Interlaboratory comparison of roughness measurement: Application of Algorithm A of ISO 13528: 2015 in determining the designated value and the standard deviation

S F Beckert$^1$ and G E Fischer
Federal University of Santa Catarina, Joinville, SC, 89.219-600, Brazil
E-mail: sueli.f@ufsc.br

Abstract. The interlaboratory comparison is a tool to assist in the performance evaluation of test and calibration laboratories and contributes in increasing the credibility of the measurement results. Usually for calibration laboratories it is used $E_n$ number to evaluate the performance of a laboratory. However, when it comes to roughness measurement, $E_n$ number is not applicable due to the difficulty in defining a reference value. Thus, the application of $z$ score test is recommended. This article makes use of Algorithm A of Annex C of ISO 13528 (2015) for determination of the assigned value ($x_{pt}$) and standard deviation ($\sigma_{pt}$) of the proficiency test and demonstrates the viability of its use through a case study.

1. Introduction

Proficiency tests are tools that seek to ensure the analytical quality of the results of the participating laboratories [1]. It is important for a calibration laboratory to participate in an interlaboratory comparison program to demonstrate the qualification (or not) of its results. Its performance is defined based on a value previously set or in consensus with the answers obtained by the group, thus providing means of comparison with similar laboratories and the verification of the influence of factors responsible for the variability of its results [2].

In general, the statistical analysis used to evaluate the performance of the laboratories participating in interlaboratory comparisons is performed through the $z$ score or the $E_n$ number. $E_n$ number is applicable when the measurement uncertainty associated with the reference value is significantly less than the uncertainty declared by the participant laboratory; otherwise it may occur of not being detected significant differences. In this way, it is common to apply $E_n$ number in results of metrological characteristics applicable to calibration laboratories.

For cases where there is no previously set value, the guidance of ISO 13528: 2015 [3], which describes about the statistical methods applicable in proficiency tests, is to use $z$ score, as it can be observed in programs performed with test laboratories.

In interlaboratory comparisons of roughness measurement it is very difficult to establish a reference value due to the existing variability of the measuring system, being then applicable $z$ score. But the variability obtained in the results of the participants can also generate constraints in determining the assigned value ($x_{pt}$) and the standard deviation of the proficiency test ($\sigma_{pt}$).

---

$^1$ To whom any correspondence should be addressed.
The purpose of this paper is to present the application of Robust Analysis through the Algorithm A of Annex C of ISO 13528: 2015 in the treatment of measurement results obtained in an interlaboratory comparisons of measurement of roughness ($Ra$ and $Rz$ parameters).

2. Statistical method for determination of $x_{pt}$ e $\sigma_{pt}$

According to ISO 13528: 2015, the $z$ score value, used for the evaluation of the laboratory performance, is calculated according to equation (1). It represents the relative distance of the laboratory measurement ($x_i$) compared to the designated value of the proficiency test ($x_{pt}$), based on the standard deviation of the proficiency test ($\sigma_{pt}$).

$$Z_i = \frac{(x_i - x_{pt})}{\sigma_{pt}}$$  \hspace{1cm} (1)

$z$ score value is evaluated as follows: if the value of $z_i$ results in $|z| \leq 2$, laboratory performance is considered acceptable; if the result is $2 < |z| < 3$, it is regarded as a warning sign to the lab in question and if the result is $|z| > 3$, the performance is unacceptable.

When the reference value of an interlaboratory comparison is estimated through the consensus of the participating laboratories, ISO 13528: 2015 presents the Algorithm A for the calculation of the average and standard deviation for a robust analysis.

Initially, the Algorithm A calculates the value assigned from the median obtained among the results of the $p$ participating laboratories, according to equation (2). The standard deviation is obtained using equation (3).

$$x^* = \text{median of } x_i \ (i = 1, 2 \ldots p)$$  \hspace{1cm} (2)

$$s^* = 1,483 \ \text{median of } |x_i - x^*|$$  \hspace{1cm} (3)

Through iterations, the values of $x_i^*$ and $s_i^*$ are updated by following the condition established in equation (4), being the value $\delta$ estimated according to equation (5).

$$x_i^* = \begin{cases} 
    x^* - \delta & \text{when } x_i < x^* - \delta \\
    x^* - \delta & \text{when } x_i > x^* + \delta \\
    x_i & \text{other cases}
\end{cases}$$  \hspace{1cm} (4)

$$\delta = 1,5 \ s^*$$  \hspace{1cm} (5)

As a result, the new estimates are calculated to $x^*$ and $s^*$ from the equations (6) and (7) to be used as $x_{pt}$ and $\sigma_{pt}$ parameters for the calculation of $z$ score in equation (1), respectively.

$$x^* = \frac{\sum_{i=1}^{p} x_i}{p}$$  \hspace{1cm} (6)

$$s^* = 1,134 \ \sqrt{\frac{\sum_{i=1}^{p} (x_i^* - x^*)^2}{p - 1}}$$  \hspace{1cm} (7)

The iterations are held until the $x^*$ and $s^*$ values converge in the third significant digit.
3. Experimental Development
The interlaboratory comparison had the participation of 11 organizations that had in their scope of activities the measurement of roughness parameters. They were named by code so that each one performed their measurements individually and had knowledge only of their own results. Following the standards: ISO 4287: 1997 [4] and the ISO 4288: 1996 [5], $Ra$ and $Rz$ parameters were measured. After data collection, they were treated according to the Algorithm A of ISO 13528: 2015.

The artifact chosen for the study is shown in figure 1, together with the roughness profile obtained. In each artifact three measurements were performed in position shown in figure 2.

Table 1 reports the values designated $x_{pt}$ and the standard deviation $\sigma_{pt}$ used for calculating the $z$ score of characteristic $Ra$ and $Rz$, respectively.

| Statistics | $Ra$ | $Rz$ |
|------------|------|------|
| $x_{pt}$   | 0.372| 2.61 |
| $\sigma_{pt}$ | 0.034| 0.21 |

Table 2 presents the deviations obtained by participating laboratories in comparison to the assigned value to the metrological characteristic $Ra$ and $Rz$, respectively. Figure 5 presents the $z$ scores of the participating laboratories, in comparison to the judging criteria used for evaluation of their performance.

| Laboratory | A  | B  | C  | D  | E  | F  | G  | H  | I  | J  | K  |
|------------|----|----|----|----|----|----|----|----|----|----|----|
| Deviation $Ra$ (in $\mu$m) | 0.011 | 0.078 | 0.031 | -0.016 | -0.072 | -0.024 | -0.012 | -0.002 | -0.002 | 0.035 | -0.024 |
| Deviation $Rz$ (in $\mu$m) | -0.09 | 0.59 | 0.13 | -0.05 | -0.60 | -0.13 | -0.07 | -0.15 | 0.08 | 0.26 | 0.05 |

Figure 1. Artifact used

Figure 2. Measuring position in the artifact
Figure 3 presents the $z$ scores of the participating laboratories, in comparison to the judging criteria used to evaluate their performance.

Due to variability from the artifact, the equipment and the conditions under which the measurement was performed, caution has been taken to reduce the area of measurement, making a mark on the artifact. Also, monitoring measurements were carried out to evaluate the stability of the artifact in the measuring region. Guidelines were also passed to the participating laboratories regarding the position and direction of measurement. This allowed for a consistent analysis. The results showed a convergence in the results of the participating laboratories, except for laboratories B and E, which had their $z$ score results in alert region.

4. Final Considerations
The objective of this work was to analyze the iterative method presented in Algorithm A described in ISO 13528 (2015) to determine the $z$ score of participating laboratories in an interlaboratory comparison of roughness measurement. This method uses all data obtained by participating laboratories, seeks convergence to determine the reference value and performs a treatment on discrepant data. The algorithm A is feasible in its application and allows a consistent analysis of the results obtained by the participants in a proficiency test.

5. References
[1] Katsuhiro S, Masanori S, Tanaka H and Ehara K 2016 Measurement 83 144-152
[2] Chui Q S H, Barros C B and Silva T D 2009 Química Nova 8 2209-2213
[3] International Organization for Standardization 2015 ISO 13528 89
[4] International Organization for Standardization 1997 ISO 4287 25
[5] International Organization for Standardization 1996 ISO 4288 8