The 57th International October Conference on Mining and Metallurgy
05 - 07 October 2026, Bor Lake, Serbia
https://ioc.irmbor.co.rs

DOI: 10.5937/IOC26315M Original Scientific Paper- Materials Science

STATISTICAL EVALUATION OF LABORATORY PERFORMANCE USING Z-SCORE IN INTERLABORATORY TESTING

Mladen Macić[1a], Slavica Miletić[1b], Biserka Trumić[1c]
[1]Mining and Metallurgy Institute Bor, Alberta Ajnštajna 1, 19210 Bor, Serbia
[1a] mladen.macic214@gmail.com, https://orcid.org/0009-0006-0928-6596
[1b] slavica.miletic@irmbor.co.rs, https://orcid.org/0000-0003-4526-4715
[1c] biserka.trumic@irmbor.co.rs, https://orcid.org/0000-0001-8328-7315

Abstract

This paper presents a statistical analysis of interlaboratory testing results from four laboratories, aimed at evaluating their performance using the z-score method. The analyzed results involve tensile testing of metallic materials in accordance with SRPS ISO 6892-1:2020, Method B. The considered parameters are tensile strength (Rm) and maximum force (Fm). Each laboratory tested eight specimens of identical material. For the statistical evaluation of the results, the mean values, standard deviations, repeatability variance, interlaboratory variance, and reproducibility variance were calculated. Laboratory performances were assessed using the z-score. The obtained results showed that all laboratories had an absolute z-score value of less than 2 for both analyzed parameters, indicating that the results were within the satisfactory range. The largest deviation for (Rm) was determined in laboratory A3, while for (Fm), the largest deviation was found in laboratory A2. The variance component analysis revealed that interlaboratory variability has a dominant influence on the overall variability of the results.

Keywords: *interlaboratory testing, z-score, tensile strength, maximum force, repeatability, reproducibility.*

1. INTRODUCTION

The reliability of laboratory test results represents one of the fundamental requirements in the field of testing metallic materials. The comparability of results obtained in different laboratories is of particular importance, as it enables the assessment of laboratory capabilities to achieve results with an acceptable degree of consensus by applying the same testing methods.

Interlaboratory testings constitute an important tool for evaluating laboratory performance and verifying the capability of a laboratory to obtain a reliable result. This approach is aligned with the concept of ensuring the validity of laboratory test results and laboratory competence. ISO/IEC 17025:2017 specifies general requirements for the competence, impartiality, and consistent operation of laboratories, while ISO/IEC 17043:2023 defines competence and performance requirements for providers of proficiency testing schemes.

For the statistical processing of interlaboratory testing results, ISO 13528:2022 is of significant importance, as it defines statistical methods for data analysis used in proficiency testing and interlaboratory comparisons.

One of the most frequently used laboratory performance indicators is the z-score. It represents the normalized deviation of a laboratory's result from the assigned, i.e., reference value. In its general form, it can be expressed as:

$$z = \frac{\bar{x} - \bar{\bar{x}}}{S_R}$$

Where: $\bar{x}$ is the laboratory result, $\bar{\bar{x}}$ is the assigned or reference value, and S_R is the standard deviation for proficiency assessment.

In this paper, in accordance with the procedure applied in the specific interlaboratory testing, the standard deviation of reproducibility (S_R) was used for deviation normalization.

The aim of this work is to evaluate the performance of four laboratories in tensile strength (Rm) and maximum force (Fm) testing using the z-score, as well as to analyze the components of variability of the obtained results.

2. MATERIAL AND METHODS

Four laboratories participated in the interlaboratory testing:

A1 – Materials Testing Laboratory – LIM, Mining and Metallurgy Institute Bor;

A2 – Laboratory for Mechanical Testing of Steel Wire, Ropes and Slings, Steel Wire Rope Factory – UNIS-USHA A.D.;

A3 – Laboratory for Testing of Materials and Machines, Department of Mechanical Engineering, Faculty of Mechanical Engineering in Niš;

A4 – Mining Institute d.d. Tuzla.

The testing was conducted on specimens of identical material, with each laboratory testing eight specimens.

The tensile testing was performed in accordance with SRPS ISO 6892-1:2020, Method B – Metallic materials – Tensile testing – Part 1: Method of test at room temperature.

The following parameters were analyzed: tensile strength (Rm) and maximum force (Fm).

The arithmetic mean of the results was calculated for each laboratory. Subsequently, the variance components were calculated: repeatability variance (Sr2), interlaboratory variance (SL2), and reproducibility variance (SR2). The z-score was used for performance assessment.

The results were evaluated based on the absolute value of the z-score:

$|z| < 2$ satisfactory result;

$2 < |z| < 3$ warning/questionable results;

$|z| > 3$ unsatisfactory results.

This interpretation is consistent with the standard criteria for evaluating the performance of participants in proficiency testing.

3. RESULTS AND DISCUSSION

3.1. Tensile strength (Rm)

Table 1 shows the mean values of the tensile strength results by laboratory.

Table 1. *Mean values of tensile strength (Rm) and z-scores*

Laboratory	Mean value (Rm)	Z-score	Evaluation
A1	1659.750	-0.39	satisfactory
A2	1596.500	-0.95	satisfactory
A3	1859.500	1.38	satisfactory
A4	1699.625	-0.04	satisfactory

The overall mean value of tensile strength was 1703.844. The lowest mean value was obtained in laboratory A2 (1596.500), while the highest was obtained in laboratory A3 (1859.500).

The largest absolute deviation from the overall mean value was determined in laboratory A3, with a z-score of 1.38. However, this value is below the limit (|z|=2), meaning the result of laboratory A3 is evaluated as satisfactory. For laboratories A1, A2, and A4, absolute z-values below 1 were obtained.

The following variance components were obtained for tensile strength:

Sr^2	Sr	S_L^2	S_L	S_R^2	S_R
167,6920	12,9496	12550,2403	112,0279	12717,9323	112,7738

It can be observed that the interlaboratory variance **(S_L^2)** is significantly larger than the repeatability variance (**Sr^2**). According to the obtained values, the interlaboratory component accounts for approximately 98.7% of the total reproducibility variance. This indicates that for this parameter, differences between laboratories are a more significant source of variability than the variation of results within individual laboratories.

3.2. Maximum force (Fm)

The results of the statistical processing of maximum force are presented in Table 2.

Table 2. *Mean values of maximum force (Fm) and z-scores*

Laboratory	Mean value (Rm)	Z-score	Evaluation
A1	5160.000	-0.11	satisfactory
A2	5012.500	-1.31	satisfactory
A3	5270.125	0.78	satisfactory
A4	5252.875	0.64	satisfactory

The overall mean value of maximum force was 5173.875. The lowest mean value was obtained in laboratory A2, and the highest in laboratory A3.

The largest absolute deviation was determined in laboratory A2, with a z-score of -1.31. Nevertheless, the result remains within the satisfactory range. Laboratories A1, A3, and A4 had absolute z-score values of less than 1.

The following variance components were obtained for maximum force:

Sr^2	Sr	S_L^2	S_L	S_R^2	S_R
1189,563	34,49	13764,46	117,322	14954,023	122,29

As with tensile strength, the interlaboratory component of variance has a dominant influence on the overall variability. Its share in the total reproducibility variance is

approximately 92.0%. The obtained result indicates that differences between laboratories have a more significant impact on result variability than repeated measurements within individual laboratories.

3.3. Comparative Analysis

By summarizing the obtained results, it can be noted that no laboratory produced an unsatisfactory result for either of the two analyzed parameters. The largest positive z-score value was obtained for (Rm) in laboratory A3 (1.38), while the largest absolute negative value was obtained for (Fm) in laboratory A2 (-1.31).

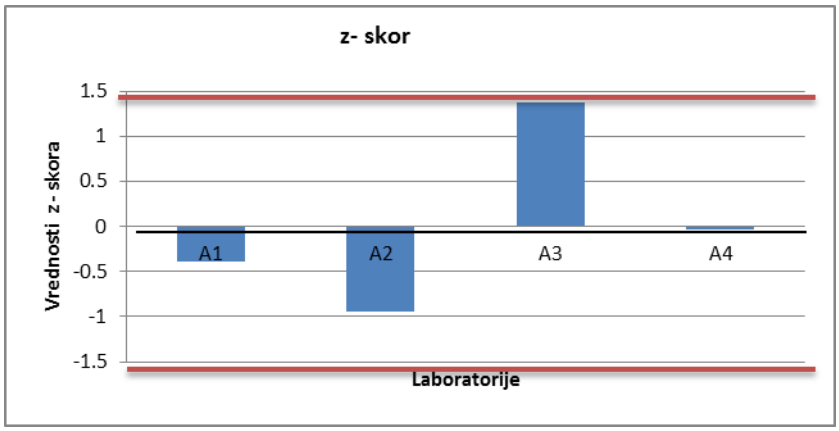


Figure 1. *Graphical interpretation of z-score for Rm*

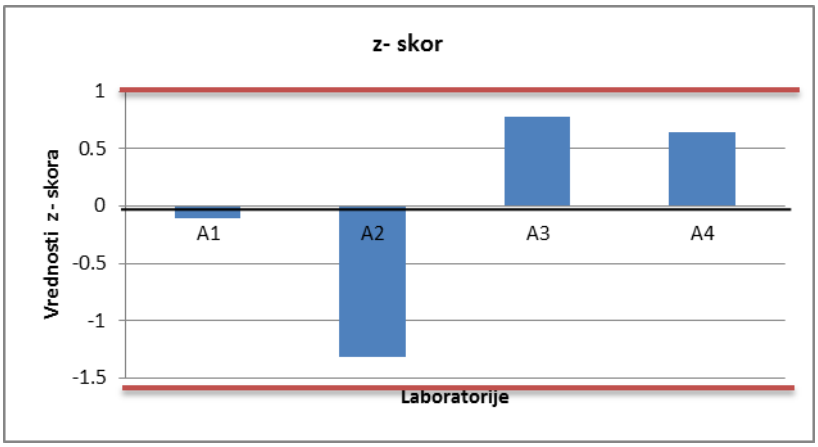


Figure 2. *Graphical interpretation of z-score for Fm*

All results fall within the limits (|z|<2), indicating satisfactory laboratory performance in this specific interlaboratory testing cycle.

However, the obtained results demonstrate that a satisfactory z-score does not mean that all laboratories achieved completely identical results. On the contrary, differences in mean values are present and can be further analyzed through variance components. For this reason, it is useful to consider the repeatability and reproducibility of results alongside the z-score.

Modern approaches to performance evaluation in proficiency testing tend to combine multiple statistical parameters to ensure a more comprehensive overview of laboratory competence.

4. CONCLUSION

Based on the results of the interlaboratory testing and the application of the z-score statistical method, it can be concluded that all four laboratories demonstrated satisfactory performance during the testing of tensile strength (Rm) and maximum force (Fm).

For tensile strength, the z-values ranged from -0.95 to 1.38, while for maximum force, they ranged from -1.31 to 0.78. No absolute value of the z-score reached the limit of the warning zone|z|=2).

The variance component analysis revealed that interlaboratory variability has a dominant influence on the overall variability for both analyzed parameters. For (Rm), the interlaboratory component accounts for approximately 98.7% of the total reproducibility variance, while for (Fm), its share is approximately 92.0%.

Based on the obtained results, it can be stated that the laboratories in this specific interlaboratory testing cycle demonstrated a satisfactory capability to apply the tensile testing method in accordance with SRPS ISO 6892-1:2020. The z-score proved to be a suitable indicator for the comparative evaluation of laboratory performance.

However, to evaluate the long-term stability of laboratory operations, it is necessary to monitor results across multiple interlaboratory testing cycles, as a single cycle provides a performance evaluation under specific testing conditions only.

ACKNOWLEDGEMENTS

This work is financially supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, Agreement on the implementation and financing of scientific research work of the National Research Institute in 2026, No. 451-03-33/2026-03/ 200052.

REFERENCES

[1] ISO 13528:2022, Statistical methods for use in proficiency testing by interlaboratory comparison. ISO, Geneva, 2022.
[2] ISO/IEC 17043:2023, Conformity assessment — General requirements for the competence of proficiency testing providers. ISO, Geneva, 2023.
[3] ISO/IEC 17025:2017, General requirements for the competence of testing and calibration laboratories. ISO, Geneva, 2017.

[4] SRPS ISO 6892-1:2020, Metallic materials – Tensile testing – Part 1: Method of test at room temperature.
[5] Cordeiro, F., Emons, H., Robouch, P. (2022), Is the z score sufficient to assess participants' performance in proficiency testing? The hidden corrective action, Accreditation and Quality Assurance, 27, 1–9.
[6] Proficiency testing report through interlaboratory comparisons, Report No. P031.240-23.001, dated January 19, 2023, Mining and Metallurgy Institute Bor, Serbia.