Int J Thermophys (2008) 29:1828–1837
DOI 10.1007/s10765-007-0354-0

# A Regional Comparison of Calibration Results for Type K Thermocouple Wire from (100 to 1,100) °C

**K. M. Garrity · D. C. Ripple · M. Araya · C. R. Cabrera · L. Cordova Murillo · M. E. de Vanegas · D. J. Gee · E. Guillén · S. Martinez-Martinez · E. Mendez-Lango · L. Mussio · S. G. Petkovic · K. N. Quelhas · G. Rangugni · O. Robatto · E. von Borries Rocha**

Published online: 5 June 2008


**Abstract** Under the auspices of the Inter-American Metrology System (SIM), the National Institute of Standards and Technology (NIST) initiated a regional comparison for type K thermocouples from (100 to 1,100) °C with 11 participating countries. The use of type K material above approximately 200 °C is considered destructive. Therefore, each participating laboratory was sent new, unused wire from a lot of material characterized by NIST. The uniformity of the lot was remarkable, especially at temperatures above 500 °C; the standard deviation of the thermocouple emf values of multiple cuts tested at NIST was 2.7 μV or less over the full temperature range. The high uniformity eliminated any need to correct for variations of the transfer standard among the laboratories, greatly simplifying the analysis. The level of agreement among the laboratories' results was quite good. Even though test procedures and equipment varied significantly among the participants, the standard deviation of all emf values at each test temperature was less than the equivalent of 0.20 °C at 200 °C and below, and less than 0.60 °C from (400 to 1,100) °C. Of the 380 total bilateral combinations of the

K. M. Garrity (✉) · D. C. Ripple
Pilot Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA
e-mail: karen.garrity@nist.gov

M. Araya
Laboratorio Custodio de los Patrones Nacionales de Temperatura,
Red Nacional de Metrología, Santiago, Chile

C. R. Cabrera
Servicio Autónomo Nacional de Normalización, Calidad, Metrología y Reglamentos Técnicos,
Caracas, Venezuela

L. Cordova Murillo · E. von Borries Rocha
Instituto Boliviano de Metrología,
La Paz, Bolivia

M. E. de Vanegas
Consejo Nacional de Ciencia y Tecnología, San Salvador, EI Salvador

data at the eight test temperatures, only 13 (i.e., 3.4% of all combinations) are outside the $k = 2$ limits, and of these 13, only 3 are outside $k = 3$ limits. All the outliers occur at temperatures of 800 °C and below, which suggests that drift of the type K wire due to high-temperature oxidation did not cause changes in the thermocouple emf comparable to or larger than the claimed uncertainties.

**Keywords** Comparison · ITS-90 · Temperature · Thermocouple · Type K

## 1 Description of the Comparison

Type K thermocouples are one of the most commonly used temperature sensors in industry. The skills, personnel, and facilities necessary for calibrating type K thermocouples are also applicable to the calibration of other base-metal thermocouples, and, to a lesser extent, calibration of platinum–rhodium alloy thermocouples. In 2004, NIST initiated a Supplementary Comparison for Type K Thermocouples from (100 to 1,100) °C, inviting all member laboratories of the Inter-American Metrology System (SIM) to participate.

The testing of type K material above approximately 200 °C is considered destructive. Therefore, each participating laboratory was sent new, unheated wire from the lot of material characterized by NIST. The wires were shipped to the participants in a coil of radius similar to the coil of the originating lot to prevent significant mechanical strain. The participating laboratories were asked to perform testing in the same manner as they normally calibrated thermocouples.

Samples were sent to a total of 11 participants listed in Table 1, including NIST (the pilot laboratory). The participating laboratories performed the measurements in the period March 1, 2004 to September 22, 2004. Table 2 contains a summary of the calibration methods used by the 11 participating laboratories.

D. J. Gee
National Research Council of Canada, Ottawa,
Canada

E. Guillén
Instituto Nacional de Defensa de la Competencia y de la Protección de la
Propiedad Intelectual, Servicio Nacional de Metrología, Lima,
Peru

S. Martinez-Martinez · E. Mendez-Lango
Centro Nacional de Metrología, Queretaro,
Mexico

L. Mussio · O. Robatto
Laboratorio Tecnológico del Uruguay, Montevideo,
Uruguay

S. G. Petkovic · K. N. Quelhas
Instituto Nacional de Metrología, Normalização e Qualidade Industrial, Rio de Janeiro,
Brazil

G. Rangugni
Instituto Nacional de Tecnología Industrial, Buenos Aires,
Argentina

**Table 1** List of laboratories participating in the comparison

| Laboratory code | Laboratory acronym | Country | Laboratory name |
|---|---|---|---|
| A | CENAM | Mexico | Centro Nacional de Metrología |
| B | CONACYT | El Salvador | Consejo Nacional de Ciencia y Tecnología |
| C | IBMETRO | Bolivia | Instituto Boliviano de Metrología |
| D | INMETRO | Brazil | Instituto Nacional de Metrología, Normalização e Qualidade Industrial, Rio de Janeiro |
| E | INTI | Argentina | Instituto Nacional de Tecnología Industrial |
| F | LATU | Uruguay | Laboratorio Tecnológico del Uruguay |
| G | LCPNT | Chile | Laboratorio Custodio de los Patrones Nacionales de Temperatura (Red Nacional de Metrología) |
| H | NIST | United States | National Institute of Standards and Technology |
| I | NRC | Canada | National Research Council of Canada |
| J | SENCAMER | Venezuela | Servicio Autónomo Nacional de Normalización, Calidad, Metrología y Reglamentos Técnicos |
| K | SNM - IN-DECOPI | Peru | Instituto Nacional de Defensa de la Competencia y de la Protección de la Propiedad Intelectual, Servicio Nacional de Metrología |

## 2 Characterization of the Transfer Standard

The NIST Thermometry Group acquired 60 m of type K, uninsulated, 1.63 mm diameter (14-gauge) thermocouple wire. The wire was then cut into 1.1 m length. To evaluate the thermoelectric inhomogeneity and the average emf versus temperature response of the wire, NIST calibrated selected cuts from the lot by comparison to a calibrated type S thermocouple in two different furnaces, and to a standard platinum resistance thermometer in stirred liquid baths [1,2].

For each of the three sets of thermocouples calibrated in the three apparatuses, the standard deviations of the emf readings at each temperature were calculated, as seen in Table 3, and the results were pooled ($s$ in Table 3) to obtain the Type A uncertainties of the NIST measurements. This component of uncertainty includes both calibration repeatability and thermoelectric inhomogeneity of the tested wire lot and may be taken as an upper limit on the standard uncertainty ($k = 1$) due to wire inhomogeneity, $u_{\mathrm{I}}$. The uniformity of the lot was remarkable, especially at temperatures above 500 °C. No trends were observed in the emf of one end of the lot versus the other end, and no outliers were seen. The participants were not informed of the high degree of lot uniformity prior to the comparison.

## 3 Analysis of Bilateral Differences

The measurement uncertainties for the participating laboratories were obtained from the survey results. To simplify the analysis, the emf values for the two cuts calibrated

**Table 2** Brief description of the methods used by participants to calibrate the wire samples throughout the temperature range

| Laboratory acronym | CENAM | CONACYT | IBMETRO | INMETRO | INTI | LATU |
|---|---|---|---|---|---|---|
| Low-temperature bath/furnace type | 3-Zone furn. | Oil bath | Oil bath | N/A[a] | N/A | Oil bath; block |
| Temp. range (°C) | 100–600 | 100–200 | 100–200 | N/A | N/A | 100–400 |
| Reference | Type S TC | PRT | SPRT | N/A | N/A | SPRT |
| High-temperature furnace type | Heat pipe | Not stated | Not stated | Single zone | Single zone | Single zone |
| Temp. range (°C) | 800–1,000 | 300–600 | 400–1,100 | 100–1,000 | 100–1,100 | 500–1,000 |
| Reference | Type S TC | Type S TC | Type S TC | Type S TC | Type S TC | Type S TC |
| **Laboratory acronym** | **LCPNT** | **NIST** | **NRC** | **SENCAMER** | **SNM-INDECOPI** | |
| Low-temperature bath/furnace type | N/A | Oil, salt baths | Oil, salt baths | Oil bath | Oil bath | |
| Temp. range (°C) | 100–200 | 100–500 | 100–500 | 100–200 | 100–200 | |
| Reference | SPRT | SPRT | SPRT | Type S TC | SPRT | |
| High-temperature furnace type | Single zone | Single zone | 3 Zone | Single zone | Not stated | |
| Temp. range (°C) | 300–1,100 | 600–1,100 | 600–1,100 | 400–1,100 | 300–1,000 | |
| Reference | Type S TC | Type S TC | Type S TC | Type S TC | Type S TC | |

[a] N/A = not applicable

**Table 3** Thermocouple inhomogeneity and repeatability; $s$: standard deviation; *df*: degrees of freedom

| | High-temp. furnace | | Tube furnace | | Stirred baths + SPRT | | Pooled | |
|---|---|---|---|---|---|---|---|---|
| Temperature (°C) | $s$ (μV) | *df* | $s$ (μV) | *df* | $s$ (μV) | *df* | $s$ (μV) | *df* |
| 100 | 0.96 | 4 | 1.59 | 2 | 0.21 | 1 | 1.1 | 7 |
| 200 | 1.91 | 4 | 1.75 | 2 | 1.31 | 1 | 1.8 | 7 |
| 400 | 2.81 | 4 | 2.67 | 2 | 2.44 | 1 | 2.7 | 7 |
| 500 | 3.19 | 4 | 0.43 | 2 | 2.86 | 1 | 2.7 | 7 |
| 600 | 2.89 | 4 | 0.83 | 2 | | | 2.4 | 6 |
| 800 | 1.84 | 4 | 1.79 | 2 | | | 1.8 | 6 |
| 1,000 | 2.16 | 4 | 1.03 | 2 | | | 1.9 | 6 |
| 1,100 | 2.48 | 4 | 1.17 | 2 | | | 2.1 | 6 |

by each laboratory were averaged. Upon taking the average, any run-to-run variance in the calibration results will be reduced, due to the statistical averaging of the two samples. The combined uncertainty for the bilateral comparison of the two laboratories is calculated as

$$u_\mathrm{c} = \left[u_{\mathrm{S,A}}^2 + \left(u_{\mathrm{I,A}}^2 + u_{\mathrm{R,A}}^2\right)/n_\mathrm{A} + u_{\mathrm{S,B}}^2 + \left(u_{\mathrm{I,B}}^2 + u_{\mathrm{R,B}}^2\right)/n_\mathrm{B}\right]^{1/2}, \tag{1}$$

where $n_\mathrm{A}$ and $n_\mathrm{B}$ are the number of calibration runs conducted by laboratories A and B; $u_{\mathrm{I,A}}$ and $u_{\mathrm{I,B}}$ account for thermocouple inhomogeneity; $u_{\mathrm{R,A}}$ and $u_{\mathrm{R,B}}$ are the standard uncertainties attributed by laboratories A and B to effects that are random from run to run; and $u_{\mathrm{S,A}}$ and $u_{\mathrm{S,B}}$ are the standard uncertainties attributed to systematic effects. The term $u_{\mathrm{I,A}}$ or $u_{\mathrm{I,B}}$ is set equal to $u_\mathrm{I}$ if a laboratory did not include thermocouple inhomogeneity as an uncertainty component; otherwise, the term is set to zero. All calculations were performed with a coverage factor $k$ of two. No attempt was made to calculate uncertainties with a confidence limit of 95%.

The measured emf values reported for the two or three wire cuts calibrated by laboratory $i$ at a nominal test temperature were averaged to obtain the quantity $E_{\mathrm{a,i}}$. The bilateral difference between laboratories $i$ and $j$ is defined as $D_{ij} \equiv (E_{\mathrm{a},i} - E_{\mathrm{a},j})/S(t)$, where $S(t)$ is the Seebeck coefficient at the nominal test temperature $t$. Because of their length, tables of the bilateral differences and associated uncertainties are omitted from the present paper and are given only in the final report for the comparison [3]. Section 5 gives a statistical summary of the results.

## 4 Choice of Reference Value

As shown in Fig. 1, three candidate reference values were calculated from the reported data: the weighted mean, the median, and the mean. At temperatures of 400 °C and 500 °C and below, the calculation of the weighted mean heavily influences the results of LCPNT, yet these results also are possible outliers. If these results are omitted from the calculation of the weighted mean, the revised weighted mean shifts from the original calculation by an amount well in excess of the $k = 2$ statistical expanded uncertainty. To a lesser extent, the same difficulty arises with the results of CONACYT at (100 and 200) °C. Because of this difficulty, the weighted mean is deemed to be a flawed reference value. Of the two other candidates, the median is chosen as the reference value because it is insensitive to outliers. Strictly speaking, we use the term median in this paper to denote the median of an assumed probability distribution, which was calculated by assuming that each reported result at a given temperature can be represented by a normal distribution, centered on the mean emf value of the two calibrated lot samples with a scale parameter equal to the standard uncertainty ($k = 1$) reported by the laboratory. The probability distributions of all laboratories were summed numerically, and the 50% point of the combined distribution was taken as the median. The uncertainty of the median was calculated using the approximate formula:

$$u(E_\mathrm{med}) = 1.253 s_\mathrm{m}, \tag{2}$$

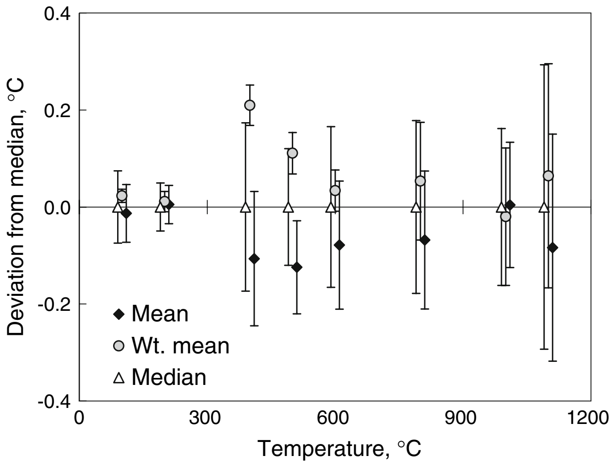


**Fig. 1** Several candidate reference values, plotted as a function of temperature. The bars indicate standard ($k = 1$) uncertainties

where the two or three emf values from each laboratory at a given temperature are first averaged, and then the standard deviation of the mean, $s_{\mathrm{m}}$, is determined from that population.

Two additional checks were made to examine the internal consistency of the data. First, the internal consistency of the calibrations within a laboratory was assessed using the method of Youden plots. Second, a third-order polynomial was used to fit the emf deviation from the reference function for each of the laboratories. The residuals of each laboratory's data were compared to the residual plots for other laboratories to look for possible anomalous patterns. These procedures identified several possible outliers; removal of these possible outliers from the calculation of the median (but not from the average emf value for each laboratory) had a negligible effect on the results presented in the figures. For all the figures, all data are included in the calculation of the median.

Figures 2 and 3 present the comparison data for each laboratory graphically, using the median simply as a baseline. In the figures, the uncertainty for laboratory $i$ is calculated from the terms of Eq. 1 applicable to laboratory $i$:

$$u_{\mathrm{c}} = \left[u_{\mathrm{S},i}^2 + \left(u_{\mathrm{I,i}}^2 + u_{\mathrm{R},i}^2\right)/n_i\right]^{1/2}, \quad (3)$$

The uncertainty of the median, which is correlated to the uncertainty of the individual measurements in a relatively complex way [4], is not included in the uncertainty bars.

In Fig. 4, Youden plots are presented for the two cuts of wire calibrated by each laboratory. Of the eight test temperatures, Youden plots are shown only for those temperatures where emf values of at least one cut of wire deviated from the median by more than two standard deviations. Three cuts were calibrated by the pilot laboratory above 600 °C; for the Youden plots, the two most discrepant cuts were used in this case. At each temperature, the median emf value for all measurements of all laboratories

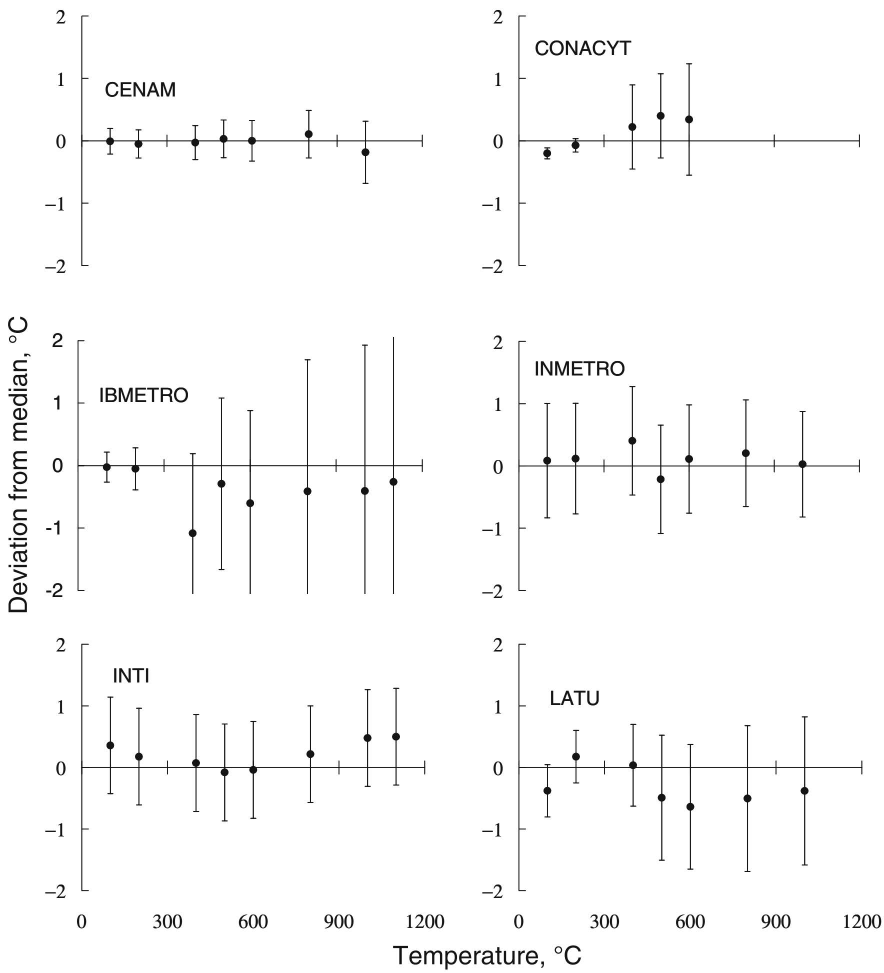


**Fig. 2** Deviation from the median of average results from the participating laboratories, expressed in equivalent degrees Celsius. The first six participants listed in Table 1 are shown

is calculated. Deviations from the median for each laboratory's two calibrated cuts (converted to equivalent degrees Celsius) are plotted as abscissa and ordinate. The dashed circles indicate deviations from the median of one, two, and three standard deviations, where the standard deviation is calculated from the emf values alone with no weighting by reported uncertainties. Examination of the plots indicates that the observed differences between laboratories are due predominantly to systematic biases (e.g., furnace effects) between the laboratories.

## 5 Discussion of Results

The level of agreement in this comparison is quite good. Of the 380 possible bilateral combinations of the data, only 13 (i.e., 3.4% of all combinations) exceed the $k = 2$

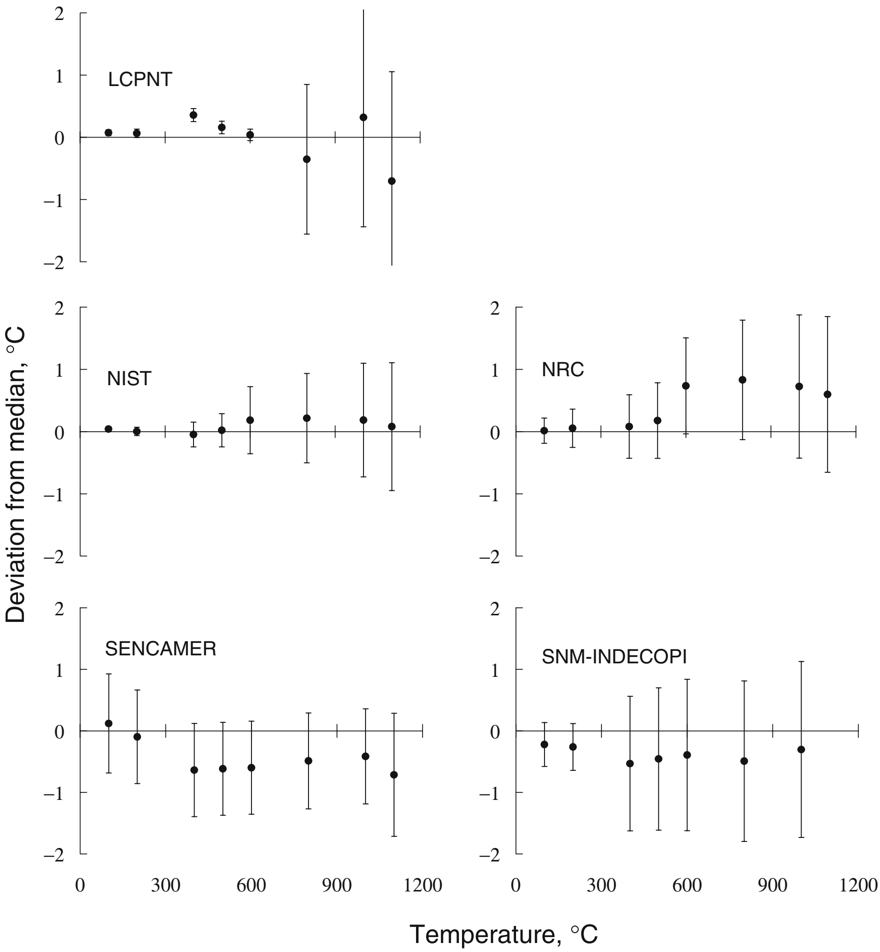


**Fig. 3** Deviation from the median of average results from the participating laboratories, expressed in equivalent degrees Celsius. The last five participants listed in Table 1 are shown

limits, and of these 13, only 3 exceed the $k = 3$ limits. All the outliers occur at temperatures of 800 °C and below, which suggests that drift of the type K wire due to high-temperature oxidation did not cause changes in thermocouple emf comparable to or larger than the claimed uncertainties.

The results of this comparison convincingly demonstrate that careful calibrations of stabilized type K thermocouples agree to better than 1 °C from 0 °C up to test temperatures of 1,100 °C for 1.63 mm diameter wire, even for a broad range of calibration protocols and equipment (see Table 2). The standard deviation of all un-averaged emf values at each test temperature was less than the equivalent of 0.20 °C at 200 °C and below, and less than 0.60 °C from (400 to 1,100) °C.

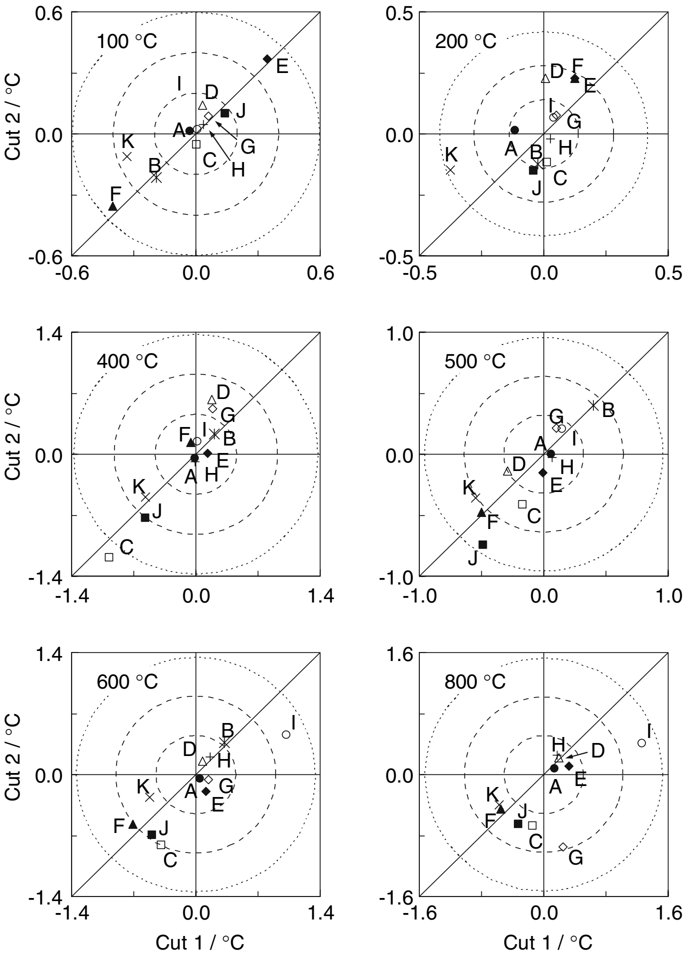


**Fig. 4** Youden plots for the two cuts of wire calibrated by each laboratory, for all temperatures where at least one emf value deviated from the median emf value by more than two standard deviations. The circles indicate deviations of one, two, and three standard deviations. Table 1 gives the laboratory codes used in the figure

The high thermoelectric uniformity of the thermocouple wire, as seen in Table 3 and by the Youden plots of Fig. 4, was much better than literature values [5] would suggest. Reference [6] discusses this point further.

## References

1. G.W. Burns, M.G. Scroger, The Calibration of Thermocouples and Thermocouple Materials, *NIST Special Pub. 250–235* (U.S. Govt. Printing Office, Washington, D.C., 1989)
2. D. Ripple, G.W. Burns, M.G. Scroger, Assessment of Uncertainties of Thermocouple Calibrations at NIST, *NISTIR 5340* (National Institute of Standards and Technology, Gaithersburg, Maryland, 1994)

3. D.C. Ripple, K.M. Garrity, M. Araya, C.R. Cabrera, L. Cordova Murillo, M.E. de Vanegas, D.J. Gee, E. Guillen, S. Martinez-Martinez, E. Mendez-Lango, L. Mussio, S.G. Petkovic, K.N. Quelhas, G. Rangugni, O. Robatto, E. von Borries Rocha, Metrologia **44,** Tech. Suppl. 03007 (2007)
4. M.G. Cox, Metrologia **39**, 589 (2002)
5. J.G. Hust, R.L. Powell, L.L. Sparks, in *Temperature: Its Measurement and Control in Science and Industry*, Vol. 4, ed. by H.H. Plumb (ISA, Pittsburgh, Pennsylvania, 1972), pp. 1525–1535
6. D.C. Ripple, K.M. Garrity, Measure **1**, 28 (2006)