ASSOCIAÇÃO BRASILEIRA DE ENERGIA NUCLEAR – ABEN
2019 International Nuclear Atlantic Conference - INAC 2019
Santos, SP, Brazil, October 21-25, 2019

Agência Brasileira do ISBN
ISBN 978-85-99141-08-3

Z-SCORE, A TOOL FOR QUALITY ASSURANCE OF ANALYTICAL RESULTS IN NEUTRON ACTIVATION ANALYSIS

Gean B. de Lima[1,3], Thiago J. S. Kaiser[2] and Edson G. Moreira[3]

[1] Instituto Federal de Educação, Ciência e Tecnologia de Rondônia – Campus Colorado do Oeste
Rodovia BR 435, KM-63, Zona Rural
76.993-000 Colorado do Oeste, RO, Brazil
gean.lima@ifro.edu.br

[2] Instituto Federal de Educação, Ciência e Tecnologia de Rondônia – Campus Cacoal
Rodovia BR 364, KM-228, Lote 2-A, Zona Rural
76.960-970 Cacoal, RO, Brazil
thiago.kaiser@ifro.edu.br

[3] Instituto de Pesquisas Energéticas e Nucleares (IPEN – CNEN/SP)
Av. Professor Lineu Prestes 2242
05508-000 São Paulo, SP
emoreira@ipen.br

ABSTRACT

The Nuclear and Energy Research Institute (IPEN - CNEN / SP), through the Research Reactor Center (CERPq), conducts several studies using Instrumental Neutron Activation Analysis (INAA) in its Activation Analysis Laboratory (LAN). The present study performed a critical analysis of the way Certified Reference Materials (CRM) are used to assess accuracy in LAN activities, from a statistical point of view. With the wide dissemination and application of metrological concepts, the use of regulatory standards (ABNT and INMETRO) became increasingly necessary. The literature defines several approaches to the statistical principles for the use of the obtained values. CRMs are used to assess the accuracy of analytical methods, being essential for trueness assessment. There are a few ways to quantify trueness, and this can be done by percent relative error and *z*-score, determining the trend of the analytical method, which is defined as the estimation of a systematic error. The use of *z*-scores is interesting in evaluating the quality of analytical methods, as it is less subject to the influence of the concentration value. Theses and dissertations developed at LAN were selected as the object of study and the calculations were performed using the concepts presented to verify the data treatment, at a 95% confidence level, presenting satisfactory results.

1. INTRODUCTION

For the treatment of data, definitions and statistical approaches are used for better decision-making. Data can be defined as information from observations, counts or measurements. The use of statistical data has been recorded in ancient Babylon, Egypt, and later in the Roman Empire. In this context, the data were related to the activities of the State, such as births and deaths. The term statistical is derived from the Latin word status, which means "state." In recent practice Statistics goes far beyond counting births and deaths, it is considered as the science that deals with the collection, organization, analysis and interpretation of data for decision making [1-2].

In Statistics there are two types of data sets: population and sample. A population is the collection of all possible outcomes of a given set; a sample is a subset or part of a population. The statistical study has two main branches, which depend on the types of data analyzed, descriptive statistics, which involves the organization, summary and re-presentation of data, and inferential statistics, which uses samples to reach conclusions about the population [1-3].

The implementation of quality management systems in laboratories and the use of standards such as ABNT NBR ISO/IEC 17025, highlights the importance of the use of certified reference materials (CRM) for statistical analysis and validation of methods [1-3].

Studies carried out at LAN use Instrumental Neutron Activation Analysis (INAA), which consists in inducing gamma radiation in a sample by irradiation with thermal neutrons generated from IEA-R1, IPEN research reactor, and to perform the measurement in the process of radioactive decay [4].

CRM is defined as a reference material (RM) accompanied by documentation issued by a body with authority to do so, which provides one or more specified property values with associated uncertainties and traceability using valid procedures. RM is defined as a material sufficiently homogeneous and stable over specific properties, prepared to suit an intended use in a measurement or examination of qualitative properties [8].

CRM and RM are increasingly being used for the purpose of increasing reliability in analytical results. However, an important question that arises is how an analytical laboratory should use the measurement results it produces from the CRM. More important than choosing which outcome evaluation approach to use will be the clear definition of the protocol used, so that users can understand the quality of results and make correct decisions from them [9-10].

The literature defines several approaches regarding the statistical principles for the use of the values obtained [9-13]. CRMs can be used for the evaluation of the accuracy of analytical methods, although they are not fundamental for this. They may be replaced by other materials which are sufficiently homogeneous for this purpose [13]. In this case, replicate measurements are obtained in a given precision condition ranging from repeatability conditions to reproducibility conditions, and the dispersion of the results is estimated, generally using the standard deviation of the results [12].

The use of the *z*-score may be more adequate in the evaluation of the quality of analytical methods, since it is less subject to the influence of the concentration value, it represents the standardization of the distance that this value is from the mean, in terms of standard deviation [15].

The use of *z*-score in quality systems first appeared in the context of analysis of proficiency programs, because it facilitates the comparison of data from different laboratories, since it standardizes the information in a way that can be evaluated, independently of factors that may influence the results of the measurement , as the z-score value does not have to be expressed in the original unit of measurement, it is normalized and described as the distance between the expected value (reference value) and the measurement result, in units of standard deviation.

The calculation of the *z*-score is based on the Gaussian or Normal distribution since it has a mean of 0 (zero) and a standard deviation of 1 (one). In this type of data distribution, the probability of the results being in the range of $z = \pm 2$ is approximately 95% (ninety five percent), based on this information, the laboratory performance is considered satisfactory or not, to perform the test [15-16].

The *z*-score means to remove the scale from the original variable, turning it into an index, in which case they are very useful because they allow the comparison between variables. A sample of actual data may not fit perfectly into a normal distribution, however if there is a reasonable approximation, one can make use of probabilities only by knowing the sample estimates of μ and σ, the mean and standard deviation of the population.

2. STATISTICAL METHODS FOR DATA ANALYSIS

CRMs are key to assessing the trueness of analytical methods. Trueness is defined as the degree of agreement between the mean of an infinite number of repeated measured values and a reference value and is inversely related to the systematic error, but it is not related to the random error [8]. Together with precision, trueness makes up the qualitative characteristic of accuracy of the analytical method. Trueness is not a quantity and therefore it cannot be expressed numerically. Thus, there are some ways to quantify trueness, which can be by means of the percent relative error and *z*-score by determining the trend of the analytical method, which is defined as the estimation of a systematic error [8].

2.1 Relative error

Traditionally, the relative error (E_r) has been used for the purpose of directly evaluating the trueness of results, even obtained in different orders of magnitude. In the recent past, in many contexts the trend was termed a systematic error or simply an error, but such terminology should be avoided [8].

In the area of trace element determination in samples of complex matrices, in general, relative errors of the order of 10% or greater can be considered satisfactory [13]. However, relative errors can be obtained which are considered too high when the concentration range is too low. Equation 1 presents the definition of E_r in percent form.

$$E_r(\%) = \frac{(\bar{x}-\mu)}{\mu} \times 100 \quad (1)$$

For practical purposes, we take $\bar{x}$ as the mean result of the measurement and μ as the certified value of the CRM used in the evaluation of the trueness of the analytical method, the smaller the relative error, the greater the trueness [13].

2.2 *z*-score

Among the statistical methods used by interlaboratory programs to evaluate performance, one of the most used is the *z*-score, which is recommended by the International Harmonized Protocol for Proficiency Testing, which is obtained according to Equation 2 [18].

$$z = \frac{(x - x_v)}{\sigma} \quad (2)$$

in which:
z is the value of the z-score for the measurement;
x is the result of the measurement;
x_v is the reference value;
σ is the standard deviation of the reference value.

When the *z*-score is calculated using a CRM, the term "σ" assumes the value of the uncertainty associated with the consensus value x_v. The ISO/IEC 13528 standard classifies the results in the measurements as satisfactory or not, as described below [22]:

$$|z| \leq 2 \text{ satisfactory result;}$$
$$2 < |z| < 3 \text{ questionable result;}$$
$$|z| \geq 3 \text{ unsatisfactory result.}$$

Nevertheless, it is also possible to calculate the *z*-score using different criteria for the definition of σ, which may bring a lot of confusion to the quality assurance context [17]. Depending on the criteria used, the assumed value for the *z*-score may change, and therefore the performance rating of the analysis can be changed. The Horwitz equation, modified by Thompson, can be used to obtain the standard deviation of reproducibility based on analyte concentration c, expressed as the dimensionless ratio of the mass, i. e., devoid of any physical unit that defines it, as follows below [21]:

- For $c > 0.138 \Rightarrow \sigma = 0.01\, c^{0.5}$
- For $1.2 \times 10^{-7} \leq c \leq 0.138 \Rightarrow \sigma = 0.02\, c^{0.8495}$
- For $c < 1.2 \times 10^{-7} \Rightarrow \sigma = 0.02\, c$

3. EXPERIMENTAL

It was defined as the object of study the master's dissertations and doctoral theses developed at LAN in the area of concentration of Nuclear Technology - Applications of the University of São Paulo, in 2018, with applications of INAA. These documents dealt with the quality and validation aspects of methods with varying degrees of depth.

In 2018, 5 theses and 5 dissertations were produced at LAN, available in the repository called USP Digital Library of Theses and Dissertations, of which 3 theses and 2 dissertations did not use the concepts, mentioned herein or did not have data to perform the calculations. In addition, 5 randomized dissertations were selected, from 2013 to 2017. The *z*-score values were recalculated using the criteria defined in the introduction, for comparison and evaluation of the analytical quality of LAN results for the various chemicals elements analyzed.

4. RESULTS AND DISCUSSION

Theses and dissertations were selected as the object of study and the calculations were performed using the concepts used in the introduction to verify the treatment of data. To assess the statistical significance of the result, the confidence level used was 95%. Table 1 shows the *z*-score values of the different experiments. The values were recalculated based on the data presented in the study, and obtained values were practically identical to those shown in the studies. There was a considerable degree of reliability in the analysis, however with some misapplications of z-scores.

Table 1: *Z*-score values for dissertations

Experiment	As	Co	Cr	Fe	Rb	Sc	Se	Zn
y1	3.30	10.00	-2.90	6.80	-0.90	0.70	0.40	5.30
y2	-4.20	0.70	-2.70	1.50	-1.90	0.70	-0.60	0.70
y3	2.30	4.80	-0.70	2.90	-0.90	0.60	1.60	0.80
y4	-1.50	0.30	-0.30	1.60	-0.40	0.10	1.50	-0.10
y5	2.80	9.60	-0.50	3.40	-0.80	0.60	0.70	4.90
y6	-3.50	0.70	-2.30	1.30	-0.70	0.50	0.10	1.00
y7	1.30	4.50	0.40	2.80	-0.70	0.70	0.80	0.20
y8	0.00	0.20	-0.40	1.40	-0.30	0.20	0.70	1.00

Fig. 1 shows the standard deviation of the z-score of each element, according to each experiment performed to obtain such values. These values were selected from the mentioned theses and dissertations, and the *z*-scores of such studies were obtained by several experiments. All values presented in theses and dissertations comply with statistical and quality management standards. Presenting the scores obtained by different studies gives an overview of the quality and accuracy of the experiments performed. The cases of values with *z*-scores above |3| were not individually treated in this study, aiming at not specifically exposing the previous studies.

Co presented the highest standard deviation, however, the study in which it was applied was not significantly influenced, since several measurements were made using some different methodologies, therefore, one of the methods used obtained a *z*-score value of 10 which influenced the final value, but these calculations were performed precisely to demonstrate which methodology was most appropriate, that in the specific study, the methodology with value outside the acceptable limits was not used.

On the other hand, Sc presented the lowest and better standard deviation results considering all experiments, which were performed in different scenarios, such as measurements after different radioactive decay times (varying from hours to days).

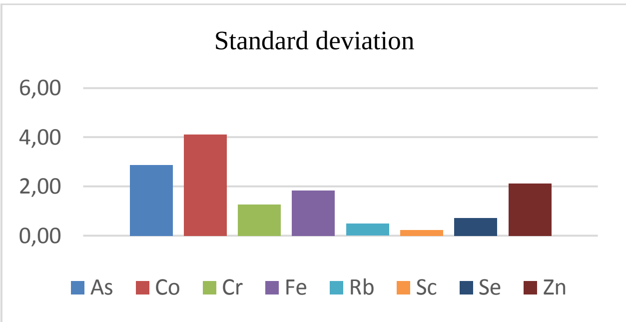

Figure 1: Standard deviation of observed *z*-scores for elements

Fig.2 shows the standard deviation of the z-score for each experiment, taking into account all the determined elements. Experiment y8 presented the lowest z-score standard deviation, showing that by analyzing *z*-scores it is possible to indicate which methodology was the most suitable.

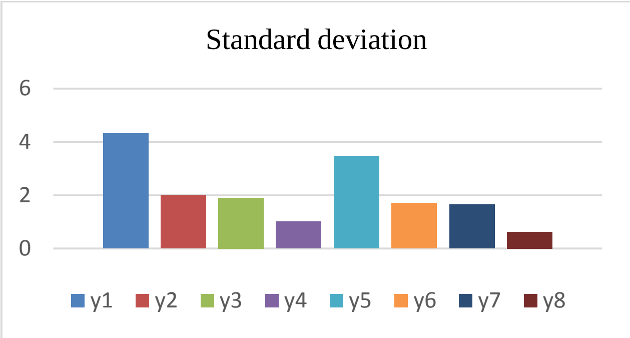

Figure 2: Standard deviation of observed *z*-scores for experiments

Table 2 shows the relative error values of the experiments. As for the *z*-score, these values were recalculated using the information available in the dissertations and thesis. The most accurate results in all experiments were for Sc, and the worst for Co, as observed previously in the *z*-score data analysis. Similarly, the experiment with the best result was y8.

Table 2: Relative error (%) of mean CRM results in dissertations and thesis

Experiment	As	Co	Cr	Fe	Rb	Sc	Se	Zn
y1	21	48	-33	30	-8.3	4.5	2.3	21
y2	-26	3.0	-30	6.7	-18	4.0	-2.9	2.9
y3	14	22	-7.3	13	-8.1	3.5	7.7	3.1
y4	-9.6	1.2	-2.4	6.6	-4.1	0.5	7.7	-0.3
y5	17	45	-5.6	15	-7.5	3.0	3.2	20
y6	-22	2.9	-27	5.7	-6.5	3.0	0.7	4.1
y7	8.1	21	4.0	12	-6.9	4.0	3.8	0.3
y8	0.00	0.8	-4.0	6.4	-2.4	1.0	3.4	-0.3

5. CONCLUSIONS

For the verification of results, as well as comparisons between laboratories, the *z*-score is an excellent tool. CRM has measurement traceability, so when comparing a CRM with the object of study, the objective is to study the degree of proximity of the results, that is, to evaluate the accuracy of the method under study with the reference material.

Based on the information extracted from only a few studies (dissertations and thesis), it was observed that it is possible to expand the study to fully verify the evolution of data processing at CERPq. The study can be expanded within the laboratory by increasing the sample, in the case, using a longer period, 15 to 20 years, analyzing all the studies done, seeking to analyze the evolution in the treatment of data at CERPq, with the intention of proposing a protocol to guide for proper use of CRM data analyzes.

ACKNOWLEDGMENTS

To the Federal Institute of Education Science and Technology of Rondônia for providing the opportunity for training. To the Nuclear and Energy Research Institute for collaboration and support.

REFERENCES

1. R. Larson, B. Farber, *Estatística aplicada*, translation: J. F. P. Gonçalves, 6 ed. Pearson, São Paulo, Brazil (2015).
2. J. L. Devore, *Probabilidade e estatística: para engenharia e ciências*, translation: J. P. N. Silva, 6 ed. Cengage Learning, São Paulo, Brazil (2006).
3. S. Vieira, *Estatística para a qualidade*, 2 ed. Elsevier, Rio de Janeiro, Brazil (2012).
4. P. Bode, *Instrumental and organizational aspects of a neutron activation analysis laboratory*, Ph.D. dissertation, Delft University of Technology, Delft, Netherlands (1996).
5. A. Zschunke, *Reference materials in analytical chemistry – a guide for selection and use***.** Springer, Berlin, Germany (2000).
6. P. Bièvre, R. Kaarls, H. S. Peiser, S. D. Rasberry, W. P. Reed, "Measurement principles for traceability in chemical analysis". *Accred. Qual. Assur***.**, **v. 1**, p. 3–13 (1996).
7. International Organization for Standardization (ISO/IEC 17025:2017), *General requirements for the competence of testing and calibration laboratories*, ISO, Geneva, Switzerland (2017).
8. Instituto Nacional de Metrologia, Normalização e Qualidade Industrial, *Vocabulário internacional de metrologia – conceitos fundamentais e gerais e termos associados* (VIM3 - JCGM 200:2008) INMETRO, Rio de Janeiro, Brazil (2009).
9. Associação Brasileira de Normas Técnicas, *Utilização de materiais de referência certificados* (ABNT ISO Guia 33), 1 ed., ABNT, Rio de Janeiro, Brazil (2002).
10. Associação Brasileira de Normas Técnicas, *Materiais de referência – princípios gerais e estatísticos para certificação* (ABNT ISO Guia 35), 1 ed. ABNT, Rio de Janeiro, Brazil (2012).
11. Instituto Nacional de Metrologia, Normalização e Qualidade Industrial, *Orientações sobre análise crítica da documentação associada aos materiais de referência adquiridos***,** ver. 00, (DOQ-CGCRE-033), INMETRO, Rio de Janeiro, Brazil (2012).
12. N. A. Majcen, "A need for clearer terminology and guidance in the role of reference materials in method development and validation". *Accred. Qual. Assur*., **v. 8**, pp.108–112 (2003).
13. N. Baccan, J. C. Andrade, O. Godinho, J. S. Barone, *Química analítica quantitativa elementar*. 2. ed. rev. ampl., Edgard Blücher, São Paulo, Brazil (1985).
14. J. F. Hair Jr., W. C. Black, B. J. Babin, R. E. Anderson, R. L. Tatham, *Análise multivariada de dados,* 6 ed., Bookman, Porto Alegre, Brazil (2009).
15. M. F. Triola. *Introdução à estatística,* translation: A. A. Farias, E. F. Soares, V. R. L. F Flores, 7 ed., LTC, Rio de Janeiro, Brazil (1999).
16. B. James. *Probabilidade: um curso em nível intermediário*. 4. ed., IMPA, Rio de Janeiro, Brazil (2015).

17. H. Bolfarine, M. C. Sandoval. *Introdução à inferência estatística.* 3 ed., SBM, Rio de Janeiro, Brazil (2010).
18. Associação Brasileira de Normas Técnicas, *Avaliação da conformidade — requisitos gerais para ensaios de proficiência* (ABNT NBR ISO/IEC 17043), ABNT, Rio de Janeiro, Brazil (2011).
19. M. Thompson, S. L. R. Ellison, R. Wood, "The international harmonized protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report)", *Pure Appl. Chem.* **v.78 n. 1**, p. 145-196 (2006).
20. Instituto Nacional de Metrologia, Normalização e Qualidade Industrial,, *Orientação sobre validação de métodos de ensaios químicos*, rev. 04 (DOQ-CGCRE-008), INMETRO, Rio de Janeiro, Brazil (2011).
21. E. G. Moreira, D. Seo, M. B. A. Vasconcellos; M. Saiki "Trace element determination in a mussel reference material using short irradiation instrumental neutron activation analysis*", J. Radioanal. Nucl. Chem.,* **v. 293**, pp. 251–254, 2013.
22. International Organization for Standardization, *Statistical methods for use in proficiency testing by interlaboratory comparisons* (ISO 13528), ISO, Geneva, Switzerland (2005).