Evaluation of the quality of measurements of the testing laboratory during interlaboratory comparative tests

Olga Mukhamedzhanova[1*] and *Anastasia* Borovkova[1]

[1]Moscow State University of Civil Engineering, Yaroslavskoe shosse, 26, Moscow, 129337, Russia

Abstract. The article presents the statistical processing of the test results of the mass fraction of the chlorine ion (Cl) obtained from two accredited testing laboratories. The stages of statistical processing of the results of interlaboratory comparison tests are as follows: elimination of gross errors, determination of the uniformity of the sample, calculation of the assigned value, evaluation of the quality of the measurement results. The presence of gross errors in the sample was determined by the Grubbs test, and the uniformity of the sample was confirmed by the Fisher's test and Student's test. The assigned value was calculated by a robust method using algorithm A with an iterative scale. The quality of laboratory measurement results was evaluated using the z-score. The z-score value for each participant showed that the results of measuring the mass fraction of chlorine ion in cement samples were acceptable. The quality of the results of these laboratories is satisfactory and does not require corrective measures. These results of the test quality assessment can be used for the organizers of interlaboratory tests, as well as for accredited laboratories that check their qualifications.
Keywords: interlaboratory comparison, proficiency testing, robust method, assigned value, measurement error, outlier, standard deviation for proficiency assessment.

1 Introduction

One of the main goals of the testing laboratories is to provide the customer with reliable and accurate results. Ensuring the reliability of the results in the testing laboratory is carried out by conducting comprehensive internal and external quality control of its activities. Internal quality control involves conducting in-lab tests, calibrating equipment using reference materials, etc. External control is the verification of qualifications by conducting interlaboratory comparison tests (ICT). [1-5,18]

In accordance with the policy of the Russian Federal Accreditation Service, interlaboratory comparison tests must be carried out by providers accredited for testing certain types of products, but currently there are no accredited providers for interlaboratory testing of building materials, products, and structures. Therefore, interlaboratory comparison tests are conducted either between testing laboratories, where the provider is one of the laboratories or an organization that takes over the functions of the ICT provider. Conducting

[*] Corresponding author: MuhamedjanovaOG@mgsu.ru

interlaboratory comparison tests consists of the following stages: sampling and preparation of the task, the ISI program, encryption of samples, sending samples for testing to testing laboratories and testing them in accordance with the appropriate methodology. Sending test reports to the ICT organizer, who processes the received tests using statistical methods. One of the critical stages of the ICT is the correct and reliable processing of test results obtained from testing laboratories. [5-18]

The purpose of the study was statistical processing of data on the results of conducting interlaboratory comparison tests.

To achieve this goal, the following tasks were solved:

- description of the main stages of statistical data processing of the results of interlaboratory comparison tests;
- determination of the assigned value based on the consensus value of the participants
- calculating the assigned value using various methods;
- analysis of the tests for evaluating the work of laboratories.

2 Method of statistical evaluation of interlaboratory comparison tests

Statistical processing of the results of interlaboratory comparison tests was carried out according to GOST R 50779.60-2017 (ISO 13528: 2015). The stages of statistical processing of the results of the ICT are as follows: checking for sharply deviating values, determining the uniformity of the sample, calculating the assigned value, evaluating the quality of the measurement results. (Figure 1)

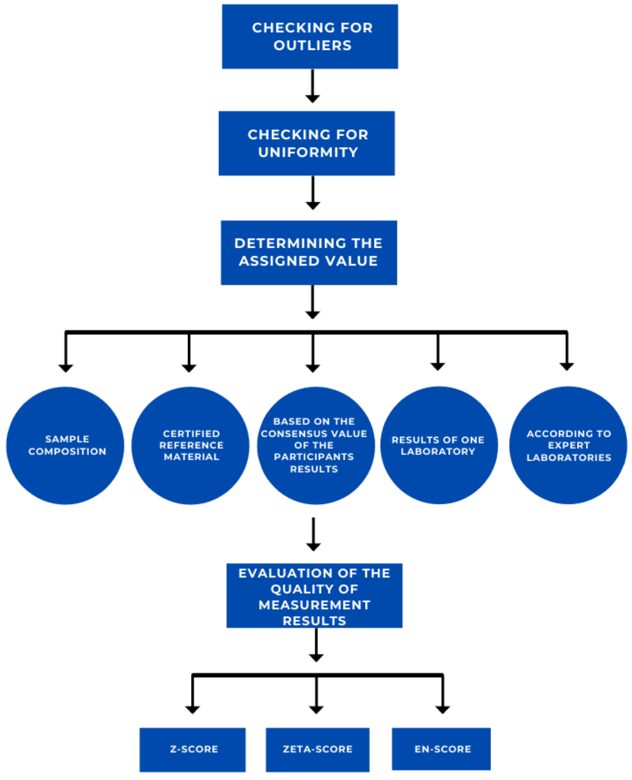

Fig. 1. Stages of statistical processing of the results of the ICT.

The first stage of processing the results of the ISI is to check the samples for outliers (misses) according to the Grubbs test. To do this, calculate the Grubbs test G1 for the maximum value and G2 for the minimum value using the following formulas:

$$G_1 = \frac{|x_{max} - \bar{x}|}{S}$$

$$G_2 = \frac{|\bar{x} - x_{min}|}{S}$$

where:

$\bar{x}$ – arithmetic average value;

xmax, xmin – maximum, minimum value in a series of measurements;

S – standard deviation.

Compare G1 and G2 with the table value Gт, which depends on the number of tests n and the confidence probability Рд. If G1≤ Gт, then Xmax is not considered a miss and is stored in a series of measurement results. If G1>Gт, then Xmax is excluded as an unlikely value and the arithmetic mean and standard deviation are calculated again, and the procedure for checking for misses is repeated until the inequality G1≤ Gт is satisfied. By the same calculation, the Grubbs test is checked for the minimum value of G2. [13-18]

The second stage of statistical data processing is to determine the uniformity of the sample using the Fisher's test for the uniformity of standard deviations and the Student's test – the uniformity of the average values. To do this, you need to calculate the calculated value of the Fisher's test and Student's test using the following formulas:

$$t_э = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{{Gx_1}^2 - {Gx_2}^2}} \times \sqrt{n}$$

$$F_э = \frac{{Gx_1}^2}{{Gx_2}^2} \times x$$

where:

n = n1= n2 – the number of partial values of the variable in the sample;

$\bar{X}_1$ – the average value of the variable for the first data sample;

$\bar{X}_2$ – the average value of the variable for the second data sample;

Gx_1^2 и ${Gx_2}^2$ – indicators of deviations of the partial values from the two compared samples from the corresponding average values.

Compare tэ and Fэ with the table values of tт and Fт, which depend on the number of tests n and the confidence probability of Рд. If the theoretical value is greater than the calculated Fт > Fэ and tт > tэ, then the hypothesis of equality of variances and averages in the tested series is not rejected. In this case, it is assumed that the dispersion of results within each series and between series performed under the same conditions reflects only random errors.

In the third stage, the assigned value x_{pt}. The assigned value x_{pt}– is the value assigned to a specific property of the sample for proficiency testing. The assigned value is determined by the following methods: the composition of the sample material; the certified standard sample; the results of one laboratory; the consensus value according to the data of the expert laboratory; based on the consensus value of the results of the participants (Figure 1), the methods for determining the assigned value are chosen by the organizer of the ICT. One of the complex methods of determining the assigned value is to determine this value based on the consensus value of the results of the participants of the ICT, which will be discussed in this article.

The method of determining the assigned value based on the consensus value of the results of the participants of the ICT is determined by a robust analysis of calculating the arithmetic mean and standard deviation of the measurement results of all laboratories participating in these tests. There are simple robust methods for algorithm A with an iterative scale, algorithm S. Elaborate methods for calculating a robust estimate Q-method, Hampel estimate, Q/Hampel. [16-18]

To determine the assigned value, a robust method was used according to algorithm A, since this method is prime in calculations and is used with a small number of participants in the ICT. [17-18]. The determination of the assigned value of x_pt according to the algorithm A with iterative calculation is presented is shown in Figure 2.

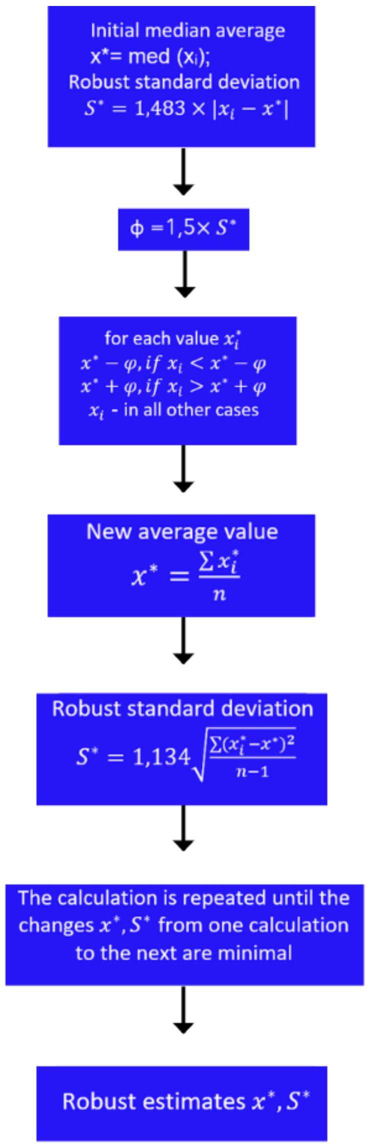

Fig. 2. Calculation of the assigned value according to the algorithm A with an iterative scale.

At the fourth stage, the quality of the measurement results is evaluated. At this stage, the deviations of the participant's measurement results are compared with the assigned value obtained. This comparison is performed using standardized methods (z-score, zeta-score, En – score). The participants ' results were evaluated by the z-score, because the assigned value is compared with the standard deviation of the participants in the qualification test. In other evaluation methods, the total standard uncertainty of the result of each participant of the ICT is taken as a comparison with the prescribed value.

The Z-scores for the result of the qualification check x^* are calculated using the following formula:

$$Z_i = \frac{(x_i - x_{pt})}{\sigma_{pt}}$$

where:

x_{pt} – assigned value;

σ_{pt}- standard deviation for proficiency assessment.

The standard deviation for proficiency assessment σ_{pt} is calculated by the formula:

$$\sigma_{pt} = \sqrt{\sigma_R^2 - \sigma_r^2 \times (1 - \frac{1}{n})}$$

where:

σR – standard deviation of reproducibility;

σr - standard deviation of repeatability;

n – the number of measurements of each participant of the ICT.

In accordance with GOST 5382-2019 " Cements and materials of cement production. Methods of chemical analysis " the standard deviation of repeatability (σr) is calculated by the formula:

$$\sigma_r = 0{,}866 \times \overleftarrow{R}$$

$\overleftarrow{R}$ – average range for all parallel tests

Standard deviation of reproducibility of measurement results (σR) under reproducibility conditions:

$$\sigma_R = \sqrt{\frac{\sum(x_i - \overleftarrow{x})^2}{n - 1}}$$

where:

xi – result of the i-analysis;

$\overleftarrow{x}$- average analysis result for all data;

n – number of tests (at least 20).

The evaluation of the obtained values from the test laboratories by the Z-score is as follows:

- the result is considered acceptable if │z│≤ 2,0;
- the result is in the warning zone (warning signal) if 2,0 <│z│< 3,0;
- the result is considered unacceptable (action signal) if │z│ ≥ 3.

A Z-score value of more than 2,0 implies the need to analyze the possible causes of what is happening, and an index value of 3,0 implies the implementation of corrective actions.

 https://doi.org/10.1051/e3sconf/202337601051

3 Results and discussions

Two test laboratories participated in the interlaboratory comparison tests, the measurements were carried out according to GOST 5382-2019 "Cements and materials of cement production. Methods of chemical analysis" by the phototurbidimetric method. The mass fraction of the chlorine ion in the cement samples was determined. The results of the tests of the mass fraction of the chlorine ion (Cl) are presented in Table 1

Table 1. Results of statistical processing of tests of the mass fraction of chlorine-ion in cement samples of interlaboratory comparison tests.

N	Mass fraction of the chlorine ion (Cl).%		Assigned value			
	participant-1	participant-2	Mass fraction of the chlorine ion (Cl) 1 iteration		Mass fraction of the chlorine ion (Cl) 2 iteration	
1	0.0164	0.0153	0.0153	0.0165	0.0153	0.0165
2	0.0164	0.0167	0.0156	0.0165	0.0156	0.0165
3	0.0164	0.0181	0.0156	0.0165	0.0156	0.0165
4	0.0165	0.0177	0.0156	0.0165	0.0156	0.0165
5	0.0163	0.0156	0.0157	0.0165	0.0157	0.0165
6	0.0165	0.0177	0.0160	0.0165	0.0160	0.0165
7	0.0166	0.0184	0.0162	0.0166	0.0162	0.0166
8	0.0165	0.0175	0.0163	0.0166	0.0163	0.0166
9	0.0164	0.0176	0.0163	0.0166	0.0163	0.0166
10	0.0163	0.0166	0.0163	0.0167	0.0163	0.0167
11	0.0165	0.0156	0.0164	0.0175	0.0164	0.0175
12	0.0164	0.0177	0.0164	0.0176	0.0164	0.0176
13	0.0165	0.0156	0.0164	0.0176	0.0164	0.0176
14	0.0165	0.0162	0.0164	0.0177	0.0164	0.0177
15	0.0163	0.0177	0.0164	0.0177	0.0164	0.0177
16	0.0164	0.0157	0.0164	0.0177	0.0164	0.0177
17	0.0165	0.0188	0.0164	0.0177	0.0164	0.0177
18	0.0165	0.0176	0.0165	0.0181	0.0165	0.0181
19	0.0166	0.0165	0.0165	0.0184	0.0165	0.0184
20	0.0164	0.0160	0.0165	0.0188	0.0165	0.0188
$\bar{X}$	0.0164	0.0169	x*	0.0165	x*	0.0165
S	0.000089	0.001069	S^*	0.0034	S^*	0.0034
G_1	2.2472	1.7774	φ	0.0031	φ	0.0031
G_2	1.1236	1.4967	$x^* - \varphi$	0.0114	$x^* - \varphi$	0.0114
$G_т$	2.7090					
$F_э$	2.33		$x^* + \varphi$	0.0216	$x^* + \varphi$	0.0216
$F_т$	2.38					
$t_э$	0.83		x^*	0.0165	x^*	0.0165
$t_т$	2.09		S^*	0.0007	S^*	0.0007
№ participant			σ_r	σ_R	σ_{pt}	Z_i
1			0.005	0.0001	0.000088	1.136
2			0.005	0.0011	0.001092	0.366

The measurement results obtained by the ICT participants were checked for the presence of sharply deviating values according to the Grubbs test. For the first participant, the Grubbs test is G1=2,2472; G2=1,1236 ≤ Gт =2,7090, hence Xmax Xmin are not misses; for the second participant, G1; G2 ≤ Gт =2,7090, so Xmax Xmin are not misses. There are no gross errors in the presented test results of the mass fraction of the chlorine ion of the two participants.

The uniformity of the average values and standard deviations obtained from the two laboratories was checked using the Student's test and Fisher's test (Table 1). The following inequalities are satisfied: the calculated value of the Student's test $t_э$=0,83 < t_T=2,09, the calculated value of the Fisher's test $F_э$=2,33< $F_т$ =2,38, therefore, the standard deviations and the average values in the tested test series are equal to each other. The results of the two samples are homogeneous and consistent, so they are combined to calculate the assigned value by a robust method.

The assigned value was determined by the robust method in accordance with the algorithm A by iterative calculation. The results of processing the assigned value of the mass fraction of chlorine ion in cement samples are shown in Table 1. Table 1 shows that for iterations 1 and 2, x^* = 0,0165%, S^* = 0,0034% are equal, therefore, the assigned value xpt= 0,0165%.

The quality of the measurement results obtained by the ICT participants is estimated by the value of the z-index, which is calculated for each laboratory (Table 1). For the first participant, Z-score1 = 1,136, for the second participant Z-score2 = 0,366. From the results, it is clear that the Z-score for two laboratories is less than 2,0, therefore, the result for determining the mass fraction of chlorine-ion in cement samples from these laboratories is acceptable. The quality of the measurement results of the laboratory data using the phototurbidimetric method of chemical analysis is satisfactory. If the value of the Z-score for individual participants is more than 2,0 or 3,0, then these results are considered unsatisfactory (in the warning zone) and this implies the need to analyze the possible causes of what is happening and implement corrective measures.

4 Conclusions

1. The stages of statistical processing of the test results of the mass fraction of the chlorine ion (Cl) are as follows: removal of gross errors, determination of the uniformity of the sample, calculation of the assigned value, evaluation of the quality of the measurement results. Checking the samples showed: the absence of gross errors according to the Grubbs test, the consistency of the average and standard deviations according to the Student’s and Fisher’s test.
2. The assigned value was determined on the basis of the consensus value of the results of the participants of the ICT and was calculated by a robust method using algorithm A with iterative calculation. The assigned value of the mass fraction of the chlorine ion in the cement samples is x_{pt} = 0,0165%.
3. The evaluation of the quality of the test results of the mass fraction of the chlorine ion of the two laboratories was carried out according to the value of the z-score. For the first laboratory Z-score 1 = 1,136, for the second laboratory Z-score 2 = 0,366, since | Z-score 1,2 $|\leq$ 2,0, the results of the laboratories are acceptable. The quality of the results of these laboratories according to the phototurbidimetric method of chemical analysis is satisfactory and does not require corrective measures.
4. The above calculations on the statistical processing of the results of participants in the interlaboratory comparison tests can be used for the organizers of the interlaboratory tests, as well as for accredited laboratories that conduct the qualification check.

References

1. A. Matveeva, B. Ezhak, B. Voida, Journal of Product Quality control **6**, 26-30 (2019) https://doi.org/10.20915/2687-0886-2020-16-2-41-55

2. O.G. Muhamedjanova, D.G. Demidov, I.G. Rekus, L.Y. Komorova, Proceedings of the Tula State University **2**, 542-547 (2021) https://doi: 10.24412/2071-6168-2021-2-542-547
3. L. Grgec Bermanec, D. Zvizdic, Int. J. Metrol. Qual. Eng. 6, 307 (2015) https://doi.org/10.1051/ijmqe/2015021
4. E. Ozhogina, M. Lebedeva, Measurement Standards. Reference Materials **13(2)** (2017) https://doi.org/10.20915/2077-1177-2017-13-2-37-47
5. O. Mukhamedzhanova, D. Dolgorukov, Transportation Researh Pcocedia **63**, 2608-2613 (2022) DOI:10.1016/j.trpro.2022.06.300
6. D. James, D. Ames, Journal of Clinical Pathology (2014) DOI:10.1136/jclinpath-2013-201621
7. Jagan Kavya, B. Alistair, Int. J. Metrol. Qual. Eng. **10**, 3-8 (2019) https://doi.org/10.1051/ijmqe/2019003
8. E. Lorenzato, F. Vianna, Int. J. Metrol. Qual. Eng. **13(4)**, 4-9 (2022) https://doi.org/10.1051/ijmqe/2022004
9. M.G. Cox, Metrologia **44**, 187–200 (2007) https://doi.org/10.1142/9789812702647_0004
10. M. Shojaei, A. Ahmadi, P. Shojaei, Int. J. Qual. Res. **13**, 349–360 (2019) DOI – 10.24874/IJQR13.02-07
11. M. Ikram, P. Zhou, S.A.A. Shah, G.Q. Liu, J. Clean. Prod. **226**, 628–641 (2019) https://doi.org/10.1016/j.jclepro.2019.03.265
12. L.R. Vega-González, R.M. Vega-Salinas, J. Technol. Manag. Innov. **13**, 90–103 (2018) http://dx.doi.org/10.4067/S0718-27242018000300090
13. F.P. Hochleitner, A. Arbussà, G. Coenders, J. Technol. Manag. Innov. **11**, 50–58 (2016) http://dx.doi.org/10.4067/S0718-27242016000300006
14. D.A. Philippi, E.A.M. Maccari, P.R. Costa, J. Technol. Manag. Innov. **13**, 54–63 (2018) http://dx.doi.org/10.4067/S0718-27242018000300054
15. R.A. Martins, P.L.O.C. Neto, Gestão Produção **5**, 298–311 (1998) https://doi.org/10.1590/S0104-530X1998000300010
16. L. Lisienkova, T. Shindina, N. Orlova, L. Komarova, Civil Engineering Journal **7(08)**, 662-675 (2021) http://dx.doi.org/10.28991/cej-2021-03091732
17. L. Lisienkova, MATEC Web Conf. **239**, 01013 (2018) DOI: https://doi.org/10.1051/matecconf/201823901013
18. L.N. Lisienkova, I.A. Savenkova, E.V Baranova, IOP Conf. Ser.: Mater. Sci. Eng. **476**, 012019 (2019) https://doi.org/10.1088/1757-899X/476/1/012019