NRC·CNRC

NRC.CANADA.CA

Highlights of CCT WG-KC activities (May 2024 to May 2026)

Andrea Peruzzi

May 21-22, 2026

WG-KC Terms of reference

- *“To oversee all aspects of key comparison documentation*
  - *Starting with the Technical Protocol*
  - *Ending with the Draft B Report and the KCDB entry*
- *Including provision of advice to pilots on:*
  - *Calculation of the Degrees of Equivalence*
  - *Key Comparison Reference Value*
  - *Linkage between RMO and CIPM key comparisons”.*
- In practice:
  - Review the initial Technical Protocols and all their subsequent iterations until approval
  - Review the Draft B Reports and all their revisions until approval
  - For all CCT KCs and RMO KCs (and, on request of the RMOs, also for RMO SCs)

Current membership

#	Member	NMI	RMO	Area of expertise
1	Megumi Akoshima	NMIJ (Japan)	APMP	Thermophysical quantities
2	Stephanie Bell	NPL (UK)	EURAMET	Humidity
3	Robert Benyon	INTA (Spain)	EURAMET	Humidity
4	Rien Bosma	Independent Researcher (NL)	EURAMET	Contact thermometry and humidity
5	Helen McEvoy	NPL (UK)	EURAMET	Non-Contact thermometry
6	Christopher Meyer	NIST (USA)	SIM	Humidity
7	Andrea Peruzzi	NRC (Canada)	SIM	Contact thermometry
8	Richard Rusby	NPL (NPL)	EURAMET	Contact thermometry
9	Gregory Strouse	NIST(USA)	SIM	Contact thermometry
10	Jianping Sun	NIM (China)	APMP	Contact thermometry
11	Andrew Todd	NRC (Canada)	SIM	Non-Contact thermometry
12	Inseok Yang	KRISS (Korea)	APMP	Contact thermometry

- 12 members
- Experts in CTh, NCTh, Hum, and TQ
- Proposed new members:

→ **Sergey Dedyulin, NRC** (replacing 8)

→ **Peter Saunders, MSL** (replacing 11)

Activities May 2026 – May 2024 (since last CCT meeting)

- We have provided our services to **26 different comparisons**

Closed comparisons (since May 2024)

- **11 approved comparisons:**
  - 3 CIPM KCs
  - 4 RMO KCs
  - 4 RMO SC
- **5 abandoned comparison:**
  - 1 CIPM KC
  - 1 RMO KC
  - 3 RMO SCs

Comparison ID	Type	Approval date
EURAMET.T-K8	RMO KC	30-08-2024
EURAMET.T-K9	RMO KC	30-08-2024
EURAMET.T-S7	RMO SC	30-04-2026
GULFMET.T-S3	RMO SC	24-10-2024
SIM.T-S4	RMO SC	Abandoned 29-10-2024
SIM.T-S8	RMO SC	Abandoned 29-10-2024
SIM.T-S9	RMO SC	Abandoned 29-10-2024
APMP.T-S16	RMO SC	21-11-2024
CCT-K8	**CIPM KC**	**23-12-2024**
COOMET.T-K6.1	RMO KC	04-04-2025
CCT-K9.2	**CIPM KC**	**16-04-2025**
CCT-K6.1	**CCT-K6.1**	**17-09-2025**
APMP.T-K4.2	RMO KC	06-01-2026
CCT-K9.4	**CIPM KC**	**09-02-2026**
COOMET.T-S5	RMO SC	11-02-2026
COOMET.T-K9.1	**RMO KC**	**19-02-2026**

Initiated comparisons (since May 2024)

- **8 initiated comparisons:**
  - 7 RMO KCs
  - 1 RMO SC

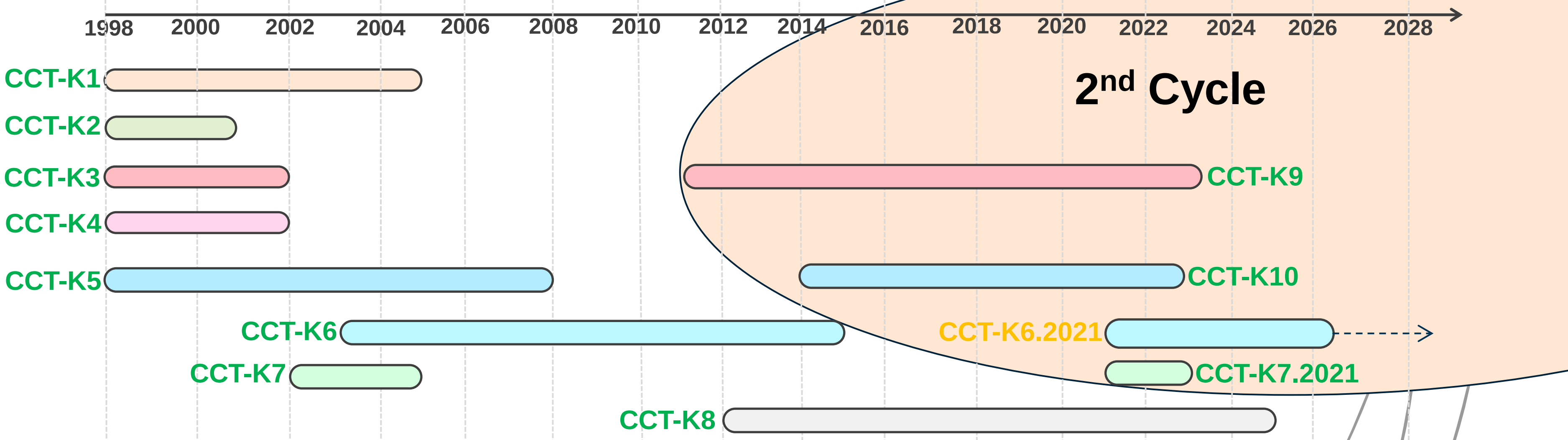
- **1 COOMET comparison:**
  - Protocol submitted but all reviewers declined to review

Comparison ID	Type	Protocol submitted
COOMET.T-K6.1	RMO KC	07-06-2024
APMP.T-K7.2021	RMO KC	10-07-2024
SIM.T-K9.4	RMO KC	18-07-2024
EURAMET.T-K10	RMO KC	18-10-2024
COOMET.T-KX	**RMO KC**	**22-07-2025**
SIM.T-S13	RMO SC	29-08-2025
SIM.T-K6/SIM.T-K8	RMO KC	12-02-2026
APMP.T-K7.2021.1	RMO KC	0.4-03-2026

Excessively sleeping CCT comparisons:

Comparison	Initiated	Status KCDB	Last communication	Planned target of completion	Pilot
CCT-K1.1	2006	Report in Progress, Draft A	04-03-2026	September 2026, if not → abandon it	NIST (USA)
CCT-K2.2	2012	Measurements in progress	20-02-2026	August 2029	INRIM (Italy)
CCT-S3	2007	Report in Progress, Draft A	23-02-2026	Submit revised Draft B in Sept. 2026	NMIJ/AIST (Japan)

Overview CCT KCs



1st Cycle:

CCT-K1: 0.65 K to 24.6 K
CCT-K2: 13.8 to 273.16 K
CCT-K3: -189.3442 °C to 660.323 °C
CCT-K4: 660.323 °C and 961.78 °C
CCT-K5: 961°C to 1700 °C
CCT-K6: -50°C to 20 °C (frost/dew point)
CCT-K7: 273.16 K
CCT-K8: 30°C to 90 °C (dewpoint)

2nd Cycle:

CCT-K9: 2nd cycle of CCT-K3
CCT-K10: 2nd cycle of CCT-K5 (up to 3000 °C)
CCT-K7.2021: 2nd cycle of CCT-K7
CCT-K6.2021: 2nd cycle of CCT-K6

Guide to the Analysis of CCT KC Results

- **Mandate:** During the last CCT meeting, the WG-KC was charged with the preparation of a document to provide guidance to CCT KC pilots on the analysis of KC results

- *I was not aware of what I was getting into…*

- **Difficulties:**
  - Ambiguities and deficiencies in the CIMP MRA documents (definition of KCRV, computation of KCRV, …)
  - Statistical science heavily involved (statistical models of the measurement, Bayesian methods, …)
  - Limited support (understandably)

- **Approach:**
  - Asked for help (Antonio Possolo, NIST Chief Statistician)
  - Draw inspiration from other CC's guidance documents (CCQM)
  - Pragmatic: try to keep it simple, relax conformity to CIPM MRA documents

The interpretation of the KCRV

- **Interpretation of the KCRV in CIPM MRA documents:**
  - Do not give a definition of KCRV (nor do they specify how it should be computed)
  - Only tell you how it should be '*considered*':
    1. MRA (1998): *"in most cases, it can be considered to be a close, but not necessarily the best, approximation to the SI value"* (of the measurand)
    2. MRA-G11 (2025): "*each KCRV is considered to be a close approximation of the true value*" (of the measurand)
  - Neither 2 nor 1 are completely appropriate for CCT: T_{90} is not exactly SI – it is an approximation to the SI and an operationally-defined quantity (no singular true value for T_{90} temperatures)
- **Interpretation of the KCRV in past CCT KCs:**
  - The CCT KC participants have always refrained from attributing the KCRV the meaning of best/close approximation to the true value (see Table)
- **Interpretation of the KCRV in the CCT Guide:**
  - We claim that, in CCT KCs, the KCRV can really be attributed the meaning of best (or close) estimate to the true value of the measurand (controversial)

Interpretation of the KCRV in the past CCT KCs

KC	Statements on the meaning attributed to the KCRV
CCT-K1	*No additional significance is attached to the KCRV with respect to 'best estimate' of* T_{90}.
CCT-K2	*The KCRV represents a common baseline against which all laboratory values can be compared.*
CCT-K3	The participants decided not to compute the KCRV.
CCT-K4	*The value of the so-defined KCRV must not be related to the ITS-90 value of the corresponding fixed point. The KCRV has no physical meaning and is used only as a notational shorthand for presenting a common baseline against which all laboratory values can be compared.*
CCT-K5	No statement on the meaning of the KCRV.
CCT-K6	*The KCRV has no absolute significance outside the comparison.*
CCT-K7	*It is clear that this KCRV is not the best possible approximation of the true SI value.*
CCT-K8	*The KCRV given in this comparison has no absolute significance, that means it does not represent a reference value in the SI, only a comparison parameter.*
CCT-K9	*The KCRV is not meant to have any intrinsic meaning – it is a computational tool allowing comparisons across labs linked through measurements at the pilot lab.*
CCT-K10	No statement on the meaning of the KCRV.
CCT-K11	No statement on the meaning of the KCRV.

Should the method to be used to determine the KCRV be included in the Technical Protocol?

- **CIPM-MRA-G11:**
  - *“For CIPM KCs, the technical protocol should include a description of the method to be used to determine the KCRV”*
- However, the appropriate method for data reduction cannot be selected before examining the measurement results from all participants: *“You need to interrogate the measurement results and force them (through statistical tests) to tell you how they would like to be analysed”*
- **Our Guide:**
  - Sets the general principles to be followed in the analysis of the results, without prescribing a specific method for computing the KCRV

General principles to be adopted in CCT KCs

1) The **exclusion of results from the computation of the KCRV** should be based on substantive reasons (not statistical).

2) No individual result should dominate the KCRV value 'automatically', simply because its uncertainty is much smaller than the uncertainties of the other results.

3) Measurement methods sufficiently well characterized to guarantee confidence in the belief that **the measured values, taken as a group, are roughly centered at the true value of the measurand**.

4) Adopt: a) a **statistical model** that explicitly relates the measured values to the true value of the measurand and b) an **optimality criterion** that the KCRV is intended to satisfy.

5) The statistical model should be able to detect, evaluate and propagate **dark uncertainty**, which manifests itself as dispersion of the measured values in excess of what the reported uncertainties suggest that it should be.

6) DoEs should express: i) the uncertainty reported for the measured value; ii) the uncertainty associated with the KCRV; iii) the correlation between the measured value and the KCRV; and iv) any dark uncertainty detected during data reduction.

Principle 4a: Selection of the statistical model

- For CCT KCs, the **random effects model** generally tends to be the most appropriate:

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

x_i = value measured by participant *i*
μ = true value of the measurand
λ_i = laboratory effect of participant *i*
ε_i = laboratory-specific measurement error of participant *i*
λ_i = non-observable outcome of independent Gaussian (or Laplace or other) random variables with mean 0 and standard deviation τ
ε_i = values of independent Gaussian random variables with mean 0 and standard deviations u_i

- These assumptions imply that **the expected value of the measured value is μ**

- The model can be easily extended to the most general comparison topology:

$$x_{ijk} = \mu_j + \delta_{ijk} + \lambda_i + \varepsilon_{ijk}$$

i = participant index
j = artefact index
k = repeat of participant's measurement
δ_{ijk} = drift

Principle 4b: Selection of the optimality criterion

- ➢ The statistical model alone does not stipulate the procedure to be used to compute the KCRV
- ➢ An optimality criterion for the KCRV needs to be chosen. **It is the choice of the optimality criterion which tells you which procedure to use to compute the KCRV (and which defines the exact meaning of KCRV).**
- ➢ How do you choose the optimality criterion for your KC?
- ➢ Your choice is driven by:
  - the **assumptions** that you regard as appropriate for your specific KC
  - the results of the statistical tests

Example 1: Simple average

- Suppose that you have chosen to compute the KCRV as the simple average of the measured values

- **Procedure:** $x_{KCRV} = \frac{1}{n}\sum_{i=1}^{n} x_i$ $u_{KCRV}^2 = \frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - x_{KCRV})^2$

- **Optimality criterion chosen:**
  - Minimization of the mean squared error of the KCRV

- **Assumptions:**
  - The measured values are a sample from a single Gaussian distribution
  - No dark uncertainty
  - All measured values x_i have the same standard uncertainty
  - All reported standard uncertainties are unreliable and are all disregarded

Example 2: Uncertainty-weighted average

- Suppose that you have chosen to compute the KCRV as the uncertainty-weighted average of the measured values

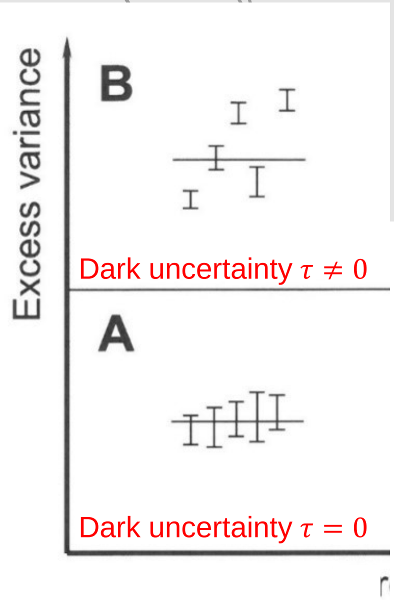
- Procedure:

$$x_{KCRV} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \qquad \frac{1}{u_{KCRV}^2} = \sum_{i=1}^{n} \frac{1}{u_i^2} \qquad w_i = \frac{1}{u_i^2}$$

- **Optimality criterion chosen:**
  - Minimization of the mean squared error of the KCRV

- **Assumptions:**
  - The measured values are outcomes of Gaussian random variables with the same mean but possibly different standard deviations that are reliably estimated by the reported uncertainties
  - No dark uncertainty
  - The formula for u_{KCRV} assumes that the reported uncertainties are based on infinite numbers of dof

Principle 5: Dealing with dark uncertainty



- The results of a KC are **mutually consistent** when the dispersion of the measured values is not larger than what their reported uncertainties suggest that they should be (A)
- If the measured values are appreciably more dispersed than what their reported uncertainties suggest (B), then the excess dispersion is quantified in what is called '**dark uncertainty**'
- Dark uncertainty is a possible composite source of uncertainty that the participants failed to identify and include in their uncertainty budgets
- The KCRV can be still computed even when the results are mutually inconsistent, but its uncertainty will reflect the presence of dark uncertainty.
- However, the reported uncertainties are not altered, only this extra contribution to the KCRV uncertainty is recognized.

Impact of dark uncertainty on DoEs (and CMCs!)

- The CCT needs to take a decision: when dark uncertainty is present, should it impact the DoEs (and the CMCs)?
- **Answer 1:** Yes, always, because this is the only way of remaining consistent with the model selected for the KC
- **Answer 2:** it depends on the purpose the DoEs serve:
  - If the DoEs serve to predict the result of a future measurement of a participant lab, then the DoEs should recognize the dark uncertainty (DoEs for prediction).
  - If the DoEs are intended to express each participant's belief about its capabilities, then they could exclude the uncertainty contribution from dark uncertainty (DoEs for trade).

Feedback from CCT members and RMOs

- Very valuable but received only in the past week!

- **MSL (Peter Saunders and Ellie Molloy):**
  - Definition of the measurand
  - Definition of the KCRV

- **EURAMET (Regina Deschermeier):**
  - Thermometry-biased and how to re-establish balance
  - Additional statistical outlier tests and examples of application of statistical tests in the Appendix

- **CCT-WG-KC Members:**
  - Robert: difference between a KC and a PT → look at document on PT analysis
  - Inseok: use dark uncertainty only in extreme cases, otherwise investigate the causes of inconsistency and remove them. Comparison uncertainty is more useful because used in the CMC review protocol.
  - Stephanie: add a section on RMO KCs/bilateral KCs. Is T_{90} SI or not?

NRC·CNRC

NRC.CANADA.CA •

THANK YOU

Andrea Peruzzi

Research Officer

andrea.peruzzi@nrc-cnrc.gc.ca