CCT-K7.2021: CIPM Key Comparison of Water-Triple-Point Cells

Interim Technical Report

A. Peruzzi[1], S. Dedyulin[1], M. Levesque[1], D. del Campo[2], B.C. Garcia Izquierdo[2], M.E. Gomez[2], K.N. Quelhas[3], M.A.P. Neto[3], B.M. Lozano[3], L. Eusebio[4], I. Yang[5], F. Sparasci[6], C. Martin[6], L. Risegari[6], P. Saunders[7], E. Molloy[7], X.K. Yan[8], J.P. Sun[8], X.J. Feng[8], J.T. Zhang[8], M.-K. Ho[9], T. Nakano[10], J.V. Widiatmo[10], I. Saito[10], E. Ejigu[11], J. Pearce[12], S. Rudtsch[13], L. Buenger[13], M. Kalemci[14], A. Uytun[14], C. Bruin-Barendregt[15], M. Panman[15], D.R. White[16], A. Possolo[17]

[1]National Research Council of Canada, Ottawa, Canada
[2]Centro Espaňol de Metrología, Madrid, Spain
[3]Instituto Nacional de Metrologia, Qualidade e Tecnologia, Rio de Janeiro, Brazil
[4]Istituto Português da Qualidade, Caparica, Portugal
[5]Korea Research Institute of Standards and Science, Daejeon, Korea
[6]Laboratoire National de Métrologie et d'Essais – Conservatoire National des Arts et Métiers, Paris, France
[7]Measurement Standards Laboratory of New Zealand, Lower Hutt, New Zealand
[8]National Institute of Metrology, Beijing, China
[9]National Measurement Institute, Australia, Lindfield, Australia
[10]National Measurement Institute of Japan, AIST, Tsukuba, Japan
[11]National Metrology Institute of South Africa, Pretoria, South Africa
[12]National Physical Laboratory, Teddington, United Kingdom
[13]Physikalisch-Technische Bundesanstalt, Berlin, Germany
[14]National Metrology Institute of Turkey, Gebze-Kocaeli, Turkey
[15]VSL Dutch Metrology Institute, Delft, the Netherlands
[16]Independent Researcher, Lower Hutt, New Zealand
[17]National Institute of Standards and Technology, Gaithersburg, USA

Version 2 (April 2023)

TABLE OF CONTENTS

1. Introduction
2. Organization of the comparison
3. Calibration of the transfer cells by the participants
4. Comparison of the transfer cells by the pilot
5. Combination of participants' results and pilot's results
6. Statistical analysis and Degrees of Equivalence
7. Link between CCT-K7 and CCT-K7.2021
8. Conclusions

References

Appendix1: Technical Protocol

Appendix 2: Participants' reports on the calibration of the transfer cells

Appendix 3: Uncertainty budgets

Appendix 4: Immersion profiles

Appendix 5: Additional statistical analysis

1 Introduction

The first CIPM Key Comparison (KC) of water-triple-point (TPW) cells, CCT-K7, was carried out in 2002-2004 [1]. In its 2017 meeting, the Consultative Committee for Thermometry (CCT) approved the initiation of a repeat of CCT-K7 and, in June 2020, NRC offered to act as coordinating laboratory of this new CIPM KC, which was named CCT-K7.2021.

The Technical Protocol for CCT-K7.2021 was drawn up by the NRC, with the support of the coordinating group, in accordance with the CIPM MRA-D-05 guide (Version 1.6) [2] taking into account the experience gained in the previous key comparison [1], its subsequent regional extensions [3-5], and current best practice. The final version of the Technical Protocol (see Appendix 1) was approved by all participants in December 2020 and by the CCT Working Group on Key Comparisons in January 2021. The initial measurements at the participants' laboratories were started in April 2021 and completed in December 2021. The measurements at the pilot's laboratory were started in July 2021 and completed in March 2022. The final measurements at the participants' laboratories were initiated in March 2022 and completed in September 2022. The final measurements at the participants' laboratories were only used to confirm the stability of the transfer cells and were not reported here. All measurement reports were delivered to the pilot by September 2022.

This Technical Report of CCT-K7.2021 is organized as follows.
Chapter 2 reproduces some parts of the Technical Protocol as an aid to the understanding of the comparison measurements and the analysis of the results.
Chapter 3 reports the results of the calibrations of the transfer cells performed by the participants.
Chapter 4 reports the results of the comparison of all transfer cells performed by NRC.
Chapter 5 determines the differences between the national reference and the NRC reference, composed of two NRC reference cells.
Chapter 6 determines the Key Comparison Reference Value and the Degrees of Equivalence.
Chapter 7 discusses the link between CCT-K7 and CCT-K7.2021.
Chapter 8 summarises the outcomes of the comparison emphasising the improvements achieved with respect to the previous key comparison (CCT-K7).
Appendix 1 presents the Technical Protocol.
Appendix 2 reports the results of the calibration of the transfer cells by the participants.
Appendix 3 reports the uncertainty budgets submitted by the participants.
Appendix 4 reports the immersion profiles of the transfer cells.
Appendix 5 provides details on supporting statistical investigations.

2 Organization of the comparison

2.1 Objectives of the comparison

The primary objective of CCT-K7.2021 is a comparison, mediated by the participant transfer cells, of the participant national realizations of the TPW temperature. A participant's realization of the TPW temperature is typically defined as the average of an ensemble of national reference TPW cells, but national realizations defined by a single reference cell are acceptable as well. While in CCT-K7 only 2 laboratories out of 21 participants based their TPW realization on VSMOW water, in this KC it was expected that, due to the "Clarification of the definition of the kelvin" of 2005 [6], all the participants would present a TPW national reference based on this document, and on the "Technical Annex for the International Temperature Scale of 1990 (ITS-90)", revised version of 2017 [7], with consequent changes in the TPW national reference of many participants. Due to this fact and to the improved measurement capabilities of the participants over the past 20 years, it was also expected that the differences between the TPW realizations would be smaller than those observed in CCT-K7 (standard deviation of 50 μK and peak-to-peak difference of 171 μK for a total of 21 participants).

In addition to the primary objective (above), two secondary objectives were pursued:
a) A direct comparison of TPW cells of the highest quality.
b) A linkage to the previous key comparison CCT-K7, mediated by the cells that were used by the participants in CCT-K7 and that are still available.

Secondary objective a) was essentially the means chosen to achieve the primary objective, but it was also expected to provide useful information on the state-of-the-art quality of TPW cell manufacturing. It was expected that, due to the improved measurement capabilities of the participants over the past 20 years and the tendency of commercial suppliers, after the clarification of the kelvin definition in 2005, to produce cells that are closer to VSMOW isotopic composition, the differences between the TPW temperatures realized by the transfer cells would be smaller than those observed in CCT-K7 (standard deviation of 50 μK and peak-to-peak difference of 163 μK for a total of 22 cells).

For secondary objective b), many participants in this KC took part also in CCT-K7 and some still had the cells used in CCT-K7, either as transfer cells or as national reference cells. With a moderate additional effort at the participating laboratories, the information on the temperature difference between the TPW realized by these cells and the local old (CCT-K7) and new (CCT-K7.2021) national reference could be obtained. This made it possible to relate the CCT-K7 KCRV to the new CCT-K7.2021 KCRV.

2.2 Participants and roles

The participant laboratories, the contact persons and email addresses are listed in Table 2.1.

The pilot assembled a subgroup of participants to aid in the harmonization of the uncertainty budgets, in the approach to the methods for analyzing the comparison results and in the use of software tools.

Institute	**Country**	**Contact person**	**Contact email**
CEM	Spain	Dolores del Campo	ddelcampo@cem.es
		Carmen Garcia Izquierdo	mcgarciaiz@cem.es
CENAM	Mexico	Enrique Martines Lopez	emartine@cenam.mx
INMETRO	Brazil	Klaus N. Quelhas	knquelhas@inmetro.gov.br
		Mario A.P. Neto	maneto@inmetro.gov.br
		Bruno M. Lozano	bmlozano@inmetro.gov.br
INRiM	Italy	Giuseppina Lopardo	g.lopardo@inrim.it
IPQ	Portugal	Liliana Eusebio	linianae@ipq.pt
KRISS	South Korea	Inseok Yang	iyang@kriss.re.kr
LNE-CNAM	France	Fernando Sparasci	fernando.sparasci@cnam.fr
		Catherine Martin	catherine.martin@cnam.fr
		Lara Risegari	lara.risegari@cnam.fr
MSL	New Zealand	Peter Saunders	peter.saunders@measurement.govt.nz
		Ellie Molloy	ellie.molloy@callaghaninnovation.govt.nz
NIM	China	Xiaoke Yan	yanxk@nim.ac.cn
		Jianping Sun	sunjp@nim.ac.cn
		Xiaojuan Feng	fengxj@nim.ac.cn
		Jintao Zhang	zhangjint@nim.ac.cn
NIST	United States	Tobias Herman	tobias.herman@nist.gov
		Antonio Possolo	antonio.possolo@nist.gov
NMIA	Australia	Mong-Kim Ho	mong-kim.ho@measurement.gov.au
NMIJ/AIST	Japan	Tohru Nakano	tohru-nakano@aist.go.jp
		Januarius V. Widiatmo	janu-widiatmo@aist.go.jp
		Ikuhiko Saito	saitou.19hiko@aist.go.jp
NMISA	South Africa	Efrem Ejigu	eejigu@nmisa.org
NPL	United Kingdom	Jonathan Pearce	jonathan.pearce@npl.co.uk
NRC	**Canada**	**Andrea Peruzzi**	andrea.peruzzi@nrc-cnrc.gc.ca
		Sergey Dedyulin	sergey.dedyulin@nrc-cnrc.gc.ca
PTB	Germany	Steffen Rudtsch	steffen.rudtsch@ptb.de
UME	Turkey	Murat Kalemci	murat.kalemci@tubitak.gov.tr
		Ali Uytun	ali.uytun@tubitak.gov.tr
VNIIM	Russia	Anatolii Pokhodun	a.i.pokhodun@vniim.ru
VSL	The Netherlands	Conny Bruin-Barendregt	cbarendregt@vsl.nl
		Matthijs Panman	mpanman@vsl.nl
Independent	New Zealand	**Rod White**	rodwhitenz@gmail.com

Table 2.1: List of participant laboratories, contact persons and email addresses. The pilot laboratory and the members of the coordinating subgroup are in bold.

2.3 Topology of the comparison

The topology of the comparison was a "*collapsed star*", executed in three phases:

1) Each laboratory selected one of its TPW cells for use as a transfer cell and directly compared it against its national reference.
In case the laboratory still possessed a cell used in CCT-K7 (either as transfer cell or as national reference cell), the laboratory additionally compared the CCT-K7 cell to the national reference.
The selected transfer cell and the measurement results were delivered to NRC.

2) NRC compared all transfer cells against two NRC reference cells.

3) Each laboratory retrieved its transfer cell from NRC and directly re-compared it against its TPW national reference.
In case the laboratory had a cell used in CCT-K7 (either as transfer cell or as national reference cell), the laboratory additionally re-compared the CCT-K7 cell to the national reference.

2.4 Measurements

The measurements at the participant laboratories commenced on April 1st 2021. Although the delivery of all the participant transfer cells to NRC was originally scheduled for the end of June 2021, some laboratories shipped their transfer cells several months later (the last cells reached NRC in November 2021).

Despite the delays, the pilot commenced measurements of the participant transfer cells according to the original schedule (June 2021), but the measurements at the pilot's laboratory extended until March 2022, as an extra set of measurements (January- March 2022) was necessary, to measure the transfer cells delivered late. Moreover, a few transfer cells were broken during transport to NRC and several transfer cells showed anomalous behaviour during the measurements at NRC. In both cases, the corresponding participants were offered the opportunity to measure and deliver a replacement cell. The extra run of measurements performed at the beginning of 2022 also allowed measurement of the replacement cells.

For different reasons, three participants (CENAM, INRiM and NIST) could not deliver a replacement cell. These participants asked to withdraw from the comparison and all the other participants accepted their withdrawal.
The return measurements at the participant laboratories were completed in September 2022.

The analysis of the results was performed from August 2022 to September 2022, and the Draft A of the report was completed in October 2022. Due to geopolitical events that occurred in 2022, many participants (including the pilot), following the mandate of their respective governments, refused to communicate directly with the Russian participant in this comparison (VNIIM) and explicitly requested the Russian participant to be excluded from the report. To minimize the consequences in the reporting of this comparison, VNIIM was excluded from this report.
The number of participants for whom the results are reported was reduced to 15.

3 Calibration of the transfer cells by the participants

In the first part of the comparison, each participant selected one of its TPW cells for use as a transfer cell and directly compared it to its TPW national reference. Table 3.1 summarizes the information available on the transfer cells. Table 3.2 describes the main features of the national references. Table 3.3 lists the equipment used by the participants. Table 3.4 shows the results of the calibration of the transfer cells by the participants, expressed as averages of measurements on two realizations (i.e., different mantles).

Further details on the national references of the participants and their calibration reports of the transfer cells are reported in Appendix 2. The detailed uncertainty budgets of all participating laboratories are reported in Appendix 3.

The uncertainty reported by each laboratory in calibrating the transfer cell against the national reference includes the uncertainty components related to the realization of the national reference.

Table 3.1: Transfer cells selected by the participating laboratories.

	NRC identification	Serial number	Manufacturer	Model	Year of manufacture	Envelope material
CEM	CEM1833Q	1833Q	Isotech	A11-50-270-Q	2020	fused silica
INMETRO	INM14405	14405	INMETRO	n.a.	2012	borosilicate
IPQ	IPQ2114	2114	Isotech	Jarrett, Type A-11	2000	borosilicate
KRISS	KRISSQ1060	1060	Fluke	5901C-Q	2019	fused silica
LNE-Cnam	LNE1747Q	1747Q	Isotech	B11-50-270-Q	2019	fused silica
MSL	MSL06/01	MSL06/01	MSL	n.a.	2006	fused silica
NIM	NIM/0	Q10	NIM	TPW-Q	2020	fused silica
NMIA	NMIA04/02	MSL04/02	MSL	Type B	2004	fused silica
NMIJ/AIST	NMIJQ1058	Q1058	Fluke	5901D	2009	fused silica
NMISA	NMISA1593Q	1593Q	Isotech	A11-50-270-Q	2017	fused silica
NPL	NPL1905Q	1905Q	Isotech	B11-50-270-Q	2021	fused silica
NRC	NRC1894Q	1894Q	Isotech	B11 65	2020	fused silica
PTB	PTBQ1175	Q1175	Fluke	5901D-Q	2021	fused silica
TUBITAK UME	UMEQ5014	5014	Fluke	5901A-Q	2006	fused silica
VSL	VSL17T048	VSL17T048	VSL	VSL Type 3	2019	fused silica

Table 3.2: Main characteristics of the national references, as defined by the participants.

	Number of cells	Fused silica cells	Borosilicate cells	Year of manufacture of cells (B = Borosilicate, Q = fused silica)	Information on isotopic composition	Correction for isotopic composition	Information on chemical impurities	Correction for chemical impurities
CEM	6	2	4	1992(B), 1994(B), 1999(B), 2002(B), 2006(Q), 2008(Q)	yes, for the 2 fused silica cells	yes, for the 2 fused silica cells	no	no
INMETRO	1	1	0	2019(Q)	yes	yes	no	no
IPQ	1	0	1	2000(B)	yes	no	no	no
KRISS	3	2	1	2002(B), 2019(Q), 2019(Q)	yes	yes, for 1 borosilicate cell and for 1 fused silica cell	no	no
LNE-Cnam	4	3	1	2009(Q), 2011(B), 2014(Q), 2018(Q)	yes	yes	no	no
MSL	2	1	1	2001(B), 2006(Q)	yes	yes	yes	yes
NIM	3	3	0	2020(Q), 2020(Q), 2020(Q)	yes	yes	no	no
NMIA	8	7	1	2006(B), 2006(Q), 2006(Q), 2006(Q), 2006(Q), 2010(Q), 2010(Q), 2020(Q)	yes	yes	no	no
NMIJ/AIST	3	3	0	2005(Q), 2012(Q), 2021(Q)	yes	yes	no	no
NMISA	2	0	2	1998(B), 1998(B)	no	no	no	no
NPL	5	1	4	2003(Q), 2003(B), 2007(B), 2007(B), 2010(B)	yes	yes	no	no
NRC	10	10	0	2003(Q), 2003(Q), 2003(Q), 2003(Q), 2003(Q), 2016(Q), 2017(Q), 2017(Q), 2020(Q), 2020(Q)	yes	yes	yes	no
PTB	3	3	0	2007(Q), 2007(Q), 2021(Q)	yes	yes	no	no
TUBITAK UME	4	1	3	2002-2016(B), 2021(Q)	yes	yes	no	no
VSL	10	10	0	2006(Q), 2006(Q), 2007(Q), 2008(Q), 2008(Q), 2008(Q), 2015(Q), 2015(Q), 2017(Q), 2017(Q)	yes	yes	yes	no

Table 3.3: Equipment used by the participating laboratories.

NMI	Model resistance bridge and current	Number of repeated measurements at each current	Sampling frequency of repeated measurements /Hz	Standard resistor model and nominal value	Reference resistor in a temperature-controlled bath: stability	Model maintenance bath for TPW cells	Model SPRT	Method preparation ice mantle
CEM	ASL F900, 1 mA	20	1/10	Tinsley 5685 A, 25 Ω	Yes: 0.0025 °C	Isotech ITL-M-18233	Fluke 5681	Crushed solid carbon dioxide
INMETRO	ASL F900, 1 mA	15	1/30	Tinsley 5685 A, 100 Ω	Yes: 0.01 °C	Ice dewar	Fluke 5683	Crushed solid carbon dioxide
IPQ	ASL F18, 1 mA	100	1/10	Tinsley 5685 A, 25 Ω	Yes: 0.005 °C	Fluke 7312	Tinsley 5187-SA	Crushed solid carbon dioxide
KRISS	ASL F900, 1 mA	16	1/17	Tinsley 5685 A, 100 Ω	Yes: <0.001 °C	Ice dewar	Fluke 5683	Crushed solid carbon dioxide
LNE-Cnam	ASL F900, 1 mA	80	1/10	Tinsley 5685 A, 25 Ω	Yes: ±0.05 °C	Fluke 7312	Rosemount 162 CG	Ethanol and aluminum rod cooled in liquid nitrogen
MSL	MI 6015T, 1 mA	20	1/30	Tinsley 5685 A, 100 Ω	Yes: 0.0003 °C	Ice dewar	Leeds & Northrup, Meyers type, Cat 8167-25	Heat pipe and crushed solid carbon dioxide
NIM	ASL F900, 1 mA	10-20	1/20	Tinsley 5685 A, 100 Ω	Yes: ±0.002 °C	Fluke 7012	NIM 58660	Liquid nitrogen flow-through cooling
NMIA	ASL F900, 1 mA	10	1/12	Guildline 9330, 100 Ω	Yes: 0.005 °C	Stirred water bath made by NMIA	Accumac 1930	Heat pipe and crushed solid carbon dioxide
NMIJ	MI 6010T, 1 mA	100	1/8	Tinsley 5685 A, 10 Ω	Yes: ±0.1 °C	Fluke 7312	Fluke 5681	Heat pipe and crushed solid carbon dioxide
NMISA	Isotech MicroK Gold 70, 1 mA	30-35	1/10	WIKA CER6000, 100 Ω	Yes: resistance corrected for air bath temperature	Fluke 7312	Chino R800-02	Heat pipe and crushed solid carbon dioxide
NPL	ASL F900, 1 mA	20	1/10	Tinsley 5685 A, 100 Ω	Yes: ±0.01 °C	Ice dewar and Fluke 7312	Chino R800-2	Crushed solid carbon dioxide
NRC	MI 6020T Premium, 2 mA	10-30	1/10	Tinsley 5685 A, 25 Ω	Yes: ±0.002 °C	Ice dewar and Isotech ITL-M-18233	Leeds & Northrup, Thermohm	Crushed solid carbon dioxide
PTB	ASL F900, 1 mA	20	1/50	Tinsley 5685 A, 25 Ω	Yes: <0.01 °C	Isotech ITL-M-18233	Tinsley 5137-SA and Yellowspring 8167-25	Electric immersion freezer (liquid-cooled rod)
TUBITAK UME	MI 6015T, 1 mA	40	1/15	Tinsley 5685 A, 25 Ω	Yes: 0.0002 °C	Fluke 7012	Fluke 5681	Crushed solid carbon dioxide
VSL	MI 6015T, 2 mA	25	1/4	Tinsley 5685 A, 25 Ω	Yes: 0.001 °C	Isotech ITL-M-18233	Leeds & Northrup, 8163	Liquid nitrogen flow-through cooling

Table 3.4: Results of calibration of the transfer cells by the participating laboratories. $u(T_i^{\text{Transfer}} - T_i^{\text{Nat Ref}})$ is the combined standard uncertainty of the temperature difference between the transfer cell and the national reference.

	Transfer cell	$T_i^{\text{Transfer}} - T_i^{\text{Nat Ref}}$ /µK	$u(T_i^{\text{Transfer}} - T_i^{\text{Nat Ref}})$ /µK
CEM	CEM1833Q	-89.7	39.1
INMETRO	INM14405	7.4	36.4
IPQ	IPQ2114	-25.1	81.8
KRISS	KRISSQ1060	5.6	31.0
LNE-Cnam	LNE1747Q	4.5	49.8
MSL	MSL06/01	-86.8	9.4
NIM	NIM/0	-25.0	31.8
NMIA	NMIA04/02	-21.7	22.6
NMIJ/AIST	NMIJQ1058	-2.9	30.5
NMISA	NMISA1593Q	6.2	61.5
NPL	NPL1905Q	7.1	23.1
NRC	NRC1894Q	-3.5	17.5
PTB	PTBQ1175	-4.0	25.0
TUBITAK UME	UMEQ5014	58.7	53.0
VSL	VSL17T048	17.2	28.1

4 Comparison of the transfer cells by the pilot

In the second part of the comparison, the differences $T_i^{Transfer} - T_{Pilot\ Ref}$ between the temperatures $T_i^{Transfer}$ realized by each transfer cell i and the pilot's reference temperature $T_{Pilot\ Ref}$, composed of the average of two NRC national reference cells (1894Q and Q1150), were measured for each participant's transfer cell. Before reporting the measured temperature differences (Section 4.2), the measurement set-up used by the pilot's laboratory is described (Section 4.1) and the uncertainty budget in measuring $T_i^{Transfer} - T_{Pilot\ Ref}$ is reported. More detailed information on the pilot's measurement set-up and uncertainty budgets can be found in [8].

4.1 Measurements in the pilot's laboratory

The measurements in the pilot's laboratory were performed using a single 25 Ω Standard Platinum Resistance Thermometer (SPRT), Leeds & Northrup, Thermohm type, s/n 1504288. This SPRT has a long history of TPW measurements, has never been annealed and has always remained at room temperature when not used for TPW measurements. Its TPW resistance value has been slowly increasing with time and the drift has been less than 0.5 mK in the past 15 years.

For each TPW cell, the SPRT resistance was measured using a DC thermometry resistance-ratio bridge, MI 6020T Premium, and a 25 Ω standard resistor, Tinsley, model 5685A, s/n 274880, maintained in a temperature-controlled bath (24.99 ºC) with temporal stability and spatial uniformity of the order of 2 mK. The standard resistor was calibrated in December 2020 and the result of the calibration was a value of 24.9997477 Ω with expanded (k = 2) uncertainty of 1.5 μΩ.

For all the measured TPW cells, the ice mantle was formed using crushed dry ice, and the cell was aged inside the maintenance bath for at least one week prior to measurements. During the measurements, the cells were maintained in two (not temperature-controlled) isothermal containers, essentially two insulated metal boxes, containing 4 tubes accommodating one cell each. The space inside the metal boxes and between the tubes was filled with crushed water ice and daily refilled. Extensive preliminary investigations demonstrated that the isothermal containers provided a satisfactory thermal environment for the cells.

For all the measured TPW cells (including the transfer cells) the distance between the bottom of the thermometer well and the free water level in the cell, to be used for the hydrostatic head correction, was measured once for each mantle at NRC.

The overall performance of the measurement setup is illustrated in Figure 4.1, where a typical SPRT self-heating measurement is shown. The noise level at both 2 mA and 2√2 mA measuring currents is of the order of 1 μK.

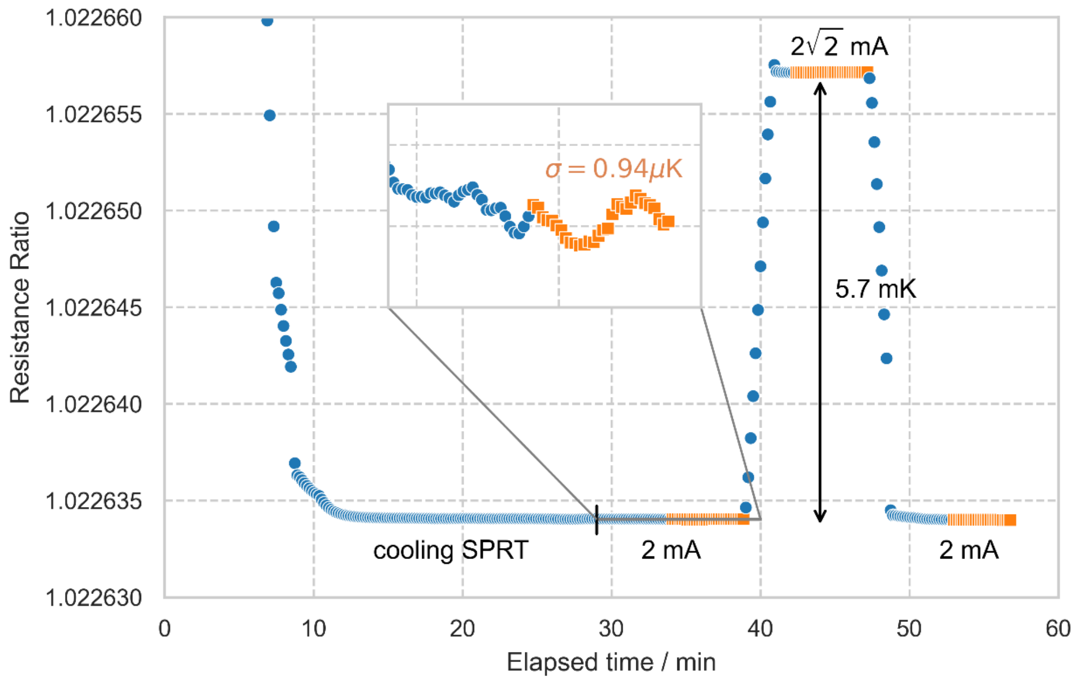


Figure 4.1: Example of SPRT self-heating measurement with the pilot's measurement setup (reproduced from [8]). The blue dots are the discarded measurement points and the orange squares are the measurement points used for calculating the self-heating correction.

4.2 Uncertainty budget for the pilot measurements

The uncertainty budget in the measurement of the temperature difference between any pair of cells in the pilot's laboratory is shown in Table 4.1. The estimated combined standard uncertainty is 13.1 μK.

Table 4.1: Uncertainty budget for measuring the temperature difference between two TPW cells at NRC. In the last column, n/a means that the uncertainty component cancels out in the difference between two cells.

Uncertainty source	Physical origin	$u(T_i - T_j)$ /μK
SPRT Self-heating correction	The resistance of the SPRT is measured at two difference currents (2 mA and 2√2 mA) and the zero current resistance is estimated from that.	6.8
Hydrostatic pressure correction	Uncertainty in identifying the position of the water level with respect to the bottom of the thermometer well (3 mm · 0.73 μK/mm)	3.1

	Uncertainty in identifying the position of the bottom of the SPRT with respect to the bottom of the thermometer well	1.0
	Uncertainty in identifying the position of the thermal centre of the sensing element of the SPRT with respect to the bottom of the SPRT	n/a
Standard resistor	Calibration uncertainty of the standard resistor	n/a
	Drift of the standard resistor from last calibration	n/a
	Temperature stability of the oil bath (included in reproducibility)	n/a
Resistance bridge	Noise (included in reproducibility)	n/a
	Differential non-linearity	3.9
	Integral non-linearity	n/a
Stray thermal exchanges	Change in room temperature, change in bath temperature, light piping	7.8
Reproducibility for a single ice mantle	Electrical noise (bridge, cables, …), temperature stability of standard resistor, SPRT changes during one day of manipulation (small mechanical shocks), different thermal contact of SPRT in the TPW cell, instability of TPW cell, instability of reference cell.	4.4
Reproducibility for different ice mantles	Morphology of different ice mantles and consequently different impurity segregation	4.4
	Combined uncertainty (*k* =1)	**13.1**
	Expanded uncertainty (*k* =2)	**26.2**

When the daily temperature difference between a participant's transfer cell $T_{\mathrm{i}}^{\mathrm{Transfer}}$ and the pilot's reference $T_{\mathrm{Pilot\ Ref}}$ is determined, we need to take into account that the pilot's reference is composed of the average of the two NRC reference cells 1894Q and Q1150:

$$T_i^{Transfer} - T_{Pilot\ Ref} = T_i^{Transfer} - \left(\frac{T_{1894Q} + T_{Q1150}}{2}\right)$$
$$= \frac{1}{2}\left[\left(T_i^{Transfer} - T_{1894Q}\right) + \left(T_i^{Transfer} - T_{Q1150}\right)\right]$$

The uncertainty in the difference is $u(T_{\mathrm{i}}^{\mathrm{Transfer}} - T_{\mathrm{Pilot\ Ref}}) = 9.3$ µK.

4.3 Measurement results in the pilot's laboratory

The pilot measurements were organized in 4 batches. Each batch included the two NRC reference cells (1894Q and Q1150) and 3 to 6 of the participants' transfer cells. Table 4.2 shows the time period and the measured cells in each batch.

Table 4.2: Time period and measured cells in each batch. For the cells marked with an asterisk, no measurement result is reported for the reasons outlined in the text.

Batch I			Batch II			Batch III			Batch IV		
Cells	**Start**	**End**	**Cells**	**Start**	**End**	**Cells**	**Start**	**End**	**Cells**	**Start**	**End**
NRC1894	20/07/2021	20/08/2021	NRC1894	20/09/2021	02/11/2021	NRC1894	12/11/2021	22/12/2021	NRC1894	31-01-2022	08-03-2022
NRCQ1150			NRCQ1150			NRCQ1150			NRCQ1150		
IMGC34*			MSL01/02*			INM14405			CENAMA005*		
IPQ2114			LNE1747Q			CEM1833Q			MSL06/01		
VNIIM0/115*			NMIJQ1058			KRISSQ1060			NIM/0		
VSL17T048			NPL1905Q			NMIA04/02			NIST1454Q*		
			PTBQ1175			NMISA1593Q			UMEQ5014		
			UME16/01*								

For the cells that appear with an asterisk in Table 4.2, no measurement result is reported because one of the following applied:
1) they showed to be unstable and were replaced by another transfer cell (MSL01/02 was replaced by MSL06/01 and UME16/01 was replaced by UMEQ5014),
2) they belonged to participants that decided to withdraw from the comparison (CENAM, INRiM and NIST), or
3) minimization of the damage from geopolitical events to the reporting of the comparison (VNIIM results excluded from the comparison).

Notably, all the transfer cells found to be unstable were older borosilicate cells (CENAMA005, IMGC34, MSL01/02 and UME16/01).

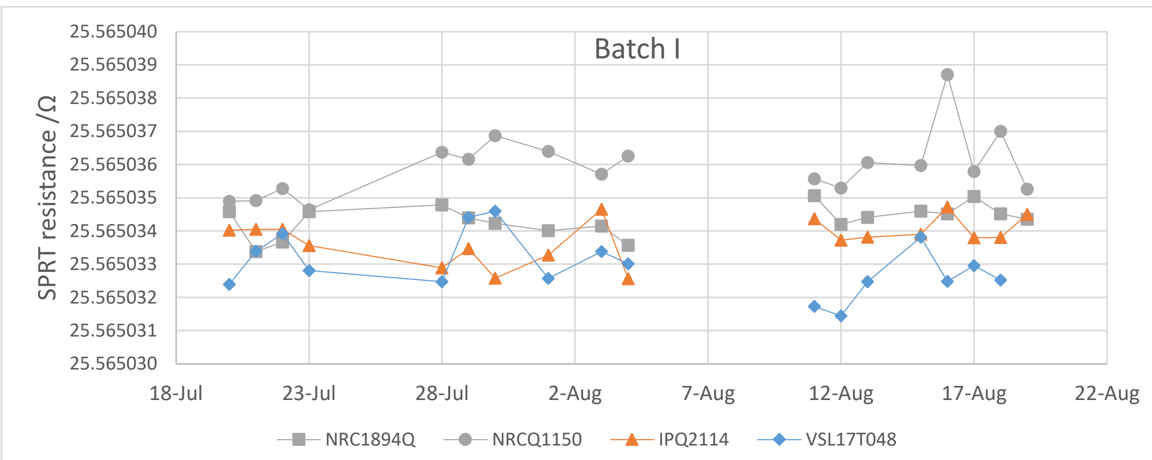
In the following tables (Table 4.3 to 4.6) and figures (Figures 4.2 to 4.5), the results obtained in each batch are reported (one batch per page). For each batch we reported (both numerically and graphically) the SPRT resistance value measured daily for all the cells of the batch and the resulting calculated temperature difference between each cell and the pilot reference (the average of the NRC reference cells 1894Q and Q1150). Note that, some daily measurements are missing for some of the transfer cells. In these cases, there was no rejection of data, but simply a lack of data because of operational problems (for example, the size of the mantle was not sufficient to guarantee a reliable measurement).

A change in the SPRT resistance occurred in the time between the first and the second mantle of batch II (see Figure 4.3) and in the time between batch II and batch III. Although the value of the resistance change was significant with respect to the daily measured temperature differences $T_{\mathrm{i}}^{\mathrm{Transfer}}$ - $T_{\mathrm{Pilot\ Ref}}$, our measurement scheme was not sensitive to changes in the SPRT resistance within each batch.

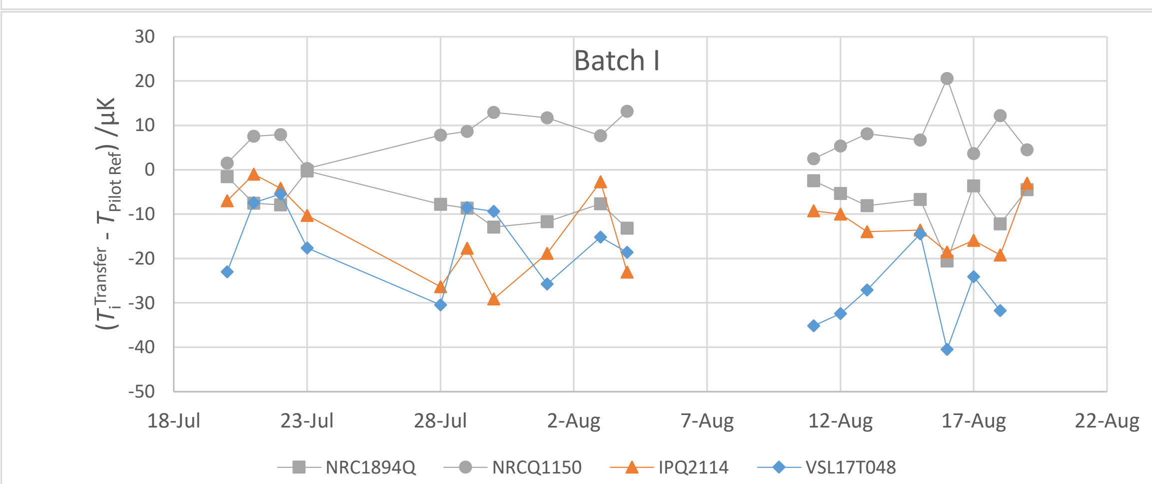
Table 4.7 and Figure 4.6 summarizes the results for each batch.

Table 4.3: Results of Batch I.

Date	Mantle	NRC1894Q	NRCQ1150	IPQ2114	VSL17T048	$T_{NRC1894Q} - T_{Pilot\ Ref}$ /µK	$T_{NRC1150Q} - T_{Pilot\ Ref}$ /µK	$T_{IPQ2114} - T_{Pilot\ Ref}$ /µK	$T_{VSL17T048} - T_{Pilot\ Ref}$ /µK
20-Jul	M1	25.56503458	25.56503489	25.56503402	25.56503239	-1.5	1.5	-7.0	-23.0
21-Jul		25.56503338	25.56503492	25.56503405	25.56503339	-7.5	7.5	-1.0	-7.4
22-Jul		25.56503367	25.56503528	25.56503405	25.56503391	-7.9	7.9	-4.2	-5.5
23-Jul		25.56503458	25.56503464	25.56503356	25.56503281	-0.3	0.3	-10.3	-17.6
28-Jul		25.56503478	25.56503637	25.56503289	25.56503247	-7.8	7.8	-26.4	-30.5
29-Jul		25.56503440	25.56503616	25.56503347	25.56503441	-8.6	8.6	-17.7	-8.5
30-Jul		25.56503423	25.56503687	25.56503258	25.56503459	-12.9	12.9	-29.1	-9.4
1-Aug		25.56503401	25.56503639	25.56503328	25.56503257	-11.7	11.7	-18.9	-25.8
3-Aug		25.56503415	25.56503571	25.56503465	25.56503338	-7.7	7.7	-2.7	-15.2
4-Aug		25.56503357	25.56503626	25.56503256	25.56503302	-13.2	13.2	-23.1	-18.6
11-Aug	M2	25.56503506	25.56503557	25.56503437	25.56503173	-2.5	2.5	-9.3	-35.2
12-Aug		25.56503420	25.56503529	25.56503372	25.56503144	-5.4	5.4	-10.0	-32.4
13-Aug		25.56503441	25.56503606	25.56503381	25.56503247	-8.1	8.1	-13.9	-27.1
15-Aug		25.56503460	25.56503597	25.56503390	25.56503381	-6.7	6.7	-13.6	-14.4
16-Aug		25.56503451	25.56503870	25.56503472	25.56503248	-20.5	20.5	-18.6	-40.5
17-Aug		25.56503504	25.56503579	25.56503379	25.56503295	-3.7	3.7	-15.9	-24.1
18-Aug		25.56503452	25.56503700	25.56503380	25.56503252	-12.2	12.2	-19.2	-31.7
19-Aug		25.56503435	25.56503526	25.56503450		-4.5	4.5	-3.0	
					Average M1	-7.9	7.9	-14.0	-16.1
					Average M2	-7.9	7.9	-12.9	-29.4
					Average M1+M2	**-7.9**	**7.9**	**-13.5**	**-22.7**



(a)



(b)

Figure 4.2: Results of Batch I: (a) SPRT resistance values measured daily for each cell and (b) calculated temperature difference between each transfer cell and the pilot reference.

Table 4.4: Results of Batch II.

Date	Mantle	NRC1894Q	NRCQ1150	LNE1747Q	PTBQ1175	NMIJQ1058	NPL1905Q	$T_{NRC1894Q}$ - $T_{Pilot\ Ref}$ /μK	T_{Q1150} - $T_{Pilot\ Ref}$ /μK	$T_{LNE1747Q}$ - $T_{Pilot\ Ref}$ /μK	$T_{PTBQ1175}$ - $T_{Pilot\ Ref}$ /μK	$T_{NMIJQ1058}$ - $T_{Pilot\ Ref}$ /μK	$T_{NPL1905Q}$ - $T_{Pilot\ Ref}$ /μK
20-Sep	M1	25.56502814	25.56502912	25.56503030	25.56502925	25.56503122	25.56503132	-4.8	4.8	16.4	6.1	25.4	26.4
21-Sep		25.56502839	25.56503028	25.56502981	25.56502928	25.56503058	25.56503117	-9.3	9.3	4.6	-0.6	12.2	18.0
22-Sep		25.56502865	25.56503055	25.56503023	25.56503064	25.56503163	25.56503094	-9.3	9.3	6.2	10.2	19.9	13.2
23-Sep		25.56502932	25.56503055	25.56503048	25.56503089	25.56503295	25.56503187	-6.0	6.0	5.4	9.3	29.6	19.0
24-Sep		25.56502906	25.56502981	25.56503138	25.56503001	25.56503013	25.56503166	-3.7	3.7	19.1	5.7	6.8	21.8
27-Sep		25.56502875	25.56503015	25.56502920	25.56503147	25.56503276	25.56503123	-6.9	6.9	-2.4	19.8	32.5	17.5
28-Sep		25.56502659	25.56502877	25.56502841	25.56502882	25.56502983	25.56502954	-10.7	10.7	7.2	11.1	21.1	18.2
29-Sep		25.56502938	25.56503007	25.56502918	25.56502997	25.56503080	25.56503125	-3.4	3.4	-5.4	2.4	10.5	14.9
30-Sep		25.56502837	25.56503009	25.56502972	25.56503044	25.56503118	25.56503217	-8.4	8.4	4.8	11.8	19.1	28.8
1-Oct		25.56503046	25.56503068		25.56502996		25.56503238	-1.1	1.1		-6.0		17.8
20-Oct	M2	25.56506699	25.56506851	25.56506872	25.56506721	25.56507039	25.56506980	-7.5	7.5	9.5	-5.3	25.9	20.1
21-Oct		25.56506847	25.56506937	25.56507180	25.56506924	25.56507216	25.56506988	-4.4	4.4	28.3	3.2	31.8	9.4
22-Oct		25.56506761	25.56506946	25.56507130	25.56507008	25.56507162	25.56507268	-9.1	9.1	27.1	15.2	30.2	40.6
25-Oct		25.56506882	25.56507110	25.56507027	25.56507046	25.56507264	25.56507189	-11.1	11.1	3.0	4.9	26.3	18.9
26-Oct		25.56507157	25.56506850	25.56507159	25.56507139	25.56507286	25.56507152	15.0	-15.0	15.3	13.3	27.7	14.6
27-Oct		25.56506788	25.56506994	25.56507054	25.56506823	25.56506914	25.56507175	-10.1	10.1	15.9	-6.7	2.2	27.8
28-Oct		25.56506682	25.56506846	25.56506886	25.56506862	25.56507081	25.56507233	-8.0	8.0	12.0	9.6	31.1	46.0
29-Oct		25.56506725	25.56506875	25.56506906	25.56506914	25.56507192	25.56507150	-7.4	7.4	10.4	11.2	38.4	34.3
1-Nov		25.56506600	25.56506689	25.56506738	25.56506890	25.56507115	25.56506994	-4.4	4.4	9.2	24.0	46.2	34.3
2-Nov		25.56506949	25.56507049		25.56507129		25.56507111	-4.9	4.9		12.7		11.0
							Average M1	-6.4	6.4	6.2	7.0	19.7	19.6
							Average M2	-5.2	5.2	14.5	8.2	28.9	25.7
							Average M1+M2	**-5.8**	**5.8**	**10.4**	**7.6**	**24.3**	**22.6**

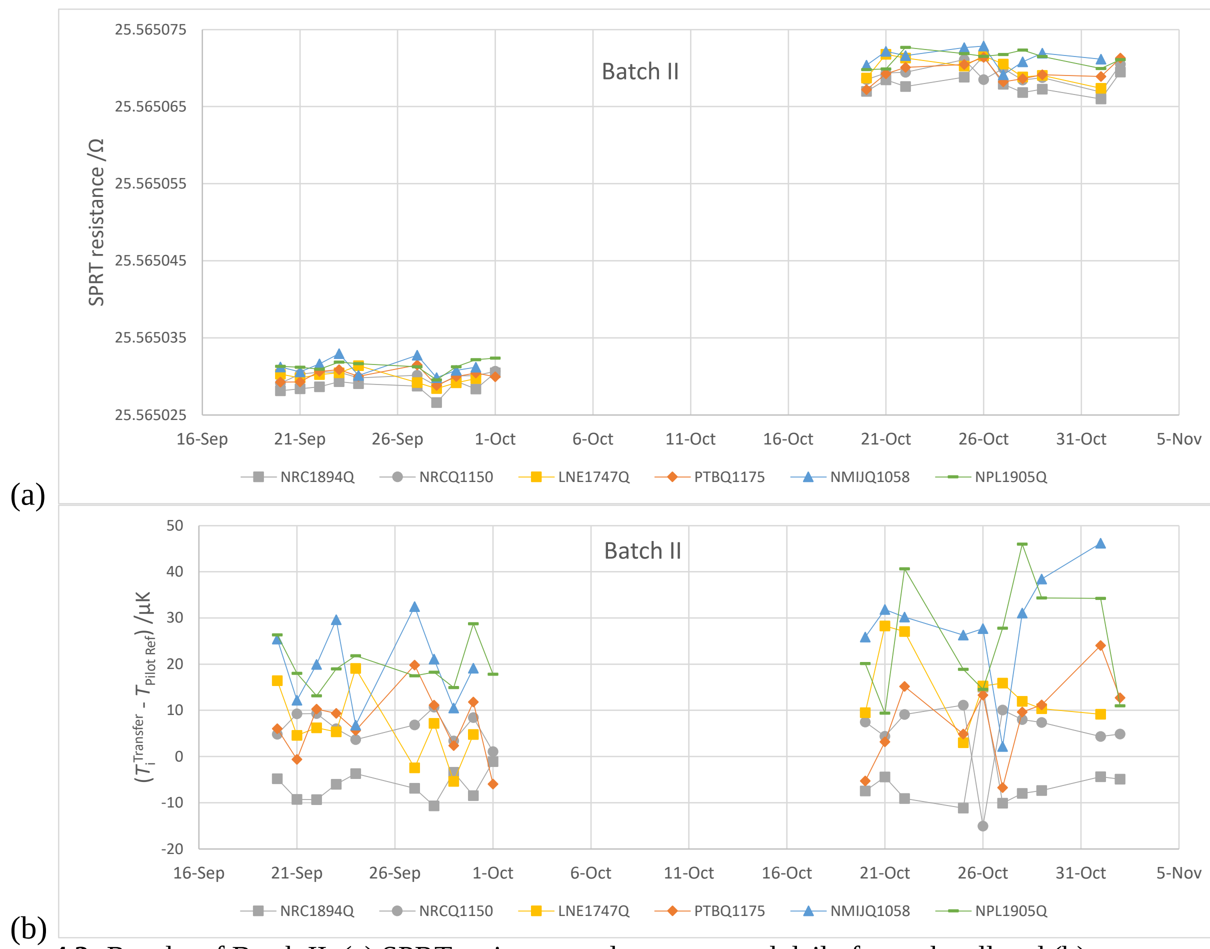


Figure 4.3: Results of Batch II: (a) SPRT resistance values measured daily for each cell and (b) calculated temperature difference between each transfer cell and the pilot reference.

Table 4.5: Results of Batch III.

Date	Mantle	NRC1894Q	NRCQ1150	CEM1833Q	INM14405	KRISSQ1060	NMIA04/02	NMISA1593Q	$T_{NRC1894Q}$ - $T_{Pilot\ Ref}$ /µK	T_{Q1150} - $T_{Pilot\ Ref}$ /µK	$T_{CEM1833Q}$ - $T_{Pilot\ Ref}$ /µK	$T_{INM14405}$ - $T_{Pilot\ Ref}$ /µK	$T_{KRISSQ1060}$ - $T_{Pilot\ Ref}$ /µK	$T_{NMIA04/02}$ - $T_{Pilot\ Ref}$ /µK	$T_{NMISA15933Q}$ - $T_{Pilot\ Ref}$ /µK
12-Nov	M1	25.56510828	25.56510882	25.56510528	25.56510685	25.56511223	25.56510615	25.56510897	-2.7	2.7	-32.0	-16.7	36.1	-23.5	4.2
15-Nov		25.56511046	25.56510917	25.56510375	25.56510538	25.56511242	25.56510653	25.56511023	6.3	-6.3	-59.5	-43.6	25.6	-32.2	4.1
16-Nov		25.56510937	25.56511082	25.56510608	25.56510901	25.56511102	25.56510762	25.56511030	-7.1	7.1	-39.4	-10.6	9.1	-24.3	2.0
17-Nov		25.56510833	25.56511268	25.56510588	25.56510875	25.56511009	25.56510595	25.56510957	-21.3	21.3	-45.4	-17.2	-4.1	-44.7	-9.1
18-Nov		25.56511107	25.56511017	25.56510555	25.56510696	25.56510987	25.56510679	25.56511160	4.4	-4.4	-49.8	-36.0	-7.3	-37.6	9.6
19-Nov		25.56510843	25.56511258	25.56510619	25.56510645	25.56511016	25.56510741	25.56511043	-20.3	20.3	-42.3	-39.8	-3.4	-30.4	-0.7
22-Nov		25.56510941	25.56511192	25.56510554	25.56510623	25.56510923	25.56510567	25.56511112	-12.3	12.3	-50.3	-43.5	-14.1	-49.0	4.5
23-Nov		25.56510951	25.56511039	25.56510698	25.56510649	25.56511047	25.56510638	25.56511140	-4.3	4.3	-29.1	-33.9	5.1	-35.0	14.2
24-Nov		25.56511258	25.56511213		25.56510646	25.56510993	25.56510567	25.56511251	2.2	-2.2		-57.9	-23.8	-65.6	1.5
25-Nov		25.56511044	25.56511218		25.56510840	25.56511004	25.56510628	25.56511276	-8.5	8.5		-28.5	-12.4	-49.4	14.3
9-Dec	M2	25.56510883	25.56511148		25.56510831	25.56511173	25.56510537	25.56510812	-13.0	13.0		-18.1	15.5	-47.0	-20.0
10-Dec		25.56510893	25.56511169	25.56510847	25.56510724	25.56511059	25.56510616	25.56511269	-13.6	13.6	-18.1	-30.1	2.7	-40.7	23.3
13-Dec		25.56510853	25.56511109	25.56510522	25.56510729	25.56511031	25.56510635	25.56511179	-12.6	12.6	-45.0	-24.7	4.9	-34.0	19.4
14-Dec		25.56510899	25.56511068	25.56510765	25.56510778	25.56511213	25.56510607	25.56511158	-8.3	8.3	-21.4	-20.2	22.5	-36.9	17.1
15-Dec		25.56510836	25.56511079	25.56510812	25.56510663	25.56511023	25.56510742	25.56511112	-11.9	11.9	-14.3	-28.9	6.5	-21.1	15.2
16-Dec		25.56510914	25.56510983	25.56510802	25.56510839	25.56511154	25.56510598	25.56511160	-3.4	3.4	-14.3	-10.7	20.2	-34.4	20.8
17-Dec		25.56511001	25.56511154	25.56510933	25.56510862	25.56511118	25.56510675	25.56511177	-7.5	7.5	-14.2	-21.1	3.9	-39.4	9.7
20-Dec		25.56510886	25.56511295	25.56510486	25.56510752	25.56511027	25.56510630	25.56511263	-20.1	20.1	-59.2	-33.2	-6.2	-45.1	16.9
21-Dec		25.56510872	25.56511103	25.56510521	25.56510832	25.56511109	25.56510646	25.56511151	-11.3	11.3	-45.8	-15.3	11.9	-33.6	16.0
22-Dec		25.56510974	25.56511132	25.56510438	25.56510808	25.56511241	25.56510697	25.56511294	-7.7	7.7	-60.3	-24.0	18.5	-34.9	23.7
								Average M1	-6.4	6.4	-43.5	-32.8	1.1	-39.2	4.4
								Average M2	-10.9	10.9	-32.5	-22.6	10.0	-36.7	14.2
								Average M1+M2	**-8.6**	**8.6**	**-38.0**	**-27.7**	**5.6**	**-37.9**	**9.3**

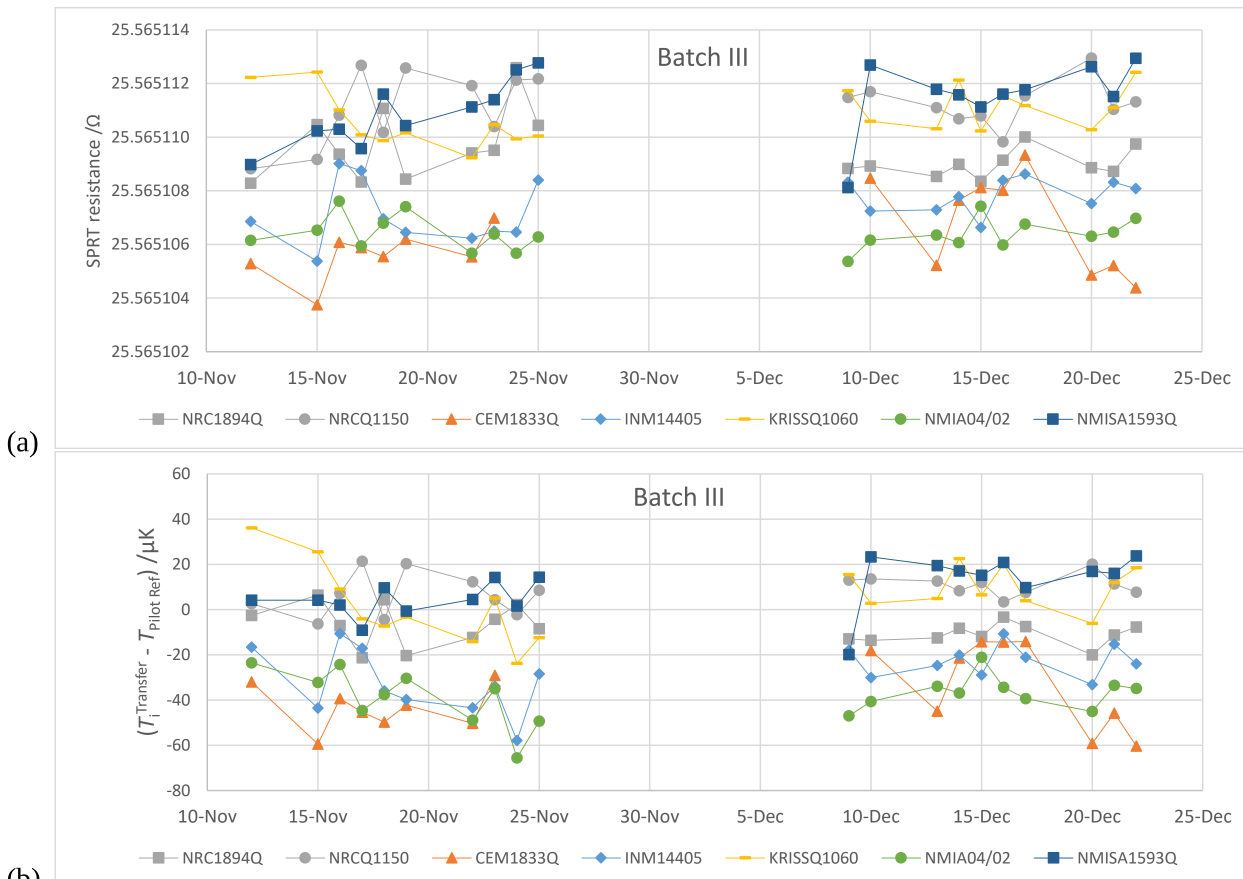
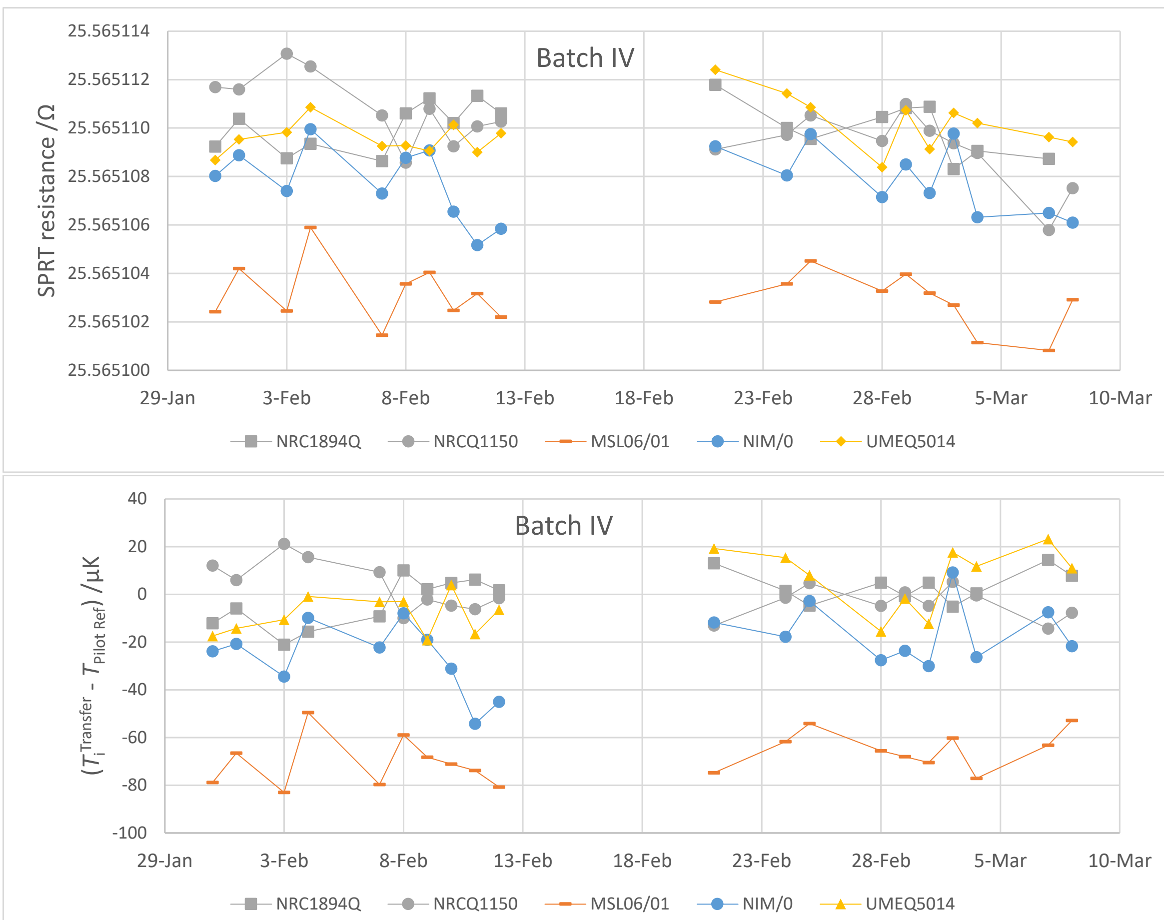


Figure 4.4: Results of Batch III: (a) SPRT resistance values measured daily for each cell and (b) calculated temperature difference between each transfer cell and the pilot reference.

Table 4.6: Results of Batch IV.

Date	Mantle	NRC1894Q	NRCQ1150	MSL06/01	NIM/0	NIST1454Q	UMEQ5014	$T_{NRC1894Q}$ - $T_{Pilot\ Ref}$ /µK	T_{Q1150} - $T_{Pilot\ Ref}$ /µK	$T_{MSL06/01}$ - $T_{Pilot\ Ref}$ /µK	$T_{NIM/0}$ - $T_{Pilot\ Ref}$ /µK	$T_{UMEQ5014}$ - $T_{Pilot\ Ref}$ /µK
31-Jan	M1	25.56510923	25.56511169	25.56510242	25.56510802	25.56511401	25.56510868	-12.1	12.1	-78.9	-23.9	-17.5
1-Feb		25.56511038	25.56511159	25.56510420	25.56510887	25.56511541	25.56510953	-5.9	5.9	-66.6	-20.7	-14.3
3-Feb		25.56510875	25.56511307	25.56510245	25.56510740	25.56511606	25.56510983	-21.1	21.1	-83.0	-34.5	-10.6
4-Feb		25.56510935	25.56511254	25.56510590	25.56510995	25.56511466	25.56511085	-15.6	15.6	-49.6	-9.8	-0.9
7-Feb		25.56510863	25.56511052	25.56510145	25.56510730	25.56511209	25.56510925	-9.3	9.3	-79.7	-22.3	-3.1
8-Feb		25.56511060	25.56510857	25.56510357	25.56510877	25.56511241	25.56510928	10.0	-10.0	-59.0	-8.0	-3.0
9-Feb		25.56511123	25.56511079	25.56510405	25.56510907	25.56511571	25.56510905	2.1	-2.1	-68.3	-19.0	-19.2
10-Feb		25.56511020	25.56510924	25.56510247	25.56510655	25.56511086	25.56511013	4.7	-4.7	-71.1	-31.2	4.0
11-Feb		25.56511133	25.56511007	25.56510317	25.56510517	25.56511416	25.56510900	6.2	-6.2	-73.8	-54.2	-16.6
12-Feb		25.56511060	25.56511027	25.56510220	25.56510585	25.56510959	25.56510978	1.7	-1.7	-80.8	-45.0	-6.5
21-Feb	M2	25.56511178	25.56510912	25.56510282	25.56510925	25.56511264	25.56511240	13.1	-13.1	-74.8	-11.8	19.2
24-Feb		25.56511000	25.56510972	25.56510357	25.56510805	25.56511131	25.56511143	1.4	-1.4	-61.7	-17.8	15.4
25-Feb		25.56510955	25.56511052	25.56510452	25.56510975	25.56511239	25.56511085	-4.7	4.7	-54.1	-2.8	8.0
28-Feb		25.56511045	25.56510947	25.56510327	25.56510715	25.56511181	25.56510838	4.8	-4.8	-65.7	-27.6	-15.5
1-Mar		25.56511083	25.56511099	25.56510397	25.56510850	25.56511184	25.56511073	-0.8	0.8	-68.1	-23.7	-1.8
2-Mar		25.56511088	25.56510989	25.56510319	25.56510732	25.56511414	25.56510913	4.8	-4.8	-70.6	-30.1	-12.3
3-Mar		25.56510830	25.56510937	25.56510269	25.56510977	25.56511116	25.56511063	-5.2	5.2	-60.3	9.2	17.6
4-Mar		25.56510905	25.56510897	25.56510114	25.56510632	25.56511081	25.56511020	0.4	-0.4	-77.2	-26.4	11.7
7-Mar		25.56510873	25.56510579	25.56510082	25.56510650	25.56511006	25.56510963	14.4	-14.4	-63.2	-7.5	23.2
8-Mar		25.56510910	25.56510752	25.56510292	25.56510610	25.56510956	25.56510943	7.8	-7.8	-52.9	-21.7	11.0
							Average M1	-3.9	3.9	-71.1	-26.9	-8.8
							Average M2	3.6	-3.6	-64.9	-16.0	7.6
							Average M1+M2	**-0.2**	**0.2**	**-68.0**	**-21.4**	**-0.6**



(a)

(b)

Figure 4.5: Results of Batch IV: (a) SPRT resistance values measured daily for each cell and (b) calculated temperature difference between each transfer cell and the pilot reference.

Table 4.7: Summary of the results obtained in the comparison of all the transfer cells with the pilot reference (average of the two NRC reference cells 1894Q and Q1150).

Batch	Transfer cell	$(T_i^{Transfer} - T_{Pilot\ Ref})$ /µK	$u\,(T_i^{Transfer} - T_{Pilot\ Ref})$ /µK
I	IPQ2114	-13.5	9.3
	VSL17T048	-22.7	9.3
II	LNE1747Q	10.4	9.3
	PTBQ1175	7.6	9.3
	NMIJQ1058	24.3	9.3
	NPL1905Q	22.6	9.3
III	CEM1833Q	-38.0	9.3
	INM14405	-27.7	9.3
	KRISSQ1060	5.6	9.3
	NMIA04/02	-37.9	9.3
	NMISA1593Q	9.3	9.3
IV	MSL06/01	-68.0	9.3
	NIM/0	-21.4	9.3
	UMEQ5014	-0.6	9.3
All	NRC1894Q	-5.7	9.3

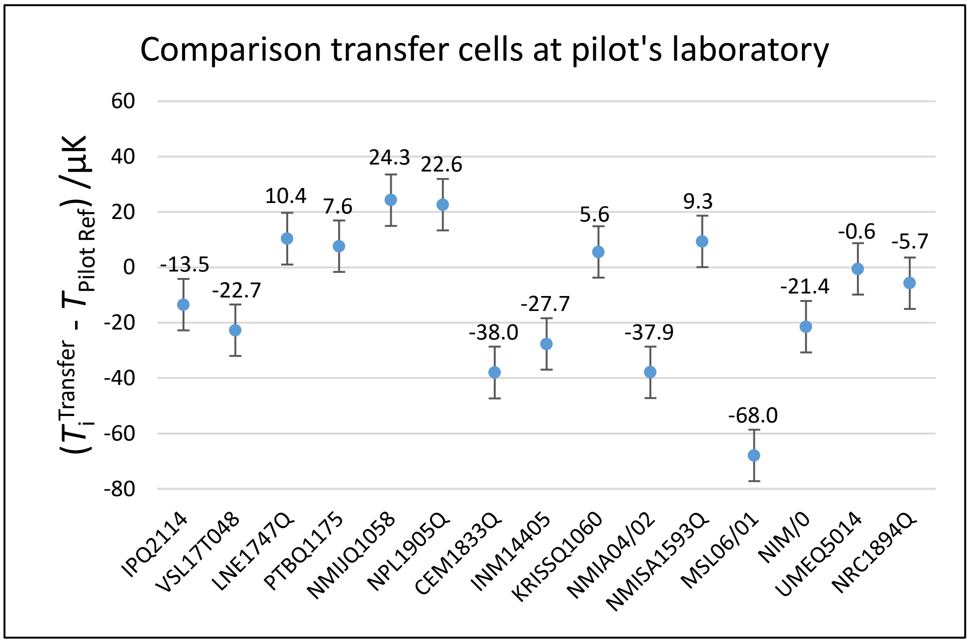


Figure 4.6: Summary of the results obtained in the comparison of all the transfer cells with the pilot reference (average of the two NRC reference cells 1894Q and Q1150). The uncertainty bars are the combined standard uncertainty of 9.3 µK.

5 Combining participants and pilot measurements

In this chapter, the measurements of the differences $T_i^{Transfer} - T_i^{Nat\,Ref}$, performed by the participants in their respective laboratories (see Chapter 3), are combined with the measurements of the differences $T_i^{Transfer} - T_{Pilot\,Ref}$, performed by the pilot (see Chapter 4).

5.1 Combination of participants' results with pilot's results

Table 3.4 provides the temperature differences $T_i^{Transfer} - T_i^{Nat\,Ref}$ and their corresponding uncertainty, as reported by the participants. Table 4.7 provides the temperature differences $T_i^{Transfer} - T_{Pilot\,Ref}$ and their corresponding uncertainty, as reported by the pilot.
The difference between the two simply provides the temperature differences $T_i^{Nat\,Ref} - T_{Pilot\,Ref}$:

$$\left(T_i^{Transfer} - T_{Pilot\,Ref}\right) - \left(T_i^{Transfer} - T_i^{Nat\,Ref}\right) = \left(T_i^{Nat\,Ref} - T_{Pilot\,Ref}\right)$$

The uncertainty in the temperature differences $T_i^{Nat\,Ref} - T_{Pilot\,Ref}$ are calculated by summing in quadrature the uncertainty in $T_i^{Transfer} - T_i^{Nat\,Ref}$ and the uncertainty in $T_i^{Transfer} - T_{Pilot\,Ref}$:

$$u^2\left(T_i^{Nat\,Ref} - T_{Pilot\,Ref}\right) = u^2\left(T_i^{Transfer} - T_i^{Nat\,Ref}\right) + u^2\left(T_i^{Transfer} - T_{Pilot\,Ref}\right)$$

The temperature differences $T_i^{Nat\,Ref} - T_{Pilot\,Ref}$ and their uncertainties are reported numerically in Table 5.1 and graphically in Figure 5.1.

The reference temperature $T_{Pilot\,Ref}$ from which the national reference temperatures $T_i^{Nat\,Ref}$ are expressed in Table 5.1 and Figure 5.1 is just a convenient stable pilot reference temperature, composed of the average of two pilot's reference TPW cells.

Table 5.1: Temperature differences $T_i^{Nat\ Ref} - T_{Pilot\ Ref}$ and their corresponding standard uncertainties.

Participant	$T_i^{Nat\ Ref}$ - $T_{Pilot\ Ref}$ /µK	$u(T_i^{Nat\ Ref}$ - $T_{Pilot\ Ref})$ /µK
CEM	51.7	40.2
INMETRO	-35.1	37.6
IPQ	11.7	82.3
KRISS	0.0	32.4
LNE-Cnam	5.9	50.7
MSL	18.8	13.2
NIM	3.6	33.1
NMIA	-16.2	24.4
NMIJ/AIST	27.1	31.9
NMISA	3.2	62.2
NPL	15.5	24.9
NRC	-2.2	19.8
PTB	11.6	26.7
TUBITAK UME	-59.3	53.8
VSL	-39.9	29.6

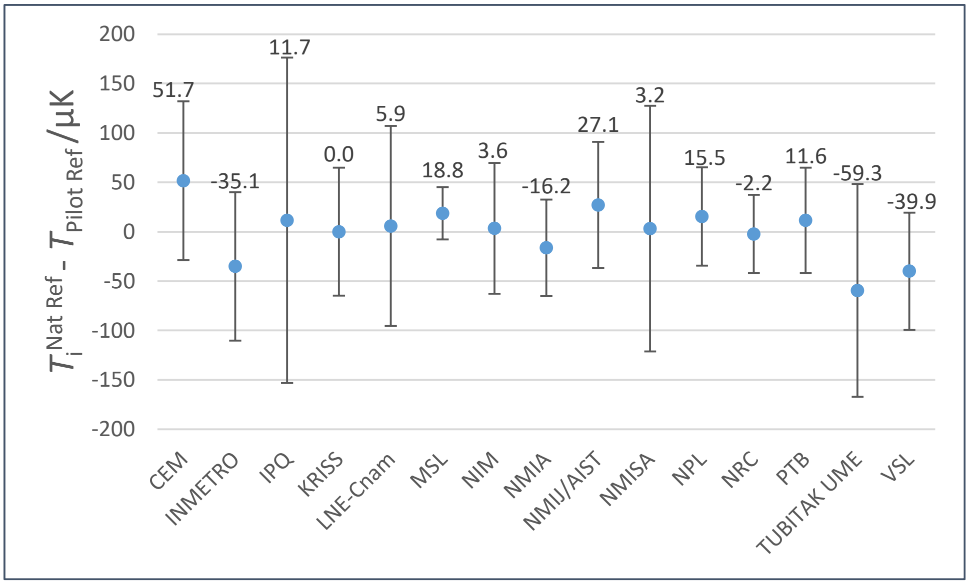


Figure 5.1: Temperature differences $T_i^{Nat\ Ref} - T_{Pilot\ Ref}$ and their corresponding expanded uncertainty bars (k = 2).

6 Statistical analysis and Degrees of Equivalence

6.1 Statistical analysis of the results and selection of the reduction method

The statistical analysis of the results, the determination of the Key Comparison Reference Value (KCRV), and the calculation of the degrees of equivalence (DoEs) were performed with the support of the NIST Decision Tree (NDT) [10].

The NDT is a web application that guides the user through a series of statistical tests (homogeneity, symmetry and normality), intended to help the user decide on an appropriate statistical model for the particular data set input by the user. Once the user, based on the results of the statistical tests, has selected the preferred statistical procedure, the NDT carries out the analysis and displays the results (including KCRV, DoEs and respective uncertainties) via plots, tables and a downloadable report.

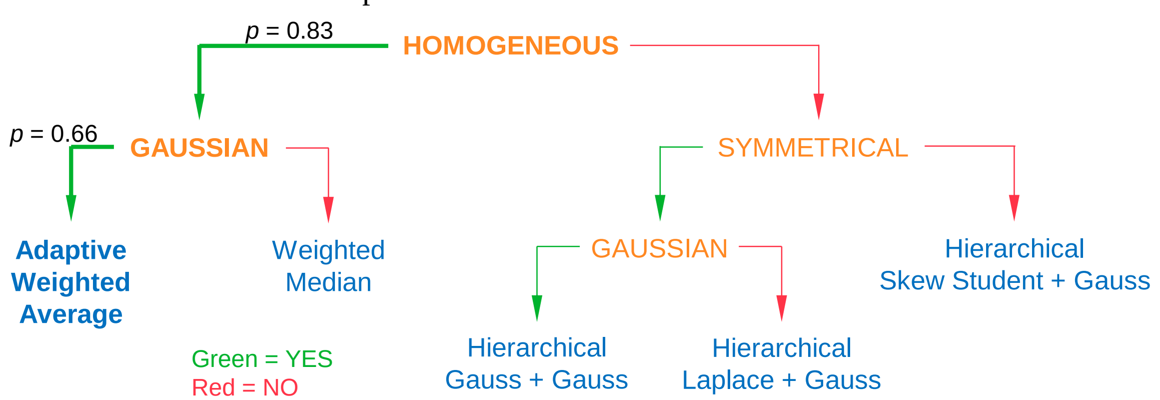


Figure 6.1: The NIST Decision Tree: it comprises 4 branching nodes (orange) and 5 leaves (blue) identifying different models for the measurement and the corresponding procedures for data reduction. A question needs to be answered at each node: if the answer is YES, then one follows the green branch (toward the left); if the answer is NO, then one follows the red branch (toward the right) until one reached a leaf. The bold green arrows show the path taken to traverse the NDT with our data set and the *p*-values above the arrows show the result of the corresponding statistical test.

The NDT initially performs a homogeneity test (see Figure 6.1), which tests the hypothesis of mutual consistency of the input data set, by applying Cochran's chi-squared (or Q) test [11]. On the assumption of homogeneity, the estimator Q of Cochran's test follows a chi-squared distribution with $n - 1$ degrees of freedom (n is the number of participants), which then serves as the reference distribution to compute the test's p-value.

For our specific data set (Table 5.1), the NDT calculated $Q = 8.95$ which, using 14 degrees of freedom ($n = 15$), returned a p-value $p = 0.83$. As we have reasons to doubt mutual consistency of our data set only when $p < 0.05$, we assumed homogeneity.

Looking at Figure 6.1, assuming homogeneity corresponds to choosing to traverse the NDT through the left branch of the tree (green arrow = YES departing from the initial homogeneity test), so the next statistical test the NDT performed was the normality (Gaussian) test, which tests the hypothesis of Gaussian shape for the distribution of the results. The statistical test applied by the NDT to test normality is the Shapiro-Wilk test [12]. The NDT returned a p-value $p = 0.66$. As consequence, we accepted the hypothesis of Gaussian shape.

For completeness, the NDT finally performed the Miao-Gel-Gastwirth test for symmetry [13]. The resulting p-value was 0.36. In fact, there are physical reasons for a slight asymmetry in the results: both isotopic depletion with respect to VSMOW water and chemical impurities in the TPW cells tend to shift the results only to lower value with respect to 273.16 K.

Looking back at Figure 6.1, having concluded that the data satisfy the requirements of homogeneity and normality, the NDT recommends the *adaptive weighted average*, AWA, (DerSimonian-Laird procedure, see [14]) for the calculation of the KCRV, the DoEs, and their respective uncertainties. The method is based on a common-mean model, and, if some heterogeneity is detected (which is not the case here), then it prevents very small laboratory-specific uncertainties from having an excessive influence on the consensus value. More details on the adaptive weighted average method are given in Appendix 5, where it is shown that, for our comparison data, the AWA procedure collapses to the conventional weighted average. This happens because AWA finds that the observed dispersion of the measured values is consistent with the reported uncertainties (i.e., there is no dark uncertainty exposed).

Given the consistency of the data, the uncertainties, and the distribution of the data, the AWA (DerSimonian-Laird) procedure was accepted for the analysis.

6.2 Key Comparison Reference Value and Degrees of Equivalence

With respect to the arbitrary reference adopted in Table 5.1, the adaptive weighted average gave the following KCRV value and standard uncertainty:

$T_{\mathrm{KCRV}} = 4.7\ \mu\mathrm{K}$,
$u_{\mathrm{KCRV}} = 7.2\ \mu\mathrm{K}$.

The DoE, $D_{\mathrm{i}} = T_{\mathrm{i}}^{\mathrm{Nat\ Ref}} - T_{\mathrm{KCRV}}$, for each participant laboratory and the corresponding expanded ($k = 2$) uncertainty $U(T_{\mathrm{i}}^{\mathrm{Nat\ Ref}} - T_{\mathrm{KCRV}})$ are reported numerically in Table 6.1 and graphically in Figure 6.1.

Table 6.1: Degrees of equivalence $D_i = T_i^{Nat\,Ref} - T_{KCRV}$ and corresponding expanded (k = 2) uncertainty $U(T_i^{Nat\,Ref} - T_{KCRV})$ of the participant laboratories. The expanded uncertainties of the D_i take into account the correlations between the measured values and the KCRV.

Participant	$T_i^{Nat\,Ref} - T_{KCRV}$ /µK	$U(T_i^{Nat\,Ref} - T_{KCRV})$ /µK
CEM	47.0	78.1
INMETRO	-39.9	71.3
IPQ	7.0	158.3
KRISS	-4.7	63.2
LNE-Cnam	1.2	97.6
MSL	14.1	25.3
NIM	-1.1	63.0
NMIA	-21.0	46.2
NMIJ/AIST	22.4	60.6
NMISA	-1.5	120.9
NPL	10.8	47.6
NRC	-6.9	36.8
PTB	6.9	50.4
TUBITAK UME	-64.1	104.7
VSL	-44.7	56.3

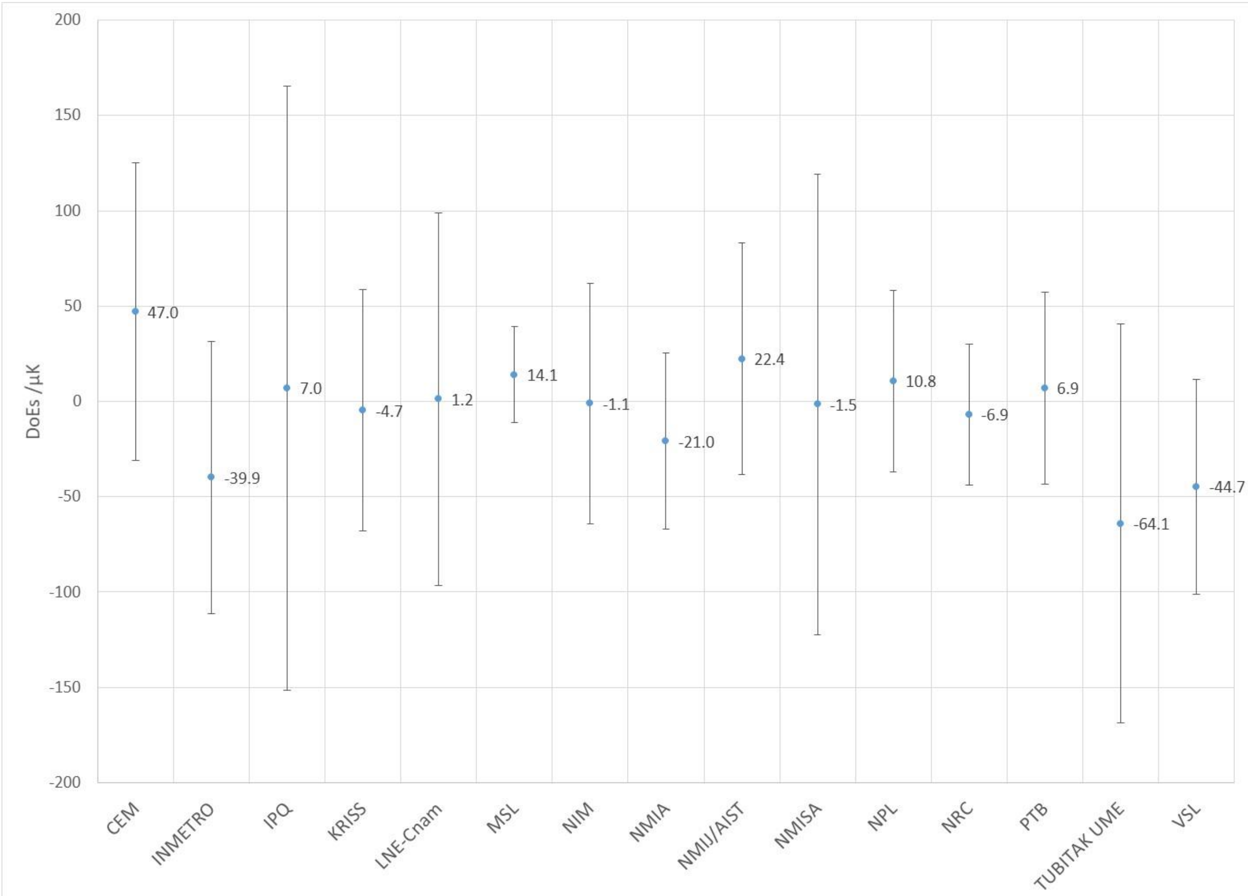


Figure 6.1: Degrees of equivalence of the participants. The zero reference here is the KCRV with standard uncertainty 7.2 µK.

In Appendix 5, the KCRV and DoEs obtained from the application of classical estimators (simple average and conventional weighted average) and from a Bayesian hierarchical method are reported, showing that, for our particular “well-behaved” data set, the choice of the reduction method is not critical and the application of different methods leads to nearly identical results.

In analogy to CCT-K7 report, Appendix 5 reports the pooled distribution of the results. Differently from CCT-K7, the pooled distribution was not bimodal and only slightly asymmetrical.

7 Link between CCT-K7.2021 and CCT-K7

7.1 CCT-K7 cells

An attempt to link the previous CCT-K7 key comparison (2002-2004) to the present CCT-K7.2021 key comparison (2021-2022) was made by asking the participants that still had the cells used in CCT-K7 (either as transfer cells or national reference cells) to measure them against their present national reference.
A total of 7 participants delivered results on cells they used in CCT-K7. The information available on these cells is shown in Table 7.1.

Table 7.1: Information available on the measured CCT-K7 cells.

Participant	Cell S/N	Manufacturer	Year	Envelope material	Cell's role in CCT-K7
CEM	2030	Jarret	1999	Borosilicate	Transfer cell
IPQ	2314	Jarret/Isotech	1999	Borosilicate	Transfer cell
KRISS	2000-5	KRISS	2002	Borosilicate	National Reference
	2002-7	KRISS	2002	Borosilicate	National Reference
NIM	1-004	NIM	2002	Borosilicate	National reference
	1-008	NIM	2002	Borosilicate	National reference and transfer cell
	2000-0025	NIM	2002	Borosilicate	National reference
NMIA	4-75	NMIA	1975	Borosilicate	Transfer cell
NPL	323	Isotech	2003	Borosilicate	Transfer cell
NRC	2063	Jarret/Isotech	2003	Borosilicate	National reference and transfer cell

7.2 Measurement of CCT-K7 cells

The results of the measurement of the CCT-K7 cells, performed by the participants during CCT-K7.2021, are shown in Table 7.2. For easy reference, the results obtained by the same participants in CCT-K7 and in CCT-K7.2021 are also shown in Table 7.2.

Table 7.2: Results obtained in CCT-K7 and in CCT-K7.2021 with the old CCT-K7 borosilicate cells that were still available during CCT-K7.2021:

- Column 4: $T_{\mathrm{Cell}}^{\mathrm{K7}} - T_{\mathrm{NatRef}}^{\mathrm{K7}}$ is the difference between the old borosilicate cell and the CCT-K7 national reference, as measured during CCT-K7;
- Column 5: $T_{\mathrm{NatRef}}^{\mathrm{K7}} - T_{\mathrm{KCRV}}^{\mathrm{K7}}$ is the difference between the CCT-K7 national reference and the CCT-K7 KCRV, as measured during CCT-K7;
- Column 6 (sum of Column 4 and 5): $T_{\mathrm{Cell}}^{\mathrm{K7}} - T_{\mathrm{KCRV}}^{\mathrm{K7}}$ is the difference between the old borosilicate cell and the CCT-K7 KCRV, as measured during CCT-K7;
- Column 7: $T_{\mathrm{Cell}}^{2021} - T_{\mathrm{NatRef}}^{\mathrm{K7.2021}}$ is the difference between the old borosilicate cell and the CCT-K7.2021 national reference, as measured during CCT-K7.2021;
- Column 8: $T_{\mathrm{NatRef}}^{\mathrm{K7.2021}} - T_{\mathrm{KCRV}}^{\mathrm{K7.2021}}$ is the difference between CCT-K7.2021 national reference and the CCT-K7.2021 KCRV, as measured during CCT-K7.2021;
- Column 9 (sum of Column 7 and 8): $T_{\mathrm{Cell}}^{2021} - T_{\mathrm{KCRV}}^{\mathrm{K7.2021}}$ is the difference between the old borosilicate cell and the CCT-K7.2021 KCRV, as measured during CCT-K7.2021;
- Column 10 (difference between Column 9 and 6): cell drift assuming KCRV time-invariant;
- Column 11 (negative of Column 10): KCRV drift assuming cells time invariant.

Participant	Cell	Cell's role in CCT-K7	Results in CCT-K7 (2002-2004)			Results in CCT-K7.2021 (2021-2022)			Cell drift (KCRV time-invariant)	KCRV drift (cells time-invariant)
			$T_{\mathrm{Cell}}^{\mathrm{K7}} - T_{\mathrm{NatRef}}^{\mathrm{K7}}$	$T_{\mathrm{NatRef}}^{\mathrm{K7}} - T_{\mathrm{KCRV}}^{\mathrm{K7}}$	$T_{\mathrm{Cell}}^{\mathrm{K7}} - T_{\mathrm{KCRV}}^{\mathrm{K7}}$	$T_{\mathrm{Cell}}^{2021} - T_{\mathrm{NatRef}}^{\mathrm{K7.2021}}$	$T_{\mathrm{NatRef}}^{\mathrm{K7.2021}} - T_{\mathrm{KCRV}}^{\mathrm{K7.2021}}$	$T_{\mathrm{Cell}}^{2021} - T_{\mathrm{KCRV}}^{\mathrm{K7.2021}}$		
CEM	2030	Transfer cell	88	-36	52	-68.4	47.0	-21.4	-73.4	73.4
IPQ	2114	Transfer cell	61	18	79	-28.7	7	-21.7	-100.7	100.7
KRISS	2000-5	National reference	-	47	-	-114.5*	-4.7	-119.2*	-166.2*	166.2
	2002-7	National reference								
NIM	1-004	National reference	-	11	-	-118.4**	-1.1	-119.5**	-130.5**	130.5
	1-008	National reference and transfer cell								
	2000-0025	National reference								
NMIA	4-75	Transfer cell	-33	-51	-84	-221.1	-21.0	-242.1	-158.1	158.1
NPL	323	Transfer cell	35	23	58	-60.9	10.8	-50.1	-108.1	108.1
NRC	2063	National reference and transfer cell	6.4	62	68.4	-37.2	-6.9	-44.1	-112.5	112.5

*For KRISS the values with asterisk are in fact measured for the average of the two cells (2000-5 and 2002-7) that defined KRISS national reference in CCT-K7.
**For NIM the values with double asterisk are in fact measured for the average of the three cells (1-004, 1-008 and 2000-0025) that defined NIM national reference in CCT-K7.

The last two columns show the relation between CCT-K7 results and CCT-K7.2021 results in two extreme cases: a) assuming that KCRV did not change across CCT-K7 and CCT-K7.2021 and the old borosilicate cells used in CCT-K7 drifted in time (KCRV time-invariant), and b) assuming that the old borosilicate cells used in CCT-K7 did not drift in time and the KCRV drifted (cells time-invariant).
Of course, the truth lies in between the two extreme cases: the KCRV moved up, because nearly all participants applied the isotopic correction in CCT-K7.2021, and the old borosilicate cells drifted down, because of borosilicate dissolution over time (nearly 20 years) [9].

In absence of additional information, it was not possible to separate the two above-mentioned effects by simply using the results of the measurements performed on the 7 available old K7 borosilicate cells.
As the estimate of the link between K7 and K7.2021 was not a primary objective of this comparison (see Technical Protocol, Section 2.2 *Secondary Objectives*), and in order to avoid delays in the publication of the report, it was decided not to pursue further the estimation of the link between the two comparisons.

8 Conclusions

8.1 Achievement of the objectives

The primary objective of this key comparison, namely a comparison of the participant national realizations of the TPW temperature, was achieved. Table 6.1 and Figure 6.1 report the degrees of equivalence and the corresponding uncertainties for the participants of the CCT-K7.2021 key comparison. The final results are reported for 15 out of the 19 initial participants because, as explained in Section 2.4, three participants (CENAM, INRiM and NIST) withdrew from the comparison, as they could not deliver a replacement transfer cell on time (their original transfer cell either got broken during the transportation or showed anomalous behavior during the measurements at NRC), and the results of one participant (VNIIM) were not included in this report because of geopolitical events that occurred in 2022. VNIIM's results could be retrieved in the future to provide a link to CCT-K7.2021 without the need to perform new measurements.

The secondary objectives of the key comparison were: 1) a direct comparison of TPW cells of the highest quality, and 2) to provide a linkage to CCT-K7, mediated by cells that were used in CCT-K7 and still available. Table 4.7 and Figure 4.6 report the results for 15 transfer cells measured in the pilot's laboratory and provide useful information on the state-of-the-art quality of TPW cell manufacturing. We had to abandon the original idea of using TPW cells measured in CCT-K7 to establish a link between the CCT-K7 and CCT-K7.2021 KCRVs, because those cells were all made from borosilicate glass and they drifted substantially over nearly 20 years between the two key comparisons.

8.2 Comparison of the CCT-K7 and CCT-K7.2021 results

The results for degrees of equivalence in Table 6.1 and Figure 6.1 of this report can be compared to Table 21 and Figure 30 in CCT-K7 report. The results for the transfer cells in Table 4.7 and Figure 4.6 of this report can be compared to Table 16 and Figure 26 in the CCT-K7 report.

The differences between CCT-K7 and CCT-K7.2021 results can be summarized as follows:

- In CCT-K7, the pilot's uncertainty in measuring the temperature difference between transfer cells was 12-13 μK ($k = 1$). In CCT-K7.2021, this uncertainty was very similar - 9.2 μK ($k = 1$). In both cases, the detailed discussion of uncertainty budgets has been published [1].
- The largest contributor to the measurement uncertainty in CCT-K7.2021 was stray thermal exchanges, which was estimated from measuring SPRT immersion profile in a TPW cell. In CCT-K7, the average of all slopes measured at the BIPM was 9.9 μK/cm (standard deviation 2.6 μK/cm), while the average of all participants' measurements, excluding the two extreme results (VSL and SPRING), was 9.7 μK/cm (standard deviation 3.5 μK/cm). In CCT-K7.2021, the average of all slopes measured at NRC was

7.6 μK/cm (standard deviation 1.3 μK/cm), while the average of all participants' measurements was 8.6 μK/cm (standard deviation 2.1 μK/cm). The CCT-K7.2021 values lie significantly closer to the theoretical value of 7.3 μK/cm which indicates smaller disturbance to the temperature environment inside the thermometer well during the measurements.

- In CCT-K7, the maximum difference between two transfer cells, all borosilicate cells of high quality as required by the comparison protocol, was 163.2 μK with a standard deviation of 49.5 μK. In CCT-K7.2021, 13 out of the 15 transfer cells were fused silica and only 2 cells were made from borosilicate glass. The maximum difference between two transfer cells was 92.3 μK with a standard deviation of 25.7 μK – almost a factor of two improvement compared to CCT-K7 results.
- In CCT-K7, the maximum difference between two TPW realizations was 171.0 μK, with a standard deviation of 49.7 μK. In CCT-K7.2021 the maximum difference between two TPW realizations was 111.1 μK, with a standard deviation of 28.0 μK – an improvement similar to the one reported for the transfer cells.
- In CCT-K7, only two participants (MSL and NRC) applied corrections for deviations of the isotopic composition from ocean water, represented by V-SMOW. One participant (CSIR) used reference cells which could be expected to be close to ocean water. Due to the two different definitions of the water triple point used, the CCT-K7 results showed a bimodal distribution with the two peaks separated by approximately 100 μK. In CCT-K7.2021, 11 out of 15 laboratories (see Table 3.2) applied isotopic corrections to their national reference cells, resulting in "warmer" national realizations. Differently from CCT-K7, the pooled distribution was no longer bimodal, but only slightly asymmetrical (see Appendix 5).

8.3 Improvements in TPW measurements since the CCT-K7

The three major improvements to TPW measurements since CCT-K7, which were evident in the CCT-K7.2021 results, are: 1) the improved quality of the TPW cells, 2) the improved definition of the national references, and 3) the improved quality of uncertainty assessments.

The improvement in the quality of the TPW cells was manifested in a smaller spread of measured temperature differences between transfer cells (Table 4.7 and Figure 4.6) and temperature stability during the two weeks of measurements for each cell at the pilot lab (Figures 4.2 – 4.5). This improvement is likely due to two major causes:

1) After the clarification of the definition of the kelvin in 2006, the manufacturers of TPW cells started producing cells that were closer to the V-SMOW definition. Some manufacturers achieved this by adding to the cell water appropriate amount of enriched water, to compensate for the depletion of the cell water due to the manufacturing process. As a result, the newly manufactured cells realize TPW temperatures that are closely grouped together.

2) Since approximately 2004, all major commercial TPW manufacturers now offer fused-silica TPW cells with potentially ten times better long-term stability than their borosilicate counterparts (see e.g. [9]). While in CCT-K7 all transfer cells were borosilicate glass, in CCT-K7.2021, 13 out of the 15 transfer cells were fused silica and only 2 borosilicate glass. In CCT-K7.2021, some, but not all, borosilicate transfer cells exhibited temperature drift during two weeks of measurements (similar to the ones reported in Figures 8-9 of the CCT-K7 report). These cells were either replaced with fused silica transfer cells or the participant had decided to withdraw from the key comparison.

One of the major observations in CCT-K7 was the bimodal distribution of the results because 3 labs out of 21 participants had their TPW realization based on V-SMOW water which resulted in "warmer" national realizations (applying isotopic correction) compared to the rest of the participants. As mentioned in Section 2.1, following this observation, CIPM issued a "Clarification of the definition of the kelvin" in 2005 [6] which specified the isotopic composition of the water to be that of V-SMOW. As expected, in CCT-K7.2021 most participants renewed their national reference ensembles with newer, higher quality and fused silica cells and applied the isotopic corrections to their respective national references. This, in turn, lead to: 1) the smaller spread of the national realizations (Table 6.1 and Figure 6.1) and 2) only slightly asymmetric distribution of the results.

Overall, the results of this KC suggest that the temperature community has a very good understanding of the behaviour of TPW cells and associated uncertainties: there are no conspicuous outliers, the overall distribution of the results is entirely consistent with the reported uncertainties (no dark uncertainty) and no uncertainty terms were knowingly omitted.

References

[1] M. Stock et al., “Final Report on CCT-K7 Key Comparison of water triple point cells”, *Metrologia* 2006, **43** Tech. Suppl. 03001

[2] CIMP MRA-D-05 “Measurement Comparisons in the CIMP MRA”, March 2016

[3] A. Peruzzi et al., “Final Report on EUROMET.T-K7: Key comparison of water triple point cells”, *Metrologia,* 2009, **46** Tech. Suppl. 03001

[4] D.T. Iosifovna, K.P. Vladimirovitch, “Final Report of COOMET.T-K7: Regional key comparison of water triple point cells”, *Metrologia,* 2012, **49** Tech. Suppl. 03003

[5] F.S. Tsai et al., “Final Report on APMP.T-K7 key comparison of water triple point cells”, *Metrologia,* 2016, **53** Tech. Suppl. 03004

[6] CGPM, Resolution 10 of the 23rd CGPM (2007): Clarification of the definition of the kelvin, unit of thermodynamic temperature, https://www.bipm.org/en/committees/cg/cgpm/23-2007/resolution-10

[7] Consultative Committee for Thermometry, CCT, “Technical Annex for the International Temperature Scale of 1990 (ITS-90)”, 2017, https://www.bipm.org/utils/en/pdf/MeP_K_Technical_Annex.pdf

[8] A. Peruzzi, S. Dedyulin, “NRC measurement set-up and preparatory work for CCT-K7.2021 key comparison of triple-point-of-water cells,” *Metrologia* **59** (2022) 045011

[9] S. Dedyulin, A. Peruzzi, “No country for old borosilicate triple-point-of-water cells,” *Metrologia* **59** (2022) 055009

[10] A. Possolo, A. Koepke, D. Newton, M.R. Winchester, “Decision Tree for Key Comparisons,” 2021 *J Res Natl Inst Stan* 126:126007

[11] W.G. Cochran (1954) “The combination of estimates from different experiments,” *Biometrics* 10(1):101-129

[12] S.S. Shapiro, M.B. Wilk (1965) “An analysis of variance test for normality,” *Biometrika* 52(3,4):591-611

[13] W. Miao, Y.R. Gel, J.L. Gastwirth (2006) “A new test of symmetry about an unknown median,” Random Walk, Sequential Analysis and Related Topics: A Festschrift in Honor of Yan-Shih Chow, eds A.C. Hsiung, Z, Ying, C.H. Zhang (world Scientific, Singapore), pp 199-214

[14] R. DerSimonian, N. Laird (1986) “Meta-analysis in clinical trials”, *Controlled Clinical Trials* 7(3):177-188

Appendix 1
CIPM key comparison of water-triple-point cells: CCT-K7.2021

Technical Protocol

A.Peruzzi and S. Dedyulin

NRC, National Research Council of Canada, Ottawa

December 2020

1. Introduction

The first CIPM Key Comparison (KC) of water-triple-point (TPW) cells was carried out in 2002-2004 [A1]. During its last meeting in 2017, the Consultative Committee for Thermometry (CCT) decided that a new key comparison of TPW cells should have first priority in the planning of the second cycle of CCT KCs, although the preparatory steps for this KC were postponed until after the redefinition of the kelvin 2019 [A2]. In June 2020 NRC offered to act as pilot of this new CIPM KC. After consultation with the Strategy and Planning Working Group of the CCT and with the Regional Metrology Organizations (RMOs), all CCT members were asked to express their interest in participating in this new CIPM KC. Seventeen CCT members expressed their intention to take part and unanimously accepted NRC as pilot of the comparison.

Although the new definition of the kelvin no longer relies on TPW cells for the realization of the kelvin, TPW cells continue to play a fundamental role in the realization of the ITS-90. Moreover, since the clarification of the definition of the kelvin in 2005 [A3], many national metrology institutes have revised their national reference for the TPW temperature.

This technical protocol has been drawn up by the NRC, in accordance with the CIPM MRA-D-05 (Version 1.6) guidance document [A4]. It takes into account the experience gained in the previous key comparison of water triple point cells [A1], its subsequent regional extensions [A5-A7] and current best practice. All participants of this KC accept the general instructions and commit themselves to follow the procedures described in this technical protocol. Once the protocol and list of participants have been agreed, no change to the protocol or list of participants may be made without prior agreement of all participants.

2. Objectives

2.1 Primary Objective

The primary objective is a comparison, mediated by the participant transfer cells, of the participant national realizations of the TPW temperature.
The participant national realization of the TPW temperature is typically defined as the average of an ensemble of national reference TPW cells but national realizations defined by a single reference cell are acceptable.
While in CCT-K7 only 3 laboratories out of 21 based their TPW realization on VSMOW water, in this KC it is expected that, due to the clarification of the definition of the kelvin in 2005, all the participants will present a TPW national reference based on VSMOW water, in accordance with the "Clarification of the definition of the kelvin" of 2005 [A3] and the "Technical Annex for the International Temperature Scale of 1990 (ITS-90)", revised version of 2017 [A8]. Due to this fact and to the improved measurement capabilities of the participants over the past 20 years, it is expected that the differences between the TPW realizations will be smaller than those observed in CCT-K7 (standard deviation of 50 μK and peak-to-peak difference of 171 μK for a total of 21 participants).

2.2 Secondary Objectives

Although for the purpose of the MRA equivalence only the primary objective described above is necessary, the following two secondary objectives will be pursued in this comparison:

a) A direct comparison of TPW cells (one transfer cell from each participant) of the highest quality.

b) A linkage to the previous key comparison CCT-K7, mediated by cells that were used by the participants in CCT-K7 and that are still available.

Although the secondary objective a) is essentially the mean chosen to achieve the primary objective, it will provide useful information on the state-of-the-art quality of TPW cell manufacturing. It is expected that, due to the improved measurement capabilities of the participants over the past 20 years and the tendency of commercial suppliers, after the clarification of the kelvin definition in 2005, to produce cells that are closer to VSMOW isotopic composition, the differences between the TPW temperatures realized by the transfer cells will be smaller than those observed in CCT-K7 (standard deviation of 50 μK and peak-to-peak difference of 163 μK for a total of 22 cells).

The secondary objective b), similarly to the secondary objective a) should be regarded as another by-product of this KC. Many participants in this KC took part also in CCT-K7 and some of them still have at their disposal cells used in CCT-K7, either as transfer cells or national reference cells. With a moderate additional effort at the participating laboratories, the information on the temperature difference between the TPW realized by these cells and the local old (CCT-K7) and new (CCT-K7.2021) national reference can be obtained. This will allow to relate the CCT-K7 KCRV to the new CCT-K7.2021 KCRV.

2.3 Distinction between national reference cells and transfer cell

National reference cell(s): is ensemble of cells or single cell used to maintain 273.16 K in the country. Accordingly, it should have all practical corrections applied and the total uncertainty reported should include all sources of uncertainty. Each participant is required to submit the uncertainty budget for the national reference as well as a detailed description of the national reference, single/ensemble, isotope corrections or not, fused silica/borosilicate, natural water or spiked, age of cells, nature of isotope corrections generic vs specific to the cell.

Transfer cell: is the single cell sent to NRC. Each participant should report the temperature difference between the temperature realised by this cell and his national reference. The temperature realised by the transfer cell should be corrected for effects associated with the measurement of the temperature difference - self heating and hydrostatic effects - and the uncertainty should include the experimental uncertainties in the measurement of the temperature difference only. Each participant is required to submit the uncertainty budget for the measurement of the temperature difference.

3. Participants and roles

The participant laboratories, the corresponding contact persons and emails are listed in Table 1. NRC, accepted as pilot by all participants, was charged with the organization of the comparison. The pilot laboratory decided to avail itself of a Coordinating Group to support him in the harmonization of the uncertainty budgets, the approach to the methods for analyzing the comparison results and corresponding software tools.

Institute	Country	Contact person	Contact email
CEM	Spain	Dolores del Campo	ddelcampo@cem.es
		Carmen Garcia Izquierdo	mcgarciaiz@cem.es
CENAM	Mexico	Enrique Martines Lopez	emartine@cenam.mx
INMETRO	Brazil	Klaus N. Quelhas	knquelhas@inmetro.gov.br
		Mario A.P. Neto	maneto@inmetro.gov.br
INRiM	Italy	Giuseppina Lopardo	g.lopardo@inrim.it
IPQ	Portugal	Liliana Eusebio	lilianae@ipq.pt
KRISS	South Korea	Inseok Yang	iyang@kriss.re.kr
LNE/CNAM	France	Fernando Sparasci	fernando.sparasci@cnam.fr
MSL	New Zealand	**Rod White**	rod.white@measurement.govt.nz
		Peter Saunders	peter.saunders@measurement.govt.nz
		Farzana Masouleh	Farzana.masouleh@measurement.govt.nz
NIM	China	Xiaoke Yan	yanxk@nim.ac.cn
		Xiaojuan Feng	fengxj@nim.ac.cn
		Jintao Zhang	zhangjint@nim.ac.cn
NIST	United States	Tobias Herman	tobias.herman@nist.gov
		Antonio Possolo	antonio.possolo@nist.gov
NMIA	Australia	Mong-Kim Ho	mong-kim.ho@measurement.gov.au
NMIJ/AIST	Japan	Tohru Nakano	tohru-nakano@aist.go.jp
		Januarius V. Widiatmo	janu-widiatmo@aist.go.jp
		Ikuhiko Saito	saitou.19hiko@aist.go.jp
NMISA	South Africa	Efrem Ejigu	eejigu@nmisa.org
NPL	United Kingdom	Jonathan Pearce	jonathan.pearce@npl.co.uk
NRC	**Canada**	**Andrea Peruzzi**	andrea.peruzzi@nrc-cnrc.gc.ca
		Sergey Dedyulin	sergey.dedyulin@nrc-cnrc.gc.ca
PTB	Germany	Steffen Rudtsch	steffen.rudtsch@ptb.de
UME	Turkey	Murat Kalemci	murat.kalemci@tubitak.gov.tr
		Ali Uytun	ali.uytun@tubitak.gov.tr
VNIIM	Russia	Anatolii Pokhodun	a.i.pokhodun@vniim.ru
VSL	The Netherlands	Conny Bruin-Barendregt	cbarendregt@vsl.nl

Table 1: List of participant laboratories, corresponding contact persons and emails. The pilot laboratory and the members of the Coordinating Group are in bold.

4. Comparison pattern

The pattern of the comparison is a "*collapsed star*", consisting of three phases:

4) Each laboratory selects one of its TPW cells for use as a transfer cell and directly compares it against its TPW national reference.
In case the laboratory still possesses a cell used in CCT-K7, (either as transfer cell or as national reference cell), the laboratory additionally compares the CCT-K7 cell to the national reference.

The selected transfer cell and the measurement results are delivered to NRC.

5) NRC compares all transfer cells against two NRC reference cells.

6) Each laboratory retrieves its transfer cell from NRC and directly re-compare it against its TPW national reference.
In case the laboratory has at its disposal a cell used in CCT-K7 (either as transfer cell or as national reference cell), the laboratory additionally re-compares the CCT-K7 cell to the national reference.

5. Timetable

The timetable of the comparison is the following:

- **April 1st, 2021** Start date: the participants select the transfer cell and compare it against their national reference. The participants disposing of a CCT-K7 cell compare it too against their national reference.
- **June 30th, 2021** Deadline delivery of transfer cell and measurement results (including uncertainty budget) to NRC.
- **July 1st, 2021 to December 31st, 2021** NRC compares all transfer cells.
- **January 1st, 2022** From this day on, the participants can retrieve their transfer cells from NRC.
- **March 31st, 2022** Deadline return measurements at participant laboratory and deliver of the return measurement results (including uncertainty budget) to NRC. The participants disposing of a CCT-K7 cell re-compare it too against their national reference.
- **June 30th, 2022** Deadline preparation of Draft A report.

6. Transfer cell

The transfer cell shall be carefully selected by the participant according to the following criteria:

- The transfer cell shall be of the highest quality and not significantly differing from the quality of the participant national reference cell(s).
- The transfer cell should be preferably a fused silica cell. If a borosilicate cell is used as transfer cell, a cell of recent manufacture is preferable.
- If the quality of a cell is suspect on simple inspection procedures or is known for any kind of abnormal behaviour, it should not be used as transfer cell.
- The following tests shall be made on the cells and shall be repeated at reception of the cells at the pilot laboratory:
    - No floating material shall be visible in the water.
    - For the cells with McLeod gauge or sufficient remnant “seal-off” tube to trap an air bubble, the compression test described in [A9] shall be performed. Prior to testing for air, the cells shall be held vertically at room temperature overnight.

- There shall be a sharp “click” audible if the cell is gently inverted, indicating very low amount of residual air (“water hammer test”).

The pilot reserves the right to reject transfer cells that do not meet the minimum selection criteria when tested on receipt. Laboratories normally using other tests are invited to apply them in addition and to describe them. Laboratories are asked to provide as soon as possible information about the dimensions (in cm) of the chosen cells. This particularly applies to cells with unusual dimensions (for examples, very large or very small cells).

The participant laboratory is free to select and measure an additional cell to keep as a back-up in case problems (or breakage) arise with the transfer cell.

7. Shipment of the transfer cell

The packing of the transfer cell, before the shipment to the pilot laboratory and back to the participating laboratory, falls under the responsibility of the participating laboratory.
The participating laboratory shall select and apply the preferred packing method before shipping the transfer cell to the pilot laboratory, and provide detailed information to the pilot for packing the transfer cell before the return shipping.

One method for packing TPW cells is described in MSL Technical Guide 44 “Shipping TPW Cells” [A10], however such method is suitable only for TPW cells not having a McLeod gauge.

Some guiding principles in packing TPW cells are reported below:

- Use very large wood crates, so that they can be only gently handled by fork lifts
- Use large soft sponge layer to reduce the g forces experienced by the cell during shipping
- The cell should be placed along the diagonal of the crate to prevent water hammer effects.

The participant laboratory is responsible also for:

- The transport of the transfer cell to and from the pilot laboratory.
- Making proper arrangements for customs formalities (e.g. ATA carnet).
- The transport costs, customs charges and any damage that may occur during transport.
- If deemed necessary, taking out insurance for the transport of the transfer cell.

8. Measurement instructions and procedures

Each laboratory must carefully select its transfer cell according to the criteria given in Section 6 and compare it against its national reference (single cell or set of cells). The measurements shall be performed on two separately prepared ice mantles of the transfer cell. The participant is free to use the preferred method for the preparation of the ice mantle but the measurements should not start earlier than 7 days after the preparation of the ice mantle. Depending on the local preparation technique, the minimum waiting time required might be longer than 7 days. A minimum of 10 measurements per mantle (one per day) shall be reported in the appropriate Measurement Report Form. Before each measurement an inner melt shall be induced. The recommended method for inducing the inner melt is the insertion of a room temperature metal or glass rod in the thermometer well for a few seconds. The ice mantle shall then be freely rotating around the well when a gentle rotational impulse is given to it. Apart from this, the measurement procedure shall be the one normally applied by the laboratory.

If a laboratory uses special parts with its transfer cell, like a bushing or a foam pad, this should also be sent to the pilot, together with a short description of its use if necessary.

For each transfer cell, an immersion profile shall be provided, to ensure that the measurement really senses the temperature of the ice/water interface. For each step of the profile, the self-heating correction shall be determined and applied. The step width shall be 1 or 2 cm, and the measurements shall be taken up to about 10 cm below the water surface. The position of the sensor at which the comparison with the reference cell(s) was made shall be indicated.
In case the laboratory has at its disposal a cell used in CCT-K7 cell (either as transfer cell or as national reference cell), the laboratory additionally compare the CCT-K7 cell to its national reference.
The same procedure described above shall be used for the comparison of the CCT-K7 cell to the national reference. Obviously, the comparison of the transfer cell to the national reference and the comparison of the CCT-K7 cell to the national reference can be performed simultaneously.

After its return from the NRC, the full set of measurements described above shall be repeated (including the measurement of the CCT-K7 cell). If the transfer cell is found to be unstable or is broken during the return travel or measurements, this information shall be immediately given to the pilot. In this case the pilot will evaluate the stability of the transfer cell during the measurements at the pilot laboratory. If the transfer cell was stable during the measurements at the pilot laboratory, only the measurements performed by the participant before delivering the transfer cell to the pilot will be used in the analysis of the results.

The pilot shall preliminarily select two reference cells and compare them against its national reference. Subsequently, all transfer cells delivered to the pilot shall be compared against the two selected reference cells. For each transfer cell, similarly to its calibration at the participant institute, a minimum of 10 measurements shall be performed on each of two separately prepared ice mantles. The ice mantles shall be prepared with the technique routinely used by the pilot. The waiting time before starting the measurement after the preparation of the ice mantle shall be at least 7 days. Each measurement day, the two reference cells and a number of transfer cells shall be measured in random order. The number of transfer cells measured daily will depend on the TPW storage capability of the pilot laboratory.

9. Reporting the results

Each laboratory must report the performed measurements by filling the appropriate Measurement Report Form, which is integrant part of this protocol. The Measurement Report Form is an Excel file, composed of 9 sheets:

- In the 1st sheet, named “**Participant**”, the participant laboratory must insert the information on the participant laboratory and the contact person(s).

- In the 2nd sheet, named "**Equipment**", the participant laboratory must insert information on the equipment used for the measurements.

- In the 3rd sheet, named "**Transfer cell**", the participant laboratory must insert information on the selected (and measured) transfer cell.

- In the 4th sheet, named "**CCT-K7 cell**", the participant laboratory must provide detailed information on the CCT-K7 cell (if available).

- In the 5th sheet, named "**National reference**", the participant laboratory must provide detailed information on the national reference.

- In the 6th sheet, named "**Results 1st mantle**", the participant laboratory must provide detailed information on the results of the measurements performed on the 1st mantle of the transfer cell and of the CCT-K7 cell (if available).

- In the 7th sheet, named "**Results 2nd mantle**", the participant laboratory must provide detailed information on the results of the measurements performed on the 2nd mantle of the transfer cell and of the CCT-K7 cell (if available).

- In the 8th sheet, named "**Immersion profile**", the participant laboratory must provide detailed information on the results of the immersion profile measurements performed on the transfer cell.

- In the 9th sheet, named "**Uncertainty budget**", the participant laboratory must provide a detailed uncertainty budget for the calibration of the transfer cell, including the uncertainty components arising from the realization of the national reference. The major uncertainty components are already listed in the excel sheet. The participant laboratory is free to modify the uncertainty budget in the excel sheet. The uncertainty budget must satisfy the following requirements:
  - Compliant with the general rules of the "Guide to the expression of uncertainty in measurement" [A11]
  - In order to avoid double-counting of uncertainty sources, each uncertainty component listed in the budget must specify its physical cause or causes.
  - Related to the previous point, repeatability and reproducibility, without clarifying exactly which physical causes are originating them, are not acceptable as uncertainty sources.

10. Communication flows

The participant laboratory must promptly communicate to the pilot:
- Any unexpected delay that does not allow the participant laboratory to deliver the transfer cell and the measurement results by the deadline reported in the comparison timetable.
- The shipment of the transfer cell to the pilot's laboratory and the expected date of the arrival of the transfer cell at the pilot's laboratory.

- Detailed instructions to the pilot for packing the transfer cell before the return shipping.

The pilot must promptly communicate to the participant laboratory:

- The reception of the transfer cell and any visible damage to the transfer cell
- The procedure in the case of unexpected delay at the participant laboratory

11. Analysis of the key comparison results

Upon the proposal of the Coordinating Group, the participants agreed on the following approach to the analysis of the key comparison results.

It is recognized that both isotope effects and impurity effects in the national reference cells used by the participants generate a one-sided distribution of errors in the national realizations of the TPW temperature.
In order to minimize these effects and obtain a Key Comparison Reference Value (KCRV in the following) that is close to the chemically pure and VSMOW definition, **for the purpose of the KCRV calculation, only national references applying isotopic corrections and including (but not limited to) fused silica cells will be used.**
Although borosilicate cells are known to suffer from much greater leaching and etching effects than fused silica cells, national references including both borosilicate and fused silica cells will be included in the KCRV calculation, as it is assumed that in this case borosilicate cells with relevant impurity content are easily identified and excluded from the national reference ensemble.

For the analysis of the results, an approach based on a statistical model for the measurement results will be adopted, as described in the "Decision Tree for Key Comparison Data Reductions" [A12], which was presented during the organizational meeting for this key comparison (September 28th, 2020).
Reliance on an explicit statistical model allows assessing the fitness of the model to the data, and selection of the most appropriate model.
That Decision Tree includes random effects models of different kinds, all of which are able to recognize, evaluate, and propagate (to the KCRV and to the DoEs), uncertainty contributions in excess of those reported by the participants that become apparent only once independent results are compared (“dark uncertainty”). But it also includes other models, which may prove best when there is no significant dark uncertainty.
The selected model will be validated using established statistical diagnostics, and ultimately will determine the procedure for data reduction in accordance with best statistical practices.

References

[A1] M. Stock et al., “Final Report on CCT-K7 Key Comparison of water triple point cells”, *Metrologia* 2006, **43** Tech. Suppl. 03001

[A2] Consultative Committee for Thermometry, CCT, “*Mise en Pratique* for the definition of the kelvin in the SI”, 20 May 2019, Appendix 2 SI Brochure, https://www.bipm.org/utils/en/pdf/si-mep/SI-App2-kelvin.pdf

[A3] Consultative Committee for Thermometry Recommendation T1 to the CIPM, CCT/05-30(rev) “Clarification of the definition of the kelvin” (2005)

[A4] CIMP MRA-D-05 “Measurement Comparisons in the CIMP MRA”, March 2016 https://www.bipm.org/utils/common/documents/CIPM-MRA/CIPM-MRA-D-05.pdf

[A5] A. Peruzzi et al., “Final Report on EUROMET.T-K7: Key comparison of water triple point cells”, *Metrologia*, 2009, **46** Tech. Suppl. 03001

[A6] D.T. Iosifovna, K.P. Vladimirovitch, “Final Report of COOMET.T-K7: Regional key comparison of water triple point cells”, *Metrologia*, 2012, **49** Tech. Suppl. 03003

[A7] F.S. Tsai et al., “Final Report on APMP.T-K7 key comparison of water triple point cells”, *Metrologia*, 2016, **53** Tech. Suppl. 03004

[A8] Consultative Committee for Thermometry, CCT, “Technical Annex for the International Temperature Scale of 1990 (ITS-90)”, 2017, https://www.bipm.org/utils/en/pdf/MeP_K_Technical_Annex.pdf

[A9] D.R. White, “Measuring the residual air pressure in triple-point-of-water cells,” Meas. Sci. Technol. **15** (2004) N15-N16

[A10] MSL Technical Guide 44 “Shipping TPW Cells”, https://measurement.govt.nz/resources/#collapse-control-1-5

[A11] Joint Committee for Guides in Metrology, “Evaluation of measurement data - Guide to the expression of uncertainty in measurement”, JCGM 100: 2008

[A12] A. Possolo, A. Koepke, D. Newton, M.R. Winchester (2021) "Decision Tree for Key Comparisons" Journal of Research of the National Institute of Standards and Technology **126** 126007

Appendix 2
Participants' reports on the calibration of the transfer cells

A2.1: CEM

First mantle				Second mantle			
Date	$T_{transf.cell} - T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell} - T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
May 4, 2021	-74.6	7.1	5.3	May 25, 2021	-91.4	6.1	6.8
May 5, 2021	-85.0	5.9	5.8	May 26, 2021	-89.6	5.0	5.2
May 6, 2021	-84.1	6.2	7.0	May 27, 2021	-103.4	7.2	5.9
May 7, 2021	-92.8	6.2	6.2	May 28, 2021	-94.4	59.1	5.8
May 10, 2021	-90.9	5.0	7.2	May 31, 2021	-91.8	7.1	6.5
May 11, 2021	-86.7	5.4	7.5	June 1, 2021	-92.1	5.5	6.9
May 12, 2021	-97.7	5.8	5.9	June 2, 2021	-80.9	6.8	7.2
May 13, 2021	-85.3	6.5	4.1	June 3, 2021	-87.8	7.0	6.8
May 14, 2021	-100.7	5.6	6.1	June 4, 2021	-80.6	5.8	6.9
May 17, 2021	-88.3	4.9	5.9	June 7, 2021	-95.5	5.2	5.9
Mean /µK	-88.6	5.8	6.1		-90.7	11.5	6.4
St. dev. Mean /µK	2.3				2.1		
$T_{transf.cell} - T_{nat.ref.}$ /µK	**-89.7**						
$u(T_{transf.cell} - T_{nat.ref.})$	**39.1**						

CEM national reference for the TPW is maintained by an ensemble of 6 cells, 2 of them made of fused silica and with an analysis of isotopic composition. There is no chemical analysis of the impurity content of the cells which are part of the national reference group, but a general statement of the purity of the water. The OME method has been used to estimate the uncertainty component due to impurities. The transfer and/or maintenance of the TPW is performed always in group of 4 cells by using 2 SPRTs. A least squares method is used to assign the cell values. They are internally compared every 5 years. The cells were purchased between 1992 and 2008.

The temperature difference between CEM transfer cell CEM1833Q and CEM national reference, and the corresponding combined uncertainty were reported as:

$T_{CEM1833Q} - T_{nat.ref.} = -89.7\ \mu K$
$u(T_{CEM1833Q} - T_{nat.ref.}) = 39.1\ \mu K$

A2.2: INMETRO

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /μK	Standard deviation of the mean $T_{transf.cell}$ /μK	Standard deviation of the mean $T_{nat.ref.}$ /μK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /μK	Standard deviation of the mean $T_{transf.cell}$ /μK	Standard deviation of the mean $T_{nat.ref.}$ /μK
April 12, 2021	3.5	3.5	4.1	May 25, 2021	-41.3	3.1	3.0
April 13, 2021	6.5	3.9	3.6	May 26, 2021	-4.2	3.1	2.9
April 14, 2021	-9.0	2.8	3.6	May 27, 2021	24.7	3.4	3.8
April 15, 2021	-14.1	4.0	3.9	May 28, 2021	-51.3	3.0	3.8
April 16, 2021	-21.5	3.7	5.9	May 31, 2021	41.4	2.9	3.8
April 19, 2021	46.6	4.1	4.1	June 1, 2021	28.2	3.1	3.3
April 20, 2021	33.9	4.0	3.8	June 2, 2021	40.0	2.9	3.0
April 21, 2021	36.2	3.5	3.1	June 3, 2021	10.7	3.5	2.7
April 22, 2021	21.0	2.6	3.5	June 4, 2021	-13.3	3.4	4.4
April 23, 2021	20.3	3.0	3.3	June 7, 2021	-10.8	3.5	3.0
Mean /μK	12.3	3.5	3.9		2.4	3.2	3.4
St. dev. Mean /μK	7.3				10.2		
$T_{transf.cell}$ - $T_{nat.ref.}$ /μK	**7.4**						
$u(T_{transf.cell} - T_{nat.ref.})$	**36.4**						

INMETRO national reference is defined by a single fused silica cell (Isotech model A11-50-270Q, s/n A11-50-1766Q) purchased in 2019. The isotopic composition of the water is: δ^2H = 8.84 ‰ and δ^{18}O = -2.45 ‰. No impurity analysis is available.

The temperature difference between INMETRO transfer cell INM14405 and INMETRO national reference, and the corresponding combined uncertainty were reported as:

$T_{\mathrm{INM14405}} - T_{\mathrm{nat.ref.}} = 7.4$ μK
$u(T_{\mathrm{INM14405}} - T_{\mathrm{nat.ref.}}) = 36.4$ μK

A2.3: IPQ

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
May 3, 2021	-5.7	3.2	2.7	May 31, 2021	-102.9	2.4	2.7
May 4, 2021	-25.0	2.6	2.4	June 1, 2021	92.6	2.4	2.7
May 5, 2021	-82.7	2.5	2.4	June 14, 2021	-25.7	2.4	2.4
May 6, 2021	-18.1	2.8	2.4	June 15, 2021	-71.7	3.2	2.7
May 7, 2021	-5.0	2.8	2.4	June 16, 2021	-16.6	2.1	1.9
May 10, 2021	-38.1	2.3	2.3	June 17, 2021	0.7	2.2	1.8
May 11, 2021	-110.8	3.6	2.3	June 18, 2021	-31.3	2.6	2.5
May 12, 2021	-38.2	3.3	2.9	June 21, 2021	-13.3	2.5	2.3
May 13, 2021	-48.0	2.4	2.7	June 22, 2021	-21.1	2.1	2.6
May 14, 2021	93.0	2.7	2.6	June 23, 2021	-34.9	2.5	2.5
Mean /µK	-27.9	2.8	2.5		-22.4	2.4	2.4
St. dev. Mean /µK	17.1				16.0		
$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	**-25.1**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**81.8**						

IPQ national reference is defined by a single borosilicate cell (Isotech model Jarret JA3, s/n A11-50-542) purchased in 2000. The isotopic composition of the water is:
δ^2H = 35.07 ‰ and δ^{18}O = 3.57 ‰. No impurity analysis is available.
The temperature difference between IPQ transfer cell IPQ2114 and IPQ national reference, and the corresponding combined uncertainty were reported as:

$T_{IPQ2114} - T_{nat.ref.}$ = -25.1 µK
$u(T_{IPQ2114} - T_{nat.ref.})$ = 81.8 µK

A2.4: KRISS

First mantle				Second mantle			
Date	**$T_{transf.cell}$ - $T_{nat.ref.}$ /µK**	**Standard deviation of the mean $T_{transf.cell}$ /µK**	**Standard deviation of the mean $T_{nat.ref.}$ /µK**	**Date**	**$T_{transf.cell}$ - $T_{nat.ref.}$ /µK**	**Standard deviation of the mean $T_{transf.cell}$ /µK**	**Standard deviation of the mean $T_{nat.ref.}$ /µK**
March 25, 2021	-7.7	2.5	2.3	May 4, 2021	-4.6	2.5	2.3
March 26, 2021	5.0	1.8	2.5	May 6, 2021	11.0	2.0	2.7
March 29, 2021	17.5	1.9	2.4	May 7, 2021	10.8	2.3	2.3
March 30, 2021	20.4	2.0	2.1	May 10, 2021	-10.5	1.5	2.0
April 1, 2021	18.3	2.4	1.9	May 11, 2021	-9.3	2.6	2.2
April 6, 2021	-9.3	1.9	2.9	May 12, 2021	10.5	3.0	2.6
April 8, 2021	3.8	2.4	2.5	May 13, 2021	15.6	2.4	2.8
April 9, 2021	8.6	1.9	2.5	May 14, 2021	8.9	3.1	2.5
April 12, 2021	13.9	2.1	2.3	May 17, 2021	5.2	2.3	2.6
April 13, 2021	7.7	2.2	1.9	May 18, 2021	17.6	2.7	2.8
April 14, 2021	3.1	2.2	1.8	May 19, 2021	3.8	2.1	2.3
April 15, 2021	-4.8	2.0	2.3	May 20, 2021	-1.9	2.5	2.3
Mean /µK	6.4	2.1	2.3		4.8	2.4	2.5
St. dev. Mean /µK	2.9				2.7		
$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	**5.6**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**31.0**						

KRISS national reference is defined by the simple average of three cells (KRISS 2000-5, manufactured in 2002, Fluke 1057, manufactured in 2019, and Isotech 1680Q, manufactured in 2019), internally intercompared every two years. The isotopic composition of the cells is measured and corrected for.

The temperature difference between KRISS transfer cell KRISSQ1060 and KRISS national reference, and the corresponding combined uncertainty were reported as:

$T_{\text{KRISSQ1060}} - T_{\text{nat.ref.}} = 5.6\ \mu\text{K}$
$u(T_{\text{KRISSQ1060}} - T_{\text{nat.ref.}}) = 31.0\ \mu\text{K}$

A2.5: LNE-Cnam

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
February 18, 2021	11.7	2.3	1.6	March 25, 2021	-8.8	1.4	2.2
February 19, 2021	41.7	1.6	2.4	March 26, 2021	-7.8	0.9	1.5
February 22, 2021	33.4	1.2	2.4	March 29, 2021	-19.0	1.7	1.5
February 23, 2021	41.2	1.3	2.4	March 30, 2021	22.1	1.2	1.7
February 25, 2021	44.9	1.2	2.4	April 1, 2021	-17.1	1.5	1.1
February 26, 2021	31.7	1.5	2.3	April 2, 2021	-26.8	0.9	1.7
March 1, 2021	29.1	1.3	2.3	April 6, 2021	-33.8	1.1	1.6
March 2, 2021	-9.0	1.0	2.9	April 7, 2021	-33.8	1.1	1.3
March 4, 2021	-22.0	1.1	2.7	April 8, 2021	-3.6	0.9	1.3
March 5, 2021	-8.9	1.8	2.6	April 9, 2021	24.4	1.1	1.3
Mean /µK	19.4	1.4	2.4		-10.4	1.2	1.5
St. dev. Mean /µK	7.8				6.5		
$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	**4.5**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**49.8**						

LNE-Cnam national reference is defined as a group of four cells, three fused silica (ages 2, 7 and 12 years) and one borosilicate (age 10 years). They are internally compared every three years. The national reference is a weighted average of the four temperatures realized by the four cells. The isotopic composition of two cells of the national reference group has been determined by the cell manufacturer. For the other two cells, the isotopic composition has been determined by an independent laboratory through the analysis of a water sample provided by the manufacturer.

The temperature difference between LNE-Cnam transfer cell LNE1747 and LNE-Cnam national reference, and the corresponding combined uncertainty were reported as:

$T_{\mathrm{LNE1747}} - T_{\mathrm{nat.ref.}} = 4.5\ \mu\mathrm{K}$
$u(T_{\mathrm{LNE1747}} - T_{\mathrm{nat.ref.}}) = 49.8\ \mu\mathrm{K}$

A2.6: MSL

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
June 21, 2021	-84	17	15	2021-12-06	-100	22	15
June 22, 2021	-86	18	15	2021-12-07	-136	23	15
June 23, 2021	-125	18	15	2021-12-08	-72	23	15
June 24, 2021	-65	18	15	2021-12-09	-109	23	15
June 24, 2021	-62	18	15	2021-12-10	-68	24	15
June 25, 2021	-94	18	15	2021-12-13	-78	24	15
June 28, 2021	-88	18	15	2021-12-14	-118	23	15
June 28, 2021	-93	18	15	2021-12-15	-66	23	15
June 28, 2021	-66	18	15	2021-12-16	-76	23	15
June 28, 2021	-92	18	15	2021-12-17	-87	22	15
Mean /µK	-84	17.9	15.0		-90	23.0	15.0
St. dev. Mean /µK	5.9				7.5		
$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	**-86.8**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**9.4**						

MSL national reference is composed of an ensemble of two cells (MSL01/4 in borosilicate glass and MSL06/01 in fused silica) and defined as the GLS weighted mean of all the measurements made with these two cells. The two cells are corrected for their isotopic composition (41.9 µK and 68.0 µK, respectively) and for their impurity content (21.4 µK and 15.4 µK, respectively).

The temperature difference between MSL transfer cell MSL06/01 and MSL national reference, and the corresponding combined uncertainty were reported as:

$T_{MSL06/01} - T_{nat.ref.}$ = -86.8 µK
$u(T_{MSL06/01} - T_{nat.ref.})$ = 9.4 µK

A2.7: NIM

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
September 19, 2021	-30	7.2	8.0	October 10, 2021	-18.0	8.0	8.0
September 20, 2021	-29	9.0	9.0	October 11, 2021	-19.2	8.0	7.0
September 21, 2021	-36	8.0	5.0	October 12, 2021	-28.5	7.0	7.0
September 22, 2021	-33	9.0	7.0	October 13, 2021	-24.7	8.0	5.0
September 23, 2021	-11	9.0	7.0	October 14, 2021	-19.6	9.0	7.0
September 24, 2021	-31	5.6	8.0	October 15, 2021	-28.4	7.0	8.0
September 25, 2021	-29	8.0	9.0	October 16, 2021	-17.8	8.0	9.0
September 26, 2021	-34	5.4	9.0	October 17, 2021	-21.1	6.0	7.0
September 27, 2021	-23	7.0	5.0	October 18, 2021	-16.0	5.0	5.0
September 28, 2021	-28	8.0	9.0	October 19, 2021	-20.9	7.8	7.0
Mean /µK	-28.6	7.6	7.6		-21.4	7.4	7.0
St. dev. Mean /µK	2.2				1.4		
$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	**-25.0**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**31.8**						

NIM national reference is defined by the average of a group of three cells (Q5, Q9 and Q18). The three cells were manufactured in 2020 and are internally intercompared every 1.5 years. The isotopic composition of the three cells is:
Q5: δ^2H = -47.6 ‰, δ^{18}O = -7.1 ‰ and δ^{17}O = -3.5 ‰
Q9: δ^2H = 32.0 ‰, δ^{18}O = 2.4 ‰ and δ^{17}O = 1.5 ‰
Q18: δ^2H = -8.4 ‰, δ^{18}O = -2.3 ‰ and δ^{17}O = -0.9 ‰
No impurity analysis is available.

The temperature difference between NIM transfer cell NIM/0 and NIM national reference, and the corresponding combined uncertainty were reported as:

$T_{NIM/0} - T_{nat.ref.}$ = -25.0 µK
$u(T_{NIM/0} - T_{nat.ref.})$ = 31.8 µK

A2.8: NMIA

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
February 8, 2021	-36.0	2.8	2.5	April 12, 2021	-18.2	2.1	1.8
February 9, 2021	25.5	3.5	3.1	April 13, 2021	-25.6	1.0	1.9
February 10, 2021	-5.4	3.0	2.9	April 14, 2021	-26.9	2.9	3.0
February 11, 2021	-30.2	3.1	2.9	April 15, 2021	-70.5	2.8	3.4
February 12, 2021	-30.0	3.3	3.2	April 16, 2021	-44.2	2.7	3.7
February 15, 2021	-24.2	3.3	3.3	April 19, 2021	-69.9	3.4	3.0
February 16, 2021	4.8	2.8	3.2	April 20, 2021	-17.8	4.4	3.5
February 17, 2021	-15.4	5.3	3.0	April 21, 2021	4.5	3.1	4.6
February 18, 2021	-57.9	3.2	2.6	April 22, 2021	1.6	3.6	3.2
February 19, 2021	-4.2	3.7	2.5	April 23, 2021	6.1	2.7	3.4
Mean /µK	-17.3				-26.1		
St. dev. Mean /µK	7.5				8.9		
$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	**-21.7**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**22.6**						

NMIA national reference is defined by a group of 8 cells, 7 fused silica and 1 borosilicate purchased at difference times. The first 5 cells were purchased in 2006, the next two in 2010 and the latest in 2020. The national reference is defined by the average of the ensemble of cells.
The temperature difference between NMIA transfer cell MSL04/2 and NMIA national reference, and the corresponding combined uncertainty were reported as:

$T_{NIM/0} - T_{nat.ref.}$ = -21.7 µK
$u(T_{NIM/0} - T_{nat.ref.})$ = 22.6 µK

A2.9: NMIJ/AIST

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
May 13, 2021	-3.4	1.2	2.3	June 24, 2021	-8.2	1.9	7.7
May 14, 2021	-4.0	2.9	1.2	June 25, 2021	0.5	2.7	1.4
May 17, 2021	-7.1	1.1	1.1	June 28, 2021	0.2	1.5	1.9
May 18, 2021	9.7	1.8	1.2	June 29, 2021	-4.9	1.5	1.5
May 20, 2021	-1.5	1.4	1.7	June 30, 2021	-19.5	1.3	3.7
May 21, 2021	-4.3	6.7	1.0	July 1, 2021	-9.8	1.4	1.4
May 24, 2021	22.7	1.1	1.2	July 2, 2021	-17.1	1.4	1.8
May 25, 2021	5.1	2.1	1.1	July 5, 2021	-4.3	1.6	1.6
May 26, 2021	12.0	1.0	1.1	July 6, 2021	-1.9	1.5	1.5
May 27, 2021	-9.7	1.1	1.1	July 7, 2021	-11.9	3.5	1.3
Mean /µK	2.0	2.0	1.3		-7.7	1.8	2.4
St. dev. Mean /µK	3.2				2.2		
$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	**-2.9**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**30.5**						

NMIJ/AIST national reference is defined by the average of a group of three fused silica cells, corrected for their isotopic composition: D-Q1008, manufactured in 2005, D-Q1103, manufactured in 2012, and D-Q1176, manufactured in 2021. The cells are internally intercompared every 5 years.

The isotopic composition of the three cells is:
D-Q1008: δ^2H = +1.7 ‰, $\delta^{18}O$ = -0.2 ‰ and $\delta^{17}O$ = -0.1 ‰
D-Q1103: δ^2H = -5.7 ‰, $\delta^{18}O$ = -1.3 ‰ and $\delta^{17}O$ = -0.7 ‰
D-Q1176: δ^2H = 1.6 ‰, $\delta^{18}O$ = -0.3 ‰ and $\delta^{17}O$ = -0.1 ‰
No impurity analysis is available.

The temperature difference between NMIJ/AIST transfer cell NMIJQ1058 and NMIJ/AIST national reference, and the corresponding combined uncertainty were reported as:

$T_{NMIJQ1058} - T_{nat.ref.}$ = -2.9 µK
$u(T_{NMIJQ1058} - T_{nat.ref.})$ = 30.5 µK

A2.10: NMISA

First mantle				Second mantle			
Date	**$T_{transf.cell}$ - $T_{nat.ref.}$ /μK**	**Standard deviation of the mean $T_{transf.cell}$ /μK**	**Standard deviation of the mean $T_{nat.ref.}$ /μK**	**Date**	**$T_{transf.cell}$ - $T_{nat.ref.}$ /μK**	**Standard deviation of the mean $T_{transf.cell}$ /μK**	**Standard deviation of the mean $T_{nat.ref.}$ /μK**
March 2, 2021	21.0	25.4	27.6	March 24, 2021	-6.0	28.0	26.1
March 3, 2021	-7.0	28.3	29.2	March 25, 2021	4.5	28.2	28.0
March 4, 2021	-6.2	28.7	27.7	March 26, 2021	-7.3	21.0	26.3
March 5, 2021	-8.3	29.5	28.7	March 29, 2021	21.9	25.5	26.7
March 8, 2021	53.8	18.3	28.1	March 30, 2021	2.9	27.8	26.3
March 9, 2021	-11.2	28.6	28.8	March 31, 2021	16.9	23.5	24.2
March 10, 2021	71.2	24.8	27.4	April 1, 2021	40.6	21.9	25.4
March 11, 2021	19.0	27.7	29.4	April 6, 2021	-23.4	20.3	23.7
March 12, 2021	-23.9	29.4	34.1	April 7, 2021	-43.2	21.4	21.9
March 15, 2021	31.0	27.9	29.7	April 8, 2021	-22.9	24.6	22.5
Mean /μK	13.9	26.9	29.1		-1.6	24.2	25.1
St. dev. Mean /μK	9.8				7.8		
$T_{transf.cell}$ - $T_{nat.ref.}$ /μK	**6.2**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**61.5**						

NMISA national reference is defined by the simple mean of two Jarrett model A11 borosilicate cells (s/n 2035 and s/n 2048, purchased in 1998).

Isotopic composition and impurity content of the two cells are unknown (and not corrected for).

The temperature difference between NMISA transfer cell NMISA1593Q and NMISA national reference, and the corresponding combined uncertainty were reported as:

$T_{NMISA1593Q} - T_{nat.ref.} = 6.2$ μK
$u(T_{NMISA1593Q} - T_{nat.ref.}) = 61.5$ μK

A2.11: NPL

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
March 22, 2021	6.6	2.8	3.5	April 19, 2021	33.5	2.2	3.0
March 23, 2021	-27.5	3.8	3.8	April 20, 2021	11.0	3.2	2.9
March 24, 2021	-2.8	2.0	4.6	April 21, 2021	-17.8	2.8	3.8
March 25, 2021	14.8	2.8	4.0	April 22, 2021	-2.7	2.5	3.8
March 29, 2021	7.6	4.3	4.3	April 26, 2021	11.5	3.1	3.8
March 30, 2021	35.0	2.8	4.7	April 27, 2021	7.4	2.3	3.3
March 31, 2021	31.6	3.1	3.0	April 28, 2021	13.3	2.8	3.0
April 1, 2021	8.2	3.1	3.7	April 29, 2021	15.0	2.5	2.7
April 6, 2021	-35.2	4.5	2.8	April 30, 2021	23.1	2.4	4.4
April 7, 2021	22.6	2.0	3.3	May 4, 2021	-12.5	3.6	3.0
Mean /µK	6.1	3.1	3.8		8.2	2.7	3.4
St. dev. Mean /µK	7.3				4.9		
$T_{transf.cell}$ - $T_{nat.ref.}$ /µK	**7.1**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**23.1**						

NPL national reference is defined by the mean of a group of 5 cells (1 fused silica cell and 4 borosilicate cells). The cells were manufactured between 2003 and 2016 by 3 different suppliers. The cells are internally intercompared every 2 years. The isotopic composition is available and corrected for.

The temperature difference between NPL transfer cell NPL1905Q and NPL national reference, and the corresponding combined uncertainty were reported as:

$T_{\mathrm{NPL1905Q}} - T_{\mathrm{nat.ref.}} = 7.1\ \mu\mathrm{K}$
$u(T_{\mathrm{NPL1905Q}} - T_{\mathrm{nat.ref.}}) = 23.1\ \mu\mathrm{K}$

A2.12: NRC

First mantle				Second mantle			
Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /μK	Standard deviation of the mean $T_{transf.cell}$ /μK	Standard deviation of the mean $T_{nat.ref.}$ /μK	Date	$T_{transf.cell}$ - $T_{nat.ref.}$ /μK	Standard deviation of the mean $T_{transf.cell}$ /μK	Standard deviation of the mean $T_{nat.ref.}$ /μK
May 3, 2021	11.3	0.4	0.3	May 25, 2021	-5.9	0.2	0.6
May 4, 2021	7.4	0.3	0.3	May 26, 2021	-15.4	0.4	0.3
May 5, 2021	2.2	0.4	0.3	May 27, 2021	-9.8	0.4	0.4
May 6, 2021	-21.3	0.3	0.2	May 28, 2021	-14.4	0.4	0.4
May 7, 2021	19.1	0.3	0.2	May 31, 2021	-29.9	0.2	0.3
May 10, 2021	-12.2	0.4	2.0	June 1, 2021	-1.2	0.3	0.6
May 11, 2021	9.3	0.4	0.3	June 2, 2021	5.7	0.2	0.4
May 12, 2021	3.0	0.4	0.4	June 7, 2021	-16.9	0.5	0.4
May 13, 2021	17.2	0.3	2.0				
Mean /μK	4.0	0.3	0.7		-11.0	0.3	0.4
St. dev. Mean /μK	4.4				3.4		
$T_{transf.cell}$ - $T_{nat.ref.}$ /μK	**-3.5**						
u ($T_{transf.cell}$ - $T_{nat.ref.}$)	**17.5**						

The NRC national reference is defined as the average of an ensemble of 10 fused silica cells, each corrected for its measured isotopic composition. The cells were manufactured between 2003 and 2020 by three different manufacturers.

The impurity content was estimated by ICPMS analysis on water samples from a subset of the ensemble. The ICPMS analysis was focused on the 8 elements that are known to constitute the large majority of impurities found in TPW cells waters (B, Na, Mg, Al, Si, K, Ca and Fe). The standard uncertainty arising from impurities, evaluated using the overall maximum estimate (OME) method, is 11 μK.

The temperature difference between NRC transfer cell NRC1894Q and NRC national reference, and the corresponding combined uncertainty were reported as:

$T_{\mathrm{NRC1894Q}} - T_{\mathrm{nat.ref.}} = -3.5$ μK
$u(T_{\mathrm{NRC1894Q}} - T_{\mathrm{nat.ref.}}) = 17.5$ μK

A2.13: PTB

First mantle				Second mantle			
Date	$T_{transf.cell} - T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell} - T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
May 5, 2021	13.2	2.19	2.78	June 22, 2021	-16.4	2.2	1.5
May 6, 2021	26.0	1.64	1.99	June 23, 2021	-2.6	2.2	2.0
May 7, 2021	-4.6	2.15	1.90	June 24, 2021	7.3	4.1	2.0
May 11, 2021	9.2	1.41	1.93	June 25, 2021	-8.5	1.2	3.7
May 12, 2021	-9.5	1.82	2.49	June 29, 2021	-17.4	1.6	2.4
May 31, 2021	-1.6	2.64	1.88	June 30, 2021	-15.4	2.1	2.3
June 1, 2021	10.2	2.04	3.12	July 7, 2021	-7.5	1.8	1.8
June 2, 2021	-23.3	2.07	2.58	July 21, 2021	6.28	1.5	2.0
June 4, 2021	5.3	1.30	1.57	July 22, 2021	-12.45	1.8	1.7
June 7, 2021	-22.3	1.73	2.09	July 23, 2021	-15.41	1.7	1.1
Mean /µK	0.3	1.9	2.2		-8.2	2.0	2.0
St. dev. Mean /µK	5.0				2.9		
$T_{transf.cell} - T_{nat.ref.}$ /µK	**-4.0**						
$u(T_{transf.cell} - T_{nat.ref.})$	**25.0**						

PTB national reference is defined by the mean of a group of 3 fused silica cells (5901D-Q 1041 (2007), 5901D-Q 1042 (2007), Fluke 5901-D-Q 1179 (2021)), corrected for their isotopic composition.
These cells are internally intercompared at least once in five years, but normally more often Supporting evidence is provided by further comparisons with a set of fused silica "replacement cells" and the calibration of new cells.

The temperature difference between PTB transfer cell PTBQ1175 and PTB national reference, and the corresponding combined uncertainty were reported as:

$T_{PTBQ1175} - T_{nat.ref.} = -4.0$ µK
$u(T_{PTBQ1175} - T_{nat.ref.}) = 25.0$ µK

A2.14: TUBITAK UME

First mantle				Second mantle			
Date	$T_{transf.cell} - T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell} - T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
November 15, 2021	46.3	15.2	16.0	December 9, 2021	48.6	21.1	23.9
November 16, 2021	84.2	20.9	23.0	December 10, 2021	91.8	24.9	21.4
November 17, 2021	44.6	18.3	23.1	December 11, 2021	72.8	23.5	23.0
November 18, 2021	55.5	17.4	19.0	December 13, 2021	63.0	16.5	22.6
November 19, 2021	94.8	14.8	21.2	December 14, 2021	50.7	20.9	18.7
November 20, 2021	17.5	19.1	19.3	December 15, 2021	78.8	17.8	11.3
November 21, 2021	-14.0	26.0	20.1	December 17, 2021	77.7	23.1	20.2
November 22, 2021	54.0	17.7	16.6	December 18, 2021	93.8	20.5	22.1
November 23, 2021	3.3	13.4	15.6	December 19, 2021	100.8	20.2	14.7
November 24, 2021	70.0	23.2	22.1	December 20, 2021	40.5	11.2	15.7
Mean /µK	45.6	18.6	19.6		71.8	20.0	19.4
St. dev. Mean /µK	11.0				6.5		
$T_{transf.cell} - T_{nat.ref.}$ /µK	**58.7**						
$u(T_{transf.cell} - T_{nat.ref.})$	**53.0**						

TUBITAK UME national reference is defined by the mean of a group of four cells. Three of them are made of borosilicate glass and the last one is made of fused silica. The three borosilicate cell were manufactured at different times between 2002 and 2016. The fused silica TPW cell was manufactured in 2021. These cells are intercompared internally every three years.

The isotopic composition of the cells is measured and corrected for. No impurity analysis is available.

The temperature difference between TUBITAK UME transfer cell UMEQ5014 and TUBITAK UME national reference, and the corresponding combined uncertainty were reported as:

$T_{\mathrm{UMEQ5014}} - T_{\mathrm{nat.ref.}} = 58.7$ µK
$u(T_{\mathrm{UMEQ5014}} - T_{\mathrm{nat.ref.}}) = 53.0$ µK

A2.15: VSL

First mantle				Second mantle			
Date	$T_{transf.cell} - T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK	Date	$T_{transf.cell} - T_{nat.ref.}$ /µK	Standard deviation of the mean $T_{transf.cell}$ /µK	Standard deviation of the mean $T_{nat.ref.}$ /µK
May 10, 2021	1	0.07	0.06	June 9, 2021	29.2	0.10	0.09
May 12, 2021	26	0.05	0.05	June 10, 2021	31.5	0.10	0.10
May 14, 2021	4	0.06	0.06	June 14, 2021	-1.0	0.08	0.10
May 17, 2021	26	0.07	0.06	June 15, 2021	28.9	0.11	0.15
May 18, 2021	21	0.07	0.08	June 17, 2021	14.0	0.09	0.08
May 26, 2021	-5	0.05	0.09	June 18, 2021	7.3	0.13	0.12
May 27, 2021	23	0.10	0.09	June 21, 2021	28.2	0.10	0.14
May 28, 2021	14	0.13	0.10	June 23, 2021	24.7	0.09	0.11
May 29, 2021	7	0.10	0.08	June 24, 2021	33.4	0.12	0.12
May 30, 2021	32	0.10	0.10	June 25, 2021	-1.4	0.10	0.09
Mean /µK	14.8	0.08	0.08		19.5	0.10	0.11
St. dev. Mean /µK	3.9				4.3		
$T_{transf.cell} - T_{nat.ref.}$ /µK	**17.2**						
$u(T_{transf.cell} - T_{nat.ref.})$	**28.1**						

VSL national reference is defined by the average of 10 cells and maintained by the known difference of each cell of the group from the mean of the group itself.

The impurity content was estimated to be less than 100 nmol·mol^{-1} by ICPMS.

The temperature difference between VSL transfer cell VSL17T048 and VSL national reference, and the corresponding combined uncertainty were reported as:

$T_{VSL17T048} - T_{nat.ref.} = 17.2$ µK
$u(T_{VSL17T048} - T_{nat.ref.}) = 28.1$ µK

Appendix 3: Uncertainty budgets of the participants

In the following tables, the uncertainty budgets of the participants are reported.

Uncertainty component	CEM		INMETRO		IPQ		KRISS		LNE-Cnam	
	Contribution (k = 1) /µK	Number of degrees of freedom	Contribution (k = 1) /µK	Number of degrees of freedom	Contribution (k = 1) /µK	Number of degrees of freedom	Contribution (k = 1) /µK	Number of degrees of freedom	Contribution (k = 1) /µK	Number of degrees of freedom
National reference (uncertainties related only to properties of the national reference cells)										
Impurity content	29.7	100	11.5	∞	37.9	∞	27	∞	30[1]	
Isotopic composition	16.0	100	1.2	∞	20.5	∞	4	∞	10	
Residual gas pressure	2.2	100	2.9	∞	2.9	∞	5	∞		
Reproducibility (for different realizations of the national reference)	8.5	100	8.7	5	35.2	50	4	3	10	
Others										
Comparison of transfer cell to national reference (Uncertainties related to the comparison of the transfer cell to the reference cells)										
Repeatability for a single ice mantle	10.9	18	10.2	9	27.7	∞	4	18	23	
Reproducibility for different ice mantles	0.6	100	5.7	∞	10.5	∞	4	3	9	
Reproducibility for different types of SPRTs	1.6	100	23.3	∞			not included	not included	2	
Hydrostatic head of transfer cell and reference cell(s)	3.0	100	2.1	∞	1.3 (ref.) 1.3 (transf.)	∞	3	∞	14	
SPRT self-heating in transfer cell and in reference cell(s)	10.5	100	15.9	9	51.4	50	1	∞	12	
Standard resistor	0.0		2.9	∞	4.9 (ref.) 4.9 (transf.)	50 (ref.) 50 (transf.)	0	∞	10	
Resistance ratio bridge	8.0	100	12	∞	29	50	6	∞	15	
Perturbing heat exchanges	2.9	100	3.9	∞	5.2	50	11	∞	10	
Others						50				
Combined standard uncertainty (k=1)	**39.1**	247	**36.4**		**81.8**		**31**	> 3000	**49.8**	

[1]Includes residual gas pressure.

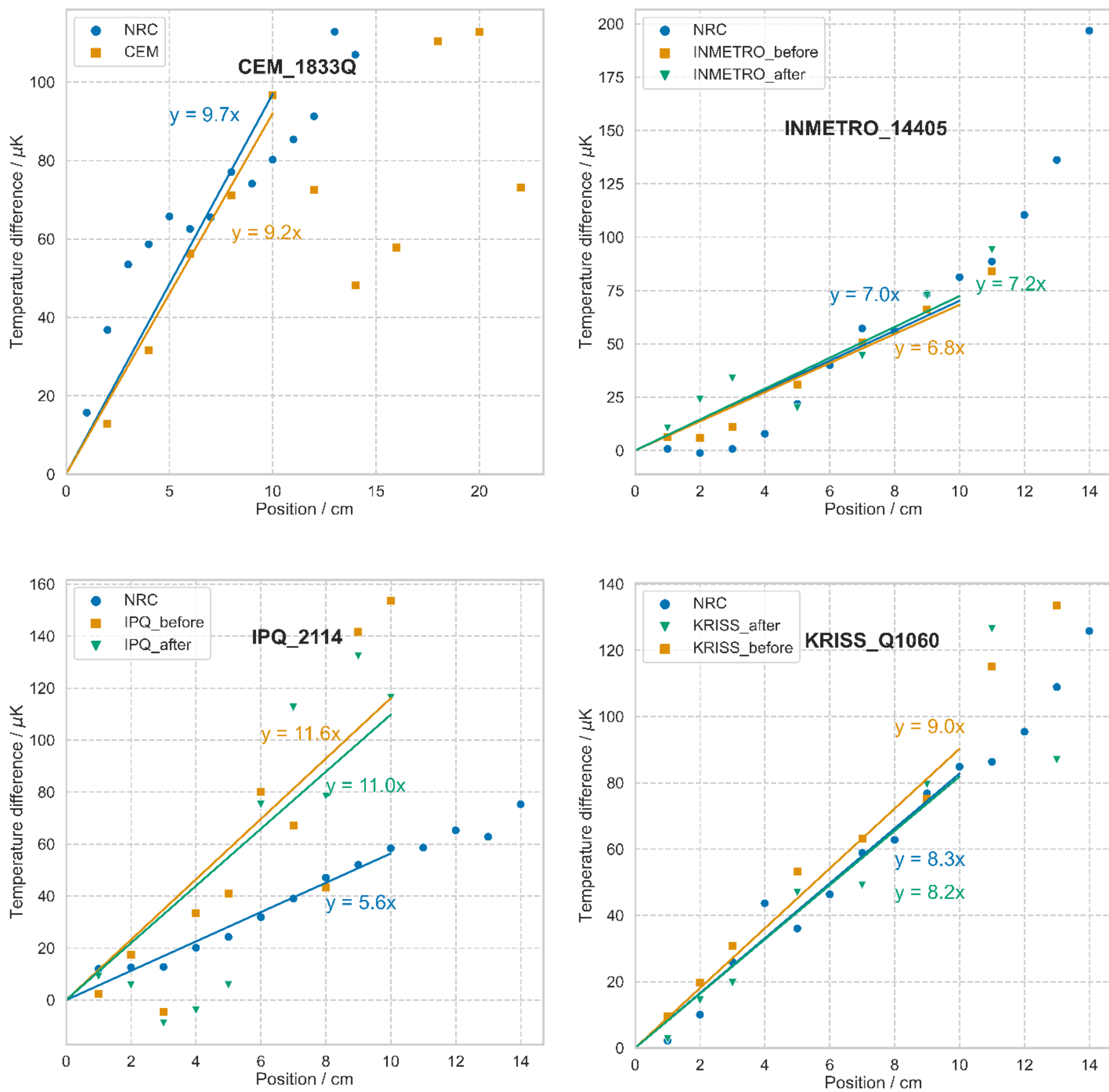
Uncertainty component	MSL		NIM		NMIA		NMIJ/AIST		NMISA	
	Contribution (k = 1) /μK	Number of degrees of freedom	Contribution (k = 1) /μK	Number of degrees of freedom	Contribution (k = 1) /μK	Number of degrees of freedom	Contribution (k = 1) /μK	Number of degrees of freedom	Contribution (k = 1) /μK	Number of degrees of freedom
National reference (uncertainties related only to properties of the national reference cells)										
Impurity content	6.3	17	10	13	5.3	7	20	∞	16.7	148
Isotopic composition	0.3	26	1	13	Included in the above		1	∞	38.1	500
Residual gas pressure	0.0	∞	25	13			1	∞	0.1	500
Reproducibility (for different realizations of the national reference)	6.4	22	2	39			7	∞	included in impurity content	
Others							4^3	∞		
Comparison of transfer cell to national reference (Uncertainties related to the comparison of the transfer cell to the reference cells)										
Repeatability for a single ice mantle	2.8	214	8	9	8.9	29	2	∞	4.4	9
Reproducibility for different ice mantles			7	19	12.6 (ref.) 10.4 (transf.)	5 (ref.) 5 (transf.)	3	19	6.3	500
Reproducibility for different types of SPRTs			3	19			2	∞	9.2	500
Hydrostatic head of transfer cell and reference cell(s)	0.1	∞	4	8	6.0	8	12	∞	6.0	500
SPRT self-heating in transfer cell and in reference cell(s)			8	79	0.1	8	7	∞	6.0	500
Standard resistor			2	8	included in repeatability		neglected		included in repeatability	
Resistance ratio bridge	0.0	∞	5	12	5.8	30	12	∞	11.6	30
Perturbing heat exchanges	0.0	∞	7	25	8.2	4	11	∞	41.2	50
Others										
Combined standard uncertainty (k=1)	**9.4**	46	**31.8**	32	**22.6**	28.7	**30.5**	1138545	**61.5**	

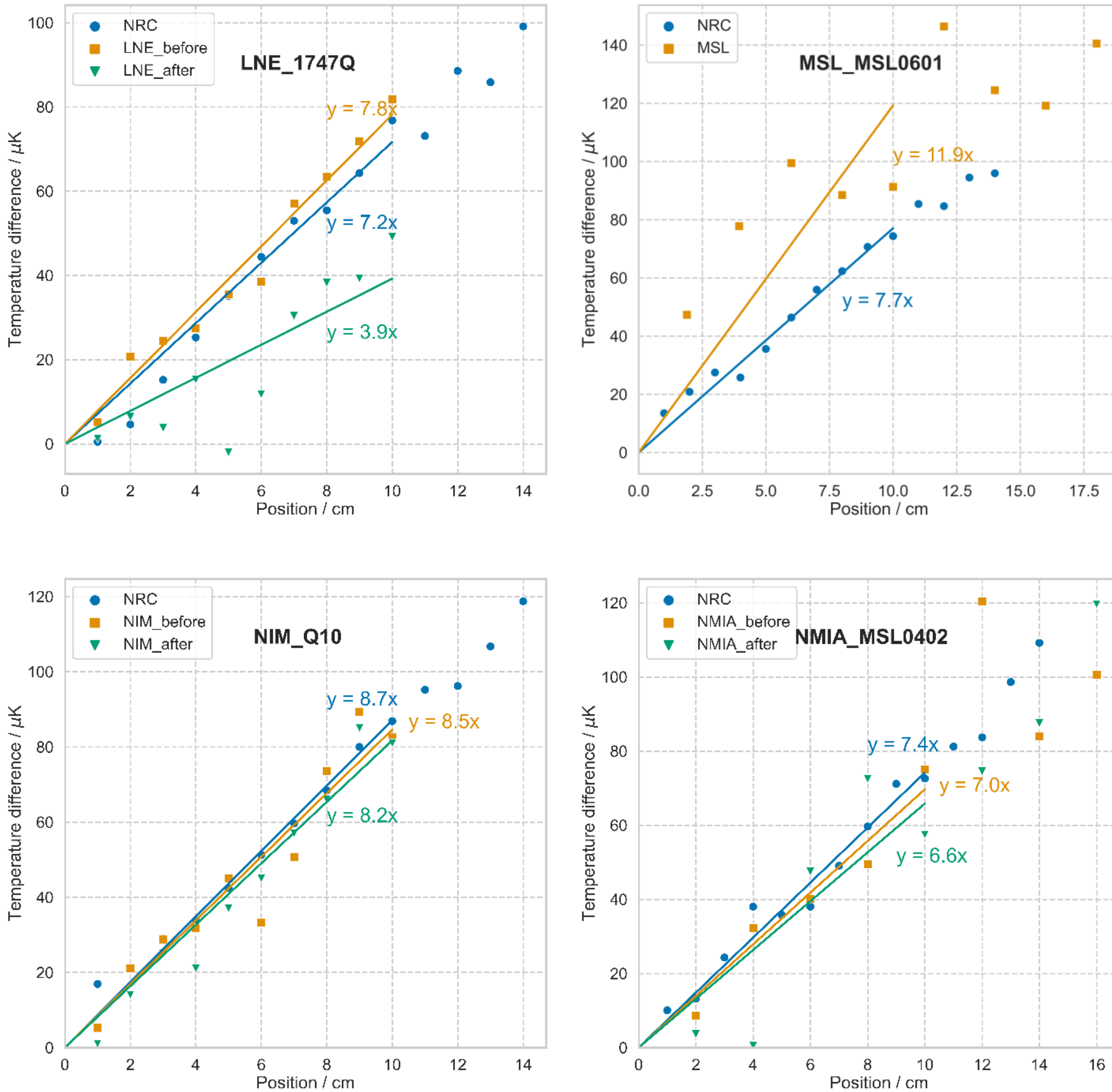
[3] Ambiguity of the definition of the triple point of water.

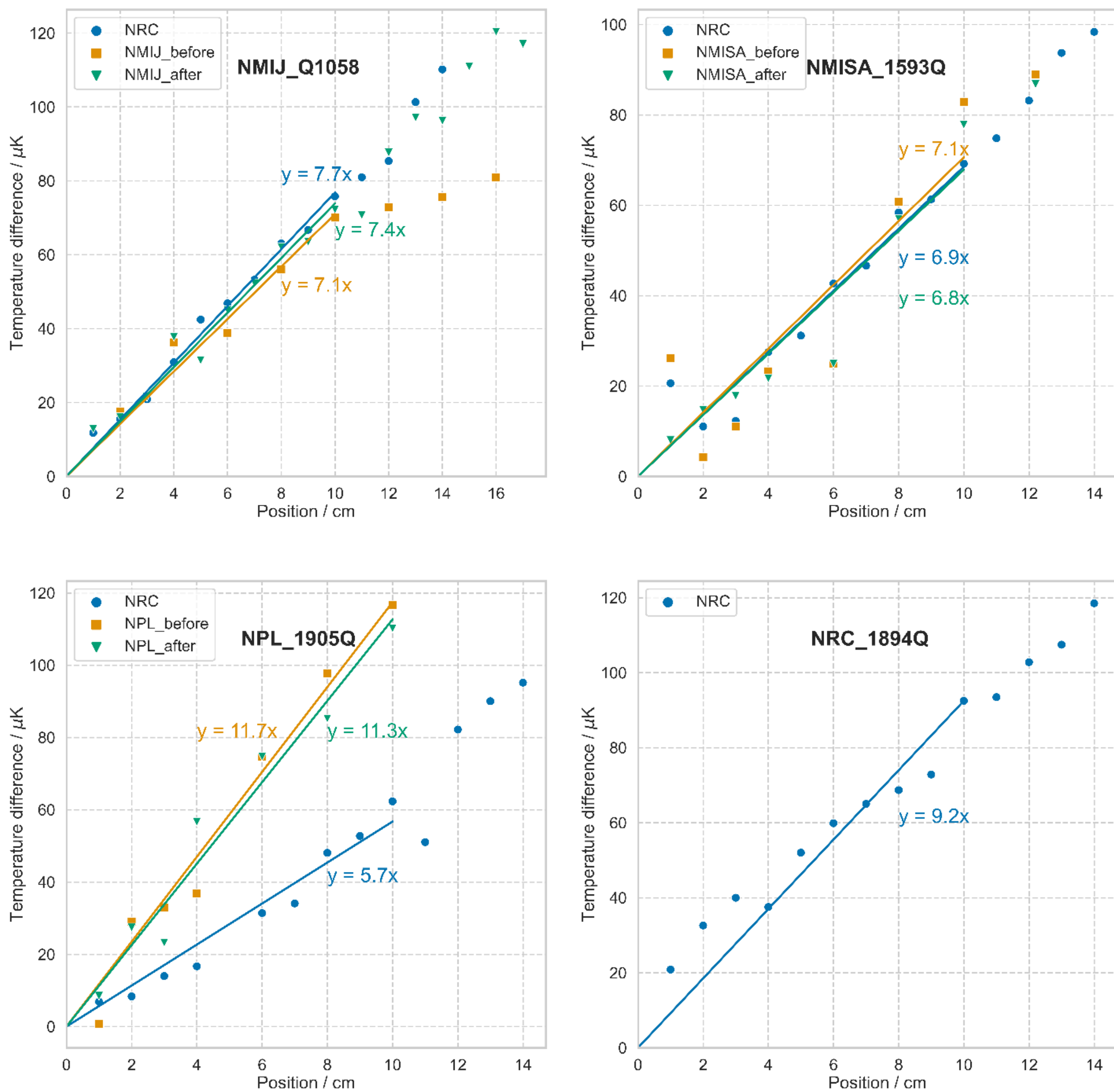
Uncertainty component	NPL		NRC		PTB		TUBITAK UME		VSL	
	Contribution (k = 1) /µK	Number of degrees of freedom	Contribution (k = 1) /µK	Number of degrees of freedom	Contribution (k = 1) /µK	Number of degrees of freedom	Contribution (k = 1) /µK	Number of degrees of freedom	Contribution (k = 1) /µK	Number of degrees of freedom
National reference (uncertainties related only to properties of the national reference cells)										
Impurity content	18	∞	11.2		10	∞	20	∞	20	
Isotopic composition	1	∞	0.4		2.5	∞	5	∞	3	
Residual gas pressure	4	∞	0.6		5	∞			2	
Reproducibility (for different realizations of the national reference)	3	∞	2.0		10	∞	20	20	5	
Others:										
Reproducibility of measurements of reference cells	5	9								
Perturbing heat exchanges reference cells	6	∞								
Hydrostatic head effect reference cells	2	∞								
Comparison of transfer cell to national reference (Uncertainties related to the comparison of the transfer cell to the reference cells)										
Repeatability for a single ice mantle	7	9	4.4		5	38	10	40	6	
Reproducibility for different ice mantles	3	∞	4.4		10	38	10	∞	14	
Reproducibility for different types of SPRTs	not applicable		not applicable		7.5	160	5	∞	3	
Hydrostatic head of transfer cell and reference cell(s)	2	∞	3.3		6	38	5	∞	5	
SPRT self-heating in transfer cell and in reference cell(s)	0	∞	6.8		10	38	20	∞	6	
Standard resistor	3	∞	not applicable		5	∞	5	∞	0	
Resistance ratio bridge	3	∞	3.9		5	∞	20	∞	5	
Perturbing heat exchanges	6	∞	7.8		5	∞	30	∞	5	
Others										
Combined standard uncertainty (k=1)	**23.1**	709	**17.5**		**25.0**		**53.0**		**28.1**	

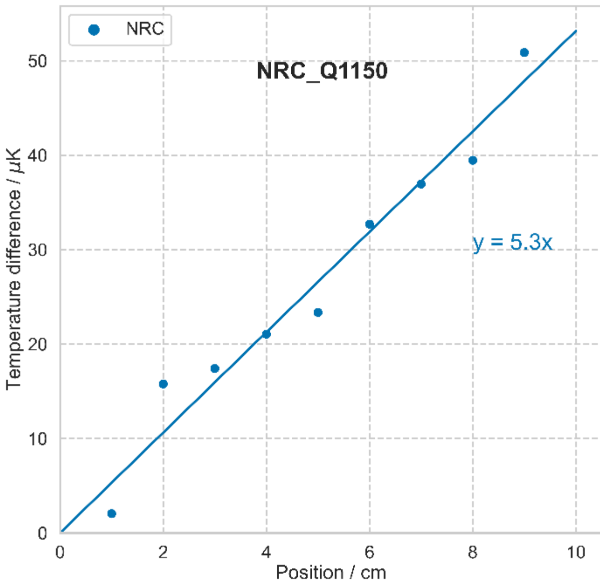
Appendix 4
Immersion profiles

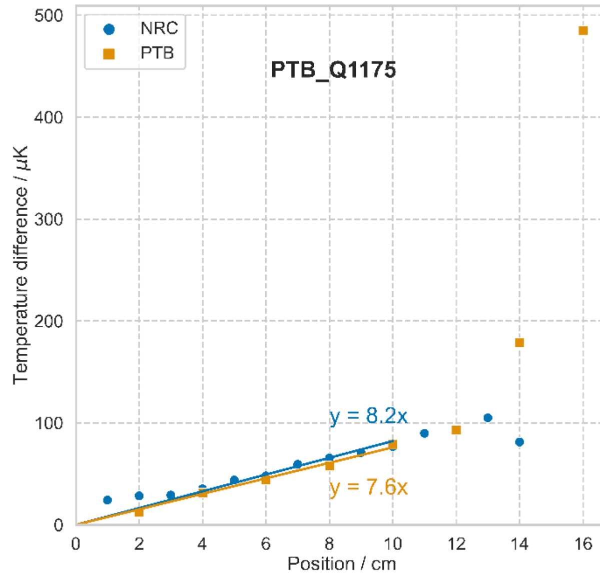
All participants were asked to provide an immersion profile of their transfer cell. These measurements were also made at NRC for each transfer cell. Most participants chose to measure the immersion profile before and after sending the transfer cell to the pilot.
At the NRC, the measurement procedure was as follows: all profiles were measured from bottom to top using a linear translation stage with an integrated controller (Thorlabs LTS300) to move an SPRT inside the thermometer well of a TPW cell. After an ice mantle was set free using the quartz rod placed in the thermometer well, the SPRT was attached to the linear stage, positioned 1 cm from the bottom of the thermometer well (corresponding to 0 in the following graphs and to the SPRT position during regular measurements) and allowed to equilibrate for 25 minutes. The linear stage was then automatically withdrawn in 1 cm increments 14 times (corresponding to a total travel distance of 14 cm). Note that the only exception was the travel distance for NRC reference cell Q1150 (9 cm) since we were still optimizing the measurement procedure at the time. The water level was not adjusted at each incremental step, instead, we made sure that the thermometer well was filled with water to the top at the start of the measurements. At each step, the computer initiated a measurement sequence consisting of 140 measurement points at 2 mA current, 60 measurement points at $2\sqrt{2}$ mA current, and 60 measurement points at 2 mA current (the rest of the resistance bridge settings remained the same as during the regular measurements). Only last 20 points at each current setting were used to evaluate SPRT self-heating at each step.
The results are shown in the following graphs. Positions and temperature differences are expressed relative to the normal measurement position. The numbers shown close to the curves give the slope (in µK/cm) of the constrained linear fit to the data ($y = kx$). To ensure consistency between the participants and NRC, the linear fits were applied only to the first 10 cm. The profiles are only shown for information, they have not been used for the data reduction of this comparison.
The slopes are in most cases larger than the theoretical value of 7.3 µK/cm. The average of all slopes measured at NRC is 7.6 µK/cm (standard deviation 1.3 µK/cm), the average of all participants' measurements is 8.6 µK/cm (standard deviation 2.2 µK/cm). For some cells both its laboratory and NRC found nearly ideal immersion curves (NIM-Q10, NMIA-MSL0402, NMIJ-Q1058, NMISA-1593Q). In other cases the measured profiles were very different from the expectation, and also differed between the originating laboratory and NRC. Certain TPW cell design features could explain some of the observed behaviour. In general, we noticed that larger cells with smaller diameter thermometer wells tend to produce better immersion curves.

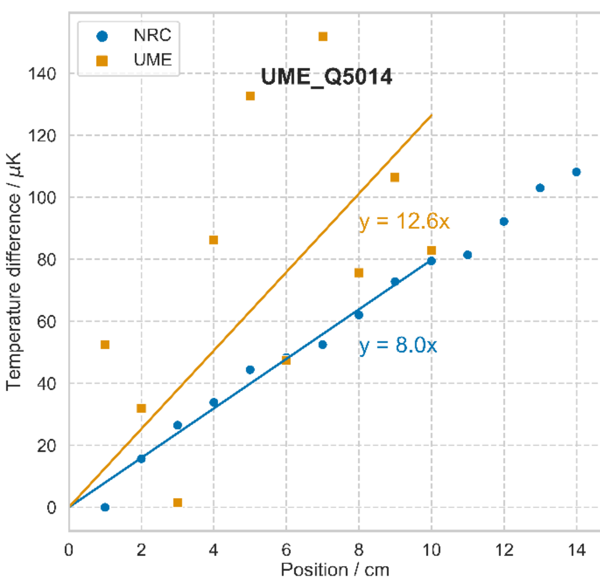

CEM_1833Q
NRC
CEM
y = 9.7x
y = 9.2x
Temperature difference / μK
Position / cm
INMETRO_14405
NRC
INMETRO_before
INMETRO_after
y = 7.0x
y = 7.2x
y = 6.8x
IPQ_2114
NRC
IPQ_before
IPQ_after
y = 11.6x
y = 11.0x
y = 5.6x
KRISS_Q1060
NRC
KRISS_after
KRISS_before
y = 9.0x
y = 8.3x
y = 8.2x

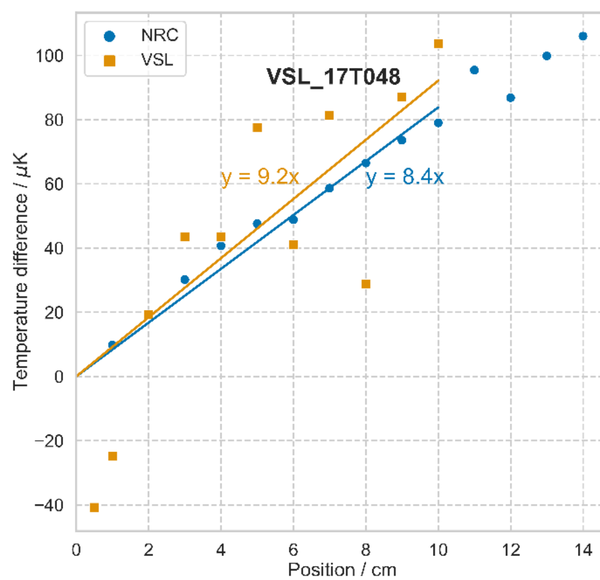

NRC
LNE_before
LNE_after
LNE_1747Q
y = 7.8x
y = 7.2x
y = 3.9x
Temperature difference / μK
Position / cm
NRC
MSL
MSL_MSL0601
y = 11.9x
y = 7.7x
Temperature difference / μK
Position / cm
NRC
NIM_before
NIM_after
NIM_Q10
y = 8.7x
y = 8.5x
y = 8.2x
Temperature difference / μK
Position / cm
NRC
NMIA_before
NMIA_after
NMIA_MSL0402
y = 7.4x
y = 7.0x
y = 6.6x
Temperature difference / μK
Position / cm


NMIJ_Q1058
NRC
NMIJ_before
NMIJ_after
y = 7.7x
y = 7.4x
y = 7.1x
Temperature difference / μK
Position / cm
NMISA_1593Q
NRC
NMISA_before
NMISA_after
y = 7.1x
y = 6.9x
y = 6.8x
Temperature difference / μK
Position / cm
NPL_1905Q
NRC
NPL_before
NPL_after
y = 11.7x
y = 11.3x
y = 5.7x
Temperature difference / μK
Position / cm
NRC_1894Q
NRC
y = 9.2x
Temperature difference / μK
Position / cm


NRC
NRC_Q1150
y = 5.3x
Temperature difference / μK
Position / cm


NRC
PTB
PTB_Q1175
y = 8.2x
y = 7.6x
Temperature difference / μK
Position / cm


NRC
UME
UME_Q5014
y = 12.6x
y = 8.0x
Temperature difference / μK
Position / cm


NRC
VSL
VSL_17T048
y = 9.2x
y = 8.4x
Temperature difference / μK
Position / cm

Appendix 5
Additional Statistical Analysis

A5.1 Introduction

In this appendix we report on additional statistical analysis performed on the CCT-K7.2021 data set (Table 5.1).
In Section 5.2 we describe the adaptive weighted average (AWA) procedure, which is the method adopted for the data reduction.
In Section 5.3 we compare the KCRV value and its uncertainty, obtained with the selected reduction method (adaptive weighted average), with the KCRV (and its standard uncertainty), obtained with the more classical estimators of simple mean and conventional weighted average.
In Section 5.4 we briefly describe an alternative Bayesian NDT method, the Hierarchical Gauss + Gauss (HGG) procedure, and its results when applied to our data set.
Finally, in Section 5.5, in analogy to CCT-K7 report, we show the joint (pooled) probability distribution, calculated from the individual participant distributions.

A5.2 The adaptive weighted average (AWA)

The adaptive weighted average (AWA) implemented in the NIST Decision Tree is the DerSimonian-Laird (DL) procedure widely used in meta-analysis. The underlying model represents each measured value x_i as an additive superimposition of three effects:

$$x_i = \mu + \lambda_i + \varepsilon_i$$

Where:
μ denotes the true value of the measurand,
λ_i is the effect of laboratory i (expressing whether it tends to measure "high" or "low"),
ε_i is measurement error.

When the λ_i do not differ significantly from 0 (that is, when there is no dark uncertainty), DL produces the conventional weighted average. This is the case of CCT-K7.2021 data. Koepke *et al.*[1] describe the details of the DL procedure, and so do many books concerned with meta-analysis, for example Borenstein *et al.*[2].

The output of the AWA can be duplicated independently of its implementation in the NIST Decision Tree as follows, using facilities available in the R programming environment:

```
Lab = c("CEM", "INMETRO", "IPQ", "KRISS", "LNE/CNAM", "MSL", "NIM", "NMIA", "NMIJ/AIST", "NMISA", "NPL", "NRC", "PTB", "TUBITAK UME", "VSL")
x = c(51.7, -35.1, 11.7, 0.0, 5.9, 18.8, 3.6, -16.2, 27.1, 3.2, 15.5, -2.2, 11.6, -59.3, -39.9)
```

[1] A. Koepke, T. Lafarge, A. Possolo, B. Toman, Consensus building for interlaboratory studies, key comparisons, and meta-analysis, *Metrologia* **54** (2017) S34-S62

[2] M. Borenstein, L.V. Hedges, J.P.T. Higgins, H.R. Rothstein, 2009 Introduction to Meta-Analysis, John Wiley & Sons.

```
ux = c(40.2, 37.6, 82.3, 32.4, 50.7, 13.2, 33.1, 24.4, 31.9, 62.2, 24.9, 19.8, 26.7, 53.8, 29.6)

library(metafor)
summary(rma(yi=x, sei=ux, slab=lab, data=z, method=”DL”))
```

A5.3 Simple average, weighted average and adaptive weighted average

In this section we compare the KCRV (and its standard uncertainty), obtained with the selected reduction method (adaptive weighted average), with the KCRV (and its standard uncertainty), obtained with the more classical estimators of simple mean and conventional weighted average. Table A5.1 and Figure A5.1 show this comparison in both numerical and PDF form.

Table A5.1: KCRV and corresponding standard uncertainty for three different estimators: simple average, weighted average and adaptive weighted average.

Statistical Estimator	T_{KCRV}	u_{KCRV}
Simple average	-0.3	10.7
Weighted average	4.7	7.2
Adaptive weighted average	4.7	7.2

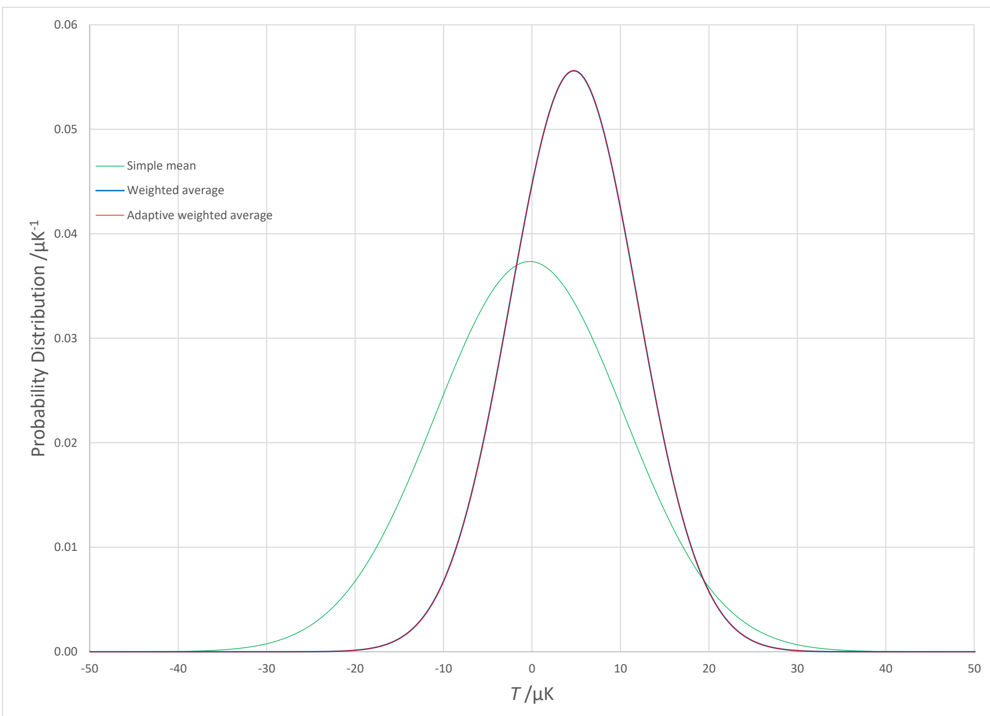


Figure A5.1: Probability distribution functions (PDFs) of different estimators: simple average, weighted average and adaptive weighted average (weighted average and adaptive weighted average are indistinguishable in this case).

A5.4 The Hierarchical Gauss + Gauss procedure

When the NDT detects significant dark uncertainty (that is, when the measurement results are mutually inconsistent) and the measured values, suitably standardized, can be regarded as a sample from a Gaussian distribution, then the NDT recommends the Hierarchical Gauss + Gauss (HGG) procedure.
The HGG procedure is a Bayesian procedure that fits the random effects model:

$$x_i = \mu + \lambda_i + \varepsilon_i$$

to the measurement results. Koepke *et al.*[1] detail the assumptions the NDT makes about the terms of this model, and explain how the model can be fitted to a set of independent measurement results for the same measurand.

When applied to CCT-K7.2021 data, the HGG procedure gave:
$x_{KCRV} = 2.3\mu K$
$u_{KCRV} = 8.0\mu K$,
which does not significantly differ from the values showed in Table A5.1.

The calculations that the NDT makes to fit this model to data can be reproduced approximately and independently using the R function "brm" defined in R package "brms"[3].

A5.5 Pooled Probability Distribution

In Section 6.1 of the CCT-K7 report, the pooled probability distribution was reported. The pooled distribution looked bimodal (see Figure 29 of CCT-K7 report), suggesting two different populations: the population of participants that applied the ocean water definition (higher temperature mode) and the larger population of participants that did not apply the ocean water definition (lower temperature mode).
The pooled distribution mentioned above is the linear pool reviewed by Koepke et al.[1] A sample from this distribution is drawn by drawing as many samples as there are participants, all of the same large size, from Gaussian distributions with means equal to the measured value, and standard deviations equal to the reported standard uncertainties, and then merging these samples into a single sample.
In Figure A5.2 we report the pooled distribution of CCT-K7.2021 results, along with the individual participant distributions. CCT-K7.2021 pooled distribution is not bimodal, though it is slightly asymmetrical.
As explained in Section 6.1 of this report, the broader distribution of the results towards lower temperatures is explained by the tendency of chemical impurities and isotopic depletion to shift the results only to lower temperatures.

[3] P.C. Burkner, 2017 "brms: An R Package for Bayesian Multilevel Models Using Stan", *Journal of Statistical Software*, **80** 1 1-28.

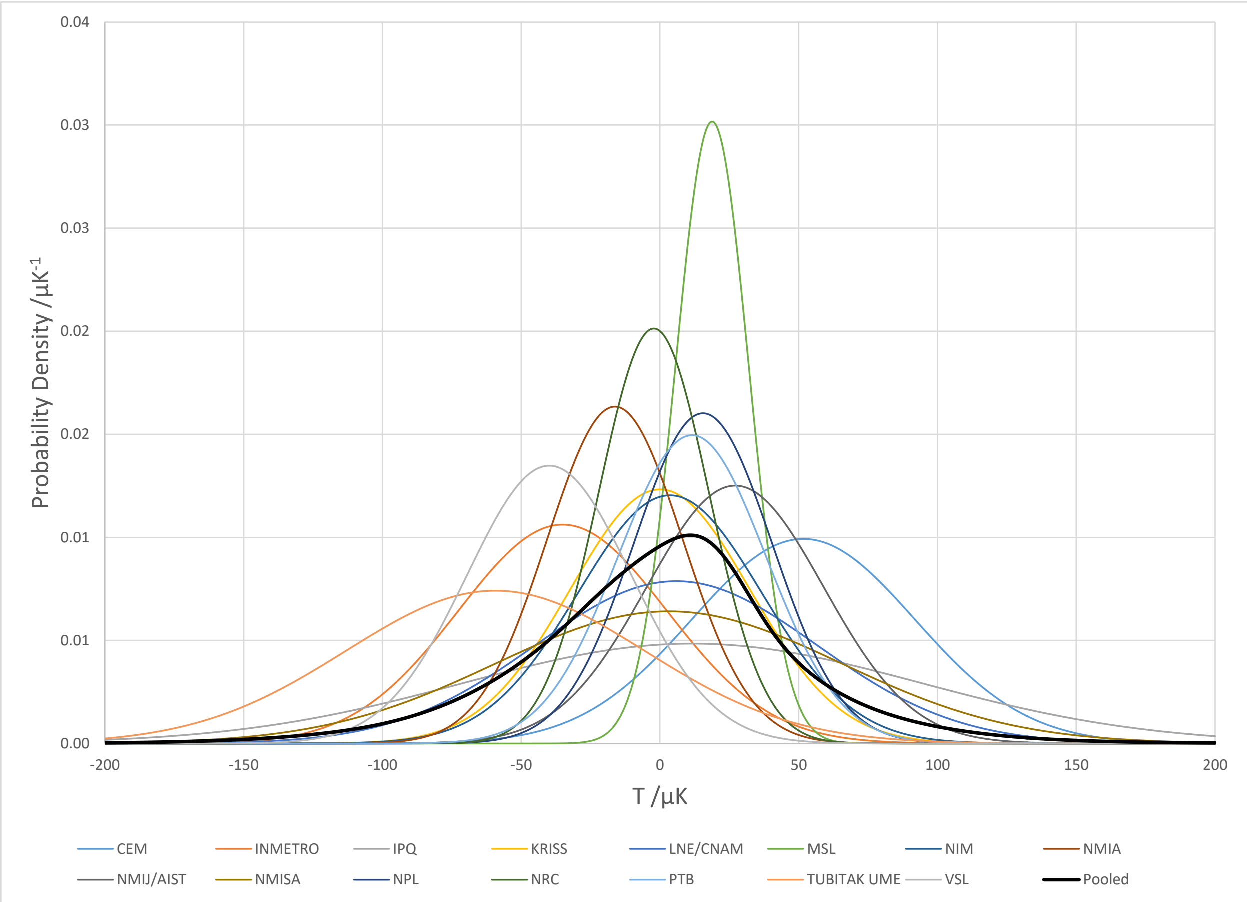


Figure A5.2: Pooled distribution (thick black) and individual participant distributions (thin curves).