

REEVALUATION OF CCM KEY COMPARISONS IN MASS METROLOGY BY THE MONTE CARLO METHOD

Sungjun Lee and J. W. Chung

Division of Physical Metrology, KRISS, Daejeon 305-340, Rep. of Korea, lsjun@kriss.re.kr

Abstract – To provide insights for analysis of Key Comparisons (KCs), we have estimated previous KCs with Monte Carlo Simulation. Two similar previous KCs in mass metrology of 50 kg scale were taken in this work as the typical example. The KC Reference Value (KCRV) was evaluated with several popular estimators as well as median in this analysis. Their uncertainties were considered with Guide to the Expression of Uncertainty in Measurement (GUM) and numerical simulation by Monte Carlo Method (MCM).

Keywords: Monte Carlo Method, Key Comparison, KCRV, Uncertainty Evaluation

1. INTRODUCTION

Among the many tasks facing international metrology today is the conduct and analysis of Key Comparisons (KCs). The goal is to quantify the degrees of equivalence among realizations of the system of units and the primary measurement techniques in each field. These comparisons are summarized in published tables that relate each participating laboratory result to a KCRV that is, in general, an aggregate statistical estimator such as the mean (weighted uniformly or otherwise) or the median of the quoted laboratory results.

The need for statistically sound analysis of on-going CCM.M-K7, organized by some of authors, motivates this paper. Several approaches are known in analysis of Key Comparisons (KCs), but at present there is no commonly accepted methodology for statistical analysis of the inter-laboratory studies for all fields of metrology as described briefly in Section 3.

Another prompt reason could be that the values of working standards of Bureau International des Poids et Mesures (BIPM) have been changed due to wear the surface of eight BIPM working standards and as result of this, as-maintained BIPM mass unit was changed by 35 ug. As discussed in the 15th CCM meeting, all certificates issued from BIPM to the national prototype of kilogram during years of 2003 ~ 2013 are going to be updated [1]. This means that some of completed KCs had better re-estimate its KCRV after reviewing necessity of updating for somewhere like Regional Metrology Organization (RMO) KCs.

As the purpose of middle term mass society's plan, some realization experiments have been performed for redefinition of kilogram unit. According to CCM recommendation on a new definition of the kilogram [2], at least three independent experiments, including work from realization experiments,

yield "consistent" values of the Planck constant with relative standard uncertainties not larger than 5 parts in 10⁸. To fulfil the condition of consistency, it should be done via several international KCs among available realizations like watt balance and x-ray crystal density (XRCD) experiments.

However, results of six CCM KCs, CCM.M-K1~K6 were published in history of mass metrology [3-8]. In this paper, we have tried to investigate with the raw data of CCM.M-K3 (K3) and CCM.M-K6 (K6) to get the insights of difference between GUM uncertainty framework and MCM, and between uncertainty evaluations according to various estimators.

2. OVERVIEW OF MONTE CARLO METHOD

JCGM GUM SP1 document provides many of practical and technical considerations of MCM including comparison with the uncertainty evaluation by GUM and excellent examples to make understood the operational principles of MCM [9].

As the first step to the crude method, to make the formulized model, define the output quantity (the mesurand) and the input quantities upon which the output quantity depends. It had better develop a model relating the output quantity to these input quantities on the basis of available knowledge and assign probability density functions (PDFs), for example, Gaussian (normal) and rectangular (uniform) to input quantities on the model. Secondly, propagate the PDFs for the input quantities through the model to obtain the PDF for the output quantity.

As a final step, we can use the PDF for the output quantity to obtain the expectation of that quantity, taken as an estimate of the quantity, the standard deviation of that quantity, taken as the standard uncertainty associated with the estimate, and a coverage interval containing the output quantity with a specified. To reduce the burden of huge number of iteration, an adaptive Monte Carlo could be used for iteration block-by-block instead of crude MCM. Its procedure involves carrying out an increasing number of Monte Carlo trials until the various results of interest have stabilized in a statistical sense.

We have made MCM codes with National Instrument's LabVIEW by faithfully following the recommendation of JCGM GUM SP1 in our simulation. The enhanced version of pseudo-random number generator (PRNG) of Wichmann-Hill (WH) routine was implemented for sampling of data because some commercial random number function like in Microsoft's Excel has been failed to pass its randomness test

[11]. WH routine has a period of 2^{121} (~36 decimal digits). Some of our results were confirmed by comparing with other commercial tool like Palisade Corporation's @RISK, which adopted Mersenne Twister's PRNG based on the Excel.

The logarithmic scaled plot on the top of Fig 1 shows the time ratio (calculation time divided by maximum value, 270 s in 6.2×10^6) depending on the number of iteration. As it expected, the computing time is the linear relationship with the number of iteration. However, we need to apply for special skills in like array management and formula node to get optimized calculation speed in LabVIEW.

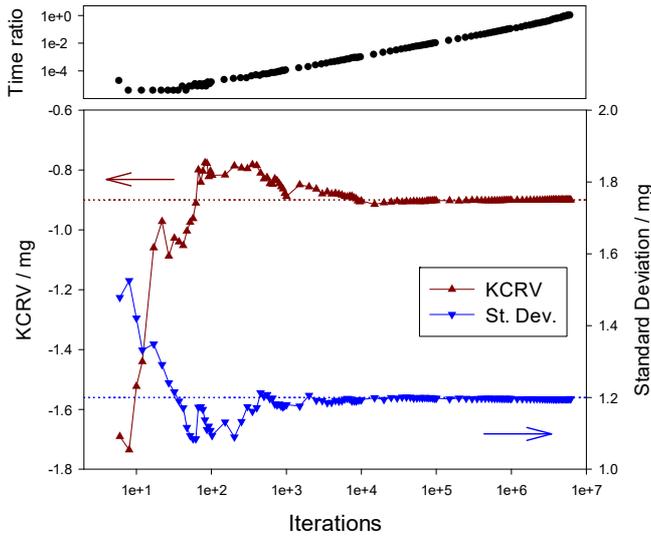


Fig. 1. (Bottom) semi-log plot. Stabilization of CCM.M-K6's KCRV and its uncertainty (standard deviation) by crude MC simulation as number of iteration increased. (Top) log-log plot. Processing time ratio of the simulation

3. KCRV IN MASS METROLOGY

3.1. Averaging Method (Estimator)

One example of using weighted mean can be found in the guideline of KC adopted by BIPM in certain limited situation [10]. However, almost all mass fields' KCs used median estimator in evaluating Key Comparison Reference Value (KCRV) except CCM.M-K4. The advantage of using median, one of great methods in measuring of central tendency, is not affected by extreme scores but it shows some difference in sets with odd numbers of observations than for even numbers of observations. Contrary to mass metrology, the weighted mean was frequently used in KCs of length metrology.

3.2. Uncertainty of Median

No definite hints for estimation of median's uncertainty were found and the way of uncertainty evaluation was slightly different in each field of metrology. The Mueller's version [12] of the median absolute deviation (MAD) shown in Eq. (1) and various forms of Eq. (2) ~ Eq. (4) including different version of MAD were used in mass metrology and electricity/magnetism metrology, respectively. The various

possible forms were found in other field like the ref. 13. σ , y_i , $u^2(y_i)$, and y_{med} can be considered as standard deviation or uncertainty, participant's reported value and variance, and median of the number of participants (N) in KCs, respectively. Here, k is a constant depending on N , which can also be found in ref. 13.

$$\sigma \approx \frac{1.858}{\sqrt{n-1}} \text{med} \{|y_i - y_{med}|\} \quad (1)$$

$$\sigma \approx k \cdot \text{med} \{|y_i - y_{med}|\} \quad (2)$$

$$\sigma \approx \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (y_i - y_{med})^2} \quad (3)$$

$$\sigma \approx \sqrt{\frac{1}{N^2} \sum_{i=1}^N u^2(y_i)} \quad (4)$$

$$\sigma \approx 1.1926 \cdot \text{med} \{\text{med}|y_i - y_j|\} \quad (5)$$

Eq. (5) is called double median deviation and the alternative to the MAD that can be used as initial or ancillary scale estimates in the same way but that are more efficient and not slanted towards symmetric distributions [14]. For each i participant, we compute the median of $\{|x_i - x_j|\}$, for $j = 1, \dots, n$. This yields n numbers, the median of which gives our final standard deviation.

Table 1 shows the estimated values according to Eq (1) ~ (5) in GUF based on the reported value of each KC. The value of Eq. (1) was underlined and recalculated value based on KC report. The value in the parenthesis of Table 1 was obtained with the 6.2×10^6 whole corrected values of resampled by MCM. This kind of way is similar with common standard deviation from the average values in MCM. We have to comment the size of sample is important when we applied for the uncertainty formula to KC evaluation.

Table 1. Calculated uncertainty according to Eq (1) ~ (5).

Equation	CCM.M-K3	CCM.M-K6
<u>(1)</u>	<u>1.0</u>	<u>1.5</u>
(2)	3.2	3.6
(3)	0.6	0.9
(4)	1.3	0.9
(5)	2.6 (1.2)	3.0 (1.2)

3.3. CCM.M-K3 and CCM.M-K6

CCM.M-K3 (K3) and CCM.M-K6 (K6) were used 50 kg transfer standards. K3 and K6 used the Guide to the Expression of Uncertainty in Measurement (GUM) uncertainty framework (GUF) and MCM using commercial package (@RISK) in calculating KCRV and its uncertainty, respectively. As shown in dashed line of Fig. 1, we have

observed the same conversion value of MCM with reported values by adopting the same procedures in evaluating.

In this result, 10^6 iterations recommended in ref 10 are quite enough to get the stabilized value but iterations even less than 10^5 is still acceptable for quick observation. The same tendency was found in other calculation for K3. In K6 analysis, some correlation was considered for multiple measurements in pilot and co-pilot labs.

Table 2. Reported (underlined) and calculated KCRV and standard uncertainty (u_{KCRV}) in the platforms of GUF and MCM using median estimator

KC Code	KCRV / mg		u_{KCRV} / mg	
	GUF	MCM	GUF	MCM
CCM.M-K3	<u>51.25</u>	50.78	<u>1.05</u>	1.4
CCM.M-K6	-0.7	<u>-0.9</u>	1.5	<u>1.2</u>

Table 2 shows the summarized result in each KC. Some slight difference between GUF and MCM's results was observed and possibly its reason could be asymmetry of median density function. Especially, the uncertainty expression of median shown in Eq. (1) for mass metrology used in most of previous KCs was derived in the assumption of normal distribution function. But, Fig. 2 presents this asymmetrical shape of the distribution function for median estimator discussed in the following. There is a minor difference in calculating the correction in two KC's pilot labs. K6 used the correction with respect to the pilot lab and the KCRV is around zero [8]. If the correction is followed by K3's procedure for the pilot lab, K6's KCRV of MCM should be about -1.02 mg but KCRV by GUF is still valid. That's the reason why two KCRVs' distribution functions, which obtained from the official report and this simulation as shown in right side of Fig 2 using the same median estimator, are slightly different.

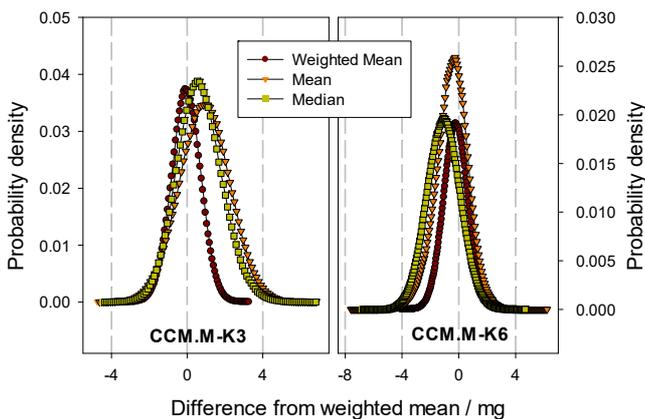


Fig. 2. Comparison of probability density functions simulated from CCM.M-K3 and CCM.M-K6 depending on the estimator.

3.4. Estimator

Some popular estimators were considered: weighted mean, arithmetic mean and median. As shown in Fig 2, MCM can simulate the probability density function from the

histogram of population obtained from much iteration of model and resampled data from participant's report. The central value is different depending on estimator. The degree of symmetry can be checked by calculating of skewness and kurtosis found in Table 3.

Skewness is a measure of symmetry in terms of the left and right of the centre point and kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. The expected values of skewness and kurtosis are 0 and 3 for symmetric normal distribution. Therefore, the arithmetic mean and the weighted mean estimators have the symmetrical shape in their PDF. However, PDF of median estimator showed long tails of right (positive) and left (negative) side for K3 and K6, respectively.

Table 3. Simulated mean of estimator, standard deviation (std. dev.), skewness and kurtosis depending on some estimators: weighted median (W. mean), arithmetic mean (A. mean) and median

		W.Mean	A.Mean	Median
CCM.M-K3	Mean of estimator (mg)	50.1	51.0	50.8
	Std. Dev. (mg)	0.6	1.8	1.4
	Skewness	0.0	0.0	0.2
	Kurtosis	3.0	3.0	3.2
CCM.M-K6	Mean of estimator (mg)	-0.7	-1.6	-1.0
	Std. Dev. (mg)	0.9	1.2	1.2
	Skewness	0.0	0.0	-0.2
	Kurtosis	3.0	3.0	3.4

4. CONCLUSIONS

We have demonstrated two previous CCM KCs were similar result by GUF and MCM. The uncertainty of median obtained by MCM is the same order with GUF in using Mueller's version among other possible schemes. The slight difference was minor issue but it will be a great concern to some participants whose value is nearby the boundary of outlier.

However, Monte Carlo methods (MCM) in tremendous works have been used as a fast and convenient tool to create the distributions for any candidate reference value in KC. For example, A. Steele et al. applied the pair-difference chi-squared tests with MCM for KCs analysis. We agree that the application of MCM can provide physical insight into some of the occasions when it is inappropriate to average measurement results statistically [15].

ACKNOWLEDGMENTS

The author would like to thank Mr. Luis Omar Becerra of CENAM for his guide to MCM world through the Key Comparison and providing @RISK data of CCM.M-K6.

REFERENCES

- [1] "CCM Recommendation to NMIs on managing the consequences of the corrections to the BIPM as-maintained mass unit" CCM/15-17, February 2015
- [2] CCM Recommendation G1 (2013) "On a new definition of the kilogram", March 2013
- [3] "Final report on CIPM key comparison of 1 kg standards in stainless steel (CCM.M-K1)", *Metrologia*, vol. 41, Tech. Suppl., pp. 07002, 2004
- [4] "Final Report on CIPM key comparison of multiples and submultiples of the kilogram (CCM.M-K2)", *Metrologia*, vol. 40, Tech. Suppl., pp. 07004, 2003
- [5] "Final Report: CCM.M-K3 comparison/50 kg mass, *Metrologia*", vol. 42, Tech. Suppl., pp. 07003, 2005
- [6] "Final report on CCM.M-K4: Key comparison of 1 kg stainless steel mass standards, *Metrologia*", vol. 51, Tech. Suppl., pp. 07009, 2014
- [7] "Report on CIPM key comparison of the second phase of multiples and submultiples of the kilogram (CCM.M-K5)", *Metrologia*, vol. 48, Tech. Suppl., pp.07008, 2011
- [8] "Final report of CCM key comparison of mass standards CCM.M-K6, 50 kg", *Metrologia*, vol. 52, Tech. Suppl., 07004, 2015
- [9] JCGM 101:2008 guide, "Guide to the expression of uncertainty in measurement", 2008
- [10] M.G. Cox, "The evaluation of key comparison data: An introduction," *Metrologia*, vol. 39, 587 – 588, 2002
- [11] B. D. McCullough, "Microsoft Excel's Not The Wichmann-Hill random number generators", *Computational Statistics and Data Analysis*, vol. 52, pp.4587-4593, 2008
- [12] Joerg W. Mueller, Possible Advantages of a Robust Evaluation of Comparisons, *Journal of Research of the National Institute of Standards and Technology*, vol. 105, n°. 4, pp.551, (2000)
- [13] J. Randa, "Update to proposal for kCRV & degree of equivalence for GTRF key comparisons", GT-RF/2005-04, 2005
- [14] P. J. Rousseeuw, C. Croux, "Alternatives to the Median Absolute Deviation" *Journal of the American Statistical Association*, vol. 88, n°. 424, pp. 1273-1283, 1993
- [15] A G Steele, K D Hill and R J Douglas, "Data pooling and key comparison reference values" *Metrologia*, vol 39, pp.269, 2002