is highly unlikely. In some cases with a sufficient number of data points, removal of a single point from the data may not significantly change the slope, but its removal will probably affect the uncertainty.



a. Linear plot



b. Standardized residuals

Figure 4: Calibration data for vertical gyroscope in pitch

4.4.1 Hypothesis test for known slope and intercept

For some calibrations, comparison with a known slope and intercept may be useful. Calibrations of some instruments are directly in physical units. In that case, a linear fit over the range of the instrument should produce a slope and intercept of 1 and 0, respectively. From Ross (2004) and Devore (2008) for the null hypothesis, H_0, for intercept accept H_0 if

$$|a-\alpha|\sqrt{n(n-2)s_{xx}/(\sum_{i=1}^{N}x_i^2 SS_R)} \le t_{\alpha/2,n-2} \quad (10)$$

and for the slope

$$|b-\beta|\sqrt{(N-2)s_{xx}/(\sum_{i=1}^{N}x_i^2 SS_R)} \le t_{\alpha/2,n-2} \quad (11)$$

where α and β are the theoretically known intercept and slope, respectively.

4.4.2 Hypothesis test for repeat calibrations

Typically, the same instruments are employed in subsequent tests. Establishment of a calibration history is important. If a calibration is reproducible, the slope and intercept should be the same statistically. Reproducibility may be determined with a hypothesis test. For the slope from Kleinbaum, et al. (2014), the null hypothesis is accepted if

$$(b_1-b_2)/s_{b1-b2} \le t_{\alpha/s,n1+n2-4} \quad (12)$$

where b_1 and b_2 are the slopes of the two calibrations and $s_{b1\text{-}b2}$ is the standard deviation of the slope difference. The variance of the slope difference is

$$s_{b1-b2}^2 = s_p^2[1/s_{xx1}+1/s_{xx2}] \quad (13)$$

where the pooled estimate is

$$s_p^2 = \left[(n_1-2)SEE_1^2+(n_2-2)SEE_2^2\right]/[n_1+n_2-4] \quad (14)$$

The number of degrees of freedom for the Student *t* comparison to equation (11) is [$n_1 + n_2 - 4$].

For the *t*-test of the intercept, the slopes are assumed to be the same. The *t*-test in this case is as follows from Armitage, et al. (2002):

$$d/s_d \le t_{\alpha/2,n1+n2-3} \quad (15)$$

where *d* is the difference in intercepts given by

$$d = \bar{y}_1 - \bar{y}_2 - b(\bar{x}_1-\bar{x}_2) \quad (16)$$

and the pooled slope, *b*, for the two calibrations is

$$b = (b_1 s_{xx1}+b_2 s_{xx2})/(s_{xx1}+s_{xx2}) \quad (17)$$

The variance of *d* is given by

$$s_d^2 = s_c^2[1/n_1+1/n_2+(\bar{x}_1-\bar{x}_2)^2/(s_{xx1}+s_{xx2})] \quad (18)$$

and the residual mean square about the lines is

$$s_c^2 = [s_{yy1}+s_{yy2}-(b_1 s_{xx1}+b_2 s_{xx2})^2/(s_{xx1}+s_{xx2})]/[n_1+n_2-3] \quad (19)$$

The number of degrees of freedom for the Student *t* comparison in this case is [$n_1 + n_2 - 3$].

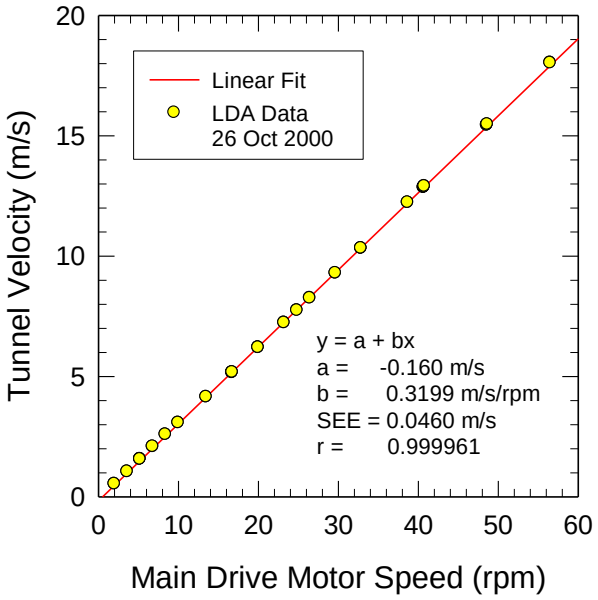
4.5 Non-linear equations

The principles in the previous section may be extended to non-linear functions. Commer-

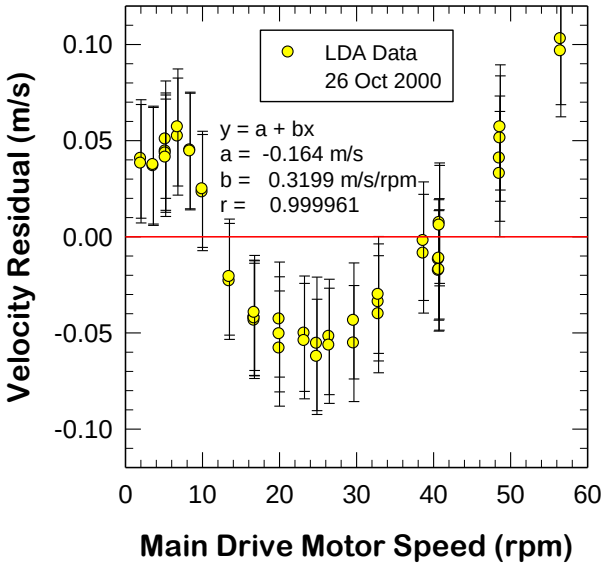
cial computer codes are available for such calculations. Many examples are available in naval hydrodynamics. The prediction limit may be applied as a measure of the uncertainty for these non-linear curve-fits or fairing. An example from Park, et al. (2005) is shown in Figure 5.

The data are for the water tunnel speed in the test section from the pump speed. The test section velocity was measured by laser Doppler velocimetry (LDV). This is also another example of the importance of a residual plot.

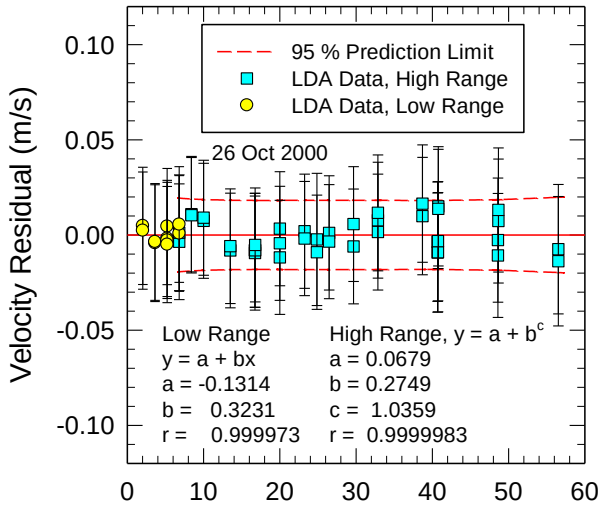
As Figure 5a indicates, all of the data lie on a straight line; however, as the residual plot for the straight line fit shows a systematic deviation of the data from zero in Figure 5b. In the residual plot, the data should be randomly distributed about zero. A plot of this form indicates that a better fit would be other than a linear fit; consequently, a non-linear fit was tried. Additionally, the figure indicates that the speed is represented by two equations. Below 2 m/s, a linear fit with a different slope is adequate.



a. Linear plot



b. Residual plot for linear curve



c. Residual plot for non-linear curve

Figure 5: Calibration data for water tunnel speed in empty test section

Above 2 m/s (6.8 rpm), a power law equation was determined to be a better representation where the power-law is given by

$$y = a + bx^c \tag{20}$$

The constants *a*, *b*, and *c* were computed by a commercial code by the least squares method. For the power-law curve fit, the *SEE* is 0.0088 m/s in comparison to 0.0460 m/s linear fit over the entire velocity range.

5. FORCE CALIBRATION

Force calibrations, including body forces, moments, and propeller thrust and torque, are usually calibrated with masses on a calibration stand. In that case, force is related to mass by the following from ASTM E74-02:

$$F = mg(1 - \rho_a / \rho_w) \tag{21}$$

where m is the mass, g is local acceleration of gravity, ρ_a is air density, and ρ_w is the density of the weight. A calibration stand may include levers for increasing the force, in which case the force multiplier should be included in the above equation and the uncertainty estimates.

The last term of equation (21) is a buoyancy correction. Local gravity can differ from standard gravity, 9.80665 m/s^2, on the order of 0.1 %, and the buoyancy correction is typically 0.017 %. Mass sets commonly applied to force calibrations have a tolerance specification on the order of ±0.01 %, such as an OIML Class M1, NIST Class F, or ASTM Class 6. The detailed characteristics for these weight classes are described in OIML R 111 (2004), NIST (2003), and ASTM E617-97, respectively. Consequently, the correction for local gravity can be 10 times the uncertainty in the reference mass.

During calibration, the force is changed by adding or removing weights. The mass in equation (21) is then given by

$$m = \sum_{i=1}^{n} m_i \tag{22}$$

The weight set is usually calibrated as a set at the same time against the same reference standard. In that case, the uncertainty in the weights is assumed to be perfectly correlated. The standard uncertainty in the total mass is then

$$u_m = \sum_{i=1}^{n} u_i \tag{23}$$

An NMI traceable laboratory report will include the actual measured mass and its expanded uncertainty for each weight. The required expanded uncertainty is

$$U \le \delta m / 3 \tag{24}$$

where $\pm\delta m$ is the maximum permissible error or tolerance per OIML (2004) or ASTM E740-02. This uncertainty estimate is recommended by OIML R 111-1 (2004). However, rather than documentation of the actual calibration of each mass and its uncertainty, δm is applied as the expanded uncertainty as a practical matter or typically ±0.010 % of the total mass.

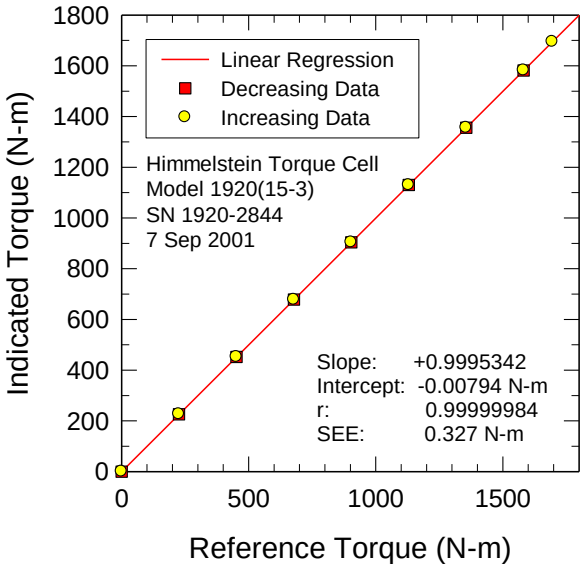
A load cell should be calibrated with random loading for avoidance of hysteresis. Before calibration commences, maximum load should be applied at least twice. During calibration, the number of force applications should total 30 of which 10 forces must be different. Additional details on the calibration of load cells are described in ASTM E74-02.

The results for a calibration of a commercial torque load cell are shown in Figure 6. This torque transducer contains a strain gage as a sensor. The load cell was calibrated with an increasing torque. At the maximum torque, the load was decreased over the same value. Due to the highly linear nature of the device, the symbols

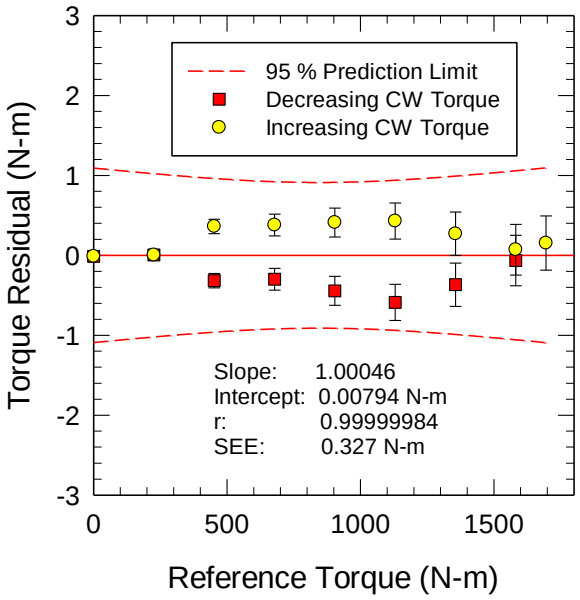
for increasing loads are directly over the decreasing values in the conventional linear plot of Figure 6a. All of the data lie on a straight line.

A residual plot of the same data is shown in Figure 6b. As the figure indicates, hysteresis of the load cell is now evident. The error bars in the figure indicate the uncertainty in the reference standard as provided by the manufacturer. The dashed lines are the uncertainty from calibration theory for the linear regression. The load cell was calibrated to provide a direct calibrated reading; however, a slight correction is required as indicated by the values of slope and intercept in the figure.

With application of the hypothesis test of Section 4.4.1, the *t*-values for a slope of 1 and intercept of 0 are 2.44 and 0.065, respectively. Consequently, the slope is not 1, but the intercept is statistically the same as 0.



a. Linear plot



b. Residual plot

Figure 6: Calibration data for commercial torque cell.

6. UNCERTAINTY IN PULSE COUNT

In naval hydrodynamic applications, rotational rate is a commonly measured parameter. In particular, two applications are shaft rotational rate in propeller performance and towing carriage speed from the rotational rate of a precision metal disk. Rotational rate is measured from a pulse-generating device such as an optical encoder or steel gear with a magnetic pick-up. These devices are inherently digital. Data acquisition cards (DAC) typically include a 16-bit analogue to digital converter, counter ports, and accurate timing. The rotational rate is measured via the equation

$$\omega = n/(pt) \tag{25}$$

where ω is the rotational rate, n the number of pulses, p the number of pulses per revolution, and t the time.

From equation (24), the uncertainty in the rotational rate is

$$u_\omega^2 = u_n^2/(pt)^2 + (n/p)^2 u_t^2/t^4 \tag{26}$$

or the relative uncertainty is

$$(u_\omega/\omega)^2 = (u_n/n)^2 + (u_t/t)^2 \tag{27}$$

The number of pulses per revolution, p, is assumed to be known precisely; therefore, the uncertainty is zero. The AD should have calibration traceability to an NMI with the uncertainty documented by a certificate with the uncertainty in time provided.

During data acquisition, either the time is fixed or the number of digital samples at a specified sample rate. The total time interval is then fixed at T

$$T = n\Delta t = n/f_s \tag{28}$$

where n is the number of samples and f_s the sample frequency.

In this case, the uncertainty in pulse count is assumed to be uniformly distributed over the interval $\pm a$ with a standard deviation of

$$u_n = a/\sqrt{3} \tag{29a}$$

from the JCGM (2008). In this case, a = ½ pulse or

$$u_n = 1/\sqrt{12} = 0.29 \tag{29b}$$

The expanded uncertainty in the number of pulses at the 95 % confidence level is then 0.58 pulses. A minimum pulse count of 1000 is recommended. In that case, the relative uncertainty is 0.058 %.

Equation (29b) can also be applied to the uncertainty in the bit resolution of an AD. However, the uncertainty in calibration of an AD is normally several times the bit uncertainty, particularly for a 16-bit AD.

Rather than fixing the time interval for sampling, the pulse count could be fixed with time starting at the first pulse and ending at the last pulse. The pulse count is then exactly known, and the uncertainty in pulse count can be assumed to be zero. Then, the only contribution to the uncertainty is the uncertainty in time.

In some cases for dynamical processes, rotational rates may be measured with a frequency to voltage converter (FV). Calibration of the FV should be performed by a direct through system calibration of the AD. The rotational rate should be determined by the method described in Section 4, Linear Regression Analysis, with changes in rotational rate like the shaft speed of the propeller in a propeller performance test. That is, a function generator is not recommended as the method for calibration of the FV. A through system calibration will include the performance of the pulse generation device. An FV may drift; consequently, repeat calibrations are recommended for a better estimate of the uncertainty.

7. DIRECT DIGITAL CALIBRATION

A DAC usually has a counter port so that frequency can be measured directly, and an FV converter is not needed. The card can be calibrated with a wave function generator for the input with the frequency measured by an NMI traceable frequency counter. The uncertainty may then be estimated with linear regression methods previously described in Section 4.2. In that case, the slope and intercept should be, respectively, 1 and 0, and the result check by the hypothesis tests in Section 4.3.

A second method is a direct difference between the reference measurements and the values from the DAC. In this case, the average difference should be zero (0). The *t*-test for zero difference is given from Devore (2008) by

$$\left|\bar{d}\right|\sqrt{n}/s \le t_{\alpha/2,n-1} \tag{30}$$

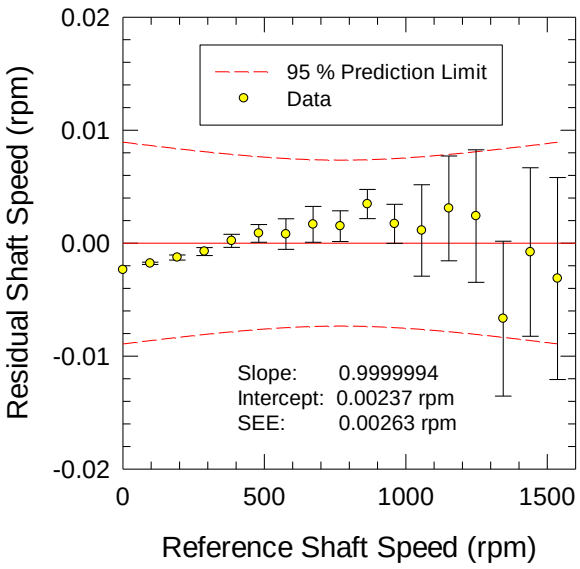
where *d* is the difference between the reference frequency and the measured frequency from the DAC, n the number of measurements and *s* the standard deviation of the average difference. The expanded uncertainty is then the prediction limit for an average from Devore (2008)

$$U = t_{\alpha/2,n-1} s\sqrt{1+1/n} \tag{31}$$

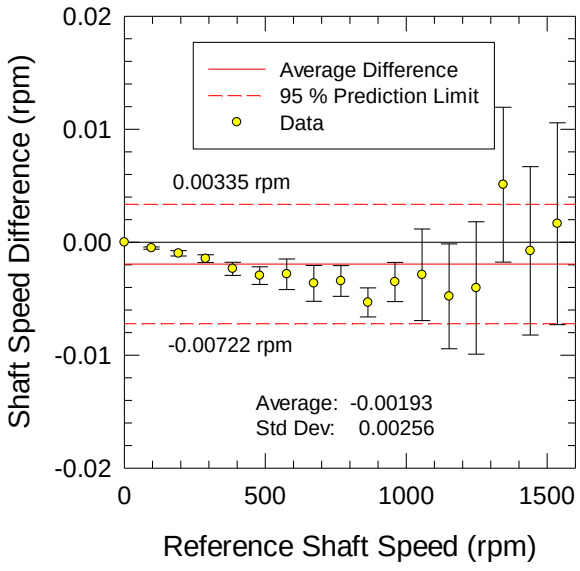
An example direct digital calibration is presented in Figure 7 for shaft speed from an optical encoder in a surface ship model test from Reynolds, et al. (2013). In this case, the input is simulated with a square-wave generator. The input values are corrected with p = 200 pulses per revolution for the optical encoder. The NMI traceable frequency measurement has a rated uncertainty of 10^{-6} × reading. The following table summarizes the results by two methods. The uncertainty in the table for the linear aggression analysis is the maximum combined expanded uncertainty and not the uncertainty in the intercept, *a*.

Test	Value	T value	Result	U95
a	0	-1.94	Pass	0.013
b	1	+0.42	Pass	
Diff.	0	-3.11	Fail	0.0052

Table 1: Hypothesis test results for calibration of surface-ship model propeller shaft speed



a. Linear Regression



b. Difference

Figure 7: Calibration data for surface-ship model propeller shaft speed.

In this example, the uncertainty from calibration theory should be applied since both the slope and intercept pass the hypothesis test and its value is larger. At the maximum shaft speed, the shaft speed is 1535.999 ±0.013 rpm. If the

hypothesis tests are not passed by either method, then engineering judgement should be applied on the selection of the uncertainty when the reading is not corrected. Otherwise, the slope, intercept, and uncertainty estimate should be applied from calibration theory.

8. SUMMARY

This procedure describes a statistical process for the establishment of the uncertainty in calibration data. Typically, a calibration is necessary in the conversion of digital Volts from an AD to physical units. For most analogue instrumentation, the calibration curve is very linear for the calibration. Usually, the uncertainty in the reference standard is small in comparison to the uncertainty indicated by the scatter in the data. The following summarizes the recommended procedures:

- Perform a calibration over the range in approximately ten (10) equal increments in a through system calibration with the same hardware and software for the test.
- Verify traceability to an NMI of any reference standard by its calibration certificate. The calibration certificate should include a statement of its uncertainty.
- Compute the mean, standard deviation, and number of samples for each data point, where the number of samples is at least 100 to 1000.
- Verify that the standard deviation, which is a measure of instrument noise, is reasonable and compute the uncertainty by the Type A method per JCGM (2008).
- Document the sample rate, cut-off frequency, number of samples for each data point, and information on the calibration reference standard. See ITTC Procedure 7.6-01-01 for additional documentation requirements.
- Perform a linear regression analysis of the calibration data for the determination of the slope, intercept, *SEE*, and correlation coefficient.
- Plot the data as a standardized residual plot, review for randomness of the data, identify any outliers, determine the cause, and remove the outliers if such removal is appropriate. If the data trend is systematic, consider a non-linear curve fit to the data.
- Repeat the regression analysis if outliers are removed and do not repeat the outlier removal process.
- Compute the uncertainty from calibration theory.
- Compare the results with any previous calibrations.
- Enter the slope and intercept in the software for the test.
- Check the calibration results by application of a through system calibration check at three points: the high, mid-range, and low values.

This revision includes the following changes from the previous version and some minor editorial changes:

- Correction of a sign error in equation (19)
- Clarification of the relation between the uncertainty in a weight and its tolerance in Section 5, Force Calibration
- Addition of Section 7, Direct Digital
- Updates to references: change GUM from ISO to JCGM and list newer editions of Armitage, et al. (2012) and Kleinbaum, et al.(2014)
- Corrections in nomenclature for 95 % prediction limit

9. REFERENCES

Armitage, P., Berry, G., and Matthews, J. N. S., 2002, Statistical Methods in Medical Research, Fourth Edition, Blackwell Scientific Publications, Oxford, pp. 322-335.

ASTM E74-02, 2002, "Standard Practice of Force-Measuring Instruments for Verifying the Force Indication of Testing Machines," American Society for Testing and Materials, West Conshohocken, Pennsylvania, USA.

ASTM E617-97, 1997, "Standard Specification for Laboratory Weights and Precision Mass Standards," American Society for Testing and Materials, West Conshohocken, Pennsylvania, USA.

Bendat, J. S., and Piersol, A. G., 2000, Random Data: Analysis and Measurement Procedures, Third Edition, John Wiley and Sons, Inc., New York, New York, USA.

Carroll, R. J., Spiegelman, C. H., and Sacks, J., 1988, "A Quick and Easy Multiple-Use Calibration-Curve Procedure," Technometrics, Vol. 30, No. 2, pp. 137 - 141.

Chirozzi, B. D., and Park, J. T., 2005, Unpublished Data, Naval Surface Warfare Center Carderock Division, West Bethesda, Maryland, USA.

Coleman, H. W., and Steele, W. G., 1999, Experimentation and Uncertainty Analysis for Engineers, Second Edition, John Wiley and Sons, Inc., New York, New York, USA.

Devore, J. L., 2008, Probability and Statistics for Engineering and the Sciences, Seventh Edition, Thomson Brooks/Cole, Belmont, California.

ITTC, 1999, "Control of Inspection, Measuring and Test Equipment", ITTC Procedure 7.6-01-01, Revision 00.

Kleinbaum, David G., Kupper, Lawrence L., Nizam, Azhar, and Rosenberg, Eli S., 2014, Applied Regression Analysis and Other Multivariable Methods, Fifth Edition, Cengage Learning, Boston, pp. 262-268.

JCGM, 2008, "Evaluation of measurement data – Guide to the expression of uncertainty in measurement," JCGM 100:2008 GUM 1995 with minor corrections, Joint Committee for Guides in Metrology, Bureau International des Poids Mesures (BIPM), Sèvres, France.

NIST Handbook 105-8, 2003, "Specifications and Tolerances for Reference Standards and Field Standards and Measures," National Institute of Standards and Technology, Gaithersburg, Maryland, USA.

OIML R 111-1, 2004, "Weights of Classes E_1, E_2, F_1, F_2, M_1, M_{1-2}, M_2, M_{2-3}, and M_3, Part 1: Metrological and technical requirements," Organisation Internationale de Métrologie Légale, Paris, France.

Otnes, R. K., and Enochson, L., 1972, Digital Time Series Analysis, John Wiley and Sons, Inc., New York, New York, USA.

Park, J. T., Cutbirth, J. M., and Brewer, W. H., 2005, "Experimental Methods for Hydrodynamic Characterization of a Very Large Water Tunnel, Journal of Fluids Engineering, Vol. 127, No. 6, pp. 1210-1214.

Reynolds, H. W., Krishen, C., and Park, J. T., 2013, Unpublished Data, Naval Surface Warfare Center Carderock Division, West Bethesda, Maryland USA.

Ross, S. M., 2004, Introduction to Probability and Statistics for Engineers and Scientists,

Third Edition, Elsevier Academic Press, Amsterdam

Scheffe, H., 1973, "A Statistical Theory of Calibration," The Annals of Statistics, Vol. 1, No. 1, pp. 1 - 37.

Strano, P. M., and Park, J. T., 2005, Unpublished Data, Naval Surface Warfare Center Carderock Division, West Bethesda, Maryland USA.