XVIII IMEKO WORLD CONGRESS
Metrology for a Sustainable Development
September, 17 – 22, 2006, Rio de Janeiro, Brazil

A COMPARISON OF THE INITIAL UNCERTAINTY BUDGET AND AN ESTIMATE OF THE POSTERIOR UNCERTAINTIES, IN THE CASE OF LARGE BATCHES OF CALIBRATION DATA: THE LHC THERMOMETERS AT CERN

Daniela Ichim [1], *Franco Pavese* [2]

[1,2] Istituto Nazionale di Ricerca Metrologica (INRIM)[a], Torino, Italy, f.pavese@imgc.cnr.it

Abstract: The paper will describe the techniques that have been used to perform the comparison on large batches of cryogenic semiconductor-type thermometers, calibrated for the CERN LHC and the main results obtained: they concern either the uncertainty of the Cernox™ thermometers under calibration and the behaviour of the standards used during the calibrations.

Keywords: calibration uncertainty, uncertainty evaluation, thermometers.

1. INTRODUCTION

When a calibration facility is set up, an estimate of the uncertainty budget is performed, on the basis of the specifications of the instrumentation used and of the requirements arising from the specifications of the sensors to be calibrated, in order to ensure that the maximum uncertainties match the needed level of the calibration uncertainty. This estimate is generally required for a 95% confidence interval, but is some cases more stringent limits are required.

When a large number of sensors have to be calibrated, the actual experimental conditions can vary from sensor to sensor, resulting in a risk that,, for some of them the calibration uncertainty exceeds the prescribed limit. On the other hand, it may result that the uncertainty budget evaluations were more severe than the calibration process actually allows, resulting in a lower calibration uncertainty.

Consequently, it is interesting to perform an *a posteriori* analysis on the calibration data, to estimate the posterior –and actual– uncertainty level of the calibrations obtained. This can be better done when the number of calibrations is large.

The results of these studies that will be reported are concerning large batches of cryogenic semiconductor-type thermometers, calibrated for the CERN LHC.

2. THE CALIBRATION PROCESS

For the thermometers to be calibrated, the CERN specifications prescribe a tolerance that should be respected by *all* 6000 thermometers: $\pm$ 5 mK (1.6–2.2 K); $\pm$ 10 mK (2.2–4.2 K); $\pm$ 15 mK (4.2–6.0 K); $\pm$ 0.5 K (6.0–25 K); $\pm$ 2.5 K (25–300 K). The experimental calibration work has been performed by IPNO (Saclay, France), the calibration assessment by INRIM.

The calibration of *M* thermometers requires several steps [1], from the experimental data, today generally acquired automatically by means of a computer-assisted system, to the sets ***a*** of parameters of the calibration function. The layered structure of this process, as designed by IMGC [2], is schematically shown in Fig. 1.

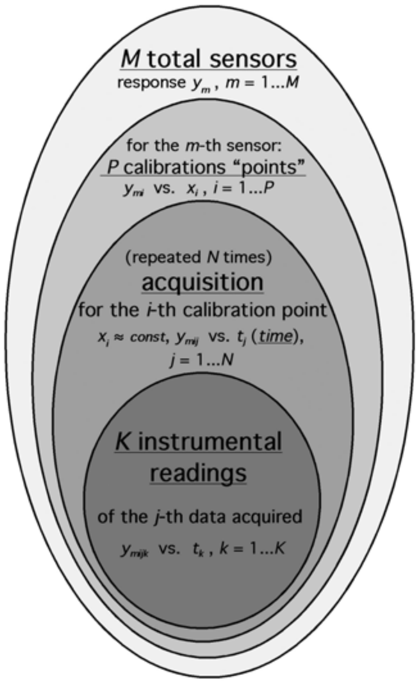


Fig. 1. The layered data structure of the calibration of *M* sensors at *P* temperatures, for each current *I* in the sensor.

The steps are described in Table 1. Initially, it is necessary to select the function best-suited for the specific thermometers: for a semiconducting type, as in the case of the LHC ones, cubic splines are a better choice than a high-order logarithmic polynomial. Only when the model function is decided is it possible to optimise the calibration point distribution and minimise their number.

[a] Until 31.12.2005, Istituto di Metrologia "G.Colonnetti" (IMGC), CNR.

Table 1. The process of calibration, data reduction and its mathematical tools.

Data treatment step	Data (Fig.1)	Mathematical tool	Ref.
1) Choice of the model for the thermometer characteristics and experimental design	$y = \mathrm{f}(\boldsymbol{a}, x)$ $x = T$	Selection of the best model function class f(x) and of the distribution of the calibration temperatures T_i and optimisation of their number P	6
2) On-line outlier rejection during automated data acquisition	$y_{mijk}\ \forall\ t_k$ t is time	Sequence-Analysis Outlier Rejection (SAODR) routine	3
3) Elaboration of the data for each calibration point, with thermal drift suppression	y_{mij} *vs* t_j $\forall\ x_i$	Least Squares Mixed Effect Method (LSME), which makes use of the whole set of (y_{mij}, t_j) data to obtain (y_{mi}, x_i), obtaining a robust compensation of thermal drift	4
4) Elaboration of the data for identification of possible anomalous data	(y_{mi}, x_i)	Use of the LSME method for estimating the group $\{y_i, x_i\}$ as a whole, for robust detection of anomalous data	5
5) Elaboration of the full calibration data for each thermometer, to obtain the calibration model function parameters	$\{(y_{mi}, x_i)\}$, $y_m = \mathrm{f}(\hat{\boldsymbol{a}}, x)$	Application of the model class (e.g., splines) to the data, including the handling of incomplete datasets, to obtain, for each thermometer the set of parameters $\hat{\boldsymbol{a}}$.	5
6) Elaboration of the calibration functions for identification of possible thermometer clusters	$\hat{\boldsymbol{a}}$	Use of the LSME method with a simplified model (polynomial) for identifying the clusters of $\hat{\boldsymbol{a}}$ sets as a whole, obtaining a robust detection of clusters of characteristics	2
7) Estimation of the overall calibration uncertainty			*This paper*

Starting from the raw data acquisition (innermost box in Fig. 1, step (2) in Table 1), each instrumental reading is performed K times [3]. This allows the detection of outlying readings and their rejection. This step was not eventually experimentally implemented.

In order to have statistical information on each calibration point, the measurement on each thermometer is repeated several times, say N. During the $N \cdot M$ measurements, the temperature generally drifts: better thermal stabilisation generally requires more expensive controllers and more time.

Consequently, there is advantage in using an algorithm that can effectively suppress thermal drift from the acquired data y_{mij} (next outer box in Fig. 1, step 3 in Table 1), in order to obtain the calibration point y_{mi} for each m-th thermometer. For the Least Squares with Mixed Effect (LSME) method to be applied, the thermometers do not need be identical, but only "similar" to a certain extent. [4]

An additional bonus is that the overall LSME easily allows the detection of anomalous data (e.g., coming from noisy acquisition channels), while providing statistical information on the *overall* batch of thermometers under calibration (step 4 in Table 1 [5]). When all the calibration points $\{(y_{mi}, x_i)\}$ are obtained, again the LSME can be used with a simplified model (polynomial) to evaluate the possibility of a grouping in clusters of the thermometers (step 5 in Table 1 [2]).

3. INITIAL UNCERTAINTY BUDGET

The initial uncertainty budget, as evaluated by INRIM, is reported in Table 2 for the most critical temperature range[1]. Note that actually a 99% confidence limit should be used, leading to an uncertainty of 6.5 mK, outside tolerance limit.

Table 2. Initial uncertainty budget (1.6 -2.2 K), 95% confidence level.

<u>On temperature</u> (reference thermometers traceable to IMGC), mK	
1) IMGC calibration	2.0
2) Overheating correction (T < 2.2 K)	1.4
3) Stability of the working reference thermometers within calibration time interval	1.0
4) Uncertainty of measurements of the reference thermometers at IPNO	2.1
5) Repeatability of measurements at IPNO	0.7 [a]
<u>On resistance</u> (of thermometers to be calibrated) , mK	
6) Uncertainty of measurements of the thermometers to be calibrated by IPNO	0.2
7) Repeatability of measurements on the thermometers to be calibrated by IPNO	0.1 [a]
8) Uniformity of the comparison block	0.5
9) Fitting of the calibration function by IMGC	3.5 [a]
Total	**4.4**

[a] It can vary from run to run, to be assessed *a posteriori*.

[1] The reason of the criticity is that the thermometers are also used for controlling the superfluid-helium bath temperature.

4. OVERALL CALIBRATION UNCERTAINTY EVALUATION: BASICS

Several of the steps in Table 1 allow an evaluation of the measurement uncertainty: step 2, when present, the uncertainty of each instrumental reading; step 3, the uncertainty of the each calibration point (for both the T value and the resistance R value of each thermometer under calibration); step 5, the uncertainty of the calibration function fit. Additionally, a screening for anomalous data (step 4, but also step 6) ensures that they do not affect the normal calibration-function determination process.

After step 5, a Monte Carlo simulation is also performed, by letting the computed calibration points randomly vary within squared boundaries determined by the uncertainties of each point on both T and R. This simulation additionally estimates the stability of the calibration function within the experimental uncertainties [7].

5. POSTERIOR CALIBRATION UNCERTAINTY EVALUATION: RESULTS AND DISCUSSION

The calibration process involves in each run about 70 thermometers, requiring overall close to a hundred thousand readings. In fact, to the default 34 calibration points obtained for lowering temperatures (Procedure A, default) from the optimization process of step 1 in Table 1, about another 50 auxiliary points have been performed in less controlled conditions for increasing temperatures (procedure B), allowing some double checks with a different procedure.

Therefore, it is not surprising that even a computer-controlled automatic data acquisition system resulted in a broad variety of actual experimental conditions, reflecting in output files showing a broad variety of non-standard outputs. This paper is reporting on the results of four (out of 110) runs, considered as representative samples of the overall population.

Table 3 is summarizing the main default outputs expected from each IPNO run, together with the actual outputs from the four test runs. In addition to uncertainty evaluation, also a classification of the thermometers in homogeneous R-T characteristics clusters has been performed according to step 6 in Table 1.

Procedure A refers to a calibration performed at a number of temperatures of decreasing value (31 is the default), with a good stability during the time required to measure all the thermometers and the references ("temperature plateau") and at least 3 repeated measurements per thermometer. This allows to fully perform the evaluations of step 2 and 3 in Table 1 (but SAODR routine (step 2) was not implemented at IPNO). The procedure also includes plateau identification and a test of the quality of the plateau, leading to occasionally discard those that do not comply with a sufficient temperature homogeneity. Therefore, an uncertainty on both R and T can be attached to each calibration point obtained with procedure A in order to compute the calibration function (step 4 in Table 1), which is a set of cubic splines.

Procedure B, on the contrary, is non standard. It refers to measurements taken at a number of temperatures of increasing value (50 is the default), with a much less controlled temperature environment and being the thermometers under calibration measured only one (one reading per thermometer). On the contrary, the reference thermometers are measured several times. No direct evaluation of the thermal conditions is possible and no uncertainty evaluation can be done on the calibration points. The only uncertainty parameter is the standard deviation of the fitting, which is performed using a logarithmic polynomial of degree 10, which does not provide residuals passing a Kolmogorov-Smirnov normality test (insufficient model).

Table 3 show quite a variety of experimental situations and of result quality. Especially in the most critical range (< 2.2 K) is can happen to have a reduced number of valid calibration points for the default current (1 μA): in this case, the fitting with the spline model is done with a number of experimental points lower than the default (e.g., …UN CASO…). The calibration with a current of 10 μA in this range will only be used in emergency in the LHC (high electrical noise): the Cernox® thermometers show quite a large overheating at this current (see Figure 1), resulting in a sharp bend in the R-T characteristics, very difficult to accurately track with the calibration function, especially with the polynomial model –this reflects in line (14) of Table 3.

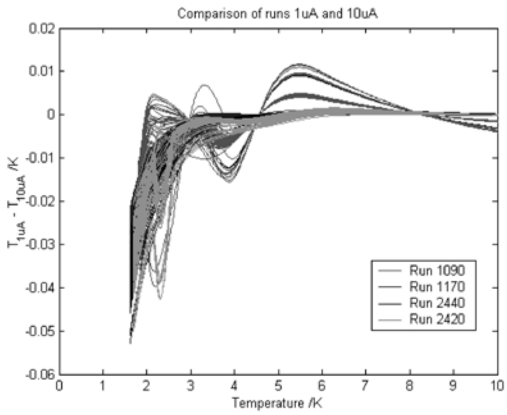
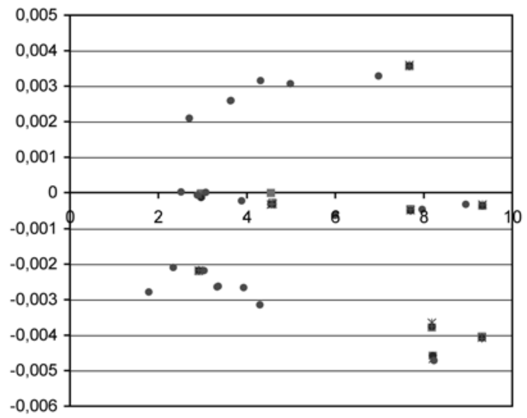
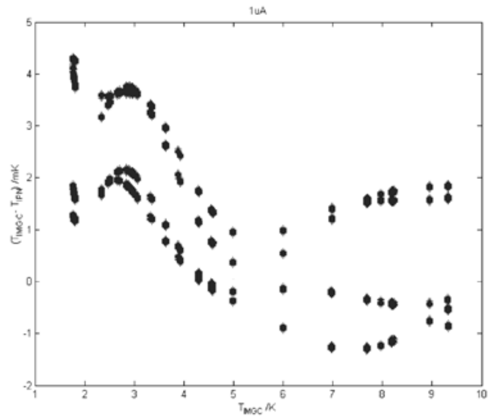


Figure 1. Overheating for I = 10 μA of Cernox® thermometers.

The procedure allows to obtain an estimate of the actual differences in calibration of the reference thermometers. The calibration uncertainty was evaluated in Table 2 (lines (1) to (5)) to be 3.5 mK: differences up to 5 mK between calibrated references were observed, which are just within the combined uncertainties, confirming the initial evaluation. An example of the sets of differences between references is reported in Figure 2.



a)



b)

Figure 2. Differences between reference thermometers measured during a calibration run #1090 at IPNO versus tempereature (K): a) difference between two of the four reference thermometers (mK); b) (IMGC – IPNO) for all reference thermometers (mK).

Table 3. Main specifications for each calibration run and the actual outcomes of four typical calibration runs.

(*per calibration run*)	CERN specification	Run #1090	Run #1170	Run #2420	Run #2440
1) Number of reference thermometers	3…4	4	3	2	2
2) Number of thermometers under calibration (Cernox®)	70…80	77*	74	52	28
3) Number of calibration points (I = 1 μA, T = < 4 K, procedure A)	20	12 T_{ini}–T_{fin} < 30 mK	10 T_{ini}–T_{fin} < 4 mK	20 T_{ini}–T_{fin} < 5 mK, except 2 (50 mK)	24 T_{ini}–T_{fin} < 4 mK
4) Number of calibration points (I = 10 μA, T > 54 K, procedure A)	28	26	24	26	30
5) Number of calibration points (I = 100 μA, T = 66–300 K, procedure A)	6	7	6	4	5
6) Calibration point sequence	1 cooldown + 1 warmup	2 cooldowns + 2 warmups	default	default	default
7) N° invalid thermometers	–	7	1	4	0
8) N° thermometers with at least one non-default feature	–	2 (1 μA) 5 (10 μA)	10 (1 μA) 2 (10 μA)	14 (1 μA) 27 (10 μA)	1 (1 μA) 19 (10 μA)
9) N° thermometers with non-typical *R-T* characteristics	–	11	10	8	2
10) N° of thermometers with no calibration anomalies (valid thermometers)	–	57 (1 μA) 54 (10 μA)	53 (1 μA) 61 (10 μA)	26 (1 μA) 13 (10 μA)	25 (1 μA) 7 (10 μA)
11) Max Monte Carlo fitting-stability boundary (95% CI), mK		3	3.5	2 (1 μA) 5 (10 μA) with some up to 11	3
12) Reference thermometers agreement (range), mK		5	5	2	1
13) Procedure A typical uncertainty (fitting with cubic splines over 31 calibration points –default) (T = 1.6–2.2 K), 95% CI, mK	< 5	< 1 (see overall estimate)	< 1	< 2 (1 μA) < 4.5 (10 μA)	2
14) Procedure B typical uncertainty (fitting with a logarithmic polynomial over 50 calibration points) (T = 1.6–2.2 K), 95% CI, mK	< 5	(1 μA: see overall estimate) (10 μA: up to 18)*	Up to 3 (1 μA) Up to 5.5 (10 μA)	Up to 4.5 (1 μA) For 10 μA only one valid calibration point in the range	Up to 8
15) Overall estimate (T = 1.6–2.2 K), 95% CI	± 5 mK (tolerance)	Unsure that all valid thermometers comply with the tolerance, due to a single valid calibration point in the range	All valid thermometers within specs	All valid thermometers within specs, with less confidence @ 10 μA	Within specs: all valid @ 1 μA, low confidence @ 10 μA

* includes the same 74 thermometers of run #1170.

The procedure also allowed to detect different types on anomalies in the thermometers under calibration: from faulty scanner channel to out-of-tolerance deviations of the fitted function; from “noisy” calibration points to anomalous uncertainty levels in step 3 of the procedure; from out of tolerance instability of the fitted model on the Monte Carlo test to critical dependance of the resulting calibration function on the model used[2]. An example of large differences in the standard deviations of the measurements performed at different calibration temperatures is reported in Figure 3.

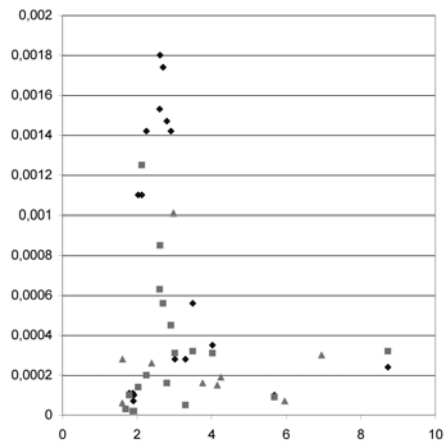


Figure 3. Standard deviations of the thermal-drift fit, made using the procedure of step 3 in Table 1, for different calibration points in run #2420.

An example of dependance of the calibration function from the model, where the tolerance is not respected by all the thermometers, is shown in Figure 4.

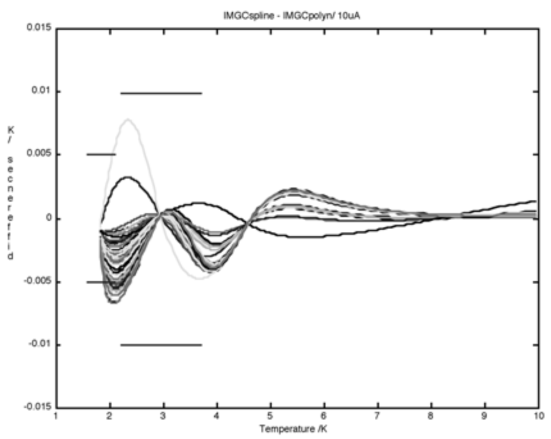


Figure 4. Differences between the calibrations in run #1090 @10 µA using the splines model and the logarithmic polynomial model.

An example of non-conforming Monte Carlo result is reported in Figure 5. The most critical temperature region is always the 2–3 K one, as also in Figure 3 and 4.

[2] Obtained not only from comparison of the calibration functions obtained with procedure A and B, but also by using as calibration points other sets evaluated with different criteria available on the CERN database for the same runs.

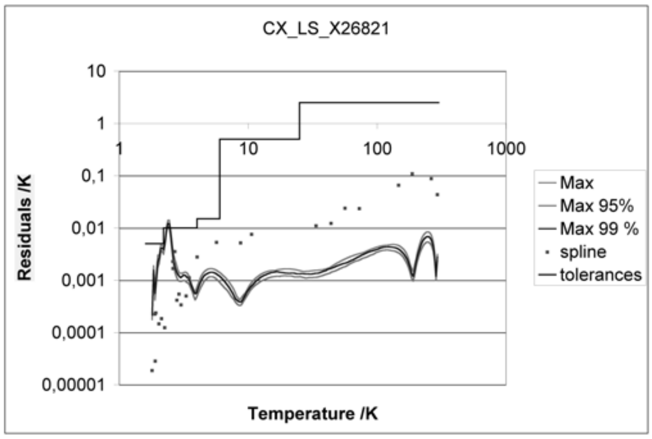


Figure 5. Monte Carlo simulation results (stability of the calibration function on the errors on both variables) for a thermometer to be calibrated.

Finally, a cluster analysis (step 5 in Table 1) detected thermometers with non-typical characteristics for the relevant run (line (9) in table 3). Figure 6 reports the dendogram of the clusters for one run.

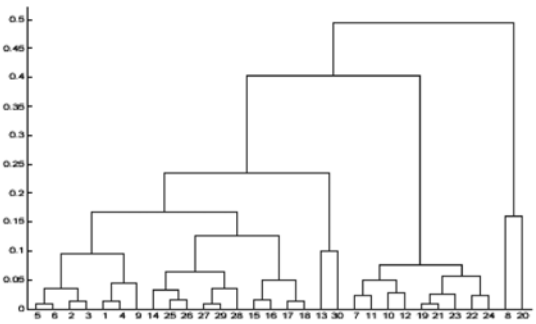


Figure 6. Dendogram of the Cernox® clusters for run #1090.

Figure 7 reports the clusters found when the procedure is repeated on all thermometers of the four runs in table 3. In this case, two large cluster of the same size has been found, with additional 5 thermometers in separate clusters, to be considered real outliers.

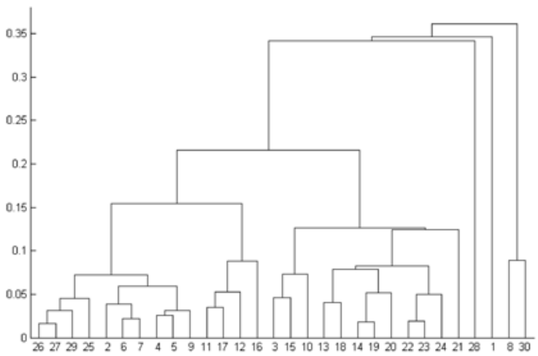


Figure 7. Dendogram of all Cernox® thermometers in all four runs.

As a result of these evaluations, all thermometers showing at least one non-conformity was considered as at risk about its use in the most critical application within the LHC machine, since an accident during the calibration or a non-typical behaviour could be considered as a source of less confidence in the quality and stability with time of those thermometers. The conservative number of first-class thermometers is reported in line (10) of Table 3.

Run #1090 calibrated the same thermometers of run #1170 with three more. It is therefore possible to obtain a real evaluation of the reproducibility of the IPNO calibrations on the same Cernox® thermometers. This is shown in Figure 8. One is expecting to find a reproducibility consistent with the tolerances, namely in the most crucial range below 2.2 K. It is the case for a current of 1 μA, where the maximum deviations are within 3 mK; on the contrary, at 10 μA there are several outlying thermometers (> 5 mK below 2.2 K, see figure) and a sparse set of differences below 3 K. This is a confirmation that there is a difficulty in correctly fitting the sharp bend due to the overheating and that the latter can be quite different from thermometer to thermometer, as already shown in Figure 1, and variable.

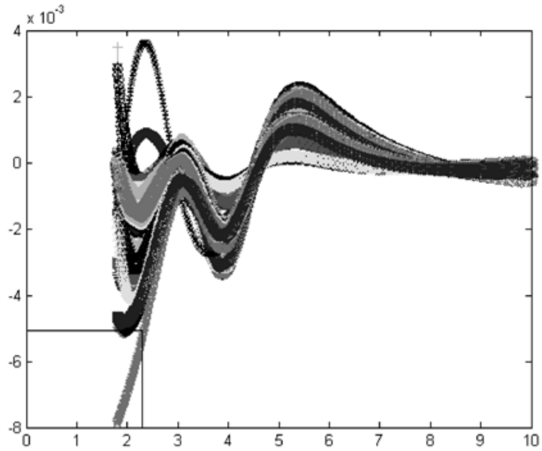


Figure 8. *T*(#1090) – *T*(#1170) for 70 thermometers, @ 10 μA using the splines model.

6. CONCLUSIONS

The results of application of the above methodologies are shown in the paper, compared to the initial uncertainty budget estimated from the thermometer characteristics, the instrumentation specifications and the results of preliminary studies –e.g., for the instrumental noise level.

A relevant variability of the experimental conditions is observed, so that the *a posteriori* analysis allows checking the initial assumptions.

The reproducibility of the measurements have been evaluated by cross-checks with several methods, thanks to the techniques embedded in the experimental procedure designed by IMGC. In the most critical range, the test runs show that not for all runs a high confidence can be attached to the respect of the tolerance, this confidence being generally lower with the calibrations with the 10 μA current, due to the large thermometer overheating (about 100 times higher than at 1 μA). Generally a sizable number of thermometers were found to be invalid or having had in some extent problems during the calibration process, so that their use in the most critical applications should not be advisable.

The accuracy of the temperature values is less important for the LHC functionality. However, a comparison of calibrations of the reference thermometers allowed drawing the conclusion that the values in the most critical range should be correct within 5 mK.

ACKNOWLEGEMENTS

The authors acknowledge the permission given by CERN for the use of the data. The work has been performed under CERN Contract K1030/AT/LHC.

REFERENCES

[1] D.Ichim, F.Pavese, C.Balle, J.Casas-Cubillos: " Tools for quality testing of batches of artifacts: the cryogenic thermometers for the LHC", in "Advanced Mathematical and Computational Tools in Metrology V", Series on Advances in Mathematics for Applied Sciences vol. **57**, World Scientific, Singapore, 2001, 208-212.

[2] F.Pavese, D.Ichim, P.Ciarlini, C.Balle, J.Casas-Cubillos: “Detection of thermometer clustering in the calibration of large batches of industrial thermometers for the LHC by automated data processing”, Temperature, Its Measurement and Control in Science and Industry, D.Ripple ed., Amer.Inst.of Phys., New York A, **8** (2003), 429-433.

[3] F.Pavese, D.Ichim, SAODR: a software tool for online outlier rejection by sequence analysis, Meas. Sci.Technol. **15** (2004), 2047-52.

[4] Ciarlini, P., and Pavese, F., “Computational and statistical analysis of the thermodynamic data which form the basis of the ITS-90 between 13.8 K and 273.16 K – I Computational basis,” in *Temperature: Its Measurement and Control in Science and Industry* **6**, edited by J. F. Schooley, AIP, New York, 1992, pp. 75-78.

[5] F.Pavese, D.Ichim, P.Ciarlini, C.Balle, J.Casas-Cubillos: “Automatic processing of large batches of experimental data for the initial calibration of LHC cryogenic thermometers”, Proc. TEMPMEKO 2001, PTB, Berlin, 653-658.

[6] D. Ichim, F. Pavese, P. Ciarlini: “Experimental design optimization: Case study – the calibration of large batches of cryogenic thermometers”, Proceedings of the International Workshop “From Data Acquisition to Data Processing and Retrieval”, Ljubljana, September 2004, CD-ROM.

[7] D. Ichim, P. Ciarlini, F. Pavese, A. Premoli and M.L. Rastello: “Comparison of LS techniques for the linear approximation of data affected by heteroschedastic errors in both variable with uncertainty estimation” in "Advanced Mathematical and Computational Tools in Metrology VI", Series on Advances in Mathematics for Applied Sciences vol. **66**, World Scientific, Singapore, 2004, 294-299.