GUIDING YOUR ORGANIZATION THROUGH LABORATORY ACCREDITATION A Technical Manager's Experience

Gregory Cenker Fluke Electronics 9045 Balboa Avenue, San Diego, Ca

ABSTRACT

Over the years, the Metrology field has been in a constant change of evolving technologies, practices, and methodologies. With the changing times, the event of Laboratory Accreditation and new ISO requirements, more, now than ever, the integration of all metrology practices must be unified. This paper will address the importance, significance, and financial considerations of integrating uncertainty analysis, process metrology and laboratory accreditation into a useable and efficient medium. By embedding technical expertise within the confines of the "digital domain", consistency, repeatable results, and a very highly integrated level of traceability are not only desirable but also a very tangible objective.

IDENTIFY BUSINESS SCOPE OF ACCREDITATION

Before beginning work on the accreditation process, it is essential to outline and target your objectives. It is always best to stand back, take a good look at the laboratory and determine the best capabilities with minimal complexities. The author took the following approach:

- 1. SPC¹ charting of standards available. Information taken from historical records and/or established SPC where none existed.
- 2. Uncertainty analysis and documentation. From previous information, a complex and "linked" uncertainty analysis was constructed.
- 3. System documentation and verification. Validated systems reporting correct uncertainties.
- 4. Identify how the organization actually worked and wrote the policies and procedures to conform to processes already established.

The first item is ultimately the most critical. Information taken from historical records was collected and placed into a software package adequately capable of producing the desired results with minimal

¹ SPC is simply an abbreviation for Statistical Process Control

"round-off" error. The author uses both Statistica[®] and Excel[®] for performing data analysis. If the latter is the software of choice, a common mistake the author has encountered, are analyses being conducted using full resolution to process statistical information. This may lead to inaccurate information due to "round-off" errors. For example, if a 10 volt reference is being analyzed, and the system reports 10 significant digits, then for a system value of 10.000 003 51V, the more correct way to record and process the information would to reduce the value to a deviation from nominal. In essence, record the information as units 3.51μ V or 0.351ppm. This will ensure accurate results when compiling statistical uncertainties.

EXPRESSION OF UNCERTAINTIES

The actual error of a reported value is usually unknowable. Limits to this error can usually be inferred -- with some risk of being incorrect -- from the precision of the measurement process that yielded the reported value, and from reasonable limits to the possible bias of the measurement process. The concept of measurement uncertainty should be used instead of error, because of the difficulty in defining "error". Uncertainty and error are not synonyms. While error is an unknown physical quantity, uncertainty expresses the metrologist's "best guess" of the interval, which contains the true value. The Heisenberg Uncertainty Principle sums it up best:

"The simultaneous measurement of two conjugate variables (such as the momentum and position or the energy and time for a moving particle) entails a limitation on the precision (standard deviation) of each measurement. Namely: the more precise the measurement of position, the more imprecise the measurement of momentum, and vice versa. In the extreme case, absolute precision of one variable would entail absolute imprecision regarding the other."

The uncertainty statement is intended to provide the customer with information, which will allow them to compute the risk of using the measurement results. The ISO Guide to the Expression of Uncertainty in Measurement redefines uncertainty as the equivalent of a standard deviation and recommends that uncertainties be placed in two categories; these categories are based on the method with which they are evaluated.

"A" those based on statistical observations

"B" those evaluated by other means such as manufacturer specs, etc.

The confidence level is the probability associated with a confidence interval. Typical confidence levels are:

1 Sigma – 68.27% 2 Sigma – 95.45% 3 Sigma – 99.73%

Assuming a 95% confidence interval, evidence leads to believing, if the same measurement were performed we would find the true value to be in the interval 95% of the time.

Of course, the stability, predictability and calibrator uncertainty must be taken into account along with the standard deviations. The author regularly hears comments referring to a magical "4:1"

 TAR^2 . It is the authors' opinion that it's more appropriate to account for complete measurement uncertainty by analysis. After all, how does one know what an acceptable risk may be if a complete analysis has not yet been performed?

Once the historical data is analyzed, process improvement may be employed to validate adequacy of test results. In essence, if making more measurements on the same parameter decreases the uncertainty, what would be sacrificed in the order of system time (or metrologist time)? Would the gain in reduced uncertainty be worth the extra measurements?

Let's look at a simple example: if you wanted to make measurements, which were twice as good as the standard deviation of the data collected, then the following equation may be employed in order to refine the measurement process. Utilizing a coverage factor of 2 (95.45%), a desired UUT standard deviation of 6μ V, and a measured standard deviation of 12μ V, this method would be as follows:

k = 2 (95.45% confidence interval) $s = 12\mu V$ (standard deviation) $d = 6\mu V$ (desired deviation)

Minimal sample size =
$$\left(\frac{k \times s}{d}\right)^2 = 16$$

Notice the significant increase in minimal sample size required as the desired confidence level increases and the multiplier of the original standard deviation increases.

Measured Standard Deviation	Desired Deviation	Multiplier of standard deviation	$Minimal \\ Sample Size \\ k = 1$	95% k = 1.96	95.45% k = 2	99% k = 2.576	99.73% k = 3		
12	12	1	1	4	4	7	9		
12	6	2	4	16	16	27	36		
12	4	3	9	35	36	60	81		
12	3	4	16	62	64	107	144		
12	2	6	36	139	144	239	324		
12	1	12	144	554	576	956	1296		
12	0.5	24	576	2213	2304	3823	5184		
	Table 1								

Once adequate process controls are established, the data has been reduced to a useable format, then the next step involves creating a "linked" uncertainty analysis. The following example is taken from actual collected data on a zener reference standard that has been imported into Excel[®].

² TAR is an abbreviation for Test Accuracy Ratio

CHANNE	L 21 RESUL	9.999 945 993	-5.40 ppm				
DATE	DIFF. FROM HAV ³	FORWARD/ REVERSAL DIFF.	MODULE VALUE	STD DEV.	Dev From Nom.	Sample size	45
						Mean	-5.373 ppm
22-Sep-01	-36.98 uV	0.08 uV	9.999946447	0.19 uV	-5.36 ppm	Standard Deviation	0.034 ppm
23-Sep-01	-37.39 uV	0.06 uV	9.999946037	0.17 uV	-5.4 ppm	Standard Error	0.032 ppm
24-Sep-01	-37.21 uV	0.08 uV	9.999946217	0.23 uV	-5.38 ppm	Noise	0.021 ppm
25-Sep-01	-37.51 uV	0.08 uV	9.999945917	0.25 uV	-5.41 ppm	Thermal EMF	0.012 ppp
29-Sep-01	-36.94 uV	0.30 uV	9.999946487	0.25 uV	-5.35 ppm	-0.29 ppm/yr. estima	ted drift
30-Sep-01	-36.68 uV	0.28 uV	9.999946747	0.24 uV	-5.33 ppm	Uncertainty	0.082 ppm
1-Oct-01	-36.87 uV	0.27 uV	9.999946557	0.31 uV	-5.34 ppm	Confidence Level 95.45 %	2.058

Table 2

The value indicated in the upper-right corner, 9.999 945 993, is representative of the value based on "today's" date and is fully dynamic. Sample size, mean, standard error and estimated drift are all dynamic values as well. As new data is collected, the sheet will automatically update to the latest true values. The value in the lower-right corner, 2.058, is the coverage factor for 44 degrees of freedom at the 95.45% confidence level. Again, this is a dynamic value, which will automatically update as more data is collected. The function built into Excel[®], TINV(), calculates the coverage factor based on available degrees of freedom.

The next step involves combining the in-house SPC uncertainty along with the data reported from either a National Laboratory or an intrinsic measurement. The data presented below is linked to the previous example and by definition is dynamic.

Type A Uncertainty Description	Uncert.	Sens.	Distrib.	DOF	Divisor	STD Unc.	Variance
Measurement Noise	0.021 ppm	1	68.27	44	1.01	0.021 ppm	0.0004 ppm
Error of Fitted Data	0.032 ppm)	1	68.27	44	1.01	0.032 ppm	0.001 ppm
Thermal EMF	0.012 ppm	1	68.27	44	1.01	0.012 ppm	0.0001 ppm
					(Student's t)		
Type B Uncertainty Description	Uncert.	Sens.	Distrib.	DOF	Divisor	STD Unc.	Variance
JJ Uncertainty (National Lab)	0.100 ppm	1	2u	10	2	0.050 ppm	0.0025 ppm
Digital Resolution	0.005 ppm	1	Rect.	100	1.732	0.003 ppm	9.0 E-6 ppm
Combined Results	Туре А	Туре В	Type AB				
Uncert.	0.040 ppm	0.100 ppm	0.108 ppm	-			
STD Unc.	0.040 ppm	0.050 ppm	0.064 ppm	⊂ ́	This value is t	he "raw" res	ult (0.064)
Variance	0.001 ppm	0.003 ppm	0.004 ppm				
Effective Degrees of Freedom:	100 (Veff)						
95.45% Coverage Factor	>						
Expanded Uncertainty:	(0.129 ppm))					

Transfer Standard Uncertainty Analysis

Table 3

The cell containing "100 (Veff)" has a custom function written into Excel[®] for calculating the Welch-Satterthwaite equation. Once "Veff" is calculated, the TINV() function is again used to

³ The "DIFF from HAV" is the difference from the individual cell as compared to the true average of all 10 cells utilized for comparison.

derive a coverage factor based on desired confidence level and effective degrees of freedom. The resultant value is multiplied to arrive at the expanded uncertainty.

Once we have an idea how well the transfer standard is performing in comparison to an intrinsic value, the next step will involve transferring the new uncertainty value to a working lab standard. Notice the uncertainty of the import standard; it was carried forward from the previous results as was the degrees of freedom and the divisor (coverage factor 2.03). We now have 212 effective degrees of freedom, a coverage factor of 2.01, and an expanded uncertainty of 0.191ppm.

Type A Uncertainty Description	Uncert.	Sens.	Distrib.	DOF	Divisor	STD Unc.	Variance
Measurement Noise	0.026 ppm	1	68.27	45	1.01	0.026 ppm	0.0007 ppm
Error of Fitted Data	0.043 ppm	1	68.27	45	1.01	0.042 ppm	0.0018 ppm
Thermal EMF	0.051 ppm	1	68.27	45	1.01	0.050 ppm	0.0025 ppm
					(Student's t)		
Type B Uncertainty Description	Uncert.	Sens.	Distrib.	DOF	Divisor	STD Unc.	Variance
Import Standard Uncertainty	0.129 ppm	1	2u	(100)	2.03	0.064 ppm	0.0041 ppm
Digital Resolution	0.005 ppm	1	Rect.	100	1.732	0.003 ppm	9.0 E-6 ppm
Combined Results	Туре А	Туре В	Type AB				
Uncert.	0.071 ppm	0.130 ppm	0.148 ppm	1			
STD Unc.	0.070 ppm	0.064 ppm	0.095 ppm	1			
Variance	0.005 ppm	0.004 ppm	0.009 ppm	1			
Effective Degrees of Freedom:	212 (Veff)						
95.45% Coverage Factor							
Expanded Uncertainty:	0.19 ppm)					

Fluke 7000 Automated DC Reference Uncertaint	v Analysis
Trake 7000 Automated Do Reference Officertaint	y Analysis

Table 4

The next step involves transferring the "in-house" reference Zener to a suitable precision voltage divider. The Zener reference uncertainty is linked to the divider worksheet with the appropriate components accounted and combined.

Ratio metric DC System Voltage Range	10V DC Ref. (Ub)	Self- Cal. (Ub)	Power Coeff. (Ub)	Null Det. (Ub)	Ub	System Total
100mV	0.19	0.50	0.00	0.20	0.33	0.65 ppm
1V	(0.19 \	0.20	0.00	0.06	0.15	0.31 ppm
10V	0.19	0.20	0.00	0.02	0.15	0.30 ppm
100V	0.19	0.20	0.05	0.02	0.15	0.30 ppm
1KV	0.19	0.50	0.30	0.02	0.30	0.61 ppm

Table	5
-------	---

Once the reference divider uncertainty is known, the next step involves tracking, analyzing, and predicting the behavior of a process standard. A process standard, in this case, is a Fluke/Wavetek 4808 Multifunction Calibrator. These standards are never adjusted; their prospective outputs are measured, usually at a 90-day interval, tracked, and a "Type-A" derived from the historical data. A predicted value is derived as well. This is accomplished by performing a time-series analysis on the historic data.

A custom function implemented into Excel[®] is a bimodal interpolation functionⁱ. The big picture for the "Interpolate" function is that two parabolas are created and averaged (weighted by the inverse distance to the interval end points) to produce the interpolated curve, a third degree polynomial. The result is an interpolated curve with a continuous first derivative. The author has not seen this method discussed or used anywhere else, yet it produces reasonable interpolants comparable to the use of natural cubic splines or clamped splines but without many of the drawbacks of those particular methods. The drawbacks namely are wild aberrations and the inability to draw a straight line.

The second half of the theory is the method used to develop the parabolas of Interpolate. The crucial point here is that the parabolas are produced without calculating the coefficients of the second-degree equations that make up the parabolas, but by a method that has far fewer computations, and never adds or subtracts values of vastly different magnitudes (X's are always subtracted from X's, and Y's are always subtracted from Y's). In addition to being faster, this method avoids the artifact of inexact computer computations, namely that of calculating and using the coefficients produced errors when the X and Y values were orders of magnitude different, or given points were very close together. If computers were of infinite precision, there would be no difference in calculating the interpolant using coefficients and using the method used in Interpolate, but as it is, the method used in Interpolate calculates the parabolas more accurately.

Armed with custom functionality, reverse interpolation is utilized to derive a "Type-A" uncertainty based on the number of days between calibrations. With the number of days between calibrations known, the function calculates two standard deviations along with the standard error of estimate and performs a root sum squared answer. This value is now the "Type-A" uncertainty for the designated calibration interval.

As previously mentioned, a predicted value is easily obtainable by performing Time-Series Analysis. One point worth mentioning is a common misconception when performing this operation at widely varying calibration intervals. In order for Time-Series Analysis to be as accurate as possible, the data being analyzed should be equidistant. Diligent as the metrologist may be, it is highly unlikely the standard will be calibrated precisely at the designated interval. If the calibration intervals are highly variable, the "Predict" function is used to create equidistant calibration interval points, otherwise the original data is utilized "as-is" and linear regression is applied. The "Predict" function is a combination of "Interpolate", creates a "Moving Average", and performs "Linear Regression", simultaneously.

When necessary to invoke "Predict", the function looks at the very first historic calibration date, the last calibration date, and determines the total number of elapsed days. Once this information is known, the functions looks at the present calibration date, compares it to the previous calibration date and determines the interval (in days). It then calculates the total elapsed time by the present calibration interval and creates a Moving Average Time-Series Analysis. Result? A 185-day moving average is created and a value predicted for the present calibration date. The function also has the option of allowing for any specific moving average, i.e. 90-day, if so desired.

Referring to the table on the following page, the "Predict" function estimated the deviation from nominal for the designated calibration interval. Although by graphing the data it doesn't appear linear, the difference between the true measured value (99) and the predicted value (99.56) was

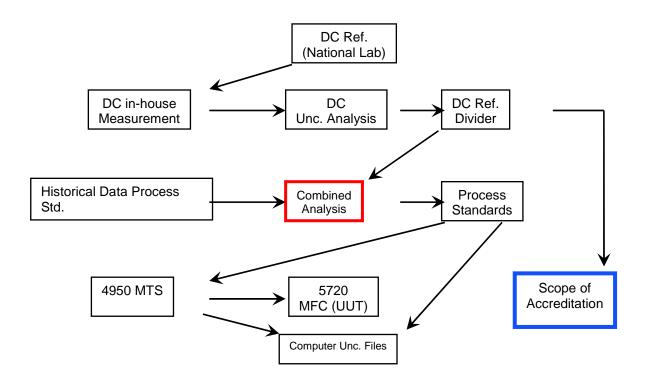
+0.56 ppm. This equates to $\pm 11.7\%$ of its "Type-A" component. By comparison, if one were to directly use the "Forecast" function in Excel[®] on the exact same data set, 12/18/96 through 2/23/01, and predicting a value for 8/27/01, the "Forecast" function result would indicate 93.82 ppm. A difference of +5.18 ppm is observed between actual and predicted when using the later method, which equates to $\pm 108\%$ of the calculated 4.8 ppm "Type-A" component.

Cal Date	Deviation	Calculate "Type-A" uncertainty from historic data
12/18/96	60	185 =(August 27 th , 2001 – February 23 rd , 2001)
5/21/97	46	Days between Calibrations
3/13/98	65	Reverse interpolate 185 days from 2/23/01 to 8/23/00
10/30/98	88	"Type-A" = 4.8 ppm
6/10/99	75	
11/05/99	80.5	Create a moving average using "Predict"
2/11/00	100.7	This function creates the moving average and predicts the value for 8/27/01
5/12/00	78.2	
9/22/00	85.0	
2/23/01	77.0	\leftarrow Historic data taken only to this point for Time-Series prediction.
8/27/01	99.0	← Estimated value for this date calculates to 99.56 ppm

Table 6

This methodology is deployed at every measured parameter for every usable function for both reference standards and process standards. The functions utilized are dynamic and crossed linked. As more history is recorded on both the reference standards and the process standards, the "Type-A" uncertainties, predicted values, and expanded uncertainties will "grow" or "shrink" accordingly in real time.

Although only a few steps have been outlined here as example, as the traceability chain grows larger, all data sheets for all systems utilized at all measured parameters are dynamic by design. Since all this work has been performed at all relevant parameters, the Scope of Accreditation is linked to the standard providing the best measurement uncertainty in the measurement process as well. Once all the information is properly linked, it leaves very little doubt as to the validity of the claimed uncertainties and Scope of Accreditation.



APPLICATION SUBMISSION

With the uncertainty analysis complete, a quality manual based on ISO 17025 needed to be constructed. The author took what he believed to be a different approach than normal when writing a quality document of this magnitude. Instead of publishing a quality manual and expecting the departments involved changing the way they carry about daily business, the employees were interviewed in order to establish the way they were already performing their tasks. The new quality manual was constructed with two primary goals in focus:

- Remain within the constraints of ISO 17025
- Write correct verbiage to cause minimal change in the way most employees already performed daily tasks while remaining in compliance with the new standard. In essence, having them conform to the new ISO standard without realizing there has been a major change.

Having said this, a quality system conforming to the new standard was written, reviewed, and put into place minimizing any confusion about the way normal business was conducted. Along with the quality manual, the following documents were submitted as well.

- List of Primary Standards used in the laboratory
- Laboratory description, staff, and operations
- Scope of Accreditation
- Sample of Calibration Certificates and Reports
- Sample of a calibration procedure

ON-SITE INTERACTION

The on-site interaction between our assessors and the relevant technicians consisted mainly of the following criteria:

- Technical Competence
- Current Training Records
- Review of Automated Test Procedures
- Review of calibration results
- Proof of software validation
- Review of incoming / outgoing equipment handling and storage

Although the majority of the systems used at our San Diego Service Facility are fully automated, the technicians are still responsible for proper test configurations, program selection, and to some extent, a small amount of data entry. An essential point worth mentioning is repeatability and consistency regardless of the operator. This accomplishment is provided by the logon ID. Once a calibration is complete, the system will automatically insert the users logon and assign a unique calibration certificate number based on UUT serial number, the present date, and the computer time. Training records were meticulously scrutinized in order to verify technician qualifications against work being performed. A history folder, chosen by the assessor at random, was recruited and checked for essential elements required of ISO 17025. The historical file contains a copy of the

original work performed, along with the responsible technician. This information was cross-verified within the technicians training matrix thus ensuring proper training and qualification requirements.

Arguably, one of the most valuable tools in the laboratory are spreadsheets constructed where complex calculations are required on a regular basis. We have several spreadsheets included as an integral part of our accreditation. They are used both for calculating measurement results and for estimating uncertainties. However, before using any type of spreadsheet, the author has found it useful to validate the results (even when internal functionality is utilized) for optimum quality assurance.

Example: If one were to use extensive statistics when creating an uncertainty budget, a simple test for function accuracy is to perform a calculation using a built-in function, perform it's inverse and calculate the reversal error.

Normal Distribution functions for $K = 1.96$ (about 2-Sigma)								
Original Excel Function	Result	Reverse function	Result	Error				
=NORMSINV(0.95) Created Custom Functions	= 1.64485300	= NORMSDIST(1.64485300)	= 0.949999951	= 0.052 ppm				
=NORM_INV(0.95)	= 1.64485363	= Cnormal(1.64485363)	= 0.950000000	= 0.000 ppm				
Normal Distribution functions for $K = 4.5$ (about 6-Sigma)								
Original Excel Function	Result	Reverse function	Result	Error				
=NORMSINV(0.9999966) Created Custom Functions	= 4.50760126	= NORMSDIST(4.50760126)	= 0.999996719	= 0.118 ppm				
=NORM_INV(0.9999966)	= 4.49985447	= Cnormal(4.49985447)	= 0.999996600	= 0.000 ppm				

Obviously, in the "grand scheme" of things, a difference of 0.052 or 0.118 ppm is not necessarily front-page news. The author only wishes to make the reader aware of the value of software validation and testing, especially where uncertainty budgets coupled with extensive statistics are concerned.

NIST has wonderful data sets along with calculated results. The web address is <u>http://www.nist.gov/itl/div898/strd/index</u>. There the reader will find exceptional information available for validating calculation accuracy of "off-the-shelf" software packages.

Below is a brief excerpt taken directly from the web site on how the reference values were calculated...

"We provide datasets with certified values for key statistics to assess the accuracy of ANOVA calculations in statistical software. For all datasets, multiple precision calculations (accurate to 500 digits) were made using the preprocessor and FORTRAN subroutine package of Bailey (1995, available from <u>http://www.netlib.org</u>). Data were read in exactly as multiple precision numbers and all calculations were made with this very high precision. The results were output in multiple precisions, and only then rounded to fifteen significant digits. These multiple precision results are an idealization. They represent what would be achieved if calculations were made without roundoff or other errors. Any typical numerical algorithm (i.e., not implemented in multiple precision) will

introduce computational inaccuracies, and will produce results, which differ slightly from these certified values."

AUDIT RESULTS

- Recommended Accreditation
- Three findings and many outstanding comments concerning the extremely high level of detail implemented ensuring the highest level of quality control and consistency

Initial Findings:

- Develop ACV proficiency test
- Develop better methodology for monitoring laboratory environmental issue
- Acquire additional data for phase parameter

CLOSING COMMENTS

- Accreditation is becoming more required by our customer base, it can be expensive from both the engineering standpoint and associated assessment costs. By performing all the required work, capabilities and quality assurance will all benefit significantly.
- Will be painful to construct a measurement quality assurance system of this magnitude however, it only need be done once. Maintenance is relatively easy to sustain once in place and established.
- Will build character and an added sense of accomplishment.

WEB SITES OF INTEREST

http://www.fluke.com http://www.a2la.org http://www.european-accreditation.org/documents.html

http://www.itl.nist.gov/div898/handbook/index.htm NIST - SEMATECH Engineering Statistics Internet Handbook

http://www.nist.gov/itl/div898/strd/index. Statistical Reference Data Sets

http://www.netrax.net/~jdavita/XlXtrFun/XlXtrFun.htm Scott Allen Rauch Excel function homepage

http://www.aip.org/history/heisenberg/ American Institute of Physics

REFERENCES

ⁱ Very special thanks to Scott Allen Rauch for his insight, support, and expertise concerning bimodal equations.

Data, Statistics, and Decision Models with Excel – Donald L. Harnett and James F. Horrell ISBN: 0-471-13398-1

Excel for Scientists and Engineers, Second Edition – William J. Orvis ISBN: 0-7821-1761-9

Microsoft® Excel 2000 Power Programming with VBA – John Walkenbach ISBN: 0-7654-3263-4

http://lib.stat.cmu.edu/apstat/

StatLib---Applied Statistics algorithms. FORTRAN programs the author found useful in this research and laboratory development.

Uncertainty Analysis for Laboratory Accreditation - Peter B. Crisp, Fluke Electronics

SOFTWARE REFERENCES AND ASSISTANCE:

- MathCad 2000 Professional® <u>http://www.mathsoft.com/</u>
- StatSoft, Inc. (2000). STATISTICA for Windows [Computer program manual]. Tulsa, OK: StatSoft, Inc., 2300 East 14th Street, Tulsa, OK 74104, phone: (918) 749-1119, fax: (918) 749-2217, email: <u>info@statsoft.com http://www.statsoft.com</u>
- TableCurve 2D 5.0 <u>http://www.jandel.com/</u>