XXI IMEKO World Congress "Measurement in Research and Industry"
August 30 – September 4, 2015, Prague, Czech Republic

METHODS FOR CONFIRMATION OF CMC, BASED ON DEGREE OF EQUIVALENCE OF MEASUREMENT STANDARDS

A.Chunovkina, N.Burmistrova

D I Mendeleyev Institute for Metrology, 19, Moskovsky pr. St Petersburg, Russia,
A.G.Chunovkina@vniim.ru

Abstract: This paper deals with evaluation of inconsistent data of key comparison. The paper considers evaluation of key comparison value and degree of equivalence of measurement standards as well as the methods for confirmation of calibration and measurement capabilities, which are based on measurement uncertainties declared by participants and key comparison results.

Keywords: inconsistent data, degree of equivalence, pairwise degree of equivalence, calibration and measurement capabilities.

1. INTRODUCTION

According to MRA [1] "The degree of equivalence is taken to mean the degree to which these standards are consistent with reference values determined from the key comparisons and hence are consistent with one another. The degree of equivalence of each measurement standard is expressed quantitatively by two terms: it's deviation from the key comparison reference value and the uncertainty of this deviation (at a 95% level of confidence) ".

Two main tasks are to be solved at evaluation of key comparison (KC) data. The first one is establishing the degree of equivalence (DoEs) of measurement standards participating in key comparisons. And the second one is confirmation of calibration and measurement capabilities (CMC) declared by participating National Metrology Institutes (NMI).

During KC participating National metrology institutes (NMI) measure the same measurand provided by travelling measurement standard. All participants submit to the pilot NMI measurement results and associated uncertainties $(x_1, u_1), \ldots, (x_n, u_n)$. If the KC data are consistent, the KCRV is calculated as a weighted mean [2-3]:

$$x_{ref} = \frac{\sum \frac{x_i}{u^2(x_i)}}{\sum \frac{1}{u^2(x_i)}} \qquad u^2(x_{ref}) = \left(\sum \frac{1}{u^2(x_i)}\right)^{-1} \qquad (1),$$

The data consistency is usually checked by applying χ^2 test.

Consequently the degree of equivalence is given by the equation:

$$d_i = x_i - x_{ref}\,, \quad u^2(d_i) = u^2(x_i) - u^2(x_{ref}) \qquad (2).$$

Measurement uncertainty declared by the participants are confirmed by KC data if the following equation is true:

$$|d_i| \le 2u(d_i) \qquad (3).$$

According to MRA calibration and measurement capabilities is expressed in terms of expanded measurement uncertainties. If the equation (3) holds a minimum acceptable CMC expressed in the form of standard measurement uncertainty can't be less then $u(x_i)$ [4].

If the equation (3) does not hold it means that the results from a particular NMI are not consistent with others (within the level of measurement uncertainties declared by this NMI). In this case according to MRA the participant may enlarge corresponding measurement uncertainties submitted as CMC to BIPM data base. Method for determining acceptable CMC is given in [4]. This method mainly corresponds to the situation when there are a few outliers and the KCRV is calculated as weighted mean using consistent subset of measurement data.

As KC reports show in case of inconsistent data the pilot NMI tries to choose a method for data evaluation which would resolve original inconsistency of the data. A lot of methods for evaluation of inconsistent KC data are suggested. Most of these methods explicitly or implicitly result in increasing the measurement uncertainties declared by the KC participants. This fact raises a question of validity of the conventional procedure (3) for confirmation of measurement uncertainties originally declared by NMIs.

Some of the methods for inconsistent data evaluation are considered in the paper and their influence on measurement uncertainties originally declared (and consequently CMC confirmation) is discussed.

2. EVALUATION OF INCONSISTENT DATA AND CONFIRMATION OF CMC

If KC data are consistent, the conclusion concerning CMC confirmation is straightforward. But if the KC data are inconsistent there are a lot of methods for evaluation of KC reference value (KCRV) and degree of equivalence (DoES). So in these cases the confirmation of CMC is not so clear as in case of consistent KC data and it needs an additional consideration.

There are two main reasons for inconsistency of KC data. It is either underestimating of measurement uncertainties or overlooking systematic biases. In practice it is very difficult to separate these two cases.

Many different approaches have been proposed for analysis of inconsistent key comparison data [5-11, 13,14,16]. In this paper we consider the methods that imply increasing measurement uncertainties declared by NMIs originally. Birge ratio method [5] suggests to multiply all uncertainties by the same factors whereas Paule and Mandel method [16] suggests additive increasing of uncertainties. To our opinion such approaches correspond mainly to the task where calculating reference value is a final aim of the data analysis. Adjustment of fundamental constant is a good example of such task. The situation is different in KC data evaluation. Calculating KCRV is only the first step for calculation DoEs and confirmation declared measurement uncertainties. So one can expect that not all participants would like to increase their uncertainty equally (not taking into account their particular measurement results). Random effect model approach is similar to Paule and Mandel method and recently it becomes widely applied for KC data evaluation [9, 12]. Below this method is considered in more details.

3. RANDOM EFFECT MODEL APPOACH

Random effect model (REM) approach is based on the following model:

$$X_i = X + B_i \qquad (4),$$

where X – quantity provided by traveling measurement standard, B_i - random effect in i-th NMI, which is modelled by Normal distribution $B_i \in N(0, \sigma^2)$, X_i - measurand in i–th NMI.

Interpretation of B_i is ambiguous. Actually they are included into model (4) to provide consistency of measurement results with KCRV which is given by:

$$X_{ref} = \frac{\sum \frac{x_i}{u_i^2 + \hat{\sigma}^2}}{\sum \frac{1}{u_i^2 + \hat{\sigma}^2}} \qquad (5),$$

where $\hat{\sigma}$ is calculated numerically, for an example, by maximum likelihood method [6,9].

Application of REM approach implies that original measurement uncertainties, u_i, are changed to $\tilde{u}_i = \sqrt{u_i^2 + \hat{\sigma}^2}$. For i-th result measurement the uncertainty is increased $k_{REM,i} = \frac{\sqrt{u_i^2 + \hat{\sigma}^2}}{u_i}$ times. For the results with small uncertainties the increasing is more significant than for the results with large uncertainties. Disadvantage of the approach is that it does not take into consideration a closeness of particular measurement result obtained by the NMI to KCRV. It fully based on the parameter $\hat{\sigma}$, which characterises joint dispersion of all measurement data.

4. MCS ALGORITHM

In this paper the method with individual increasing of measurement uncertainties is advocated [13]. It is called metrological compatible set (MCS) algorithm because it proposes to increase uncertainties $\alpha_i \geq 1$ times in order to provide compatibility of all measurement results:

$$\left| x_i - x_j \right| \leq 2\sqrt{\alpha_i^2 u_i^2 + \alpha_j^2 u_j^2} \qquad \text{for all } i,j \ (6)$$

To obtain the unique solution for α_i it is proposed to minimise the following function: $\sum \alpha_i^2 u_i^2$.

Similar approach was earlier proposed in [14] but the idea of calculating α_i was different and it admited decrease of uncertainties. The idea of the given algorithm [13] is very simple and it is based on two statements:

- DoEs of measurement standards can be reasonably calculated only for metrological compatible set of data,
- For metrological compatible set of data the conventional approach (1)-(3) is applicable.

Performance of this approach is studied in [13] and it should be noted that it demonstrates good robustness. Namely excluding some measurement results does not significantly influence on other parameters α_i .

5. COMPARISON OF DIFFERENT METHODS

In this section REM and MCS algorithms are used for evaluating real data of KC with the aim:

- to compare the results of measurement uncertainty increasing and
- to analyze what measurement uncertainties can be recommended as minimum acceptable uncertainties in CMC.

The results of two algorithms application are compared with the results obtained by using the procedure in [4].

Measurement results with associated expanded uncertainties for CCQM-K5 [15] are presented in Figure 1.

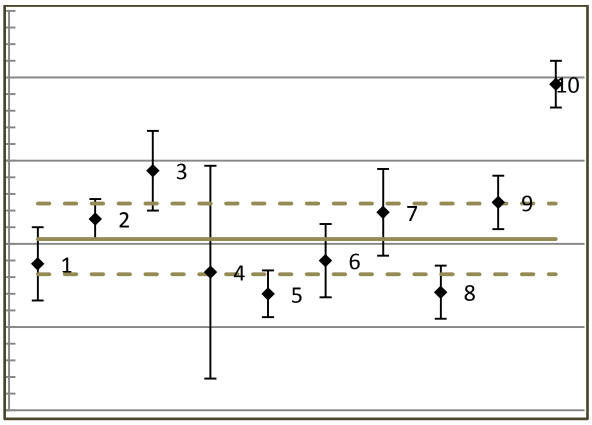

Fig.1. CCQM-K5 measurement data

The horizontal lines present KCRV and associated measurement uncertainty. In the report [15] KCRV was estimated as a simple mean and associated uncertainties were calculated as sample variance:

$$x_{ref} = \bar{x} = \frac{1}{n}\sum x_i, \; u^2(\vec{x}) = \frac{\sum (x_i - \vec{x})^2}{n(n-1)} \qquad (7).$$

This choice is caused by the fact that original data are inconsistent and measurement uncertainties declared seem to be underestimated for several results. It means that these uncertainties can be used as weights in calculating KCRV. Using (7) leads to situation when uncertainty associated with KCRV is larger than some uncertainties associated with measurement data and consistency with KCRV (3) does not any longer mean consistency of measurement results with each other. So in spite that (3) is true for some results it does not mean that originally measurement uncertainties are confirmed. Therefore they should be increased.

In Table 1 measurement results and expansion factors for REM (k_{REM}), MCS(k_{MCS}) algorithms and for two procedure (k_I, k_{II}), considered in [4], are given.

Table 1. Measurement data and expansion factors for CCQM-K5.

	x_i	u_i	k_I	k_{II}	k_{REM}	k_{MCS}
1	1,498	0,011	1,0	**1,2**	**3,5**	**1,1**
2	1,525	0,006	1,0	1,0	**6,2**	1,0
3	1,554	0,012	**1,5**	**1,3**	**3,2**	**2,4**
4	1,493	0,032	1,0	1,0	**1,5**	1,0
5	1,480	0,007	1,0	**3,2**	**5,3**	**3,5**
6	1,500	0,011	1,0	1,1	**3,5**	1,0
7	1,529	0,013	1,0	1,0	**3,0**	1,0
8	1,481	0,008	1,0	**2,8**	**4,7**	**2,9**
9	1,535	0,008	1,0	1,0	**4,7**	**1,8**
10	1,606	0,007	**6,5**	**5,8**	**5,3**	**8,3**

MCS expansion factors are less for all measurement result except the result 10 which is outlier. MCS expansion factors are close to factors obtained by procedure [4]. Finally, MCS algorithm proposes to increase uncertainties for 6 measurement results and 4 from these results originally have uncertainties that are significantly less than the others. Uncertainty increasing can be consider as a cost for compatibility/equivalence obtained for all measurement results.

In Figure 2 measurement data of CCM.P-K12 are presented. Traveling measurement standard drift was estimated and corrected data are given in Table 2.

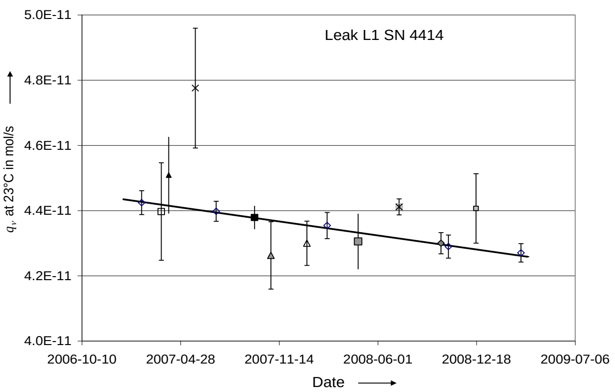

Fig.2. CCM.P-K12 measurement data

Table 2. Measurement data and expansion factors for CCM.P-K12

	x_i	u_i	REM		MCS	
			k_{REM}	E_n	k_{MCS}	E_n
1	4,348E-11	1,7E-13	2,6	0,50	1,0	0,12
2	4,326E-11	7,5E-13	1,1	0,30	1,0	0,18
3	4,441E-11	5,9E-13	1,2	0,51	1,0	0,76
4	4,719E-11	9,2E-13	1,1	**1,84**	2,4	0,83
5	4,346E-11	1,8E-13	2,5	0,45	1,0	0,19
6	4,238E-11	5,2E-13	1,3	**1,21**	1,8	0,60
7	4,291E-11	3,4E-13	1,6	1,00	1,8	0,52
8	4,317E-11	4,3E-13	1,4	0,57	1,0	0,41
9	4,440E-11	1,2E-13	3,4	**1,14**	3,6	1,00
10	4,347E-11	1,6E-13	2,7	0,42	1,0	0,17
11	4,455E-11	5,3E-13	1,3	0,79	1,2	0,81

In the report CCM.P-K12 [12] REM was chosen as a basic approach (other methods were also discussed) . The result 4 is an outlier and it influences the parameter $\hat{\sigma}$ significantly, but its corresponding expansion factor is very close to 1. Three of eleven results did not pass criterion (3) and consequently did not confirm uncertainties (in spite of increasing their uncertainties).

6. CONCLUSION

The paper discusses the methods applied for evaluation of inconsistent data of KC and the issue how these methods addresses the task of confirmation measurement uncertainties acceptable for CMC.

We share the opinion that if the method applied for evaluation of inconsistence data implies increasing the initial measurement uncertainties declared by participants it should be taken into account at CMC review.

The following items should be pointed out:

- for inconsistent data measurement uncertainty associated with KCRV is quite often not calculated using only the measurement uncertainties declared by the NMIs. But it also incorporates observed dispersion of measurement data. It leads to the situation when KCRV measurement uncertainty equals or even exceeds measurement uncertainties originally associated with some measurement results;
- in such cases the consistency of measurement results with KCRV does not any longer provides the mutual consistency (compatibility) of measurement results. So in this context consideration of pair-wise degrees of equivalence becomes a crucial issue for confirmation uncertainties declared .

We share the opinion that DoEs can be established only for a group of measurement standards that provide consistent data in key comparison. Otherwise measurement uncertainties should be increased. The paper advocates the method for inconsistent data evaluation (MCS), which on the first step propose procedure for forming compatible data set and on the second step it realises a conventional approach for evaluation consistent data. This method is compared with REM method and the procedures for increasing measurement uncertainties in [4], which implies uncertainty modification after comparison report approval. Consideration is illustrated by the examples of key comparisons from BIPM key comparison data base.

REFERENCES

1. CIPM MRA. Bureau International des Poidset Mesures 1999
2. Cox M G "The evaluation of key comparison data: An introduction" 2002 *Metrologia* 39 587-8
3. Cox M G "The evaluation of key comparison data" 2002 *Metrologia* 39 589-95
4. Cox M G,, Harris P, Milton M "Methods for determining acceptable CMC" http://www.bipm.org/wg/AllowedDocuments.
5. Lira I. "Combining inconsistent data from interlaboratory comparisons", 2007, *Metrologia* 44, pp. 415-21.
6. J.Willink R. "Statistical determination of a comparison reference value using hidden errors", 2002, *Metrologia* 39, pp. 343-54.
7. Cox M. G. "The evaluation of key comparison data: determining the largest consistent subset", 2007,*Metrologia* 44, pp. 187-200
8. Elster C., Toman B."Analysis of key comparisons: estimating laboratories' biases by a fixed effects model using Bayesian model averaging",2010, *Metrologia* 47, pp. 113-119
9. Toman B, Possolo A (2009) Laboratory effects models for interlaboratory comparisons Accreditation and Quality Assurance, 14, 553-563
10. C. Elster and B. Toman, Analysis of key comparison data: Critical assessment of elements of current practice with suggested improvements, *Metrologia* 50, 549 (2013).
11. A.G.Chunovkina, N.D.Zviagin, N.A.Burmistrova Evaluation of inconsistent data of key comparisons of measurement standards, ACTA IMEKO, December 2013, Volume 2, Number 2, 28 – 33
12. CCM.P-K12 http://kcdb.bipm.org/AppendixB/appbresults/ccm.p-k12/CCm.p-k12.pdf
13. N.A.Burmistrova Development and study of algorithms for processing inconsistent data in key comparison of measurement standards, Measurement Techniques 57 (10) (2015) 1103-1112
14. W Bremser , Self-consistent reference value estimation, Advanced mathematical and computational tools in metrology and testing IX (2012) 58-65
15. CCQM-K5 http://kcdb.bipm.org/AppendixB/appbresults/ccqm-k5/ccqm-k5_final_report.pdf
16. R.C. Paule, J. Mandel, Consensus values and weighting factors. J Res Nat Bur Std 87 (1982) 377–385