Int J Thermophys (2007) 28:1832–1842
DOI 10.1007/s10765-007-0304-x

Overcoming Inhomogeneity and Hysteresis Limitations of Type R Thermocouples in an International Comparison

Ferdouse Jahan · Mark Ballico

Published online: 6 November 2007

Abstract Type R thermocouples are widely used and convenient high-temperature transfer standards; however, the achievable accuracy is limited by the effects of inhomogeneity and hysteresis. In this article, we summarize the results of the recent international comparison APMP-T-S1-04 and discuss the results of the thermoelectric scanning, spatially resolved over the length of the thermocouples. The thermoelectric signatures show both reversible (hysteresis) and irreversible inhomogeneities introduced by the calibration processes used by the participants. The results demonstrate that although the reversible hysteresis of Type R thermocouples limits their performance as a transfer standard in thermometry, this can be managed by appropriate design of the comparison protocol. By performing all calibrations from lower to higher temperatures from an initial 450°C anneal state, a pilot laboratory reproducibility of typically 0.03°C ($k = 2$) and a reference value uncertainty of 0.03–0.06°C (at $k = 2$) over 0–1,100°C were achieved. This allowed statistically significant testing of the calibration capabilities of all the participants.

Keywords Annealing state · Inhomogeneity · Intercomparison · Stability · Type R thermocouple

F. Jahan (✉) · M. Ballico
National Measurement Institute, Bradfield Road, West Lindfield, P.O. Box 264, Lindfield, NSW 2070, Australia
e-mail: Ferdouse.Jahan@nmi.gov.au

M. Ballico (✉)
e-mail: Mark.Ballico@nmi.gov.au

1 Introduction

The platinum–rhodium thermocouple, Type R (Pt-Pt13%Rh) or S (Pt-Pt10%Rh), is a precision temperature measuring sensor that is widely used in many industries to measure temperatures as high as 1,600°C. Many calibration laboratories also use this thermocouple as a standard to calibrate other industrial thermocouples. These thermocouples can be calibrated by using ITS-90 defined metal fixed points (Ga, Sn, Zn, Al, and Ag), a comparison method using a calibrated standard sensor (SPRT, or Type S or R thermocouple), and a variable temperature enclosure, or a combination of both.

In recent years, the mutual recognition arrangement (MRA) coordinated by the BIPM has been developed to ensure mutual recognition of calibration and measurement certificates issued by national metrology institutes (NMIs), and this is supported by a series of international key comparisons. In thermometry, in the temperature range of 0–962°C, the CCT-K3 [1] and CCT-K4 [2] comparisons used SPRTs and fixed points as transfer standards. However, many NMIs use Pt–Rh thermocouples as the predominant transfer artifact of the temperature scale to industry, and the dominant uncertainties in the calibration process relate to techniques not evaluated by the CCT-K3 and CCT-K4 comparisons.

In 2005, the National Measurement Institute of Australia (Pilot Lab) organized a regional intercomparison of the Type R thermocouple calibration from 0 to 1,100°C for national laboratories in the Asia/Pacific region. The objective of this APMP intercomparison is to assess the equivalence of the calibration results obtained by the various procedures and methods of calibration of rare metal thermocouple up to 1,100°C. Twelve laboratories of the Asia Pacific region, including NMIA, Australia, took part: NIM—China, SCL—Hong Kong, NPLI—India, KIMLIPI—Indonesia, NMIJ—Japan, KRISS—Korea, SIRIM—Malaysia, SPRING—Singapore, CSIR—South Africa, CMS—Taiwan, and NIMT—Thailand.

The stability or drift of the artifacts is one of the important criteria for a successful intercomparison. It is well known that Pt–Rh thermocouples suffer reversible hysteresis due to Rh oxidation and lattice defects at high temperatures, and thus, in use, they develop significant inhomogeneity, typically up to 0.04% [3], which limits their performance. However, the protocol of this comparison was designed in such a way that it overcame this limitation, and thus achieved a low comparison uncertainty, allowing the statistically significant testing of the calibration capabilities of all the participants.

2 Comparison Process

The artifacts used in this intercomparison were 11 Type R thermocouples, serial numbers APMP-01 to APMP-11. Wires from the same reel of 0.5 mm diameter of platinum and platinum-13% rhodium wires (Reference Grade), purchased from Sigmund Cohn Corp. (USA) were used for the thermocouples. The insulators used were high-purity alumina (99.8%), purchased from Ceramic Oxide Fabricators (Australia) and baked at 1,100°C for 6 h. The thermocouple wires were bare-wire annealed at 1,400°C for 1 h and at 1,100°C for 1 h. After being assembled into the prebaked insulator, a further 1h

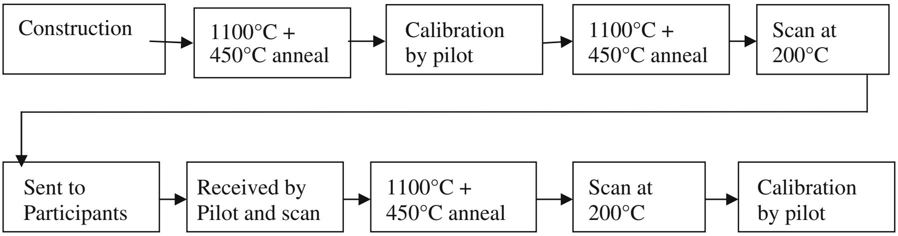


Fig. 1 Annealing and measurement sequence of the thermocouples

anneal at 1,100°C and 16 h at 450°C was applied. The thermocouple wires emerging from the alumina tube were insulated with PVC sleeves.

The inhomogeneities of the thermocouples were measured by scanning in an oil bath at 200°C [4], and then calibrated by the pilot lab by comparison with a standard platinum resistance thermometer (SPRT) in a salt bath up to 550°C followed by Ag and Au fixed points [5]. After calibration, the thermocouples were given 1 h anneals at 1,100°C and 16 h anneals at 450°C to bring them to the same '450°C annealed state' and to anneal out any inhomogeneity imposed by the calibration (see Sect. 5). They were then sent to the participating laboratories—one thermocouple for each lab.

The calibration procedures of different labs varied widely. Some used fixed-point calibration only, others used furnace calibration against an SPRT at lower temperatures and against a Type R thermocouple at higher temperatures, and some used a combination of both [6]. The participating labs performed the calibration of the thermocouple using their own test method—however, participants were asked to ensure that the calibration would be from lower to higher temperatures. No annealing of the thermocouple was done by the participating lab. The participating labs sent the calibration results, $E - E_{\text{ref}}$ (where E is the measured EMF of the thermocouple and E_{ref} is the reference EMF at the same temperature), and calculated uncertainty at 100°C steps from 0 to 1,100°C to the pilot lab. The participants provided their estimated calibration uncertainties using the inhomogeneity value provided by the pilot lab, with no allowance for any short-term drift.

After receiving the thermocouples back from the participating labs, they were scanned in an oil bath 'as received' to determine the change in the thermoelectric signature of the thermocouple due to the calibration by the participants. Then, they were again annealed at 1,100°C for 1 h and at 450°C for 16 h to bring the thermocouples to the same 450°C annealed state. All thermocouples were scanned again in an oil bath at 200°C, to confirm that the participant-induced inhomogeneity was annealed out, and then calibrated by the pilot lab. The measurement sequence and the annealing of a particular thermocouple are shown in Fig. 1.

3 Analysis of Comparison Results

3.1 Comparison Data Analysis

All thermocouples were calibrated twice by —initially, before sending to the participating labs and finally, after they came back. Both calibrations were done at the same

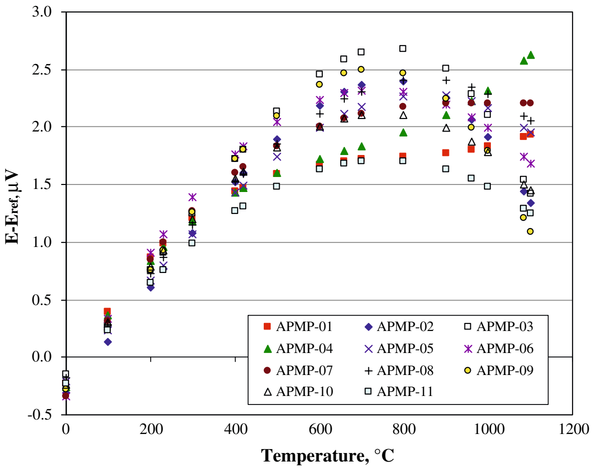


Fig. 2 Initial calibration results of 11 thermocouples used for the intercomparison

annealing state and from lower-to-higher temperatures. The initial calibration data of all thermocouples is shown in Fig. 2. The values of $E - E_{\text{ref}}$ of all thermocouples were within $\pm 0.5\,\mu\text{V}$ of the mean (standard deviation $\leq 0.2\,\mu\text{V}$) for temperatures up to 1,000°C. Above 1,000°C, the standard deviation is slightly larger, about $0.4\,\mu\text{V}$. These results are consistent with the typical measured inhomogeneities within each thermocouple (0.008–0.01%), and the fact that all thermocouples were made in the same way and from the same reels of wire. In the comparison data analysis, the calibration of each thermocouple was considered individually. The calibration results, $E - E_{\text{ref}}$ of a particular thermocouple, given by a participant are compared with the calibration result of that thermocouple by the pilot lab at 100°C steps from 0 to 1,100°C. The pilot lab results are taken as the average of the initial and final calibrations.

The differences between the participating lab results and the pilot lab results are plotted in Fig. 3 at three out of the 17 comparison temperatures [6], where

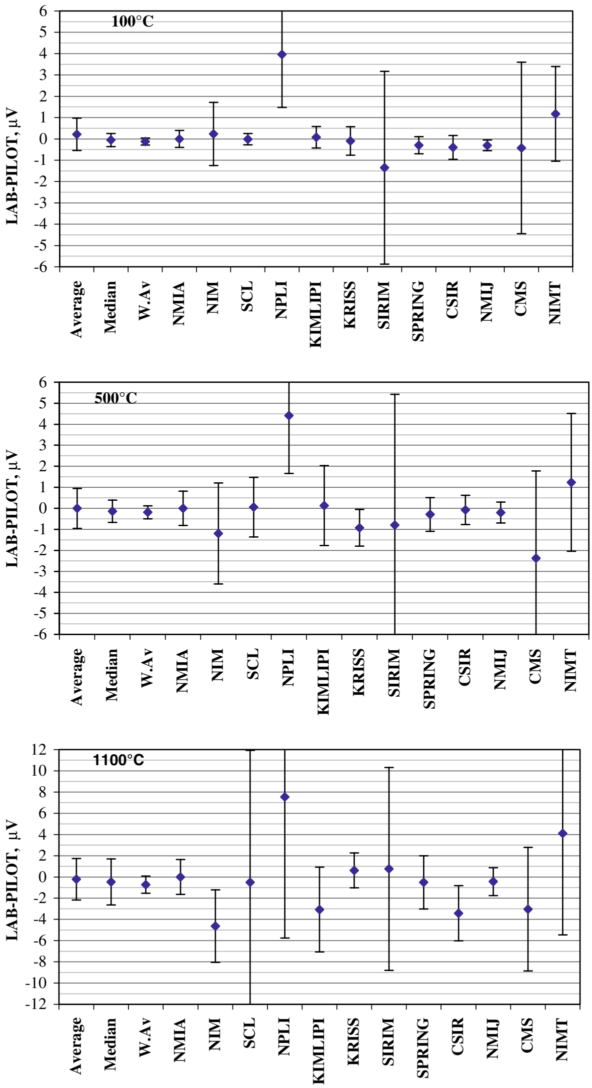
$$X_i = (E - E_{\text{ref}})_{\text{lab}} - (E - E_{\text{ref}})_{\text{pilot}} = E_{\text{lab}} - E_{\text{pilot}} \tag{1}$$

For the pilot lab, the average of its 11 initial uncertainties is plotted.

3.2 Pilot Lab Reproducibility and Drift of the Artifacts

In the analysis of the comparison results, it is important to know the reproducibility of the pilot lab measurements, as the participating labs are linked to the reference value via the pilot lab. The pilot lab calibration uncertainty [5] includes both the reproducibility of the pilot lab measurements and the systematic errors of the pilot laboratory scale. A calculation of the pilot lab reproducibility would necessitate accurate knowledge of the correlation between terms. A more direct approach is used, instead, to estimate the pilot lab reproducibility and any drifts in each thermocouple by using data from 22 calibrations performed by the pilot lab.

Fig. 3 Difference of $E - E_{ref}$ between the pilot laboratory and 11 participants together with the calculated values of simple mean, weighted mean, and the median. Uncertainties (at $k = 2$) for each lab are as reported by each participant. For NMIA, the average of its 11 reported initial uncertainties is plotted



The difference, $drift_i$, between initial and final calibrations of the 11 thermocouples by the pilot lab is shown in Fig. 4. The initial and final calibrations of most thermocouples agree to within $\pm 0.4\,\mu V$, except for one thermocouple, APMP-04. For this thermocouple, after the initial calibration, the insulator broke during packing, and it was replaced by another insulator of unknown origin. The replacement of the insulator may have some impact on the final Au-point measurement and, hence, affects the results above 1,000°C.

The measured difference between the initial and final calibrations plotted in Fig. 3 differs from zero because of (a) drifts of the thermocouple and (b) the reproducibility of the pilot lab measurements: the spread of values is due to both these effects and it is

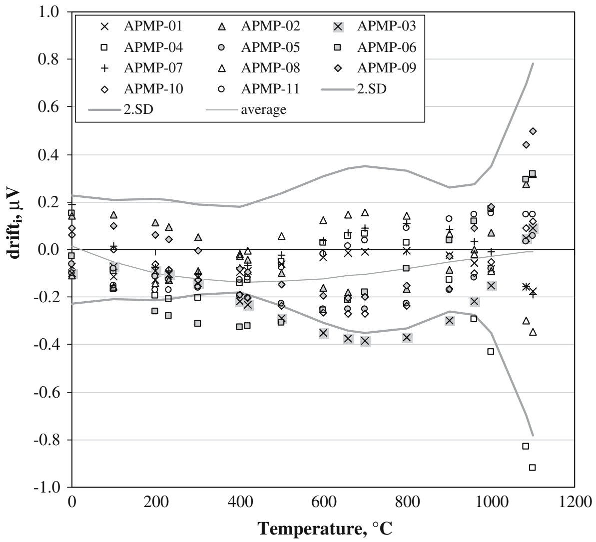


Fig. 4 Difference, *drift$_i$*, between the initial and final calibrations of 11 thermocouples by the pilot lab, plotted as a function of temperature

not possible to separate them. If drifts did not contribute, the pilot lab reproducibility would be given by

$$u_{\text{pilot·reprod}} = \underset{\text{(all but APMP04)}}{STDEV} \quad [drift_i]/\sqrt{2} \tag{2}$$

The pilot lab reproducibility is 0.03°C ($k = 2$) over the range up to 1,100°C, calculated using Eq. 2 and data from Fig. 3. This will be a slight overestimate if drifts are significant. When we plot a $k = 2$ envelope for this term in Fig. 4, data for several thermocouples fall just outside this envelope at some of the temperature points. One of these thermocouples, APMP-03, is known to have experienced damage, so a real drift is likely. In order not to underestimate the artifact stability in such cases, the drift term is retained for all the thermocouples as an uncertainty component, even though it will result in a slight uncertainty overestimate in some cases.

4 Calculation of a Reference Value

A comparison reference value was calculated by three different methods: the simple mean, the median, and the weighted mean, as given by the three following equations:

(i) Simple Mean:

$$X_{\text{simple}} = \sum X_i/n \tag{3}$$

$$u(X_{\text{simple}}) = STDEV(X_i)/\sqrt{n} \tag{4}$$

(ii) Median: Computed using the MEDIAN function of Microsoft Excel. The uncertainty was calculated using the equation given in Ref. [7]

$$X_{\text{median}} = median\{x_i\} \tag{5}$$

$$u(X_{\text{median}}) \cong \frac{1.9}{\sqrt{n-1}} median\{\,|X_{\text{median}} - X_i\,|\} \tag{6}$$

(iii) Weighted mean:

$$X_{\text{weighted}} = \sum X_i \cdot u(X_i)^{-2} / \sum u(X_i)^{-2} \tag{7}$$

$$u(X_{\text{weighted}})^2 = 1/\sum u(X_i)^{-2} \tag{8}$$

$$\text{where} \quad u(X_i)^2 = u(E_{\text{lab},i})^2 + (drift_i/2\sqrt{3})^2 + (u_{\text{pilot·reprod}})^2 \tag{9}$$

In the cases of the simple mean and median, each of the laboratories contributes equally to the calculation of the reference value; that is all values of the X_i are given equal weight. The uncertainities for the simple mean and median provide an estimate of the combined effect of lab uncertainties, thermocouple drift, and pilot lab reproducibility, but note that the lab uncertainty $u(X_i)$ does not contribute to these two reference values. These measured variances are statistically sound because the X_is are fully uncorrelated: each $X_i = E_{\text{lab}} - E_{\text{pilot}}$ arose from separate calibrations of a unique thermocouple by both the pilot and participating labs. Any scale uncertainty of the pilot lab is obviously fully correlated between all the X_is, and so cancels out.

For the weighted mean, the contributions from each lab are weighted according to the estimated uncertainty of their contribution to the mean. For the pilot lab, this is just their calibration uncertainty. For the other participants, the uncertainty due to the link to the pilot should be included—they are: (i) drift of the thermocouple, *drift*$_i$, and the measured difference between the initial and final calibration of the particular thermocouple, and (ii) the reproducibility of the pilot lab calibration given by Eq. 2. All three reference values are consistent with one another; however, the weighted mean offers the lowest uncertainty, as the weighted mean includes the extra data inherited from the lab uncertainty estimates.

However, before the weighted mean can be used, the internal self-consistency of the uncertainty estimates provided by the participants must be checked. The Birge ratio is a measure of how well the estimated measurement uncertainties explain the measured dispersion of the actual data values. The Birge ratio and the statistical criterion for internal self-consistency for *n* laboratories [8] are given by,

$$\text{Birge Ratio} = \sqrt{\left[\sum (X_i - X_{\text{weighted}})^2 u(X_i)^{-2}/(n-1)\right]} < \sqrt{[1 + \sqrt{8/(n-1)}]} \tag{10}$$

In this comparison, the Birge ratio was between 0.40 and 1.08, which is well below the statistical criterion of 1.38 for 11 laboratories. As the Birge criterion was satisfied and it offered the lowest uncertainty, the *weighted mean* was used *as the reference value* for the comparison [6].

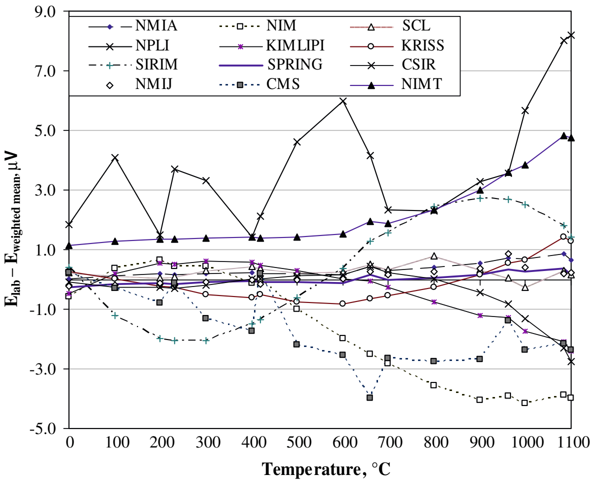


Fig. 5 Difference between the lab value and the comparison reference value (weighted mean) for each participant is plotted as a function of temperature

The differences between the participating lab value X_i and the comparison reference value $X_{weighted}$ are plotted in Fig. 5 as a function of temperature. The figure shows that most of the labs are within $\pm 1.0\,\mu V$, which is equivalent to 0.1°C at 1,100°C. The uncertainty of the reference value calculated as the weighted mean ranges from 0.14 to 0.76 µV ($k = 2$), equivalent to 0.03–0.06°C over the temperature range of 0 to 1,100°C. The low comparison uncertainty together with the good reproducibility of the pilot lab, typically 0.03°C ($k = 2$, as calculated by Eq. 2), allowed this intercomparison to be useful for statistically significant testing of the lowest calibration uncertainties of the participants.

5 Annealing State and Reversible Hysteresis

It is well known that Type R and S thermocouples suffer hysteresis during use, mainly due to the Pt–Rh alloy thermoelement. The Seebeck coefficient of the thermocouple wires changes as they experience high temperatures; as a result, inhomogeneity grows in the wires. The formation of rhodium oxide in the Pt–Rh leg of the thermocouple in the temperature range of 600–800°C decreases the Seebeck coefficient. At higher temperatures (above 800°C), rhodium oxides dissociates, and the various lattice vacancies and other defects establish higher equilibrium levels. If a thermocouple is "quenched" by quickly removing it from a furnace above 800°C, this higher defect concentration is "trapped," lowering the Seebeck coefficient. If then used at temperatures above about 400°C (where atomic mobility is higher), short range reordering can occur and the Seebeck coefficient will slowly increase [3]. Although both these processes are reversible, they are unavoidable. This reversible hysteresis limits the performance of Type R or S thermocouples as transfer standard artifacts. NMIs generally calibrate thermocouples in the "quenched" or "450°C annealed" state, but there is presently no international agreement.

This reversible hysteresis is a fundamental limitation of the performance of Type R and S thermocouples. Although the intercomparison protocol was designed to minimize the effects of hysteresis, its effects could be clearly seen in the inhomogeneity measurements of the thermocouples at the different stages of the comparison. The initial scan in the 450°C annealed state and the scan after calibration by the participants are plotted in Fig. 6A, B, and C for three of the thermocouples.

In Fig. 6A, the EMF along the length of wire heated by the participant (tip to 380 mm) decreased by about 0.9 μV at 200°C, compared to the unheated wire past 500 mm. This participant used both fixed points and a comparison furnace, and the overlapping effects led to the small "peak" observed at 400–450 mm. The heat treatment experienced during calibration increased the inhomogeneity (over 550 mm of length) from ±0.008% to 0.034%, which is equivalent to 3.6 μV ∼ 0.3°C at 1,000°C. However, after an 1,100°C anneal for 1 h and 16 h at 450°C, the inhomogeneity value returned to the initial value of ±0.008%.

The 'as received' scan in Fig. 6B showed a similar change, but for a shorter length than that in Fig. 6A. The high-temperature effect can be seen over a length of 250 mm from the tip. This particular thermocouple was calibrated in a furnace up to 1,000°C, against a Type R standard thermocouple, at a shorter immersion and the emf changed by 0.6 μV for the calibrated section. The inhomogeneity of this thermocouple increased from 0.008% to 0.027% (less than Fig. 6A) after calibration and returned to the initial value after annealing. All thermocouples showed this similar nature. Most of the changes of the thermoelectric signature of Type R thermocouples due to exposure to high temperatures up to 1,100°C are reversible. The final inhomogeneity values of most of the thermocouples, measured after annealing, are close to the initial values.

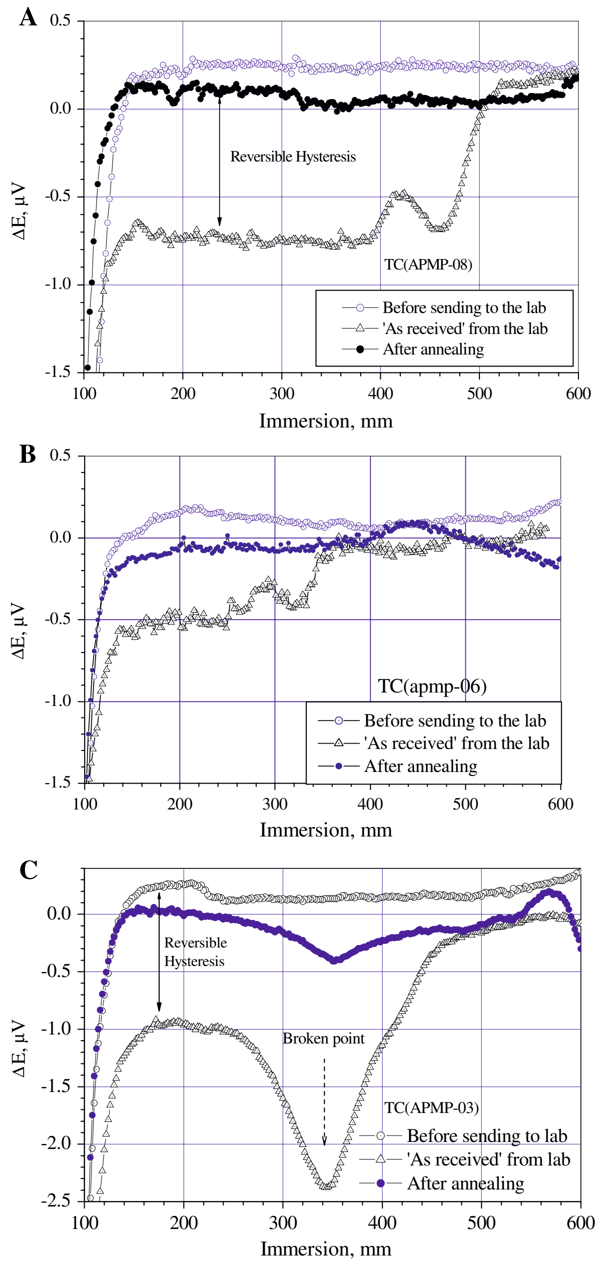
However, the final inhomogeneity values increased for two of the 11 thermocouples, one of which is shown in Fig. 6C. The insulators of this thermocouple broke during transport and the lab did the calibration with the broken insulator. During calibration, the wires were extensively cold worked, as evident from the 'as received' scan shown in Fig. 6B. The 'as received' inhomogeneity of this thermocouple increased to ±0.08% compared to the initial inhomogeneity of ±0.010%. The cold work, induced by bending of the wire, affects the Seebeck coefficient and, hence, the emf, which is not fully recovered by the 1,100°C anneal. As the initial and final calibration by the pilot lab and the calibration by the other participants were carried out with the thermocouples in the same "450°C annealed state," this reversible hysteresis did not affect the comparison process.

6 Conclusion

This comparison demonstrated that Type R or S thermocouples can be used to transfer a calibration with an accuracy of 0.06°C, provided the annealed state of the thermocouples and the sequence of calibration temperatures are known and controlled.

However, although the Type R or S thermocouple can be calibrated with this level of uncertainty, because of the reversible hysteresis due to rhodium oxidation and trapped vacancy concentration at high temperatures, there can be major systematic errors (±0.3°C at 1,000°C), even with short term use of these thermocouples. Before

Fig. 6 Scan of three APMP thermocouples (**A**) TC (APMP-08), (**B**) TC (APMP-06), and (**C**) TC (APMP-03); (i) –○– 450°C annealed state, (ii) –Δ– 'As received' from the lab after calibration, and (iii) –●– after annealing at 1,100 and 450°C



measurement, these thermocouples should always be put in the same annealing state as when calibrated in order to attain the best uncertainties.

References

1. B.W. Mangum, G.F. Strouse, W.F. Guthrie, R. Pello, M. Stock, E. Renaot, Y. Hermier, G. Bonnier, P. Marcarino, K.S. Gam, K.H. Kang, Y.-G. Kim, J.V. Nicholas, D.R. White, T.D. Dransfield, Y. Duan, Y. Qu, J. Connolly, R.L. Rusby, J. Gray, G.J.M. Sutton, D.I. Head, K.D. Hill, A. Steele, K. Nara, E. Tegeler, U. Noatsch, D. Heyer, B. Fellmuth, B. Thiele-Krivoj, S. Duris, A.I. Pokhodun, N.P. Moiseeva, A.G. Ivanova, M.J. de Groot, J.F. Dubbeldam, Metrologia **39**, 179 (2002)
2. H.G. Nubbemeyer, J. Fischer, Metrologia **39**, Tech. Supp. 03001 (2002)
3. R.E. Bentley, T.P. Jones, High Temp.–High Press. **12**, 33 (1980)
4. F. Jahan, M.J. Ballico, in *Temperature: Its Measurement and Control in Science and Industry*, vol 7, Part I, ed. by D.C. Ripple (AIP, New York, 2002), pp. 469–474
5. F. Jahan, *NMI Quality System PM—EADA—8.2.25* (National Measurement Institute, Australia, 2001).
6. F. Jahan, M.J. Ballico, APMP T-S1, Metrologia **44**, Tech. Supp. 03004 (2007)
7. E. Andras, *EUROMET Project No. 635- Final Report* (National Office of Measures, Budapest, 2003)
8. R. Kacker, R. Datla, A. Parr, Metrologia **39**, 279 (2002)