Guidance for carrying out key comparisons within the CCAUV

Key Comparison Working Group

for the

Consultative Committee for Acoustics, Ultrasound and Vibration

November 2015

Table of Contents

1 Purpose

This Document covers a number of aspects that becomes important when conducting measurement comparisons in the framework of the CIPM-MRA. The objective is to support the overall quality and reliability of the reported data, as well as to harmonize the methods of planning, analysis and reporting. The information given here is complementary to the guidelines given by the International Committee for Weights and Measures (CIPM) in CIPM-MRA D-05, *Measurement comparisons in the CIPM MRA* [1], to be applied within the Consultative Committee for Acoustics, Ultrasound and Vibration (CCAUV).

2 Technical Protocol

A key comparison (KC) is usually initiated within the CCAUV or the respective Regional Metrology Organization (RMO) under which auspices the comparison is conducted.

The potential pilot laboratory drafts a technical protocol to prepare the KC. The TP defines the scope of the comparison work in terms of measurements as well as the subsequent analysis and reporting. The comparison is ***started only after*** that the TP has been approved by the Key Comparison Working Group (KCWG) of the CCAUV and registered in the Key Comparison Data Base (KCDB) using the appropriate registration form [2].

For approval, the TP should include ***at least*** the following information:

- Task and purpose of the comparison.
- Information on the Pilot Laboratory and contact person.
- Description of the circulated devices under test.
- Scope of the measurements (result quantities, methods to be applied, frequencies, amplitudes, intensities, etc.).
- Relevant documentary standards (ISO, IEC, ...).
- Environmental conditions to be fulfilled.
- Special instructions to be followed if necessary.
- Circulation type (Ring, Star, mixture), means of monitoring stability.
- Means of transportation, responsibilities, absorption of costs, insurance.
- List of participants with planned time schedule.
- Specific measurement, data analysis and reporting instructions for the participants.
- Reporting template for the data (possibly as attached digital spread-sheet).
- Reporting instructions for the measurement system, methodology, measurement conditions and measurement uncertainty including the number of significant digits.
- Individual results should be reported by participants with one more significant digit than directly related to the relative expanded uncertainty, which shall be presented with two significant digits.
- The intended procedure for the analysis of the KC and the linking to other KCs where applicable.
- Include reference to previous related KCs or relevant publications (lessons learnt).

- Consider the inclusion (in an annex to the reporting spread-sheet) of information on the staff resources used (optional but helpful for future strategic planning).

3 Analysis and Report

The analysis is, in many details, dependent on the specific conditions under which a KC is performed. However, there are still a number of general principles that have to be considered for a proper analysis. These will subsequently be explained in relation to the reported result quantities of KCs.

The reader should be able to reproduce the analysis, i.e. the calculations should be performed on the basis of the documented values. This has to be considered in conjunction with a reasonable rounding and the documented number of digits.

All necessary formulas should be available in the report or in publicly available and referenced publications.

3.1 Reference Value

The key comparison reference value (KC-RV) is the reference value derived from the set of all consistent measurements of all participants. In the case of the CCAUV there is typically one KC-RV for each frequency and artifact.

In order to calculate a KC-RV, two steps have to be considered:

1. To determine of a set of consistent measurements; this may also be called investigation and involves a possible elimination of outliers.
2. To calculate a KC-RV based on a consistent set of results.

For a consistency test, the CCAUV has usually applied χ^2 methods, but other methods may be found in literature. When calculating the KC-RV, a weighted mean method shall be applied in order to derive an unbiased estimator.

It should be noted that ***the term KC-RV is strictly reserved for CIPM or BIPM KCs***; this term should not be used in RMO KCs, nor in Supplementary Comparisons or Pilot Studies. However, there may of course be an interest to establish a reference value in the latter comparisons and studies. In this case, the term “weighted mean" (WM) should be used. In formulas, this will result in indices like X_{KC} for CIPM comparisons and X_{wm} for other KCs.

By definition, the ***KC-RV is a tuple*** of the weighted mean and its associated combined expanded uncertainty. The calculation of the uncertainty part is dependent on the specific method of calculation of the mean value. The following formulas were used in several approved CCAUV KCs:

$$X_{\mathrm{KC}} = \left(\sum \frac{1}{u_i^2}\right)^{-1} \cdot \sum \frac{x_i}{u_i^2}$$

$$u_{\mathrm{KC}} = \left(\sum \frac{1}{u_i^2}\right)^{-\frac{1}{2}}$$

Generally, a RMO comparison which is linked to a CIPM KC does not need a reference value on its own as it inherits the CIPM-KC-RV by the linking procedure (see below).

3.2 Degree of Equivalence

The Degree of Equivalence (DoE) quantifies to what extent two measurement results may be considered equivalent. There are two approaches for this:

- unilateral DoE; equivalence with respect to the KC-RV or WM.
- bilateral DoE (Bi-DoE; a mutual equivalence between two participants.

In both cases a single ***DoE is a tuple of two values***: a deviation and its uncertainty.

For a unilateral DoE, the deviation is represented by the difference between the respective participant's result and the KC-RV (or WM); the uncertainty component in this case includes the correlation introduced by the weighted mean. Hence, a participant having consistent results the calculation has to take the correlation into account; for a participant that is not represented in the consistent set of results, the correlation is typically zero.

For a Bi-DoE, the deviation is represented by the difference between the results of the two considered participants. Typically there is no systematic correlation to be considered in this case. There may be an issue with correlation if there is a traceability relationship between those two participants; i.e. one claims traceability to the SI via a calibration performed by the other.

Note: According to a decision of the 9th CCAUV meeting [3], the explicit reporting of Bi-DoEs is discouraged. However, the formulas and data necessary for the calculation of correlations shall be included in the report.

3.3 Linking

RMO KCs are to be linked to previously performed CIPM-KCs, allowing the RMO KC results to become comparable to those of the CIPM-KC. As a consequence, the KC-RV becomes directly applicable to the RMO KC: ***no RMO reference value should be reported***.

The DoEs can be calculated in relation to the participants of the RMO as well as the CIPM KC.

A prerequisite for a proper linking is the existence of linking laboratories that took part in both comparisons and whose results are considered consistent in both cases. The linking itself is usually achieved by applying a shifting (additive) or scaling (multiplicative) transformation of the RMO KC results (including uncertainties) to the level of the CIPM KC results. The magnitude of the transformation is calculated by linking the results of the linking labs between both KCs. There is a multitude of proposals for linking procedures published in relevant metrology journals.

3.4 Comparisons without Linking

Sometimes the linking of a comparison (or of parts of its results) to a CIPM comparison is not possible. In such cases an individual reference value may be of interest. The calculation of such values should follow the guidelines for the KC-RV; however, due to the defined terminology (see 3.1) the value should only be named “weighted mean” or similar.

4 Reporting of KC results:

4.1 Representation of Results in Tables

The report of the comparison results should include ***at least*** the following quantitative information:

- Stability data of the artifacts during the measurement period of the KC and, if available, prior to the KC.
- Individual results of the participants (including uncertainties and relevant measurement conditions). Reasonable rounding should be applied.
- The measurand and the associated uncertainty placed in a double column, side by side.
- Results from linking calculations (factors, shift, …) including uncertainties.
- KC-RV and DoEs wherever applicable. Double columns are used for the tuples. The DoEs should not have better resolution than necessary to achieve two significant digits for the smallest uncertainty. The KC-RV should be documented with one significant digit more than for the results.
- In the result-table, the cells containing non-consistent results should be marked, color-coded or otherwise made identifiable.
- In the DoE tables, those cells containing non-consistent results should be marked, color-coded or otherwise made identifiable. In addition, the tuples where the deviation exceeds the expanded uncertainty should be marked in a different way.
- The TP shall be referenced or added as an appendix.
- The measurement uncertainty budgets of the participants, applicable to the reported measurements, should be added as appendix or referenced on the CCAUV website.
- A description of the system and methodology should be added as appendix.

4.2 Graphical Representation of Results

Typically stability data and unilateral DoEs are to be represented in graphical form as well as in tables.

These graphical charts should always include error bars in order to visualize the reliability or significance of the data. That is:

- Stability data should include the applicable expanded measurement uncertainty associated with the measurements given in the charts.
- DoE charts should include the respective expanded measurement uncertainty of the plotted DoEs.

Other items to consider are:

- Axis labels that should be given in the correct units (SI).
- Axis scaling that should be such that the error bars are within the chart and clearly visible.
- A limited number of curves in one chart in order to maintain readability.

This list is not exhaustive.

4.3 Autorship of the Final Report

Following the CIPM MRA Guidelines for authorship of Key, Supplementary and Pilot Study Comparison Reports [4], it is recommended to include in the front page of the final report not just one sole author, but also a co-author from each participating NMI/DI.

5 References

[1] CIPM MRA – D-05, *Measurement comparisons in the CIPM MRA,* *http://www.bipm.org/utils/common/CIPM_MRA/CIPM_MRA-D-05.pdf*

[2] *Key and supplementary comparison registration form* *http://www.bipm.org/utils/common/documents/jcrb/registration_form.doc*

[3] *CCAUV 2013 Report of the 9th meeting, to be published*

http://www.bipm.org/en/committees/cc/ccauv/publications_cc.html

[4] *CIPM MRA - G-04 VERSION 1, CIPM MRA Guidelines for Authorship of Key, Supplementary and Pilot Study Comparison Reports*

http://www.bipm.org/utils/common/documents/CIPM-MRA/CIPM-MRA-G-04.pdf