UDC 006.91:053.088 DOI:

Data processing specific features of supplementary bilateral comparisons

I. Zakharov[1,2], O. Botsiura[2], I. Zadorozhnaya[1]

[1] National Scientific Centre "Institute of Metrology", Myronosytska Str., 42, 61002, Kharkiv, Ukraine
newzip@ukr.net
[2] Kharkiv National University of Radio Electronics, Nauky Ave., 14, 61166, Kharkiv, Ukraine
olesia.botsiura@nure.ua

Abstract

The data processing specific features of supplementary bilateral comparisons are considered. Recommendations are given for checking the stability of the travelling standard, as well as the procedure for processing data for two types of comparisons of the measurement standards: comparisons of the measurement standards that haven't been included in the list of key comparisons and those that take the unit size from the participants in key comparisons. As an example, comparisons of Koesters interferometers for measuring the gage blocks in the range from 0.1 to 100 mm are considered. It is shown that the obtained discrepancy of the measurement results confirmed a good degree of equivalence of the installations being compared.

Keywords: supplementary bilateral comparisons; measurement uncertainty; Koesters interferometer; gage block.

Received: 04.03.2020 Edited: 19.05.2020 Approved for publication: 28.05.2020

Introduction

As is it known [1], supplementary comparisons of the national measurement standards are carried out in order to confirm the Calibration and Measurement Capabilities (CMC) of National Metrological Institutes (NMI). When evaluating these supplementary comparisons, it is advisable to distinguish two types of comparisons [1]:

1) comparisons of the measurement standards that have not been included in the list of key comparisons;

2) comparisons of the measurement standards that take the size of the unit from participants in key comparisons.

In metrological practice, situations are often encountered where supplementary comparisons are bilateral. Such comparisons are allowed by Recommendation [1], but have their own specific features, the consideration of which is the purpose of this article.

Main part

1. Checking the stability of the travelling standard

When carrying out both types of comparisons, the necessary condition is the stability of the travelling standard (TS).

To check the stability of the TS, the pilot laboratory determines the arithmetic mean values of two groups of multiple measurements of the value X_s reproduced by the TS at the beginning and at the end of comparisons: $\bar{X}_{sb}$, $\bar{X}_{se}$, and standard uncertainties of type A of these values $u_A(\bar{X}_{sb})$, $u_A(\bar{X}_{se})$. Student's test is applied to the obtained results [2]:

$$t = \frac{\left|\bar{X}_{sb} - \bar{X}_{se}\right|}{\sqrt{u_A^2(\bar{X}_{sb}) + u_A^2(\bar{X}_{se})}} \le t_\alpha(\nu_{eff}), \qquad (1)$$

where $t_\alpha(\nu_{eff})$ is the Student's coefficient for a given level of significance and the effective number of degrees of freedom ν_{eff}, determined by the Satterthwaite formula [3]:

$$\nu_{eff} = \frac{[u_A^2(\bar{X}_{sb}) + u_A^2(\bar{X}_{se})]^2}{\dfrac{u_A^4(\bar{X}_{sb})}{n_b - 1} + \dfrac{u_A^2(\bar{X}_{se})}{n_e - 1}} \qquad (2)$$

with unequal variances in groups, and according to the formula:

$$\nu_{eff} = n_b + n_e - 2 \qquad (3)$$

with equal variances in groups. In formulas (2) and (3) n_b and n_e – the number of measurements in groups at the beginning and end of comparisons.

The equality of variances in groups is determined by the Fisher criterion [4]:

$$\psi = \frac{u_A^2(\bar{X}_{sg})}{u_A^2(\bar{X}_{sl})} \le \psi_0, \qquad (4)$$

where $u_A(\bar{X}_{sl}) = \max[u_A(\bar{X}_{sb}); u_A(\bar{X}_{se})]$,
$u_A(\bar{X}_{sg}) = \min[u_A(\bar{X}_{sb}); u_A(\bar{X}_{se})]$
are the larger and smaller standard uncertainties of type A from the comparisons determined at the beginning and at the end; ψ_0 – a critical point from the Fisher table for numbers of degrees of freedom $\nu_b = n_b - 1$, $\nu_e = n_e - 1$, and a given significance level α.

Example 1. Between June 2009 and August 2009 supplementary international comparisons were carried out according to the COOMET 265/UA/02 Project – "Comparisons of high-accuracy gage block interferometers" [5]. As TS, quartz and steel gage blocks (GB) with a nominal size of 100 mm were used. In order to determine the stability of TS, the pilot laboratory (NSC "Institute of Metrology") made ten measurements of the deviations of the central length of each GB from the nominal size at the beginning Δ_b and at the end Δ_e of comparisons. The arithmetic mean of these deviations $\bar{\Delta}$ and their standard uncertainties of type A $u_A(\bar{\Delta})$ are presented in Table 1.

Table 1

TS measurement results at the beginning and end of comparisons

Parameter	Deviation of the central length at the beginning of comparisons, Δ_b, μm		Deviation of the central length, at the end of comparisons Δ_e, μm	
	Quartz GB	Steel GB	Quartz GB	Steel GB
—	1.4367	0.05174	1.4392	0.05218
$u_A(\bar{\Delta})$	0.0033	0.0004	0.0047	0.0005

Let us verify the equality of variances according to the Fisher criterion (4):

$$\psi \le \psi_0,$$

where ψ_0 is the critical point from the Fisher table for the number of degrees of freedom $\nu_1 = \nu_2 = n - 1 = 9$ and a given significance level α. For the data in Table 1, values ψ and ψ_0 for different significance levels α are given in Table 2.

Table 2 shows that the variance values at the beginning and end of comparisons can be considered equal for all significance levels.

Table 2

Determination of equality of variances of GB at the beginning and end of comparisons

Gage blocks	Values ψ	Critical values ψ_0		
		$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
Steel GB No.423	1.111	2.44	3.18	5.35
Quartz GB No.11	2.028			

Table 2 shows that the variance values at the beginning and end of comparisons can be considered equal for all significance levels.

Using the Student test, we find:

$$t = \frac{|\bar{\Delta}_b - \bar{\Delta}_e|}{\sqrt{u_A^2(\bar{\Delta}_b) + u_A^2(\bar{\Delta}_e)}} \le t_\alpha(\nu_{eff}),$$

where the effective number of degrees of freedom is $\nu_{eff} = n_1 + n_2 - 2 = 18$.

For the data in Table 1, values t and $t_\alpha(\nu_{eff})$ for different significance levels α are given in Table 3.

Table 3 shows that both gage blocks can be considered stable for any significance level.

Table 3

The stability of GB determination

Gage blocks	Values t	Values $t_\alpha(18)$		
		$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
Steel GB No.423	0.682	1.734	2.101	2.878
Quartz GB No.11	0.433			

2. Data processing of the supplementary bilateral comparisons of the first type

With this type of comparisons, it is not necessary to know the real value reproduced by TS. What is important, is that this value should be stable. Unlike CIPM key comparisons, the degree of equivalence of the measurement standards is not established for this type of comparisons; only the consistency of data from different NMIs is determined, which is an objective confirmation of the declared uncertainties.

Each NMI participating in supplementary comparisons represents the result of measuring x_j by TS and the corresponding combined standard uncertainty $u(x_j)$. Based on these data, the value of the Pearson criterion χ^2 is calculated:

$$\chi^2 = \sum_{j=1}^{m} \frac{(x_j - \overline{x}_{ref})^2}{u^2(x_j)} \le \chi^2_{0,95}(m-1), \qquad (5)$$

where m is the number of participants in comparisons; $\overline{x}_{ref}$ is the reference value of supplementary comparisons, calculated by the formula:

$$\overline{x}_{ref} = \frac{\sum_{j=1}^{m} \frac{x_j}{u^2(x_j)}}{\sum_{j=1}^{m} \frac{1}{u^2(x_j)}}; \qquad (6)$$

$\chi^2_\alpha(m-1)$ is critical value for the significance level α and the number of degrees of freedom $m-1$, taken from the Table [6]. For bilateral comparisons, expressions (5) and (6) take the following form:

$$\chi^2 = \frac{(x_1 - \overline{x}_{ref})^2}{u^2(x_1)} + \frac{(x_2 - \overline{x}_{ref})^2}{u^2(x_2)} = \frac{(x_1 - \overline{x}_{ref})^2 u^2(x_2) + (x_2 - \overline{x}_{ref})^2 u^2(x_1)}{u^2(x_1) u^2(x_2)}, \qquad (7)$$

$$\overline{x}_{ref} = \frac{\frac{x_1}{u^2(x_1)} + \frac{x_2}{u^2(x_2)}}{\frac{1}{u^2(x_1)} + \frac{1}{u^2(x_2)}} = \frac{x_1 u^2(x_2) + x_2 u^2(x_1)}{u^2(x_1) + u^2(x_2)}, \qquad (8)$$

and the values $\chi^2_\alpha(1)$ for different significance levels α are given in Table 4.

When inequality (5) is fulfilled, the data of different NMIs can be recognized as consistent, which is an objective confirmation of the declared uncertainties.

If criterion (5) is not met, in accordance with [1], it is necessary to identify inconsistent data. For this purpose, the NMI is determined that provides the maximum criterion E_n:

$$\max_i E_n = \frac{|x_i - x_{ref}|}{2\sqrt{u^2(x_i) - u^2(x_{ref})}}, \qquad (9)$$

where

$$u^2(\overline{x}_{ref}) = \frac{1}{\sum_{j=1}^{m} \frac{1}{u^2(x_j)}} \qquad (10)$$

is the variance of the reference value (6).

Further, the data of this NMI are temporarily excluded from consideration, and the procedure described above is repeated. Successive data exclusion is repeated until condition (5) is satisfied.

It should be noted that in case of bilateral comparisons, it is impossible to identify inconsistent data, because at the same time, each of their NMIs provides the same E_n criterion value:

$$E_{n1} = E_{n2} = \frac{|x_1 - x_2|}{2\sqrt{u^2(x_1) + u^2(x_2)}}. \qquad (11)$$

However, in this case, it becomes possible to evaluate the fulfillment of the E_n criterion:

$$E_n \le 1, \qquad (12)$$

confirming the consistency of the results of comparisons.

Example 2. The results of comparisons of COOMET 265/UA/02 [5] of high-accuracy gage block interferometers (Koesters interferometers) belonging to the NSC "Institute of Metrology" (Ukraine) and South Kazakhstan branch of the RSE "Kazakhstan Institute of Metrology" (Kazakhstan) are deviations of the measurement results of TS values from the nominal value and their combined standard uncertainties $u_c(\Delta)$ (Table 5).

Table 4

Values $\chi^2_\alpha(1)$ for different significance levels α

α	0.01	0.025	0.05	0.1
$\chi^2_\alpha(1)$	6.635	5.024	3.841	2.706

Table 5

Comparison data of Koesters interferometers [5]

Gage blocks	NSC "Institute of Metrology"		RSE "Kazakhstan Institute of Metrology"	
	Δ_1, μm	$u_c(\Delta_1)$, μm	Δ_2, μm	$u_c(\Delta_1)$, μm
Steel GB No.423	0.05218	0.007	0.06169	0.0177
Quartz GB No.11	1.4392	0.006	1.4315	0.0172

We find by the formula (8):

$$\overline{x}_{ref\text{st}} = \frac{0.0522\cdot 0.00177^2 + 0.062\cdot 0.007^2}{0.00177^2 + 0.007^2} = 0.053466\ \mu\text{m};$$

$$\overline{x}_{ref\text{qu}} = \frac{1.439\cdot 0.00172^2 + 1.432\cdot 0.006^2}{0.00172^2 + 0.006^2} = 1.438365\ \mu\text{m}.$$

Hence, by the formula (7) we have:

$$\chi^2_{\text{st}} = \frac{(x_1-\overline{x}_{ref})^2 u^2(x_2) + (x_2-\overline{x}_{ref})^2 u^2(x_1)}{u^2(x_1)u^2(x_2)} = 0.25;$$

$$\chi^2_{\text{qu}} = \frac{(x_1-\overline{x}_{ref})^2 u^2(x_2) + (x_2-\overline{x}_{ref})^2 u^2(x_1)}{u^2(x_1)u^2(x_2)} = 0.18.$$

Since the obtained values χ^2 are significantly less than $\chi^2_\alpha(1)$ indicated in Table 4 for all significance levels, the NMI data were recognized as consistent, which was an objective confirmation of the uncertainties declared by them.

In addition, the values of the E_n criterion were determined by the formula (11):

$$E_{n\text{st}} = \frac{|0.05218 - 0.06169|}{2\sqrt{0.007^2 + 0.177^2}} = 0.25;$$

$$E_{n\text{qu}} = \frac{|1.4392 - 1.4315|}{2\sqrt{0.006^2 + 0.0172^2}} = 0.21.$$

Since the criterion (12) was fulfilled for both GBs, the comparisons were considered consistent.

3. Data processing of the supplementary bilateral comparisons of the second type

When performing supplementary comparisons of this type, it is necessary to know the real value being reproduced by TS x_{ref} and its uncertainty $u(x_{ref})$, which are determined by a reference laboratory that is a participant in key comparisons in this type of measurement. Supplementary comparisons of this type are carried out strictly in order to confirm CMC, therefore, in this case, the measurement procedure for comparisons and the calibration procedure are identical.

When evaluating data for this type of comparisons, it is important to take into account the correlations of the measurement results arising from taking the size of the unit [1]. Each NMI participating in COOMET supplementary comparisons represents the measurement result x_j and the corresponding combined standard uncertainty $u(x_j)$, as well as the uncertainty budget.

Based on this data, to check the consistency of these comparisons, the E_n criterion value for each NMI is calculated:

$$E_{nj} = \frac{|x_j - x_{ref}|}{2\sqrt{u^2(x_j) + u^2(x_{ref}) - 2\,\text{cov}(x_j, x_{ref})}}, \quad (13)$$

where $\text{cov}(x_j, x_{ref})$ is the covariance of the measurement results and the reference value of the supplementary comparisons, due to the taking of the unit size from the reference laboratory. To calculate covariance, it is necessary to analyze the uncertainty budget of the participant in comparisons and the reference laboratory and identify those components that are common and unchanged $u_0^2(x_i)$:

$$\text{cov}(x_j, x_{ref}) = u_0^2(x_j). \quad (14)$$

Since, as a rule, the common component for all laboratories is the reference value x_{ref}, its variance $u^2(x_{ref})$ is equal to the covariance $\text{cov}(x_j, x_{ref})$.

In this case, expression (12) takes the form:

$$E_{nj} = \frac{|x_j - x_{ref}|}{2\sqrt{u^2(x_j) - u^2(x_{ref})}}. \quad (15)$$

If condition (12) is satisfied for the obtained value E_{nj}, then this confirms the declared uncertainties of the j-th NMI. Those NMIs for which inequality (14) is not satisfied should analyze the reasons for the loss of their results. As a result of the analysis, the following can be established:

- the measurement result is an error, and the NMI decides to exclude its result. In this case, the declared uncertainties are not confirmed in the course of these supplementary comparisons, and to confirm them, participation in other similar supplementary comparisons is necessary;
- evaluation of NMI uncertainty is underestimated, which requires identification and presentation of reasons for understating to the pilot laboratory and other participants in comparisons; participants in the comparisons agree with the explanations presented, after which the initially declared uncertainty is increased so that condition (12) is satisfied.

Example 3. Table 6 shows the values and combined standard uncertainties of steel GB, taken from its calibration certificate, which participated as a travelling standard in COOMET 265/UA/02 comparisons of Koesters interferometers [5].

Table 6

TS characteristics and comparison results

Gage blocks	l_{ref}, mm	$u(l_{ref})$, μm	NSC "Institute of Metrology"		SKB of the RSE "Kazakhstan Institute of Metrology"	
			l_1, mm	$u_c(\Delta_1)$, μm	l_2, mm	$u_c(\Delta_1)$, μm
Steel GB No.423	100.00000018	0.004	100.00005218	0.007	100.00005218	0.007

Substituting the values given in Table 6 into expression (15), we obtain:

$$E_{n1} = \frac{|x_1 - x_{ref}|}{2\sqrt{u^2(x_2) - u^2(x_{ref})}} = 0.835;$$

$$E_{n2} = \frac{|x_2 - x_{ref}|}{2\sqrt{u^2(x_2) - u^2(x_{ref})}} = 0.919.$$

Since the obtained values $E_n < 1$, this confirms the declared uncertainties of each NMI.

Conclusions

1. To check the stability of the traveling standard, the pilot laboratory determines the arithmetic mean values of two groups of multiple measurements of the reproducible TS at the beginning and at the end of comparisons and the standard uncertainties of type A of these values, and according to these data it should check the equality of variances in the groups using the Fisher criterion and the insignificance of the bias of arithmetic mean groups using the Student's test.

2. When processing the data of supplementary bilateral comparisons of the first type, it is not necessary to know the real value reproduced by the TS. In this case, using the Pearson criterion, it is necessary to determine the degree of consistency of the NMI data, which is an objective confirmation of the declared uncertainties. In bilateral comparisons, it is impossible to identify inconsistent data, because in addition, each of their NMIs provides the same E_n criterion value. In this case, the fulfillment of the E_n criterion is an additional confirmation of the consistency of these NMIs.

3. When processing data of supplementary bilateral comparisons of the second type, it is necessary to know the real value reproduced by the TS and its uncertainty. When evaluating the data for this type of comparisons, it is important to take into account the correlations of the measurement results arising from taking the size of the unit.

Особливості обробки даних додаткових двосторонніх звірень

І.П. Захаров[1,2], О.А. Боцюра[2], І.М. Задорожна[1]

[1] *Національний науковий центр "Інститут метрології", вул. Мироносицька, 42, 61002, Харків, Україна*
newzip@ukr.net
[2] *Харківський національний університет радіоелектроніки, пр. Науки, 14, 61166, Харків, Україна*
olesia.botsiura@nure.ua

Анотація

Досліджено алгоритм обробки даних додаткових звірень. Наведено рекомендації з перевірки стабільності еталона, що транспортується, на основі критеріїв Фішера та Стьюдента. Розглянуто порядок обробки даних для двох

типів звірень еталонів: звірення еталонів, які не були внесені до списку ключових звірень, і звірення еталонів, які запозичують розмір одиниці в учасників ключових звірень.

При обробці даних додаткових двосторонніх звірень першого типу не є необхідним знання дійсного значення, відтвореного еталоном, що транспортується. При цьому необхідно за допомогою критерію Пірсона визначити ступінь узгодженості даних національних метрологічних інститутів, що є об'єктивним підтвердженням калібрувальних і вимірювальних можливостей. Показано, що при двосторонніх звіреннях виявлення неузгоджених даних провести неможливо. Однак у цьому випадку виконання критерію E_n є додатковим підтвердженням узгодженості даних національних метрологічних інститутів.

При обробці даних додаткових двосторонніх звірень другого типу необхідно знати дійсне значення величини, відтвореної еталоном, що транспортується, та її невизначеність. При оцінюванні даних для цього типу звірень проводиться облік кореляцій результатів вимірювань, що виникають внаслідок запозичення розміру одиниці.

Як приклад розглянуто двосторонні звірення інтерферометрів Кестерса для вимірювання кінцевих мір довжини в діапазоні 0,1–100 мм, що дозволили підтвердити калібрувальні та вимірювальні можливості національних метрологічних інститутів. Рішення провести звірення прийнято на засіданні ТК 1.5 "Довжина і кут" КООМЕТ. Пілотною організацією звірень був ННЦ "Інститут метрології" (Україна), другим учасником звірень виступав Казахстанський РДП ПКФ "КазІнМетр".

Отримані результати дозволили зробити висновок про узгодженість даних національних метрологічних інститутів, що підтвердило заявлені ними невизначеності вимірювань.

Ключові слова: додаткові двосторонні звірення; невизначеність вимірювань; інтерферометр Кестерса; кінцева міра довжини.

Особенности обработки данных дополнительных двусторонних сличений

И.П. Захаров[1,2], О.А. Боцюра[2], И.Н. Задорожная[1]

[1] *Национальный научный центр "Институт метрологии", ул. Мироносицкая, 42, 61002, Харьков, Украина*
newzip@ukr.net
[2] *Харьковский национальный университет радиоэлектроники, пр. Науки, 14, 61166, Харьков, Украина*
olesia.botsiura@nure.ua

Аннотация

Рассмотрены особенности обработки данных дополнительных двусторонних сличений. Приведены рекомендации по проверке стабильности транспортируемого эталона, а также порядок обработки данных для двух типов сличений эталонов: сличения эталонов, которые не были внесены в список ключевых сличений, и сличения эталонов, которые заимствуют размер единицы у участников ключевых сличений. В качестве примера рассмотрены сличения интерферометров Кестерса для измерения концевых мер длины в диапазоне 0,1–100 мм. Показано, что полученное расхождение результатов измерений подтвердило хорошую степень эквивалентности сличаемых эталонов.

Ключевые слова: дополнительные двусторонние сличения; неопределенность измерений; интерферометр Кестерса; концевая мера длины.

References

1. COOMET R/GM/19:2008. Guideline on COOMET supplementary comparison evaluation. Available at: http://www.coomet.org/EN/doc/r19_2008.pdf
2. Student. The Probable Error of a Mean. *Biometrika*, 1908, vol. 6(1), pp. 1–25.
3. Satterthwaite F.E. An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 1946, vol. 2(6), pp. 110–114.
4. Pearson K. Tables for Statisticians and Biometricians. Cambridge, Eng., Printed at the University Press, 1914–31.
5. COOMET.L-S3. Final report. Supplementary comparisons. COOMET Project 265/UA/02. Comparisons of high-accuracy gage block interferometers. *Metrologia*, 2011, 48(1A):4002-04002. doi: 10.1088/0026-1394/48/1A/04002
6. Fisher R.A. On the interpretation of χ^2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society*, 1922, vol. 85(1), pp. 87–94.