Demonstrating Competency and Equivalency of Two Commercial SPRT Calibration Facilities

T. J. Wiandt[1,2]

[1] Fluke Corporation, Hart Scientific Division, American Fork, Utah United States.

[2] E-mail: tom.wiandt@hartscientific.com

Preprint: Tempmeko 2007 (Presented 24 May 2007)

ABSTRACT

The Hart Scientific Division of Fluke Corporation operates two accredited SPRT calibration facilities that provide SPRT calibrations at very low levels of uncertainty; one in the US and one in the UK. Competency and equivalency must be demonstrated for both facilities. However, because of the low levels of uncertainty involved, the required experiments are expensive and very challenging. In the US a proficiency test is available through NVLAP based on the long standing NIST measurement assurance program (MAP) to accomplish this purpose. Although needed, a PT of this level is not available elsewhere in the world.

Consequently, an alternative approach is required. This paper will describe the approach taken in an effort to show both competency and equivalency of these two facilities and a logical link to the US NVLAP proficiency test conducted at the US facility. Additionally, the description of the tests and establishment of performance criteria will disclose the seriousness and rigor to which this activity was held. Finally, the data will demonstrate that not only are such tests possible, but the degree of equivalence attained can be very high.

KEY WORDS: accreditation; calibration; competency; equivalency; fixed point cell; interlaboratory comparison; NVLAP; SPRT; UKAS.

1. INTRODUCTION

To achieve and maintain accreditation, the requirement exists for laboratories to successfully pass one or more proficiency tests (PT) intended to demonstrate competency. Generally, the PT consists of the laboratory under test performing a calibration on one or more measurement standards for which the characteristic(s) are known. The results are evaluated and pass/fail criteria are established based on the normalized error, E_n. Because of the expense, difficulty, and expertise required, PTs for SPRT calibrations at the lowest levels of uncertainty are not generally available.

The Hart Scientific division of the Fluke Corporation operates two accredited SPRT calibration facilities, one at the Hart Scientific factory in Utah, USA, and the other at a service facility in Norwich, UK. The US facility is accredited through NVLAP and the UK facility is accredited through UKAS. Both provide SPRT calibrations using similar equipment and procedures, and at similar levels of uncertainty. These uncertainties are among the lowest commercially available. Consequently, the PT requirements are extremely demanding. In the US, a suitable PT is available through NVLAP based on the long standing NIST SPRT Measurement Assurance Program (MAP). Thus, a PT was conducted in the US facility. No such PT is offered in the UK or among Euromet. Consequently, it was understood that the goals of the PT had to be met using a different approach.

To further complicate matters, a requirement exists for the sealed fixed point cells used in the SPRT calibration process to be verified periodically. Due to the delicate nature of these cells and the difficulty in transport, a traditional approach of returning the cells to a central location to accomplish the verification is extremely inconvenient. Again, it was determined that this requirement had to be met using a different approach.

Finally, in addition to demonstrating competency for the purpose of accreditation, the additional challenge faced by these two laboratories is one of demonstrating equivalency to interested customers. It is important from the customer's viewpoint that calibrations provided by the two laboratories be equivalent relative to the stated uncertainties. Such is necessary to satisfy the purpose of the SPRT user while maintaining operational efficiency.

2. MEASUREMENTS

2.1. SPRT Calibration Comparison

2.1.1. Strategy

For the SPRT portion it was determined that a bilateral interlaboratory comparison between the US and UK facilities, when taken in conjunction with the NVLAP PT conducted at the US facility, could accomplish the purpose. This section will describe the interlaboratory comparison, the results of the comparison, and the results of the NVLAP PT upon which the comparison was based.

2.1.2. Measurements

The NVLAP PT is available for several ranges of temperature. The range selected for this test should cover the range of accreditation. In our case, we needed the widest span available; from the triple point of argon to the freezing point of aluminium, approximately –190 °C to 660 °C. This corresponds to ITS-90 ranges 4 and 7. It made sense to apply this same range to the bilateral comparison. The NVLAP PT utilizes three SPRTs, calibrated over the entire range. Since SPRTs are available with glass or steel sheaths, and it reasonable to assume that calibration lab performance might be different as the calibration applies to differing types, it was decided to include both types in the MAP. These SPRTs belong to NIST and were not all manufactured by Hart Scientific. The use of multiple SPRTs proved impractical for the bilateral comparison, consequently, one artefact was used (as is common with most PTs). We desired the most conclusive of results so we elected to use the best available SPRT from our product portfolio.

First, to ensure that the SPRTs were stable and arrived without damage, the initial R_{TPW} was measured and compared to the previous value. In the case of the NVLAP PT, the values were provided to NIST and NIST performed the check. Once the stability of the SPRTs was demonstrated, calibration commenced in the conventional manner. The procedures used for the actual calibration were essentially conventional SPRT calibration procedures with the exception that all of the inclusive ITS-90 fixed points were included. In the case of the MAP, additional checks were employed to verify proper application of the corrections and validate the mathematical operations. These additional steps were not necessary for the bilateral comparison because both laboratories employ the same internally written software. After calibration was completed, the SPRTs and calibration results were returned to the reference laboratory. In the case of the NVLAP PT, NIST recalibrated the SPRTs to demonstrate that stability was maintained throughout the process. This was not possible in the case of the bilateral comparison because the SPRT was required for another project.

2.1.3. Results

The MAP report includes the measurement results, details pertaining to the ITS-90 fixed point cell corrections and mathematics, information relating to the redundant fixed points, and data relating to the stability of the SPRTs involved. This paper will describe only the measurement results including redundant fixed points.

PT results are generally evaluated using the normalized error, denoted E_{normal} or E_n. The normalized error is a ratio of the difference in the measurement results relative to the combined measurement uncertainties. E_n is calculated as follows:

$$E_n = {(x_i - x_r)} \Big/ {\sqrt{(U_i)^2 + (U_r)^2}} \qquad (1)$$

Where:	E_n	=	normalized error
	x_i	=	measurement result from laboratory under evaluation
	x_r	=	measurement result from reference laboratory
	U_i	=	expanded uncertainty of measurement under evaluation (k=2)
	U_r	=	expanded uncertainty of reference measurement (k=2)

When $|E_n| \leq 1$, the comparison is deemed successful. In most cases, this outcome is considered conclusive even when the two uncertainties are similar in magnitude. However, when $E_n > 1$ problems can arise. When this occurs and the uncertainty in the reference value is small relative to the uncertainty in the unknown value, it can be logically concluded that the unknown value is suspect because the reference uncertainty does not contribute much to the combined uncertainty. However, when $E_n > 1$ occurs and the two uncertainties are similar in magnitude, the result may be inconclusive because one may not be able to determine which value is the correct value without additional evidence.

In the case of the NVLAP PT, the NIST uncertainties are significantly smaller than the uncertainties in the unknown, therefore, a conclusive result (successful or unsuccessful) is expected. In the case of the bilateral comparison, the uncertainties of the two labs are essentially identical and an inconclusive result is possible.

Finally, since both Fluke laboratories are traceable through one set of cells and apparatus, it is expected that some of the individual components of uncertainty may be correlated. If this is the case, the E_n calculation may under represent the actual errors. Consequently, we attempted to identify the components that might be correlated and remove them from the result. The components of uncertainty along with the correlation assumptions are shown in Table I. Since equivalence of the two laboratories is very important to our customers, we decided to take a very conservative position on the value of E_n. Although values of E_n between -1 and 1 are considered as passing results, we decided to evaluate any conditions where E_n is between -1 and –0.5 or 0.5 and 1 with an effort to improve the equivalence.

The results of the comparison experiments are shown in Table II and III and graphically in Figure 1. The results of the redundant fixed point measurements at the Ga MP and In FP are shown graphically in Figure 2. The comparison experiments demonstrate definitive agreement both between NIST and the Hart Scientific US facility and between the Hart Scientific US and UK facilities. When taken in conjunction with our rigorous cut-off criteria of ± 0.5 for E_n we conclude that the PTs were successful and no further action is required. Additionally, in both cases, the non-uniqueness is consistent with expectations, demonstrating good internal consistency in the calibration process.

2.2. Fixed Point Cell Comparison

2.2.1. Strategy

The US facility maintains three sets of fixed point cells and the UK facility maintains two sets of fixed point cells. In the US facility, these cells function as the primary set, working set, and SPRT calibration set. In the UK facility, these cells function as SPRT calibration set and backup and cross-check set. The US primary set has been tested at NIST. The use of these cells is restricted to the certification of the working cells, the SPRT calibration cells, the cells for the UK facility, and newly purchased cells for selected customers who require the lowest possible uncertainties (primarily NMI customers). The US working cells are used in the routine certification of customer fixed point cells and as cross-check cells for the SPRT calibration cells. Since the primary set was tested at NIST, the uncertainties assigned to these cells are smaller than the other sets. The uncertainties assigned to the other sets are comprised mainly of the uncertainties attributed to the primary cells and the uncertainties of the comparison process. As a result, the uncertainties assigned to these sets are nominally identical. If the uncertainties in the comparison process are kept as small as possible, equivalency among the cells should be fairly straightforward to demonstrate.

Although the fixed point cells can be assigned an uncertainty based upon purity, construction, and other characteristics, we find it simpler to treat the cells as calibrated artefacts. This approach is somewhat unconventional but makes the traceability and uncertainty analysis more direct. In line with this approach, the certification includes the ΔT observed during the comparison experiment, corrected to ΔT from the ITS-90 nominal value, along with the uncertainties. The uncertainties propagate from the uncertainties of the original NIST certification and the uncertainties of the various comparison experiments. For most cells, the observed ΔT is small relative to the uncertainties of the comparison experiment. Traceability to NIST is established through an unbroken chain of comparisons in the conventional manner for calibrated instruments.

Finally, both NVLAP and UKAS require periodic verification that the cells in every day use are stable over time. It was determined that this requirement could be met by alternating semi-annual plateau evaluation and semi-annual comparison of the working cells with new cells or the backup cells.

2.2.2 Measurements

As described previously, the various sets of cells have been certified with the ΔT from the nominal ITS-90 temperature provided. Therefore, the difference in observed temperature of any two cells can be calculated and compared. The results can then be evaluated in context with the uncertainties to determine the success of the comparison. If the ΔT exceeds the expectation, the test can be considered unsuccessful. If the ΔT is smaller than the expectation, the test can be considered successful.

The measurements were conducted using an ASL F18 or MI 6010T bridge (or both), SPRTs known to be stable at the temperatures of interest, thermally regulated reference resistors, and appropriate realization apparatus. In all cases, multiple SPRTs were used for each cell. For direct comparison, it has been suggested that R_{T90} is superior to W_{T90} for detecting small differences. However, to ensure SPRT stability during the comparison process, R_{TPW} was measured at the opening and closing of each fixed point cell comparison experiment. The bridges were controlled using software both to reduce the possibility of operator error and to improve the resolution and reproducibility of the results. The software used with ASL F18 was written in-house. The software used with the MI 6010T is commercially available software purchased with the bridge. To ensure that the plateaus were evaluated at the identical percent of sample frozen (or melted), the initiations of the plateaus were offset by the time interval required to complete one measurement sequence. The measurements were executed at two levels of current with the results extrapolated to zero power. The zero power values were taken as the values representing R_{T90}. To ensure the achievement of thermal equilibrium, the measurement sequence consisted of three elements; nominal power, double power, and nominal power. Thermal equilibrium and thermal stability were verified before the measurement was accepted. Once the data were obtained, the ΔT values were compared to the calculated ΔT values.

2.2.3 Results

The normalized error parameter, E_n, will be used to demonstrate equivalence. However, unlike the comparison of SPRTs, the comparison of fixed point cells within the individual laboratories should not contain significant correlated uncertainties. However, E_n was calculated both using the combined uncertainties, U_C, and with the uncertainties of only one lab, $U_{Norwich}$.

The results are shown in Table IV and graphically in Figure 3. In both cases, the comparison experiments demonstrate definitive agreement between the two Hart Scientific facilities. When taken in conjunction with our rigorous cut-off criteria of ± 0.5 for E_n we conclude that the comparison tests were successful and, with the exception of including the Al FP and LN_2 comparison later this year, no further action is required.

3. CONCLUSIONS

Through the experiments described above, it can be concluded that the two Fluke Corporation laboratories show excellent equivalence both in SPRT calibrations and fixed point comparison tests. Furthermore, the NVLAP PT results demonstrate excellent equivalence between the US facility and NIST.

ACKNOWLEDGMENTS

The author would like to thank Miss Alison Shrimpling of our Norwich facility for her assistance and participation in conducting the majority of the measurements in the Norwich laboratory, as well as Greg Strouse of NIST for his guidance in designing the experiment and tutelage in interpreting the NVLAP PT results.

REFERENCES

1. B.W Mangum, P. Bloembergen, M.V. Chattle, B. Fellmuth, P. Marcarino, and A.I. Pokhodun, *Metrologia* **36**:79-88 (1999).

1. A.G. Steele and R.J. Douglas, *Metrologia* **436**:S235-S243 (2006).

3. ISO/IEC Guide 43:1997 Part 1: Development and operation of proficiency testing schemes (1997).

4. ILAC-G22:2004 Use of proficiency testing as a tool for accreditation in testing (2004).

5. G.F. Strouse, in *Tempmeko 2004 Proceedings*, Vol. 2, D Zvizdic, ed. (Laboratory for Process Measurement, Zagreb, Croatia, 2004), pp. 879-884.

6. G.F. Strouse, in *Temperature: Its Measurement and Control in Science and Industry*, Vol. 6, Part 1, J. F. Schooley, ed. (AIP, New York, 1992), pp. 175-178.

7. NVLAP Calibration Laboratories Accreditation Program Proficiency Test Report, Test folder number 836/270208-04 (2004)

8. NVLAP Lab Bulletin LB-10-2004 Thermometry Proficiency Tests (2007).

Table I. SPRT Uncertainty Components and Correlation Assumptions

Uncertainty Component	Type	Correlation Assumption
Process variability as observed by check standard SPRT	A	Uncorrelated
Precision of measurement (procedure limit n = 40)	A	Uncorrelated
Fixed point value (reference cell certification)	B	Partially Correlated
SPRT self-heating correction	B	Fully Correlated
Hydrostatic head correction	B	Fully Correlated
Non-ideal immersion profile	B	Fully Correlated
R_{TPW} propagation	B	Uncorrelated
Shunt losses	B	Fully Correlated
Bridge nonlinearity	B	Uncorrelated
Reference resistor instability during process	B	Uncorrelated

Table II. NVLAP SPRT PT Results

ITS-90 Fixed Point	$\Delta W(t_{90})$ (mK)	U_{NIST} (k=2) (mK)	U_{HART} (k=2)[a] (mK)	U_C (k=2) (mK)	E_{normal}[b]
LN_2	-0.28	0.14	0.60	0.62	-0.5
Hg TP	-0.04	0.15	0.40	0.43	-0.1
Ga MP	-0.02	0.07	0.40	0.41	0.0
In FP	-0.04	0.18	0.90	0.92	0.0
Sn FP	0.00	0.28	0.90	0.94	0.0
Zn FP	-0.53	0.51	1.10	1.21	-0.4
Al FP	-0.88	0.79	2.10	1.24	-0.5

[a] The uncertainties shown are those on the laboratory scope of accreditation (NVLAP lab code 200348)
[b] The E_{normal} values shown were calculated with the accredited uncertainties rather than the preliminary uncertainties estimated at the time of the PT. Therefore, the E_{normal} values are negligibly different from the values shown on the PT report.

Table III. SPRT Bilateral Comparison Results

ITS-90 Fixed Point	$\Delta W(t_{90})$ (mK)	$U_{HART\ AF}$ (k=2)[a] (mK)	$U_{HART\ UK}$ (k=2)[a] (mK)	U_C (k=2) (mK)	E_{norma}
LN_2	0.23	0.55	1.55	1.64	0.1
Hg TP	0.18	0.30	0.30	0.42	0.4
Ga MP	-0.07	0.35	0.35	0.49	-0.1
In FP	0.13	0.69	0.69	0.98	0.1
Sn FP	-0.25	0.83	0.83	1.17	-0.2
Zn FP	-0.51	1.05	1.05	1.48	-0.3
Al FP	-0.52	1.83	1.93	2.66	-0.2

[a] The uncertainties shown represent those on the respective scopes of accreditation with the correlated components removed. (NVLAP lab code 200348, UKAS certificate number 0775)

Table IV. Bilateral Fixed Point Cell Comparison Results

Fixed Point	$\Delta t_{calculated}$ (mK)	$\Delta t_{measured}$ (mK)	$\Delta t_{difference}$ (mK)	U_{AF} (k=2) (mK)	U_{UK} (k=2) (mK)	U_C (k=2) (mK)	E_n[a]	E_n[b]
Hg TP	-0.14	-0.20	-0.06	0.20	0.20	0.28	-0.2	-0.3
Ga MP	-0.07	-0.11	-0.04	0.08	0.08	0.11	-0.4	-0.5
In FP	-0.47	-0.65	-0.18	0.50	0.50	0.71	-0.3	-0.4
Sn FP	-0.24	-0.09	0.15	0.60	0.60	0.85	0.2	0.3
Zn FP	-0.65	-0.44	0.21	0.80	0.80	1.13	0.2	0.3

[a] E_n calculated using combined uncertainty, U_C

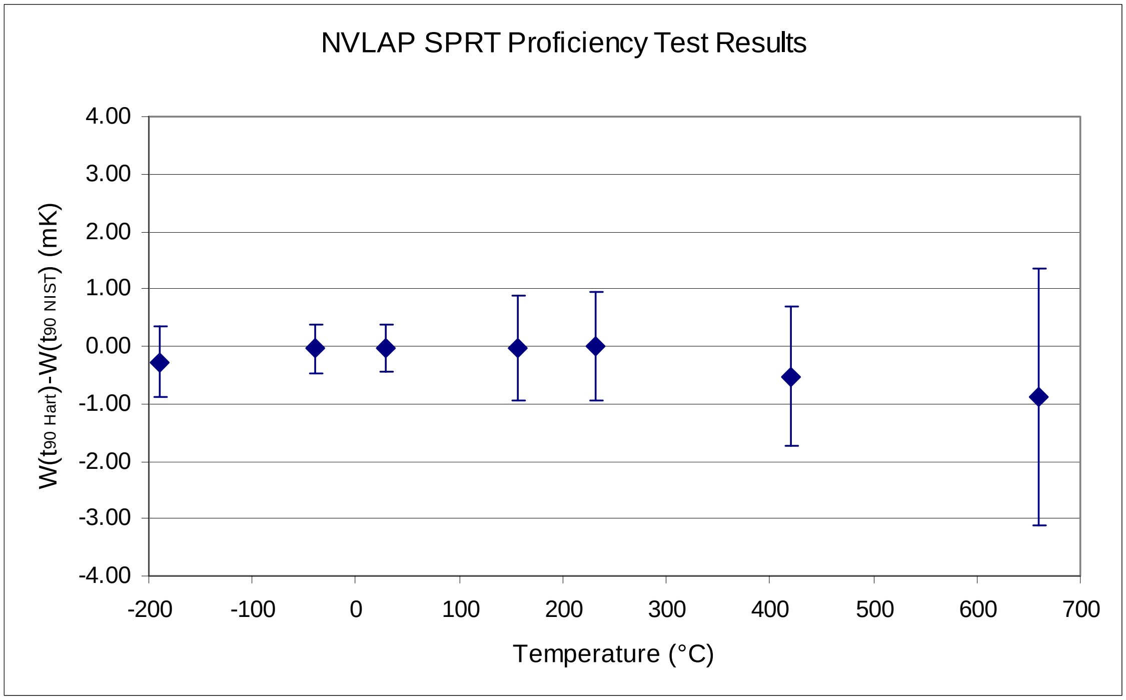
[b] E_n calculated using individual uncertainty of Norwich laboratory, U_{UK}

FIGURE CAPTIONS

Fig. 1. Proficiency test results at 0 mA covering the range -200 °C to 660 °C.

Fig. 2. Error/non-uniqueness in the SPRT calibrations at the Ga MP and In FP relative to the RSS propagated subrange uncertainty.

Fig. 3. Fixed point cell comparison results covering the cells Hg TP through Zn FP.


NVLAP SPRT Proficiency Test Results
W(t90 Hart)-W(t90 NIST) (mK)
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-200
-100
0
100
200
300
400
500
600
700
Temperature (°C)

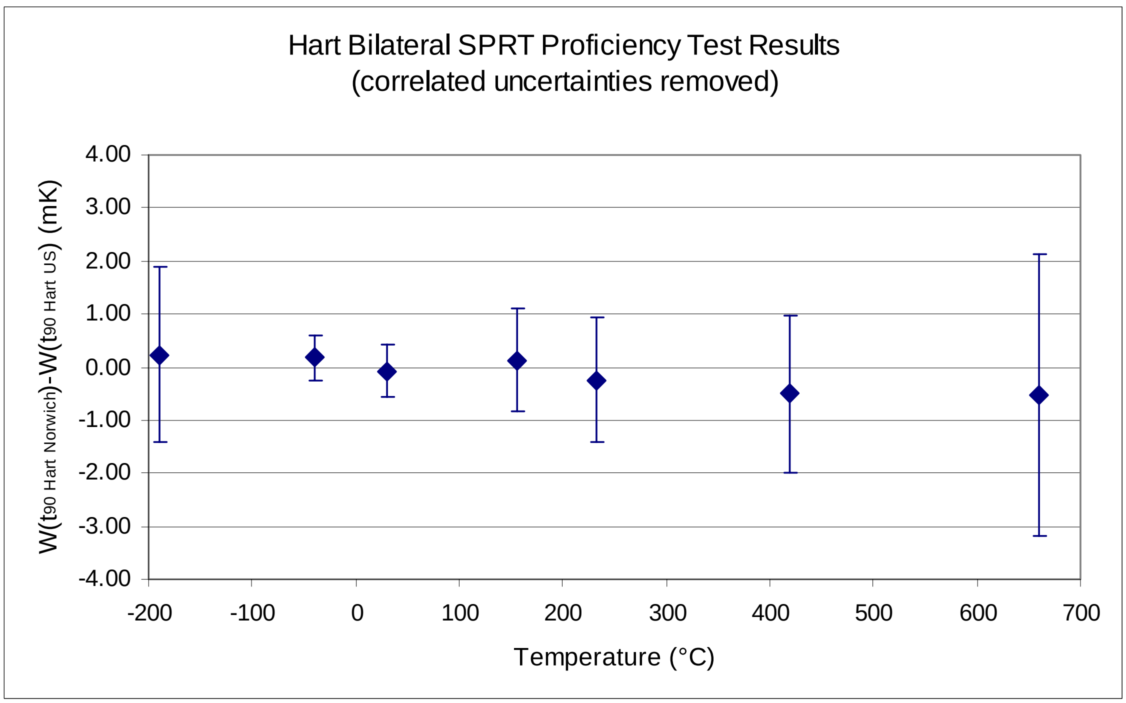
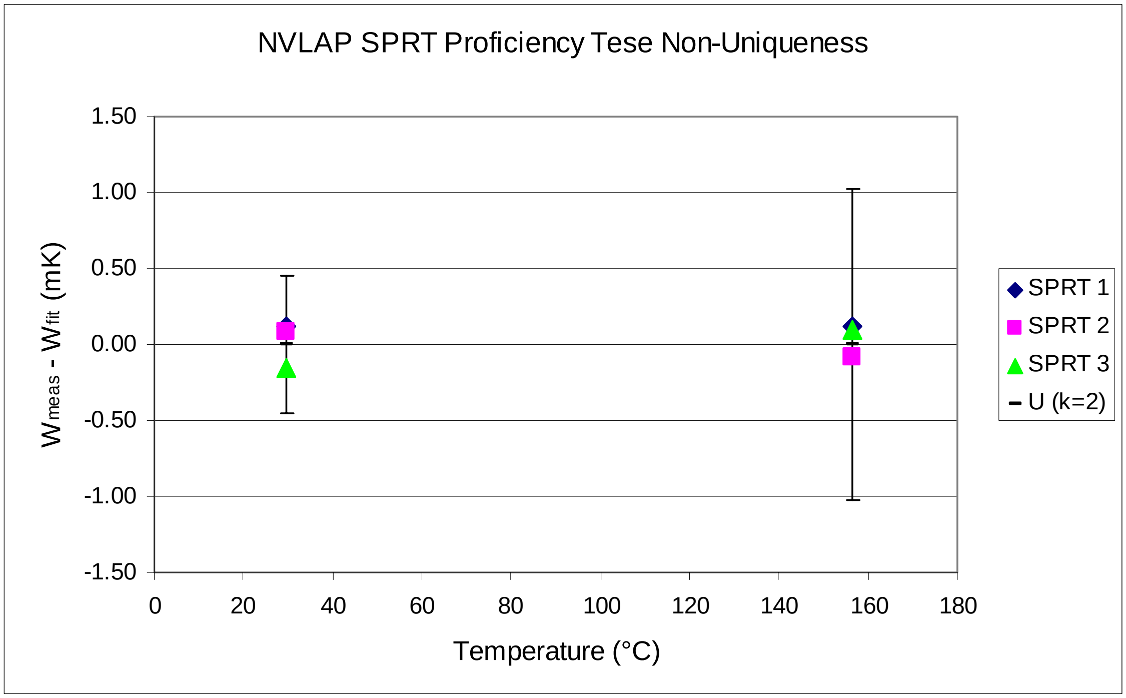

Hart Bilateral SPRT Proficiency Test Results
(correlated uncertainties removed)
W(t90 Hart Norwich)-W(t90 Hart US) (mK)
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-200
-100
0
100
200
300
400
500
600
700
Temperature (°C)

Figure 1


NVLAP SPRT Proficiency Tese Non-Uniqueness
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
0
20
40
60
80
100
120
140
160
180
Temperature (°C)
Wmeas - Wfit (mK)
SPRT 1
SPRT 2
SPRT 3
U (k=2)

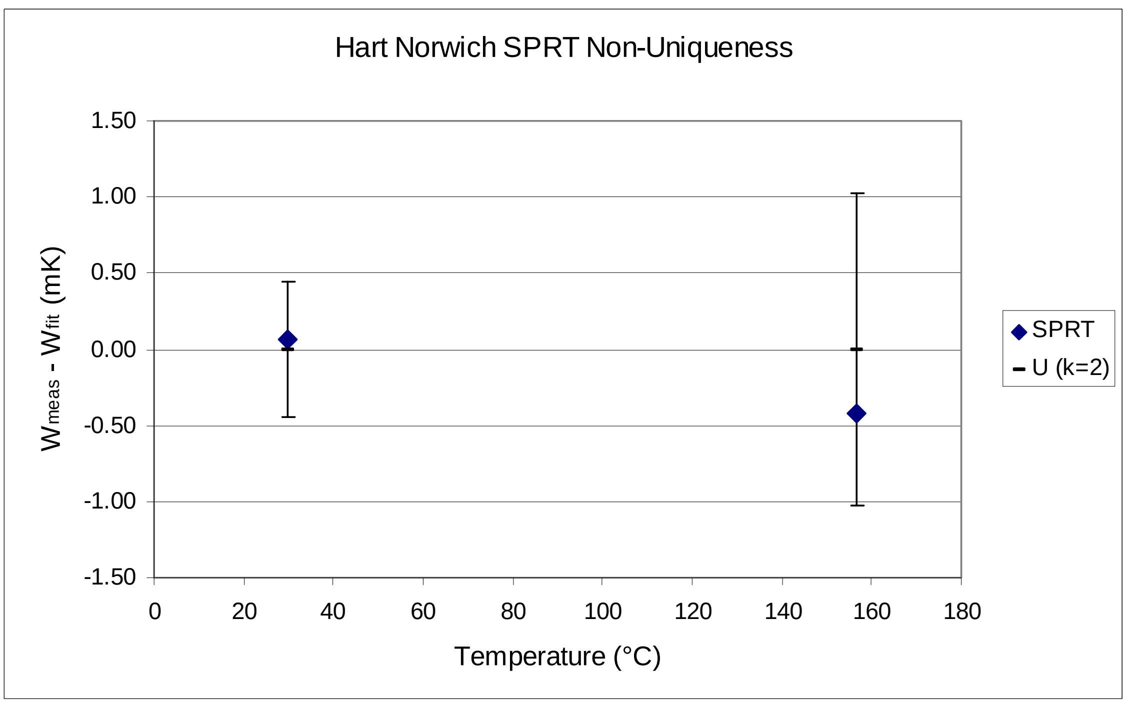
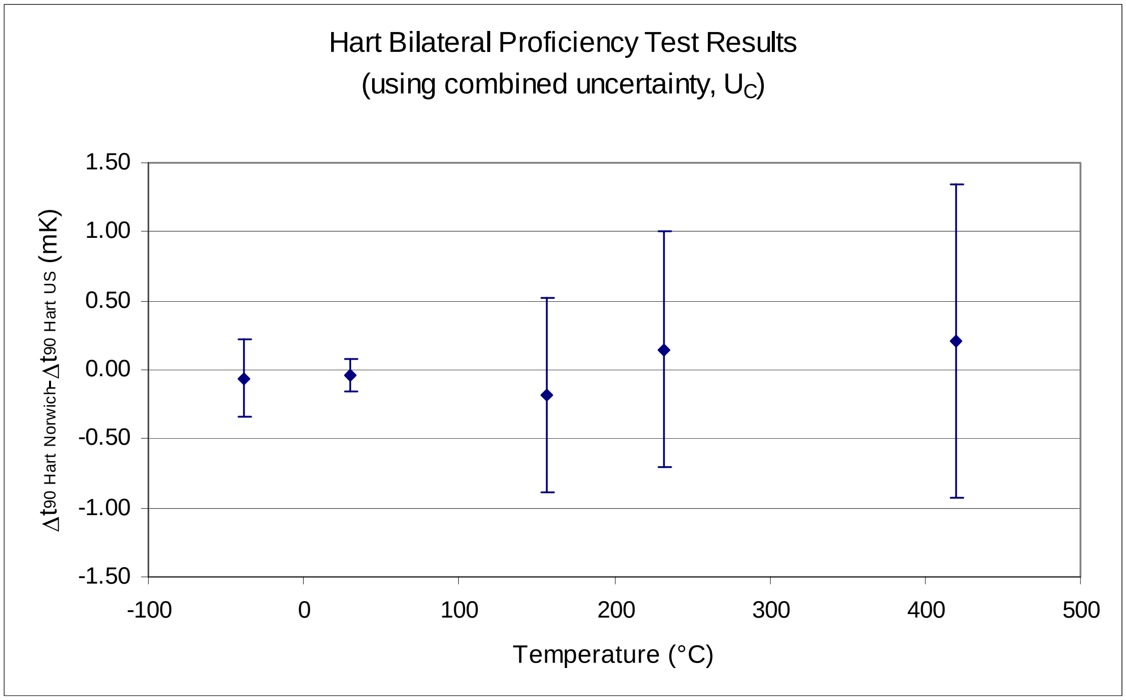
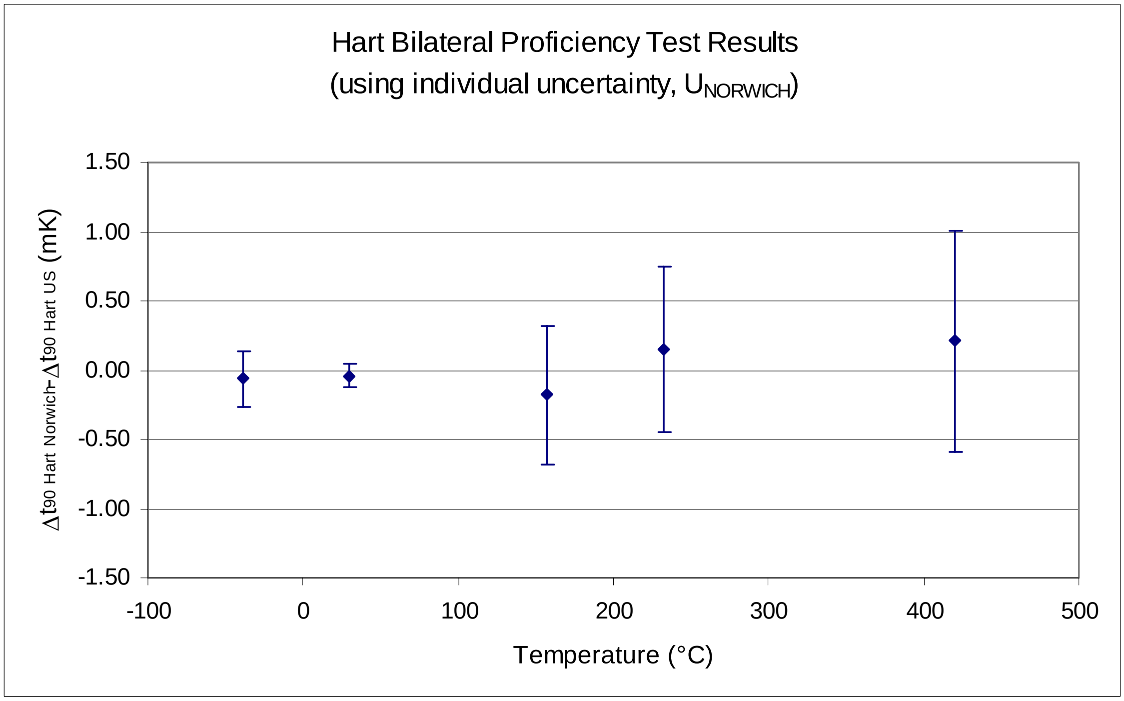

Hart Norwich SPRT Non-Uniqueness
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
0
20
40
60
80
100
120
140
160
180
Temperature (°C)
Wmeas - Wfit (mK)
SPRT
U (k=2)

Figure 2



Hart Bilateral Proficiency Test Results
(using combined uncertainty, UC)
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-100
0
100
200
300
400
500
Temperature (°C)
Δt90 Hart Norwich-Δt90 Hart US (mK)



Hart Bilateral Proficiency Test Results
(using individual uncertainty, UNORWICH)
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-100
0
100
200
300
400
500
Temperature (°C)
Δt90 Hart Norwich-Δt90 Hart US (mK)

Figure 3