UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

ADRIANO VELAME BRANCO CHAGAS

PRODUCTION AND CHARACTERIZATION OF A CANDIDATE SOIL REFERENCE MATERIAL FOR INORGANIC CHEMICAL ANALYSIS

SALVADOR
2025

ADRIANO VELAME BRANCO CHAGAS

PRODUCTION AND CHARACTERIZATION OF A CANDIDATE SOIL REFERENCE MATERIAL FOR INORGANIC CHEMICAL ANALYSIS

Dissertação apresentada ao Colegiado dos Cursos de Pós-Graduação em Química, Instituto de Química, Universidade Federal da Bahia, como parte do requisito para obtenção do grau de mestre em Química.
Área de concentração: Química Analítica

Orientador: Prof. Dr. Sérgio Luís Costa Ferreira

SALVADOR
2025

SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

TERMO DE APROVAÇÃO

ADRIANO VELAME BRANCO CHAGAS

"Produção e caracterização de um candidato a material de referência de solo para análises químicas inorgânicas"

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre em Química, Universidade Federal da Bahia, pela seguinte banca examinadora:

Prof. Dr. Sergio Luís Costa Ferreira ____
Doutorado em Química, Pontifícia Universidade Católica do Rio de Janeiro, PUC
Universidade Federal da Bahia

gov.br Documento assinado digitalmente
JEZ WILLIAN BATISTA BRAGA
Data: 11/09/2025 10:02:14-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Jez Willian Batista Braga____
Doutorado em Química, Universidade Estadual de Campinas, UNICAMP
Universidade de Brasília

gov.br Documento assinado digitalmente
WALTER NEI LOPES DOS SANTOS
Data: 12/09/2025 11:42:59-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Walter Nei Lopes dos Santos______________________________________
Doutorado em Química, Universidade Federal da Bahia, UFBA
Universidade do Estado da Bahia

Salvador, 10 de setembro de 2025.

Ficha catalográfica elaborada pela Biblioteca Universitária de Ciências
e Tecnologias Prof. Omar Catunda, SIBI – UFBA.

C433 Chagas, Adriano Velame Branco

Production and characterization of a candidate soil reference material for inorganic chemical analysis/ Adriano Velame Branco Chagas. – Salvador, 2025.

77 f.

Orientador: Prof. Dr. Sérgio Luís Costa Ferreira

Dissertação (Mestrado) – Universidade Federal da Bahia. Instituto de Química, 2025.

1. Material de referência - Solo. 2. Homogeneidade. 3. Quimiometria. I. Ferreira, Sérgio Luís Costa. II. Universidade Federal da Bahia. III. Título.

CDU: 546:631.416

AGRADECIMENTOS

Gostaria de começar agradecendo aos meus pais, Jorge Ferreira Chagas e Maria Izabel Velame Branco Chagas, por sempre acreditarem em mim, pelo apoio incondicional e pela força que me deram em todos os momentos que me trouxeram até aqui.

À minha esposa, Natália Erdens Maron de Freitas, que esteve ao meu lado em todos os dias dessa caminhada, tornando-os mais leves e alegres. Obrigado pelo companheirismo, pela compreensão, pelo amor e por transformar cada desafio em algo mais fácil de enfrentar.

Ao meu orientador, Professor Dr. Sérgio Luís Costa Ferreira, pela dedicação, paciência, amizade e confiança, que foram fundamentais nesta etapa da minha vida acadêmica e pessoal. Sua orientação sempre esteve acompanhada de zelo e incentivo, o que fez toda a diferença.

Ao Professor Dr. Francisco Antônio da Silva Cunha pela contribuição ao trabalho e pela amizade durante essa jornada.

À Professora Dra. Liz Oliveira dos Santos, pela colaboração e acompanhamento dos resultados do trabalho.

Aos meus irmãos, Anderson Velame Branco Chagas e Alexandre Velame Branco Chagas, que compartilham comigo não apenas a vida, mas também experiências, aprendizados e apoio desde sempre.

Aos meus padrinhos, José Vianna Lopes e Antônia Velame Branco, que sempre estiveram presentes e foram como segundos pais para mim, acompanhando minha trajetória com carinho e cuidado.

Aos meus familiares Roseli Lopes, Lídia Beatriz Lopes, Maria da Conceição Velame Branco, Gabriel Velame Branco, Katyuscia Paes, Jessica Fortuna, Thaina Paes Chagas, Thalita Paes Chagas e Gabriel Velame Paes Chagas, pelo carinho, apoio e incentivo em todos os momentos. A presença de vocês na minha vida sempre foi motivo de força e alegria, tornando essa caminhada mais leve e especial.

Aos membros da banca, Dr. Jez Willian Batista Braga e Dr. Walter Nei Lopes dos Santos, pela generosidade em aceitarem o convite e contribuírem com suas valiosas observações para o meu trabalho.

Aos amigos que considero irmãos, Bruno Araújo, Diego Barreto, João Brasil e Patrick Passos, que caminham ao meu lado desde a época de colégio e continuam sendo parte essencial da minha vida.

Aos colegas do GRPQQ, que estiveram presentes nessa caminhada desde a graduação, compartilhando não só conhecimento, mas também amizade e momentos de alegria: Adílio, Adriana, Alda, Caio Felix, Caio Vitor, Cássio, Ícaro, Lânio, Luana, Lucas Cintra, Matheus, Rafael, Ravena, Rodrigo, Saulo, Vanessa, Tiago e William. Minha gratidão a cada um de vocês, que foram parte indispensável dessa trajetória.

Agradeço aos 29 laboratórios que contribuíram para o estudo interlaboratorial. A participação de cada um foi essencial para o desenvolvimento deste trabalho. Muito obrigado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro que possibilitou a realização deste trabalho. Assim como, a Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento de equipamentos ao laboratório.

E, finalmente, a todos aqueles que torceram por mim, me apoiaram, acreditaram e enviaram boas energias ao longo do caminho. Minha sincera gratidão a cada gesto e palavra de incentivo que recebi.

RESUMO

O desenvolvimento de materiais de referência é fundamental para assegurar a qualidade e a comparabilidade dos resultados obtidos em análises químicas. Nesta dissertação, é apresentado o estudo de um candidato a material de referência de solo para análises inorgânicas, coletado em uma área impactada por atividades metalúrgicas de chumbo no município de Santo Amaro da Purificação, Bahia, Brasil. O solo foi submetido a processos de preparo, incluindo secagem, peneiramento, homogeneização e esterilização por radiação gama, resultando na produção de 100 frascos de 30 g destinados ao processo de certificação. A homogeneidade do material foi avaliada por meio de digestão assistida por micro-ondas e determinação multielementar (As, Ca, Cd, Cu, Mg, Mn, Na, Ni, Pb, Sb, Se, Sr e Zn) utilizando espectrometria de emissão óptica com plasma acoplado indutivamente (ICP OES). Foram realizados testes de massa mínima, homogeneidade intra e entre frascos, validados pela análise de variância (ANOVA), conforme recomendações da ISO 35, além da aplicação da técnica quimiométrica DD-SIMCA, que corroborou os resultados. Como inovação, também foi proposta a aplicação do Z-score robusto de forma multielementar para avaliação da homogeneidade. A etapa seguinte consistiu na condução de um ensaio interlaboratorial envolvendo 28 laboratórios de instituições de ensino superior brasileiras, a fim de certificar os elementos químicos presentes. Os resultados foram avaliados pelos testes de Cochran, Grubbs, Z-score e Z-score robusto, assegurando a confiabilidade dos valores atribuídos. Foram certificados os elementos As, Cd, Ca, Cu, Mg, Mn, Ni, Pb e Zn, com estimativas de incerteza considerando estudos de homogeneidade, estabilidade e interlaboratorial. Os resultados obtidos confirmam que o material desenvolvido apresenta propriedades adequadas para ser utilizado como material de referência, contribuindo para a melhoria da confiabilidade metrológica em análises de solos.

Palavras-chave: Material de referência; Solo; Homogeneidade; ICP OES; Quimiometria

ABSTRACT

The development of reference materials is essential to ensure the quality and comparability of analytical results. This dissertation presents the study of a candidate soil reference material for inorganic analysis, collected from an area impacted by lead metallurgical activities in Santo Amaro da Purificação, Bahia, Brazil. The soil underwent preparation procedures, including drying, sieving, homogenization, and sterilization by gamma radiation, resulting in 100 vials of 30 g each for the certification process. Homogeneity was assessed through microwave-assisted digestion and multi-element determination (As, Ca, Cd, Cu, Mg, Mn, Na, Ni, Pb, Sb, Se, Sr, and Zn) using inductively coupled plasma optical emission spectrometry (ICP OES). Minimum sample mass, within-bottle, and between-bottles tests were performed and validated by analysis of variance (ANOVA), according to ISO 35 guidelines, and further confirmed by the chemometric technique Data-Driven Soft Independent Modeling of Class Analogy (DD-SIMCA). Additionally, the innovative use of a multi-element robust Z-score approach was proposed as an alternative tool for homogeneity assessment. Subsequently, an interlaboratory study involving 28 analytical chemistry laboratories from Brazilian higher education institutions was carried out to certify the chemical elements. The results were evaluated using Cochran, Grubbs, Z-score, and robust Z-score tests, ensuring the reliability of the assigned values. The certified elements were As, Cd, Ca, Cu, Mg, Mn, Ni, Pb, and Zn, with uncertainty estimates considering homogeneity, stability, and interlaboratory studies. The findings confirm that the developed material exhibits suitable properties to be used as a reference material, contributing to the improvement of metrological reliability in soil analysis.

Keywords: Reference material; Soil; Homogeneity; ICP OES; Chemometrics

SUMMARY

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AAS – Atomic Absorption Spectrometry

ACR – Absolute Residuals

AES – Atomic Emission Spectrometry

AFS – Atomic Fluorescence Spectrometry

ANOVA – Analysis of Variance

CID – Charge Injection Device

CRM – Certified Reference Material

DD-SIMCA – Data-Driven Soft Independent Modeling of Class Analogy

EPA – Environmental Protection Agency

FAAS – Flame Atomic Absorption Spectrometry

GF AAS – Graphite Furnace Atomic Absorption Spectrometry

HCA – Hierarchical Cluster Analysis

ICP OES – Inductively Coupled Plasma Optical Emission Spectrometry

IPEN – Instituto de Pesquisas Energéticas e Nucleares

IQR – Interquartile Range

ISO – International Organization for Standardization

IUPAC – International Union of Pure and Applied Chemistry

LOD – Limit of Detection

LOQ – Limit of Quantification

MR – Multiple Response

PCA – Principal Component Analysis

RMs – Reference Materials

RSD – Relative Standard Deviation

SIMCA – Soft Independent Modeling of Class Analogy

U_{bb} – Between-Unit Uncertainty

U_{lts} – Uncertainty Long-Term Stability

U_{char} – Uncertainty Characterization

Introduction

Inorganic chemical analysis is an essential tool across a wide range of fields and productive sectors, including industry, agriculture, healthcare, environmental monitoring, and scientific research (1). These determinations allow for the identification, quantification, and monitoring of chemical elements in various matrices, supporting product quality control, compliance assessment with regulations, and decision-making based on reliable data. The growing demand for accurate, precise, and comparable results is closely linked to advances in academic research, which requires dependable data to develop new knowledge, validate hypotheses, and consolidate scientific models (2).

In this context, reference materials (RMs) play a central role in ensuring the quality of chemical measurements. According to ISO Guide 30 (3), an RM is a material that is sufficiently homogeneous and stable with respect to one or more specified properties, used for instrument calibration, method verification, procedure validation, and ensuring metrological traceability. In laboratories performing macro- and micronutrient analyses, for example, the use of RMs is essential to ensure the comparability of results obtained using different methods, at different times, and in different locations.

However, the availability of soil RMs in Brazil remains limited, with most being imported. In addition to their high cost, these materials often do not reflect the physical, chemical, and mineralogical characteristics of tropical soils that dominate the country's territory (4). This mismatch can affect the accuracy and applicability of locally developed analytical methods, compromising both comparability and traceability of results.

Producing reference materials within the laboratory itself is an important strategy to optimize resources, reduce operational costs, and increase control over the development and characterization stages. This practice can also facilitate participation in interlaboratory proficiency testing programs, support the advancement of academic research, and contribute to the standardization of analytical methods tailored to the specific characteristics of the samples of interest (5).

This work proposes the development and characterization of a candidate soil reference material suitable for inorganic chemical analysis. The study includes sampling, preparation, homogenization, characterization, and the evaluation of both homogeneity and stability, aiming to ensure that the final product meets the requirements established by applicable standards.

Chapter 1: Literature Review

1. Soil (contextualization and importance)

Soil is a non-renewable natural resource on a human timescale, formed through the complex interaction of minerals, organic matter, water, air, and living organisms over thousands of years. It plays a fundamental role in sustaining terrestrial life, serving as a medium for plant growth, regulating biogeochemical cycles, and acting as a natural barrier against pollutants. Beyond its ecological role, soil holds economic and social importance, supporting agricultural, forestry, and construction activities, and providing the foundation for the production of food, fibers, and raw materials (6).

From an environmental perspective, soil functions as a filter and buffer, retaining and transforming chemical substances and preventing the contamination of aquifers and aquatic ecosystems. However, these functions are intrinsically linked to its composition and structure, which can be altered by natural processes or human activities. Unsustainable land-use practices - such as intensive agriculture, deforestation, excessive application of fertilizers and pesticides, and the improper disposal of industrial waste - can lead to degradation, loss of fertility, and chemical contamination (7).

The analysis of inorganic chemical parameters in soils is essential for understanding their quality and functionality. Elements such as phosphorus, potassium, calcium, magnesium, and metallic micronutrients are key to agricultural productivity (8), while heavy metals such as lead, cadmium, nickel, and chromium, when present at elevated concentrations, pose risks to human health and the environment(9). Determining the concentration and distribution of these elements is therefore critical for sustainable land management and environmental risk assessment.

Continuous chemical monitoring of soils makes it possible to detect trends in degradation, contaminant accumulation, and changes in fertility over time. Such monitoring is essential for meeting environmental regulations, guiding remediation strategies, and preventing irreversible impacts. Without systematic surveillance, issues such as salinization, acidification, heavy metal contamination, and nutrient depletion may progress unnoticed, ultimately undermining agricultural productivity and food security (7).

The reliability of these assessments depends directly on the quality of the analytical data obtained, which in turn requires standardized methods and metrological traceability. In this context, the availability of soil reference materials suited to the matrix and target parameters becomes a key element in ensuring accurate and comparable results across laboratories and over time.

2. Reference Materials

2.1. Reference Materials and Candidate Reference Materials

Reference materials (RMs) are materials whose properties are sufficiently homogeneous, stable, and well-characterized to be used for the calibration of instruments, the validation of analytical methods, the assessment of measurement performance, and the assurance of metrological traceability. According to ISO 17034 and ISO Guide 35, an RM must present one or more property values established using sound metrological principles and accompanied by associated uncertainties. In the context of inorganic chemical analysis, RMs are essential for ensuring the comparability, accuracy, and reliability of analytical results across laboratories and over time (10, 11).

Certified reference materials (CRMs) are a special category of RMs that are accompanied by a certificate issued by a competent producer. This certificate provides the assigned property value(s), associated measurement uncertainty, and evidence of traceability to the International System of Units (SI), when applicable. The production of a CRM requires rigorous procedures to assess and ensure material homogeneity, short- and long-term stability, and accurate characterization using validated methods, often with interlaboratory participation (12). Despite their importance, CRMs are not available for all types of matrices, particularly for soils with specific chemical and mineralogical characteristics, such as tropical soils in Brazil.

Candidate reference materials - in this context, the term refers to “candidate” in the sense of a material under development - are materials prepared and subjected to preliminary studies with the potential to become RMs or CRMs. Prior to certification, these materials undergo preparation, preliminary characterization, and tests for homogeneity and stability to determine whether they meet the necessary quality requirements (13-15). The development of candidate RMs is a critical step in filling gaps in the availability of reference materials for specific applications, particularly when suitable certified materials do not exist in the market.

For soil analysis, the need for matrix-matched RMs is particularly acute. Soils are complex and heterogeneous matrices whose chemical composition and physical properties can vary significantly over short distances (16). This intrinsic variability poses challenges for obtaining representative, stable, and homogeneous materials suitable for use as reference materials. In many cases, laboratories resort to importing soil CRMs produced under conditions very different from those of local environments, which may compromise the validity of method validation and quality control. In this scenario, the development of candidate soil reference materials tailored to local conditions—such as those reflecting the mineralogical and chemical profiles of Brazilian tropical soils—is a strategic approach to improving analytical reliability and supporting environmental monitoring and sustainable land management (17, 18).

2.2. Production of soil reference materials

The production of soil reference materials (RMs) involves a series of carefully controlled steps to ensure the representativeness, homogeneity, and long-term stability of the material. The process begins with the selection of the sampling area and soil type, considering factors such as representativeness with respect to the environmental or geographical conditions of interest, mineralogical and chemical stability, and the absence of undesirable external contaminants that could compromise its use as a standard (11). Sample collection must follow strict protocols to avoid changes in the soil's natural composition, preferably using inert and thoroughly cleaned tools and containers to minimize the risk of contamination. (19, 20).

After collection, the material undergoes preparation steps that include controlled drying - typically at room temperature or under mild heating to prevent the loss of volatile compounds - followed by grinding and sieving to achieve a uniform and suitable particle size distribution. Homogenization is a critical stage, as it ensures the even distribution of analytes within the matrix, thereby reducing variability among subsamples. Common approaches include mechanical mixing, successive quartering, or the use of specialized mixers. The prepared soil is then packaged in appropriate containers that protect it from moisture, light, and temperature fluctuations and that are chemically inert to avoid any interaction with the material (21).

2.3. Steps in the production of reference materials

After the sampling stage, the production of a soil reference material involves a series of studies to ensure its metrological quality and confirm that it meets the requirements for use in chemical analyses. One of the first steps is the homogeneity test, which aims to verify whether the analytes of interest are evenly distributed throughout the entire batch. This step is critical because significant variations between units could compromise the material's reliability as a standard. In practice, the process involves dividing the bulk material into small portions (usually placed in individual containers), then randomly selecting a number of these units for analysis. The selected samples are tested under repeatable conditions, and the results are evaluated statistically - commonly using analysis of variance (ANOVA) - to determine whether the variation observed between units is significantly smaller than the measurement variability (10, 11).

Once homogeneity has been confirmed, the next step is the stability assessment, which determines whether the material's chemical and physical properties remain unchanged over time and under different storage conditions. This evaluation is typically divided into short-term and long-term stability studies. Short-term testing simulates transport conditions by exposing the material to variations in temperature and humidity, while long-term testing monitors the material over months or years under controlled storage. Periodic analysis of stored samples is carried out to detect any changes in analyte levels, ensuring that the material retains its properties within the established uncertainty limits (10, 11).

The final step before making the material available is the interlaboratory study, also referred to as a proficiency test for characterization. In this stage, a group of participating laboratories—preferably those with proven expertise in the relevant measurements—receive identical units of the material and apply their own analytical methods to determine the values of the target analytes. The results are then compiled and statistically processed in accordance with guidelines such as ISO Guide 35 (11) and ISO 17043 (22), to establish the reference value and its associated uncertainty. This approach minimizes potential bias from individual methodologies and ensures that the assigned value is robust and widely accepted.

Taken together, the homogeneity tests, stability studies, and interlaboratory characterization form essential steps in ensuring the reliability and metrological traceability of a soil reference material. These processes allow the material to be used safely and consistently over long periods in a wide range of analytical applications.

2.4. International Guidelines for the Production of Reference Materials

Creating RMs requires careful procedures to ensure the material's quality and traceability. To this end, the International Organization for Standardization (ISO) has published a set of internationally recognized guides that provide recommendations to producers. In Brazil, some of these documents (ISO Guides 30 to 34) have already been officially translated and made available by the Brazilian Association of Technical Standards (ABNT).

These guides address different aspects related to the preparation, use, and certification of reference materials(3, 10, 11, 23-26):

- **ISO Guide 30 (1992)** defines the essential terminology applied to RMs.
- **ISO Guide 31 (2000)** describes the mandatory items to be included in reference material certificates.
- **ISO Guide 32 (1997)** provides general guidance on calibration issues in chemical analyses and the use of internal RMs and certified reference materials (CRMs).
- **ISO Guide 33 (2000)** presents practical recommendations for the use of RMs in evaluating the accuracy of analytical methods.
- **ISO Guide 34 (2009)** establishes requirements for producers of CRMs, including quality system criteria, infrastructure, preparation, storage, and post-distribution procedures.
- **ISO Guide 35 (2017)** focuses on the characterization of materials, as well as the assessment of homogeneity and stability.
- **ISO Guide 13528 (2015)** details statistical methods applied to proficiency testing and interlaboratory comparisons.

When planning a new RM, it is recommended that the producer adopt practices capable of reducing potential sources of variation, particularly those affecting batch homogeneity. According to ISO Guide 34, steps such as milling, mixing, sieving, sample splitting, and qualitative analysis may be incorporated when relevant. Additional procedures may also be necessary, such as particle size distribution assessment, drying, sterilization, and packaging in containers representative of the entire produced material (10).

2.5. Calculation of Uncertainty

2.5.1. Homogeneity uncertainty

Measurement uncertainty is a parameter associated with the result of a measurement that describes the dispersion of values that could reasonably be attributed to the measured (27). In other words, it reflects the level of confidence in the reported value, taking into account all sources of variation that may influence the measurement. In the context of reference material production, accurately estimating the uncertainty is essential to ensure metrological traceability and the reliability of certified values, enabling researchers to understand the variation associated with the reference value.

For soil reference materials, the uncertainty calculation involves combining different components obtained from the homogeneity, stability, and characterization studies. Each of these stages provides a term that contributes to the final combined uncertainty (11).

The uncertainty associated with homogeneity—also referred to as between-bottle uncertainty—represents the variation among different units within the batch. This component is estimated from the results of the homogeneity test, in which a set of bottles is randomly selected and analyzed under repeatable conditions (28).

Analysis of variance (ANOVA) is the most common statistical approach used to separate between-bottle variability from within-bottle variability. The homogeneity uncertainty, u_{bb}, can be calculated as:

$$u_{bb} = \sqrt{\frac{MS_{among} - MS_{within}}{n_0}} \quad \textbf{\textit{eq. 1}}$$

$$u_{bb} = \sqrt{\frac{MS_{among}}{n_0}} \cdot \sqrt[4]{\frac{2}{v_{MS_{within}}}} \quad \textbf{\textit{eq. 2}}$$

MS_{among} refers to the mean square between the units of the RM batch (obtained via ANOVA), while MS_{within} corresponds to the mean square within those units. The term $v_{MSwithin}$ represents the degrees of freedom associated with MS_{within}, and n (or n_0) denotes the number of measurements performed in the study. When MS_{among} exceeds MS_{within}, both equations for estimating the between-unit uncertainty (u_{bb}) are applied, and the higher resulting value is adopted. However, if MS_{among} is less than MS_{within}, only Eq. 2 is considered valid for estimating u_{bb} (13).

2.5.2. Uncertainty stability

According to ISO Guide 35, the simple linear regression model (univariate model) is one of the most widely applied approaches for assessing the stability of a candidate reference material (RM). This method examines the relationship between two variables: the property value of interest (for example, the analyte concentration) as a function of time. If the regression slope or any deviation from linearity is not statistically significant - meaning that the concentration does not change, or changes only within an acceptable range over time - the material is considered stable. In practical terms, stability is confirmed when the absolute value of the slope, $|b_1|$, is smaller than the product of the critical Student's t value and the standard deviation of the slope $s(b_1)$ that is, $|b_1| < t_{(0.95,n-2)} * s(b_1)$. Both the t value and $s(b_1)$ are obtained according to equations (3) to (6), where x refers to the months of study duration, and y refers to the analyte concentration:

$$b_0 = \bar{y} - \overline{b_1}\bar{x} \quad \textbf{\textit{eq. 3}}$$

$$s(b_1) = \frac{s}{\sqrt{(\Sigma^{n}{}_{i=1}(x_i -)^2)}}\bar{x} \quad \textbf{\textit{eq. 4}}$$

$$s^2 = \frac{\Sigma^{n}{}_{i=1}(y_i - b_0 - b_1 x_i)^2}{(n-2)} \quad \textbf{\textit{eq. 5}}$$

$$s(b_0) = s(b_1)\sqrt{\left(\frac{\Sigma^{n}{}_{i=1}\ x_i{}^2}{\Sigma^{n}{}_{i=1}(x-\bar{x})^2}\right)} \quad \textbf{\textit{eq. 6}}$$

The uncertainty related to long-term stability (u_{lts}) is assessed from the regression curve that relates the analyte concentration to the duration of the stability study. When no significant trend is observed, the uncertainty in the determination of the regression slope $s(b_1)$ can be considered as the uncertainty associated with evaluating the material's degradation. Therefore, the uncertainty for long-term stability (u_{lts}) can be estimated using Equation (7) (29).

$$u_{lts} = s(b_1)\ .\ t_{(0.95,n-2)} \quad \textit{\textbf{eq. 7}}$$

For this estimation, a shelf life (t) must be defined. The longer the shelf life, the greater the contribution of uncertainty to stability.

2.5.3. Characterization uncertainty

The estimation of the uncertainty associated with the assigned value in an interlaboratory program is a key step to ensure both the reliability and the metrological traceability of the results obtained. In this type of study, different laboratories perform independent measurements of the same sample, and the assigned value is usually derived from a statistical consensus, such as the mean or a robust median of the reported results. However, this assigned value is not free from variability, as it is influenced both by the dispersion among participating laboratories and by the number of valid results included in the evaluation (30, 31).

In this context, calculating the uncertainty of the assigned value makes it possible to quantify the degree of confidence in this consensus, reflecting how well the collective result can adequately represent the measurand. Moreover, this calculation is fundamental for a correct interpretation of laboratory performance, since the comparison between a reported result and the assigned value must take into account not only the observed deviation but also the uncertainty associated with that reference value (32).

ISO 13528:2005 (31) recommends that the uncertainty of the assigned value be estimated using the robust standard deviation derived from the results of the interlaboratory test. The expression used is in equation 8:

$$u_x = \frac{1.25 \, . \, s^*}{\sqrt{p}} \quad \textbf{\textit{eq. 8}}$$

Where u_x represents the uncertainty of the assigned value, s^* is the robust standard deviation calculated from the laboratories' results, and $\sqrt{p}$ is the number of laboratories considered after the statistical treatment of the data. The factor 1.25 is applied as a conservative correction, ensuring that the estimated uncertainty is sufficiently broad to cover additional variation that may occur in the process.

2.6. Analysis of variance (ANOVA) in Reference Material Production

Analysis of variance (ANOVA) is a widely used statistical tool to assess whether there are significant differences between the means of different groups. In the context of reference material (RM) production, the use of ANOVA is even recommended by international normative documents, such as the ISO guidelines, especially in homogeneity and stability studies (11). This recommendation is based on the fact that ANOVA is a well-established method, easy to apply, and relatively straightforward to interpret. However, despite its importance, the technique presents some limitations that need to be critically considered.

One of the main limitations is its strict reliance on statistical assumptions, such as data normality and homogeneity of variances. When these assumptions are not met — which is relatively common in homogeneity or stability studies — results may become biased, reducing the reliability of the conclusions (33).

Another critical aspect is that ANOVA, by itself, does not provide quantitative information about the magnitude of variation between groups. It only tests the hypothesis of mean equality. For RM characterization, this can be insufficient, as the process requires not only detecting statistical differences but also evaluating the impact of these variations on the overall uncertainty of the assigned value (34).

In this context, the use of chemometric tools has emerged as a promising alternative to ANOVA. Techniques such as principal component analysis (PCA) (35), partial least squares discriminant analysis (PLS-DA), and other multivariate calibration methods allow for the simultaneous exploration of multiple variables, the identification of hidden patterns in data, and a more robust assessment of material homogeneity and stability (36, 37). Moreover, these approaches are better suited to deal with complex data sets and with deviations from the assumptions required by classical univariate statistics.

Therefore, although ANOVA is recommended by ISO guidelines and remains a useful first step in RM evaluation, its application should be interpreted with caution. Whenever possible, it should be complemented — or even replaced — by chemometric approaches. This integration provides more reliable conclusions, both statistically and metrologically, ultimately strengthening the quality of the produced reference materials.

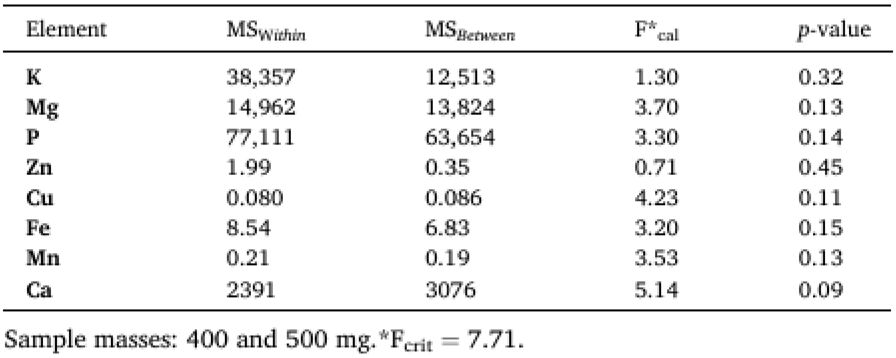
3. Chemometric tools

3.1. Use of chemometric tools for the production of reference material

The use of chemometric tools has proven to be an effective strategy in the development of reference materials, particularly in the evaluation of homogeneity and stability. In the study conducted for Azevedo et al. (2025) (13), the preparation of a reference material from pumpkin seed flour, Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) were applied alongside Analysis of Variance (ANOVA) to enhance the robustness of the results.

The application of PCA enabled the visualization of sample distribution according to the evaluated masses, highlighting the formation of consistent clusters for subsamples prepared with 500 mg, whereas smaller masses (100 to 300 mg) showed greater dispersion and reduced precision. These results supported the definition of the minimum representative mass of the material, ensuring the reliability of subsequent analyses. HCA confirmed these findings, revealing two main groups: one formed by the 500 mg samples and another encompassing the smaller masses, thus reinforcing the recommendation to use larger sample masses to obtain accurate and precise results. Figures 1, 2 e 3 shows the results obtained by Azevedo et al. (2025).

Figure 1 - ANOVA results for the minimum mass study (400–500 mg).

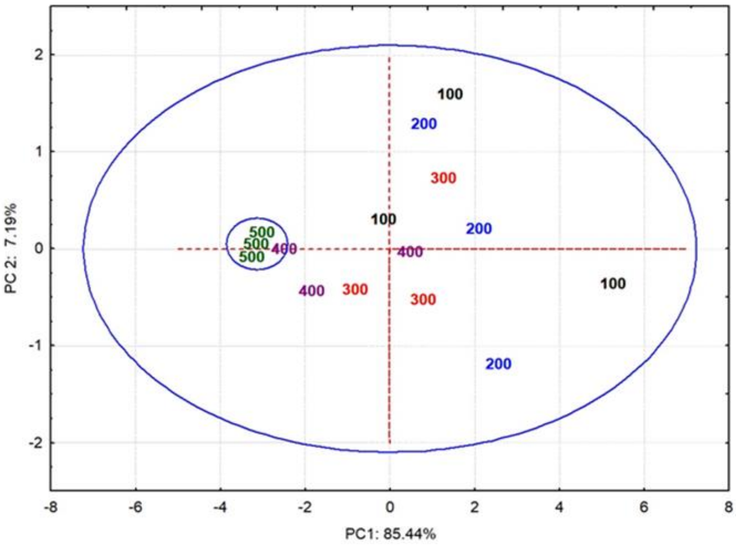


Element	MS_{Within}	$MS_{Between}$	F^*_{cal}	*p*-value
K	38,357	12,513	1.30	0.32
Mg	14,962	13,824	3.70	0.13
P	77,111	63,654	3.30	0.14
Zn	1.99	0.35	0.71	0.45
Cu	0.080	0.086	4.23	0.11
Fe	8.54	6.83	3.20	0.15
Mn	0.21	0.19	3.53	0.13
Ca	2391	3076	5.14	0.09

Sample masses: 400 and 500 mg.$^*F_{crit} = 7.71$.

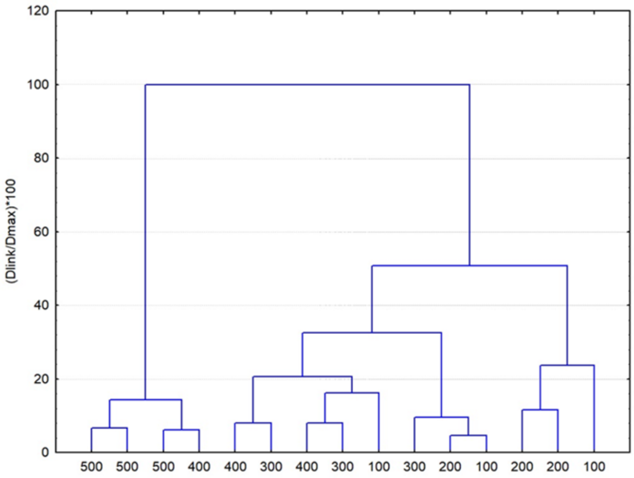
Source: AZEVEDO et al. (2025).

Figure 2 - Score graph of the first two principal components (PC1 x PC2), for the study of the minimum mass of the pumpkin seed flour candidate for reference material



Source: AZEVEDO et al. (2025).

Figure 3 - Dendrogram of the study of the minimum mass for the candidate reference material of pumpkin seed flour



Source: AZEVEDO et al. (2025).

Furthermore, the same chemometric techniques were applied in the stability study, confirming that the samples remained consistent over time and under different storage conditions. The random distribution of scores in PCA and the absence of differentiated clusters in the HCA dendrograms demonstrated that the matrix remained stable for up to 12 months, even when subjected to temperature variations.

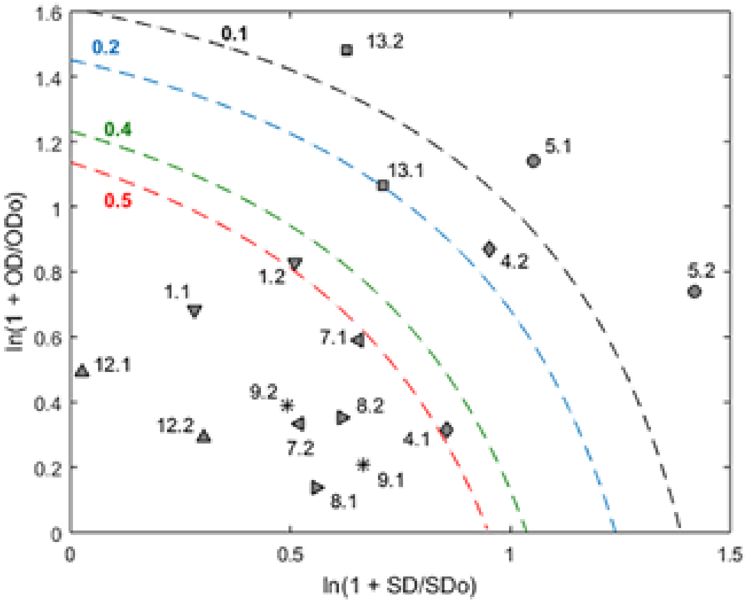
The conclusions of the study indicate that pumpkin seed flour is a viable matrix for the production of reference material, not only due to its availability and composition rich in essential elements, but also because of the satisfactory performance obtained through statistical and chemometric analyses. The integrated use of ANOVA, PCA, and HCA provided greater confidence in the evaluation of homogeneity and stability, offering an innovative and robust approach to the characterization of reference materials.

In the study by dos Santos et al. (2022) (36), the authors aimed to develop a corn flour laboratory reference material for inorganic analysis, assessing both its stability and the performance of an interlaboratory comparison. To support this evaluation, they applied advanced chemometric tools, with particular emphasis on DD-SIMCA (Data-Driven Soft Independent Modeling of Class Analogy), a recent extension of the traditional SIMCA method.

The DD-SIMCA approach was employed to process the results generated by 13 participating laboratories, which determined the concentrations of elements such as calcium, potassium, magnesium, phosphorus, copper, iron, manganese, and zinc. By modeling the expected class of results, the method defined an acceptance region based on statistical distances, allowing the authors to check whether the laboratories' outcomes were consistent with the overall dataset.

The analysis revealed that results from laboratories 5 and 13 displayed clear anomalies, with their replicates falling outside the acceptance boundaries, can be seen in figure 4. These outcomes were considered unsatisfactory and were excluded from the final assignment of mass fraction values. In addition, one replicate from laboratory 4 showed questionable behavior. Importantly, the conclusions drawn with DD-SIMCA were in agreement with those obtained through the conventional robust z-score recommended by ISO Guide 35 (11), but the multivariate approach offered clearer visualization and a more comprehensive interpretation of the data.

Figure 4 - One-class-modeling plot showing the acceptance area for different alpha values: α =0.1, 0.2, 0.4 and 0.5.



Source: dos SANTOS et al. (2022).

The authors concluded that DD-SIMCA is a valuable tool for the evaluation of interlaboratory studies, as it facilitates the identification of inconsistent results in a straightforward and reliable manner. They further highlighted that the method can serve as a robust alternative to the univariate procedures prescribed in ISO 13528 (31) and ISO Guide 35 (11), strengthening the certification process of reference materials.

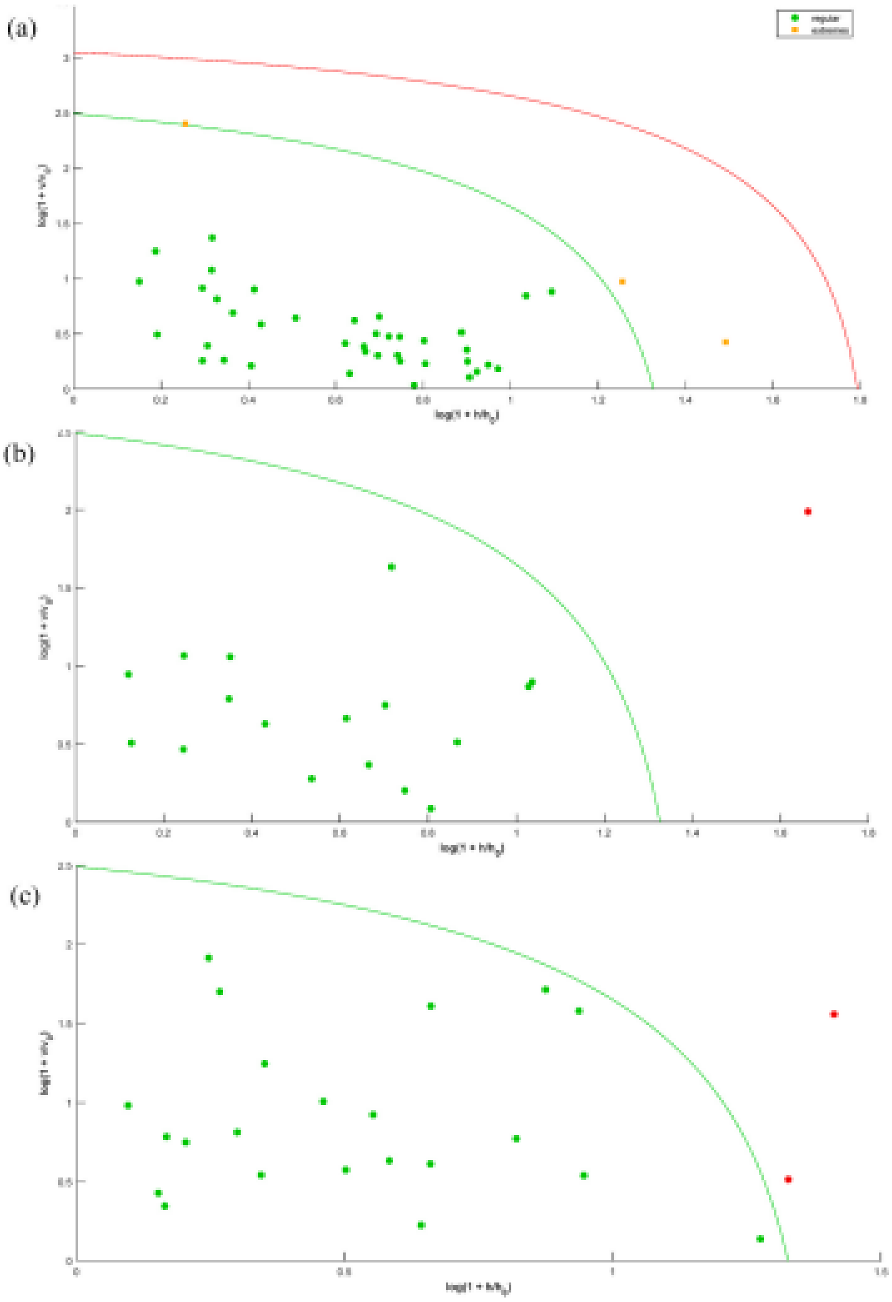
Overall, the work demonstrates how the integration of chemometric methods, particularly DD-SIMCA, enhances the reliability of reference material production by ensuring that only consistent and representative results are considered in the calculation of assigned values and associated uncertainties.

In the work of Gomes et al. (2025) (37), the authors investigated the short-term stability of a marine sediment candidate reference material (CRM) by applying DD-SIMCA as an alternative tool to conventional statistical approaches. The study sought to verify whether DD-SIMCA could provide a reliable and more informative perspective on stability assessment, complementing or replacing the traditional univariate methods recommended by ISO Guide 35.

The experimental design involved exposing subsamples of the sediment candidate to different temperature conditions (−20 °C, 4 °C, 20 °C and 40 °C) over a period of three months, followed by elemental analysis. While stability evaluation is typically carried out through ANOVA or regression-based tests, the authors proposed that DD-SIMCA could reveal systematic trends and highlight deviations in a multivariate and visually intuitive manner.

The application of DD-SIMCA allowed the construction of class models for the reference temperature (−20 °C), against which the other conditions were compared (which can be seen in figure 5). The results showed that the material remained within the model boundaries at 4 °C and 20 °C, indicating no significant degradation. However, samples stored at 40 °C presented a clear deviation from the acceptance area, suggesting a loss of stability under this condition. These findings were consistent with traditional statistical analysis, but DD-SIMCA provided a clearer and more straightforward interpretation of the data.

Figure 5 - (a) Model built using the DD-SIMCA method with quintuplicates of samples from time zero. (b) Acceptance plot of the test set to verify the viability of the constructed model. (c) Acceptance plot of triplicates of samples subjected to a temperature of 60 °C



Source: GOMES et al. (2025).

The authors concluded that DD-SIMCA is a powerful complementary tool for stability studies of reference materials, as it is capable of detecting changes in sample behavior in a simple and robust way. They emphasized that its multivariate nature offers advantages over purely univariate approaches, reinforcing its potential as an alternative method for the assessment of reference material stability in accordance with international guidelines.

In summary, this study demonstrated that incorporating DD-SIMCA into stability evaluation contributes to a more reliable certification process by confirming the robustness of the material under realistic storage conditions, while also offering an innovative multivariate perspective.

3.2. Z-score and robust Z-score

The Z-score is a widely used statistical tool in interlaboratory studies and proficiency testing to evaluate the performance of laboratories against an assigned reference value. It is a standardized index that expresses the difference between the result reported by a laboratory and the assigned value of the measurand, in units of the performance assessment standard deviation (38). Mathematically, the definition of z-score is in equation 9:

$$Z\ score = \frac{x_i - X}{\sigma} \quad \textbf{\textit{eq. 9}}$$

Where x_i is the result reported by the participating laboratory, X is the reference value (usually obtained from the mean of the participants' results), and σ is the the standard deviation of the values obtained from each laboratory. Interpretation is straightforward: values of z-score between -2 and +2 are considered satisfactory, values between ±2 and ±3 are regarded as questionable, and values greater than ±3 indicate unsatisfactory performance. While simple and effective, this approach can be sensitive to the presence of outliers, since both the mean and the standard deviation are influenced by extreme results (38).

To overcome this limitation, the robust Z-score is often used, this technique is a statistical approach frequently applied to data sets with large internal variability. Its main advantage is that it is less affected by extreme values compared to the conventional Z-score. In this method, the arithmetic mean is replaced by the median, and each data point receives a Z-score that reflects its deviation from the central value of the distribution. Based on the obtained Z-score, results can be classified as acceptable, doubtful, or outliers (28).

The calculation follows equation 10, where Z_a is the robust Z-score for the measurement X_a, M represents the dataset median, and NIQR is the normalized interquartile range:

$$Z_a = \frac{X_a - M}{NIQR} \quad \textbf{\textit{eq. 10}}$$

The interquartile range (IQR) is defined as the difference between the third and first quartiles. To normalize it, the IQR is multiplied by a constant factor of 0.7413. This constant comes from the standard normal distribution (mean = 0, standard deviation = 1), where the theoretical interquartile range equals 1.34898. The normalization factor is therefore obtained as 1 / 1.34898 = 0.74131 (28).

3.3. Data-Driven Soft Independent Modeling of Class Analogy (DD-SIMCA)

DD-SIMCA (Data-Driven Soft Independent Modeling of Class Analogy) is a chemometric method based on multivariate analysis, mainly applied to sample classification and authentication. It can be seen as an extension of the traditional SIMCA approach (Soft Independent Modeling of Class Analogy), which builds statistical models from a target class defined by a training set of samples. The main idea is to determine whether new samples belong to that previously modeled class or not (39).

The method uses Principal Component Analysis (PCA) as the main tool to describe the variability of the samples within the class of interest. PCA allows the dimensionality of the data to be reduced, condensing information from many variables into a few principal components, without significant loss of relevant information. In this way, each sample is projected into a multivariate space defined by these components, and its position is evaluated in relation to the class model.

In traditional SIMCA, the assessment of whether a sample belongs to the model is based on two main parameters:

- **Orthogonal distance (Q residual):** measures how far the sample deviates from the model in terms of unexplained variance;
- **Hotelling's T^2 distance:** measures how far the sample is from the center of the class within the principal component space.

The DD-SIMCA methodology enhances this approach by introducing rigorous statistical limits derived from theoretical distributions (such as the chi-squared distribution) to establish acceptance and rejection zones (39). These limits allow each sample to be clearly classified into three categories:

1. **Accepted** – the sample belongs to the class model with a high degree of confidence.
2. **Suspect** – the sample lies in an uncertainty region, close to the class boundary.
3. **Rejected** – the sample does not fit into the class model.

One of the advantages of DD-SIMCA is that it is data-driven, meaning it does not rely on strict assumptions about the distribution of the original variables, but instead on the statistical properties observed in the training samples. This makes it particularly useful in fields such as food analysis, authentication of biological matrices, and quality control, where variables may be highly correlated and data often exhibit complex behaviors (40).

In addition, the method provides acceptance curves based on significance levels (α and γ), which can be adjusted depending on the need for greater sensitivity (avoiding false negatives) or greater specificity (avoiding false positives). In this way, DD-SIMCA offers a robust and flexible solution for class analysis in real-world problems, and has been widely applied in studies of authenticity, verification of geographical origin, plant species discrimination, and detection of adulteration (40, 41).

Chapter 2: Homogeneity of the Soil Reference Material Candidate

1. Introduction

Guaranteeing the reliability of analytical measurements has long been a central challenge in analytical chemistry. In areas such as food safety and environmental monitoring, the preparation of official reports and risk assessments depends on the use of validated methods, which in turn require appropriate reference materials. Ideally, these materials should closely match the chemical composition of the target matrix and exhibit analyte concentrations comparable to those found in real samples. However, the limited commercial availability of such materials and their high cost often restrict their use in routine validation procedures. As a result, increasing attention has been directed toward the development of laboratory-produced reference materials.(42).

The production of a reliable reference material is a demanding, multi-step process that follows strict metrological and statistical principles. It begins with the careful selection of a bulk sample representative of the target matrix, followed by preparation steps such as drying, milling, sieving, and homogenization to reduce variability and standardize particle size. Candidate materials are then subjected to rigorous testing to demonstrate their homogeneity, ensuring that subsamples provide consistent analytical results across the batch. According to ISO Guide 35:2017, additional stability studies are also essential to confirm that the material maintains its chemical integrity under different storage and handling conditions. Once these requirements are met, the material can be characterized through interlaboratory studies involving multiple laboratories and validated analytical methods. Homogeneity testing plays a particularly critical role, as it ensures that the uncertainty assigned to the reference value properly accounts for any potential material heterogeneity. (10, 11, 43).

Soft Independent Modeling of Class Analogy (SIMCA) is a chemometric technique proposed by Wod in 1976 (44), which has been widely employed by researchers in the environment, food, fuels, etc (41, 45). The DD-SIMCA version, Data-Driven Soft Independent Modeling of Class Analogy (39, 46), has been used in the preparation stages of laboratory reference materials. Santos et al. used it in the stability and interlaboratory steps (36). Azevedo et al. evaluated the homogeneity and stability of a sample of pumpkin seed flour (47).

Gomes et al. also used DD-SIMCA in the stability study of a marine sediment matrix (37). DD-SIMCA is a supervised one-class classification method that uses a PCA to calculate the distance in scores (Hi) and the class members' orthogonal distance (Qi). Thus, the computed distances determine the class limit and classification of extreme members and outliers at the chosen confidence level (37).

The robust Z-score is a statistical test that detects outliers in interlaboratory programs while preparing laboratory reference materials (48-50). Its value is calculated in a one-elementary manner, where the Z values denote the quality of a laboratory's result as a function of the median of all received data and the difference between the third and first quartiles of the results. Z values with a module lower than or equal two denote acceptable results (50-52). In addition, the use of a multi-elementary robust Z-score has been proposed in the evaluation of analytical data from a homogeneity study involving the quantification of fourteen chemical elements present in a contaminated soil matrix. The multiple response (MR) proposed by Ferreira (53, 54) is applied to the results of the elements, enabling a multi-elementary evaluation of the data obtained.

2. Objectives

2.1. Main Objective

Evaluate the homogeneity of a candidate soil reference material contaminated with metals, aiming at its application in inorganic chemical analyses.

2.2. Specific Aims

- To quantify thirteen chemical elements in the candidate soil reference material using inductively coupled plasma optical emission spectrometry (ICP OES).
- To determine the minimum sample mass required to ensure representativeness in the analyses.
- To assess within-bottle and between-bottle homogeneity of the prepared material.

- To apply analysis of variance (ANOVA) in order to verify the homogeneity of the material according to established statistical criteria.
- To use the chemometric technique Data-Driven Soft Independent Modeling of Class Analogy (DD-SIMCA) as a complementary tool for homogeneity evaluation.
- To apply the multi-element robust Z-score, through the Multiple Response (MR) approach, to provide an integrated assessment of the analytical results.

3. Experimental

3.1. Collection and preparation of the soil

Soil samples of approximately 4.0 kg were collected from *Santo Amaro da Purificação, Brazil*, an area known for its contamination by toxic metals (55). This material was transported to the laboratory in plastic bags. Firstly, this soil was processed using a 50 µm mesh sieve to exclude stone and leaf residues. Later, it was dried at 40°C for two weeks until constant weight. Afterward, the soil was again carefully processed through a 10 µm mesh sieve and then homogenized by a liquid paint mixing system, kindly provided by a commercial wall paint company. Afterward, the material was divided into 100 amber bottles containing approximately 30 g. To preserve the original characteristics of the material intended for certification, the material was sterilized using 15 kGy gamma radiation at the IPEN laboratory in São Paulo, Brazil.

3.2. Reagents and solutions

A Milli-Q purification system (Millipore, Bedford, MA, USA) was employed to obtain water with a resistivity of 18.2 MΩ cm. All reagents used had analytical grades from (Merck, Darmstadt, Germany). The calibration curves of the chemical elements were prepared using analytical standards of concentration 1000 mg L^{-1} (Merck, Darmstadt, Germany). The ranges of the calibration solutions were from 0.05 to 50.00 $mg{\cdot}L^{-1}$ for magnesium, calcium, and sodium and from 0.05 to 5.00 mg·L^{-1} for arsenic, cadmium, copper, manganese, nickel, lead, tin, selenium, strontium, and zinc, all in 1.5 mol·L^{-1} nitric acid medium. Blanks were prepared in the same way but without the addition of the sample.

3.3. Instruments

The inorganic species extraction from soil samples was performed using a Milestone Ethos Easy microwave digestion system, following the EPA 3051A method. The procedure was performed in two stages: initially, samples were subjected to 1800 W and 180 °C for 20 minutes, followed by the same conditions for an additional time of 10 minutes.

Multi-element determination was performed using inductively coupled plasma optical emission spectrometry (ICP OES, iCAP PRO XP Duo model, Thermo Scientific, Waltham, MA, USA), operating in axial and radial views. The instrument features a high-resolution echelle spectrometer with a charge injection device (CID) detector, a peristaltic pump, a concentric nebulizer, a cyclonic spray chamber, and a demountable quartz plasma torch with a 2.0 mm i.d. injector. Sample introduction was automated using an ASX-560 autosampler (Teledyne CETAC Technologies).

The operating conditions included an RF power of 1.25 kW, with plasma generation using high-purity argon gas (99.999%, White Martins Praxair, Salvador, Brazil) at a flow rate of 13.5 L·min^{-1}, auxiliary flow of 0.50 L·min^{-1}, and nebulizer flow of 0.70 L·min^{-1}. The determined elements and their respective emission wavelengths (nm) were: As (189.042), Ca (393.366), Cd (226.502), Cu (324.754), Mg (279.553), Mn (257.610), Na (588.995), Ni (231.604), Pb (220.353), Sb (206.279), Se (206.279), Sr (407.771), and Zn (202.548).

3.4. Extraction procedure employing microwave radiation

Masses of 0.5 g of the sample were weighed into microwave digestion vessels. Then, 5 mL of concentrated nitric acid was added, and the tubes were left to stand for 30 minutes to initiate digestion. After this period, 5 mL of ultrapure water was added, and the samples were subjected to microwave digestion. Once cooled, the digests were filtered, and brought to a final volume of 25 mL with ultrapure water. A two-step dilution was performed to adjust the final acidity, reaching a final nitric acid concentration of 1.5 mol·L^{-1}.

3.5. Homogeneity study

Homogeneity tests involve three steps, which are minimum mass determination within-bottle and between-bottle checks.

A minimum mass test was performed using a single bottle, and the chemical elements were determined in triplicate with sample portions of 0,1; 0,2; 0,3; 0,4 e 0,5g.

The between-bottle test checks consistency among different bottles from the same batch. To do this, a small percentage of bottles—typically between 2% and 5%—are randomly selected. From each bottle, two independent subsamples are taken and analyzed for the target analytes using validated methods. The results are then compared using statistical techniques, such as ANOVA or calculating the coefficient of variation, to determine whether any differences fall within acceptable limits. This step is essential to confirm that all bottles in the batch are equivalent (43).

The within-bottle test is used to assess the uniformity within a single bottle. This study randomly selects one bottle and performs ten independent analyses of its contents. The same analytes used in the within-bottle test are measured, and the data are statistically evaluated—often by calculating the standard deviation or coefficient of variation. A low variation suggests good internal consistency, vital for ensuring reliable and repeatable results during routine use of the RM (11, 43, 56).

The experiments were carried out using twelve bottles, and all analyses were quantified in triplicate in the following scheme:

(i) The minimum sample mass was determined using 0.1, 0.2, 0.3, 0.4, and 0.5 g samples taken from a single bottle.

(ii) The between-bottle homogeneity was evaluated using ten material bottles with 0.5 g samples.

(iii) Within-bottle homogeneity was assessed using ten subsamples of 0.5 g from the same bottle.

Conventionally, the data obtained in these experiments have been evaluated using analysis of variance (ANOVA) to assess the material homogeneity. Also, the ANOVA allows the calculation of the uncertainties associated with between-bottle variations, according to the guidelines of ISO Guide 35 (2017) (11, 43, 49).

3.6. Homogeneity assessment using DD-SIMCA

The homogeneity assessment using the DD-SIMCA model involved 42 samples with 13 variables, using autoscaling as a pretreatment process. The Kennard-Stone (KS) algorithm allowed the division of the samples into training (70%) and testing (30%) subsets (57, 58). The DD-SIMCA homogeneity model was developed and evaluated regarding sensitivity, considering the significance levels of 0.01, applied to the distance score (SD), the embedded absolute residuals (ACR), and Type I and II errors. The optimal number of principal components (PCs) was determined using a compatible approach, considering the best compromise between the number of true positives and true negatives.

All chemometric procedures were performed in the Matlab® 2018 environment (MathWorks Inc.). The DD-SIMCA models were developed using the graphical interface developed by Zontov *et al* (46), accessible at https: //github.com/yzontov/dd-simca.

3.7. Robust Z score calculation

The robust Z-score statistical tool has been used to evaluate data series with wide internal variations. This technique has the advantage of not being strongly influenced by the extreme values of the data set. In the robust Z-score test, the mean is replaced by the median. For each point of the data series, a Z-score is assigned, indicating the deviation of the point with the median value. The value of the Z-score obtained classifies the point as acceptable, questionable, or outlier (50, 51, 59). The Z-score is calculated by the equation 9 (can be seen in chapter 1, section 3.2).

The homogeneity assessment using the robust multielement z-score was processed in a matrix containing 30 rows that come from three replicates of 10 samples from ten different bottles and nine columns consisting of the concentrations of the chemical elements arsenic, calcium, cadmium, copper, magnesium, manganese, nickel, lead, and zinc. First, to calculate the robust multielement z-score, a multiple response (MR) was determined, which results from the sum of the concentrations of the nine chemical elements determined in the material, normalized by the median values of each chemical element. The MR values are used directly in equation 1 to determine the Z values.

4. Results and discussions

4.1. Validation of the analytical method

The limits of detection (LOD) and quantification (LOQ) were determined following the guidelines established by the International Union of Pure and Applied Chemistry (IUPAC) (60). The calculations were based on the standard deviation (σ) of ten blank measurements and the slope (S) of the analytical calibration curve, using the expressions LOD = 3σ/S and LOQ = 10σ/S. The LOD and LOQ values, calculated for a sample mass of 0.5g and expressed in milligrams of element per kilogram of the sample, were as follows: 0.05 and 0.17 for As; 3.24 and 10.79 for Ca; 0.0042 and 0.014 for Cd; 0.017 and 0.056 for Cu; 0.55 and 1.82 for Mg; 0.0023 and 0.0076 for Mn; 11.15 and 37.17 for Na; 0.011 and 0.036 for Ni; 0.014 and 0.047 for Pb; and 0.11 and 0.36 for Zn, respectively.

During the validation process, the accuracy of the method was assessed through the analysis of three certified reference materials: CRM-Agro E2002a (sandy soil), PACS-2 (marine sediment), and MESS-3 (marine sediment). As shown in Table 1, the results demonstrate statistical agreement between the concentrations obtained for ten chemical elements using ICP OES and the certified values of the reference materials.

Table 1: Validation parameters of the ICP OES method used for the determination of the chemical elements

Chemical elements	**Certified value (mg/kg)**	**Achieved value* (mg/kg)**	**Recovery (%)**	**CRM's used**	**Repeatability (RSD, %, n = 7)**	**Calibration curve ranges (mg/L)**
As	59.3 ± 7.2	53.6 ± 1.5	90	CRM-Agro E2002a	3.60	0.01 – 5.00
Ca	250 ± 70	237 ± 8	95	CRM-Agro E2002a	1.40	0.54 – 10.00
Cd	94.0 ± 11.4	85.4 ± 1.8	91	CRM-Agro E2002a	6.78	0.01 – 5.00
Cu	8.8 ± 4.0	7.1 ± 0.2	81	CRM-Agro E2002a	2.69	0.01 – 5.00
Mg	330 ± 60	300 ± 19	91	CRM-Agro E2002a	5.95	0.01 – 10.00
Mn	130 ± 20	114 ± 7	88	CRM-Agro E2002a	5.90	0.01 – 5.00
Na	80 ± 20	101 ± 32	126	CRM-Agro E2002a	12.36	1.85 – 10.00
Ni	39.5 ± 2.2	37.3 ± 0.8	94	PACS-3	5.17	0.01 – 5.00
Pb	188.0 ± 7.4	177.6 ± 3.6	95	PACS-3	4.68	0.01 – 5.00
Sb	1.07 ± 0.16	1.19 ± 0.16	111	MESS-4	9.30	0.01 – 5.00
Se	0.92 ± 0.22	0.93 ± 1.14	101	PACS-2	7.49	0.06 – 5.00
Sr	132 ± 8	103 ± 8	78	MESS-4	4.34	0.01–5.00
Zn	147 ± 6	156 ± 13	106	MESS-4	8.50	0.02 – 5.00

* - Results expressed as a confidence interval at a 95% level; Calibration curves prepared with 10 points

4.2. Minimum mass study

The minimum sample mass study aimed to determine the smallest representative portion of the material. For this purpose, sample masses ranging from 0.1 to 0.5 g were evaluated. All experiments were performed in triplicate. The data evaluation using analysis of variance (ANOVA) revealed that this mass range lacks homogeneity at the 95% confidence level. For seven of the thirteen elements analyzed, the calculated F-value exceeded the critical F-value, and the corresponding p-values were below 0.05, indicating statistically significant differences within this mass interval (61). This can be seen in Table 2.

Table 2: ANOVA for the minimum mass study (0.1 to 0.5g)

	MS_{among}	MS_{within}	F_{cal}	P-value
As	67.18	12.49	5.38*	0.01714328
Ca	71225.30	2455.61	29.01	0.00001763
Cd	434.85	251.13	1.73	0.21924429
Cu	31.87	144.20	0.22	0.92059056
Mg	68195.11	564.89	120.72	0.00000002
Mn	767.88	548.28	1.40	0.30235491
Mo	163.95	128.11	1.28	0.34085432
Ni	25.21	3.23	7.80	0.00403167
Pb	6472.91	3709.96	1.74	0.21651186
Sb	152.10	0.72	210.67	0.00000000
Se	38.57	3.51	10.98	0.00111464
Sr	105.05	3.77	27.85	0.00002120
Zn	982.07	2009.61	0.49	0.74436571

MS = mean square; F_{crit} = 3.478; * F_{crit} = 3.633

Degrees of freedom = 14

Given this, another assessment was performed using the test results obtained with the mass range from 0.2 to 0.5 g. In this second evaluation, the ANOVA demonstrated that the calculated F-values were lower than the critical F-values for all elements, and the p-values were greater than 0.05 for all samples, indicating the absence of statistically significant differences (61). The results are presented in Table 2. It can be concluded that the material is homogeneous for sample masses greater than 0.2 g.

Table 3: ANOVA results for the minimum mass study (0,2 -0,5 g)

	MS_{among}	**MS_{within}**	**F_{cal}**	**P-value**
As	30.1274	6.9472	4.3366*	0.0503
Ca	3.0616	10.3144	0.2968	0.8269
Cd	551.9318	212.5587	2.5966	0.1248
Cu	0.9284	1.9702	0.4712	0.7107
Mg	2.4983	44.3441	0.0564	0.9811
Mn	948.9459	569.0453	1.6676	0.2501
Na	5785.9364	11909.7838	0.4858	0.7015
Ni	2.1709	2.2966	0.9453	0.4632
Pb	8567.2125	3176.8999	2.6967	0.1165
Sb	0.7095	0.5334	1.3302	0.3309
Se	6.5745	3.7824	1.7382	0.2364
Sr	3.1689	1.9550	1.6209	0.2598
Zn	1309.3353	1431.4032	0.9147	0.4760

MS = mean square; F_{crit} = 4.066; * F_{crit} = 4.346

Degrees of freedom = 11

4.3. Within-bottle study

The within-bottle homogeneity was evaluated by analyzing twelve subsamples from a single, randomly selected bottle of the candidate reference material. All measurements were carried out under repeatability conditions. All experiments were performed in triplicate. The ANOVA results for the elements analyzed are presented in Table 4. For all chemical elements evaluated, the calculated F-values (F_{cal}) were lower than the critical F-values (F_{crit}) at the 95% confidence level, and the corresponding p-values were greater than 0.05. These results indicate no statistically significant differences among the replicate measurements (61). Therefore, the bottle selected for this study exhibits sufficient homogeneity throughout its contents.

Table 4 - ANOVA results for the within-bottle test

	MS_{among}	MS_{within}	F_{cal}	P-value
As	9.48	3.54	2.68	0.09
Ca	17.94	525.63	0.03	0.97
Cd	1.51	147.78	0.010	0.98
Cu	1.19	1.59	0.75	0.48
Mg	87.75	2302.54	0.04	0.96
Mn	88.45	478.25	0.18	0.83
Na	15.91	1000.37	0.016	0.98
Ni	0.04	2.26	0.02	0.98
Pb	16.71	573.96	0.03	0.97
Sb	0.78	1.45	0.53	0.59
Se	0.28	3.07	0.09	0.91
Sr	0.30	4.57	0.066	0.94
Zn	136.24	1347.56	0.10	0.90

MS = mean square; F_{crit} = 3.316

Degrees of freedom = 32

4.4. Between-bottle study

Between-bottle homogeneity was assessed by analyzing ten randomly selected bottles of the candidate reference material. Three independent subsamples of 0.5 g were taken from each bottle, and thirteen chemical elements were determined. The data were evaluated using ANOVA. For all chemical elements, the calculated F-values (F_{cal}) were lower than the corresponding critical F-values (F_{crit}) at the 95% confidence level, and the associated p-values exceeded 0.05 (61). These results indicate no statistically significant differences among the bottles, confirming the between-bottle homogeneity of the candidate reference material. The detailed ANOVA results are presented in Table 5.

Table 5 - ANOVA results for the between-bottle test

	MS_{among}	MS_{within}	F_{cal}	P-value
As	54.00	26.02	2.0747	0.0835
Ca	1724	1117	1.5440	0.1999
Cd	59	111	0.5298	0.8358
Cu	3.45	2.92	1.1823	0.3575
Mg	384	339	1.1338	0.3854
Mn	611	318	1.9226	0.1072
Na	2470	4273	0.5780	0.7994
Ni	1.32	1.51	0.8749	0.5622
Pb	11494	16131	0.7126	0.6916
Sb	0.18	0.49	0.3623	0.9399
Se	4.20	1.88	2.2380	0.0641
Sr	3.95	2.47	1.5949	0.1838
Zn	2288	1448	1.5799	0.1884

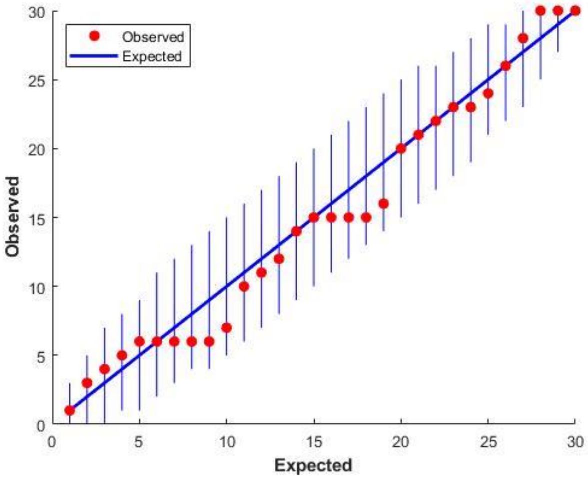
MS = mean square; F_{crit} = 2.39.

Degrees of freedom = 29

4.5. Evaluation of the homogeneity test using the DD-SIMCA method

The data obtained in the within-bottle and between-bottle homogeneity tests were combined, resulting in a matrix with 42 samples and 13 variables. These data were then processed using the Kennard-Stone (KS) algorithm, which divided the samples into two sets: training (70%) and testing (30 samples) (30%, 12 samples). The preprocessing used was autoscaling in both sets. Then, the DD-SIMCA algorithm was applied to the training set. Two PCs were chosen to build the model, which presented the largest number of principal components with the smallest number of objects considered extreme and without outliers, as shown in the extremes graph in Figure 4.

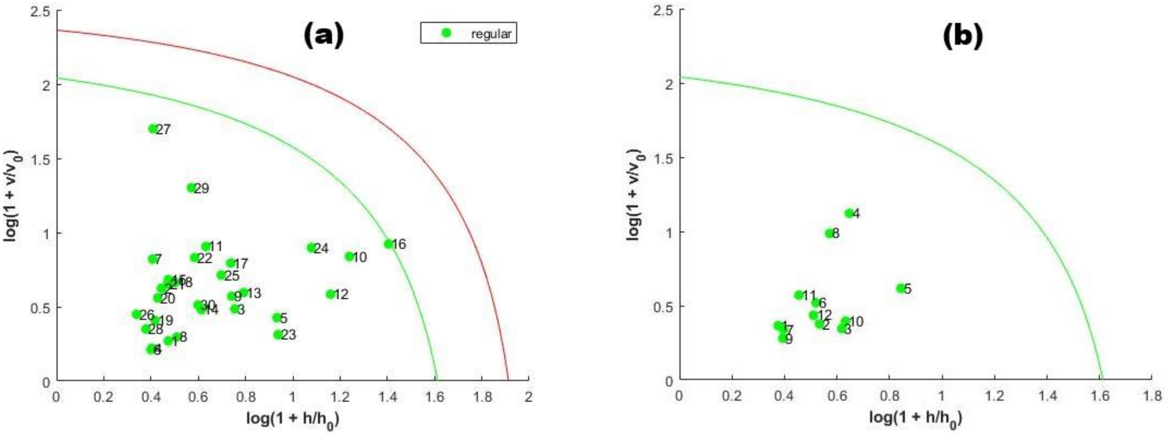
Figure 6 - Extreme plot of the training set.



Source: Own elaboration

The acceptance graph shown in Figure 5a shows that all samples in the training set are in the acceptance region below the green line. It can also be seen that no sample was considered extreme (located between the green and red lines) or outlier (located above the red line), and the limit highlighted in red has a significance level of α = 0.01. Then, the DD-SIMCA model obtained was applied to the test set (12 samples) through a controlled prediction. The result, as seen in Figure 5b, demonstrates that all samples were correctly classified, indicating good homogeneity of the sample set. The 42 samples showed acute sensitivity, reaching a rate of 100% in both the training and test sets.

Figure 7 - (a) Model built using the DD-SIMCA method with training set. (b) Acceptance plot of the test set to verify the homogeneity of the constructed model.



4.6. Use of robust multivariate Z-score in the assessment of homogeneity

The robust multivariate Z-score statistical test evaluated the data from the between-bottle assay consisting of the analysis of 10 bottles performed in triplicate. For the multivariate evaluation, a multiple response function consisting of the sum of the normalized concentrations was used. Then, the normalization was performed in two ways: (i) processing using the median of the concentrations. (ii) processing using autoscaling that proceeds by dividing the individual values by the standard deviation of the series. The analysis involved determining 13 chemical elements whose concentrations are distinct, as seen in Table 6, and therefore, normalization would necessarily need to be performed.

The results of the multivariate robust Z-score, using the normalization of the concentrations by the median, presented Z values ranging from -1.45 to 1.53, indicating no statistically significant difference in the 30 results obtained, proving the homogeneity of the material. The data obtained in this test are presented in Table 7.

The application of the multivariate robust Z-score processing and the normalization of the concentrations by the autoscaling method indicated Z values ranging from -1.47 to 1.42, demonstrating that all 30 values obtained are acceptable, also confirming the homogeneity of the analyzed soil sample. This can be seen in Table 8.

Z values less than or equal to -2 and greater than or equal to 2 indicate questionable values, and greater than or equal to 3 and less than or equal to -3 indicate outliers (50, 51, 59).

Table 6 - Between-bottle study data

	Conc. (mg/kg)												
	As	**Ca**	**Cd**	**Cu**	**Mg**	**Mn**	**Na**	**Ni**	**Pb**	**Sb**	**Se**	**Sr**	**Zn**
1	47.30	1299.04	29.81	31.54	1082.51	174.92	135.57	13.36	460.49	5.66	19.33	24.67	289.23
2	47.15	1260.98	31.80	25.54	920.39	179.31	72.08	9.06	383.11	4.03	18.32	18.76	268.32
3	59.79	1309.46	38.52	29.85	1012.15	185.55	162.24	10.41	436.34	4.94	22.35	21.68	253.62
4	64.59	1311.21	28.44	29.24	1051.30	192.40	129.20	10.48	496.31	4.51	19.40	21.85	355.31
5	47.44	1296.66	46.37	28.63	1028.22	165.37	225.72	11.02	414.01	5.74	17.13	21.99	269.67
6	42.35	1276.31	22.92	28.14	1029.18	176.79	211.56	11.10	452.04	5.10	17.76	21.52	260.95
7	52.06	1345.04	45.09	30.57	1042.20	196.32	88.98	10.54	493.37	4.56	20.92	21.94	315.78
8	51.49	1350.39	39.79	29.42	1047.44	164.61	153.09	11.04	469.43	4.77	21.56	25.43	264.58
9	52.47	1309.33	33.16	30.87	1067.96	222.62	150.06	12.44	449.58	5.24	20.27	23.69	294.28
10	42.74	1273.17	28.10	28.79	1012.56	154.85	312.59	12.53	388.84	6.11	20.93	23.59	206.97
11	44.65	1309.50	26.57	27.85	926.31	169.95	76.66	8.48	445.25	3.92	20.70	19.21	282.81
12	53.11	1269.39	38.61	27.68	956.98	179.81	169.43	9.94	388.07	5.36	19.52	20.37	209.85
13	54.84	1350.31	45.15	29.81	1058.72	210.02	159.36	11.50	503.05	4.95	21.09	21.86	346.52
14	51.58	1324.64	46.56	30.48	1015.55	193.04	176.58	10.55	468.90	4.99	19.13	21.97	229.80
15	54.76	1293.04	31.46	30.25	1009.02	178.42	272.01	10.96	468.21	4.91	20.34	21.80	249.41
16	58.77	1270.05	34.91	27.37	926.56	157.02	194.08	9.78	379.86	4.48	22.06	19.92	211.24
17	55.33	1276.42	43.78	29.39	1017.77	180.78	245.85	11.61	380.43	5.42	21.04	21.88	210.94
18	53.53	1320.24	33.02	28.52	999.92	177.62	150.61	10.39	446.68	4.35	18.08	21.10	260.53
19	58.62	1322.39	39.76	29.49	1003.67	184.06	288.22	11.21	410.23	6.44	19.34	22.61	223.65
20	59.34	1306.90	35.06	30.11	969.67	183.01	141.00	9.78	427.81	4.53	19.08	20.89	273.30
21	55.00	1292.21	46.12	26.30	896.55	183.20	87.26	8.68	360.52	3.98	19.19	18.83	249.38

22	59.62	1291.74	36.94	28.95	1004.46	207.52	229.02	11.29	423.10	5.09	20.95	21.61	267.70
23	61.85	1297.82	33.26	30.90	959.46	226.15	114.00	10.12	406.16	5.16	23.48	20.76	218.11
24	54.33	1262.36	41.77	28.42	912.72	195.62	194.30	9.33	375.27	4.13	20.94	20.02	189.51
25	47.19	1272.21	36.94	29.75	979.77	183.09	155.80	10.48	421.50	3.96	15.85	17.44	234.64
26	46.59	1309.81	28.89	27.14	922.74	164.05	175.30	8.41	422.19	4.67	19.50	19.74	254.25
27	42.95	1188.78	20.28	23.11	870.38	137.00	155.80	8.65	333.69	3.96	15.85	17.44	182.99
28	54.71	1201.81	27.12	27.75	925.40	164.21	163.38	10.01	395.28	4.49	18.43	20.57	237.07
29	56.39	1313.19	69.79	28.46	998.34	157.66	149.84	9.71	403.92	4.57	18.94	20.73	285.52
30	51.89	1276.90	25.09	27.72	931.92	208.80	135.96	9.62	462.42	4.19	19.99	20.58	277.91

Table 7 - Results of the multivariate robust Z-score (median)

	As	Ca	Cd	Cu	Mg	Mn	Ni	Pb	Sb	Se	Sr	Zn	Multiple Response (MR)	Z –score robust
1	0.88	1.00	0.85	1.09	1.08	0.97	1.28	1.09	1.19	0.99	1.15	1.11	12.67	0.66
2	0.88	0.97	0.91	0.88	0.92	0.99	0.87	0.91	0.84	0.94	0.87	1.03	11.00	-1.45
3	1.12	1.01	1.10	1.03	1.01	1.03	0.99	1.03	1.03	1.14	1.01	0.97	12.48	0.42
4	1.21	1.01	0.81	1.01	1.05	1.06	1.00	1.17	0.94	0.99	1.02	1.36	12.64	0.63
5	0.89	1.00	1.32	0.99	1.02	0.91	1.05	0.98	1.20	0.88	1.02	1.04	12.30	0.20
6	0.79	0.98	0.65	0.97	1.03	0.98	1.06	1.07	1.07	0.91	1.00	1.00	11.51	-0.81
7	0.97	1.04	1.29	1.06	1.04	1.09	1.01	1.17	0.96	1.07	1.02	1.21	12.91	0.97
8	0.96	1.04	1.13	1.02	1.04	0.91	1.05	1.11	1.00	1.10	1.18	1.02	12.57	0.54
9	0.98	1.01	0.95	1.07	1.06	1.23	1.19	1.06	1.10	1.04	1.10	1.13	12.91	0.98
10	0.80	0.98	0.80	0.99	1.01	0.86	1.20	0.92	1.28	1.07	1.10	0.79	11.80	-0.44
11	0.83	1.01	0.76	0.96	0.92	0.94	0.81	1.05	0.82	1.06	0.89	1.09	11.15	-1.27
12	0.99	0.98	1.10	0.96	0.95	0.99	0.95	0.92	1.12	1.00	0.95	0.81	11.72	-0.55
13	1.02	1.04	1.29	1.03	1.05	1.16	1.10	1.19	1.04	1.08	1.02	1.33	13.35	1.53
14	0.96	1.02	1.33	1.05	1.01	1.07	1.01	1.11	1.05	0.98	1.02	0.88	12.49	0.44
15	1.02	1.00	0.90	1.04	1.01	0.99	1.05	1.11	1.03	1.04	1.01	0.96	12.15	0.00
16	1.10	0.98	1.00	0.95	0.92	0.87	0.93	0.90	0.94	1.13	0.93	0.81	11.45	-0.89
17	1.03	0.98	1.25	1.02	1.01	1.00	1.11	0.90	1.13	1.08	1.02	0.81	12.34	0.25
18	1.00	1.02	0.94	0.99	1.00	0.98	0.99	1.06	0.91	0.93	0.98	1.00	11.79	-0.46
19	1.09	1.02	1.13	1.02	1.00	1.02	1.07	0.97	1.35	0.99	1.05	0.86	12.57	0.54
20	1.11	1.01	1.00	1.04	0.97	1.01	0.93	1.01	0.95	0.98	0.97	1.05	12.02	-0.15
21	1.03	1.00	1.32	0.91	0.89	1.01	0.83	0.85	0.83	0.98	0.87	0.96	11.48	-0.84

22	1.11	1.00	1.05	1.00	1.00	1.15	1.08	1.00	1.07	1.07	1.00	1.03	12.56	0.53
23	1.16	1.00	0.95	1.07	0.96	1.25	0.97	0.96	1.08	1.20	0.96	0.84	12.39	0.31
24	1.01	0.97	1.19	0.98	0.91	1.08	0.89	0.89	0.87	1.07	0.93	0.73	11.53	-0.79
26	0.88	0.98	1.05	1.03	0.98	1.01	1.00	1.00	0.83	0.81	0.81	0.90	11.28	-1.10
27	0.87	1.01	0.82	0.94	0.92	0.91	0.80	1.00	0.98	1.00	0.92	0.98	11.14	-1.28
28	1.02	0.93	0.77	0.96	0.92	0.91	0.96	0.93	0.94	0.94	0.96	0.91	11.15	-1.27
29	1.05	1.01	1.99	0.98	0.99	0.87	0.93	0.95	0.96	0.97	0.96	1.10	12.77	0.80
30	0.97	0.98	0.72	0.96	0.93	1.15	0.92	1.09	0.88	1.02	0.96	1.07	11.65	-0.64

Mean (MR) = 12.06; Standard Desviation (MR) = 0.65; First quartile (MR) = 11.51; Third quartile (MR) = 12.57; median (MR) = 12.15

Table 8 - Results of the multivariate robust Z-score (autovalor)

	As	**Ca**	**Cd**	**Cu**	**Mg**	**Mn**	**Ni**	**Pb**	**Sb**	**Se**	**Sr**	**Zn**	**Multiple Response (MR)**	**Z-score robust**
1	8.31	42.08	3.17	22.14	20.55	9.36	11.34	11.70	8.83	11.84	14.21	7.29	170.81	0.94
2	8.28	40.85	3.38	17.93	17.47	9.59	7.69	9.74	6.28	11.22	10.80	6.76	149.99	-1.47
3	10.50	42.42	4.09	20.96	19.21	9.93	8.83	11.09	7.70	13.69	12.48	6.39	167.30	0.54
4	11.35	42.48	3.02	20.53	19.95	10.29	8.89	12.61	7.03	11.89	12.58	8.96	169.58	0.80
5	8.33	42.01	4.92	20.10	19.52	8.85	9.35	10.52	8.95	10.49	12.67	6.80	162.50	-0.02
6	7.44	41.35	2.43	19.76	19.53	9.46	9.42	11.49	7.95	10.88	12.40	6.58	158.67	-0.46
7	9.15	43.57	4.79	21.46	19.78	10.50	8.95	12.54	7.11	12.82	12.64	7.96	171.26	1.00
8	9.05	43.75	4.23	20.65	19.88	8.81	9.37	11.93	7.44	13.21	14.65	6.67	169.62	0.81
9	9.22	42.42	3.52	21.67	20.27	11.91	10.56	11.42	8.17	12.41	13.64	7.42	172.63	1.16
10	7.51	41.24	2.98	20.21	19.22	8.28	10.63	9.88	9.52	12.82	13.59	5.22	161.11	-0.18

11	7.84	42.42	2.82	19.55	17.58	9.09	7.20	11.31	6.11	12.68	11.06	7.13	154.81	-0.91
12	9.33	41.12	4.10	19.44	18.16	9.62	8.44	9.86	8.36	11.96	11.73	5.29	157.41	-0.61
13	9.63	43.74	4.79	20.93	20.09	11.24	9.76	12.78	7.72	12.92	12.59	8.73	174.94	1.42
14	9.06	42.91	4.94	21.40	19.27	10.33	8.96	11.92	7.78	11.72	12.65	5.79	166.73	0.47
15	9.62	41.89	3.34	21.24	19.15	9.55	9.30	11.90	7.65	12.46	12.55	6.29	164.93	0.26
16	10.32	41.14	3.71	19.22	17.59	8.40	8.30	9.65	6.99	13.51	11.47	5.32	155.62	-0.82
17	9.72	41.35	4.65	20.63	19.32	9.67	9.85	9.67	8.44	12.89	12.60	5.32	164.11	0.17
18	9.40	42.77	3.51	20.02	18.98	9.50	8.82	11.35	6.78	11.07	12.15	6.57	160.93	-0.20
19	10.30	42.84	4.22	20.71	19.05	9.85	9.51	10.42	10.04	11.84	13.02	5.64	167.44	0.55
20	10.42	42.34	3.72	21.14	18.40	9.79	8.30	10.87	7.07	11.68	12.03	6.89	162.66	0.00
21	9.66	41.86	4.90	18.47	17.02	9.80	7.37	9.16	6.20	11.75	10.84	6.29	153.32	-1.08
22	10.47	41.85	3.92	20.33	19.06	11.10	9.58	10.75	7.94	12.83	12.44	6.75	167.03	0.51
23	10.86	42.04	3.53	21.70	18.21	12.10	8.59	10.32	8.04	14.38	11.95	5.50	167.23	0.53
24	9.54	40.89	4.44	19.95	17.32	10.47	7.91	9.54	6.44	12.83	11.53	4.78	155.64	-0.81
26	8.29	41.21	3.92	20.89	18.60	9.80	8.89	10.71	6.18	9.71	10.05	5.91	154.15	-0.99
27	8.18	42.43	3.07	19.06	17.51	8.78	7.14	10.73	7.28	11.95	11.37	6.41	153.90	-1.02
28	9.61	38.93	2.88	19.49	17.56	8.79	8.50	10.04	7.01	11.29	11.84	5.98	151.91	-1.25
29	9.91	42.54	7.41	19.99	18.95	8.43	8.24	10.26	7.12	11.60	11.94	7.20	163.59	0.11
30	9.12	41.37	2.66	19.47	17.69	11.17	8.16	11.75	6.54	12.25	11.85	7.00	159.03	-0.42

Mean (MR) = 162.37 ; Standard Desviation (MR) = 6.89; First quartile (MR) = 155.643; Third quartile (MR) = 167.296; median (MR) = 162.66

4.7. Uncertainty associated with the homogeneity test

The uncertainty associated with the between-bottle homogeneity test was calculated using the data obtained from ANOVA, as described in chapter one section 2.5.1, according to Equations 1 and 2. This uncertainty will be incorporated into the calculation of the expanded uncertainty for the thirteen certified chemical elements in the soil reference material. The results are presented in Table 9.

Table 9 - Uncertainty values obtained in the homogeneity test

Analyte	U_{bb}	Equation used
As	3.05	1
Ca	10.85	2
Cd	3.42	2
Cu	0.42	1
Mg	3.89	1
Mn	5.79	2
Na	21.22	2
Ni	0.40	2
Pb	11.44	2
Sb	0.23	2
Se	0.44	2
Sr	0.70	1
Zn	12.36	2

5. Conclusion

The precision assessment of the ICP OES method revealed relatively higher RSD values for sodium and zinc compared to the other elements, with sodium being particularly challenging for this technique. Nevertheless, these variations did not compromise the homogeneity of the soil reference material under study.

The homogeneity evaluation of the candidate material proved effective through the combined application of conventional statistical analyses and chemometric approaches, ensuring a uniform composition throughout the batch. The minimum mass assessment indicated that portions above 0.2 g are representative of the material. Both within-bottle and between-bottle homogeneity tests, performed using ANOVA, confirmed the even distribution of analytes across the prepared vials.

The DD-SIMCA model demonstrated robust performance, achieving 100% sensitivity in both the training and test sets, with no outliers detected. Additionally, the application of the robust multielement Z-score in conjunction with the Multiple Response (MR) approach enabled a comprehensive evaluation of homogeneity across multiple chemical elements simultaneously. In this approach, the Z-score was calculated based on the median rather than the mean, minimizing the influence of extreme values. Variability was assessed using the interquartile range (IQR), scaled by a correction factor, providing a more stable representation of data dispersion. These combined strategies allowed for a more accurate and reliable interpretation of the results, particularly in cases involving non-symmetric distributions or outliers.

This statistical framework aligns well with established methodologies, offering a consistent and thorough assessment of the material’s homogeneity. The findings from these analyses confirm that the candidate soil material meets the necessary criteria to serve as a reliable laboratory reference material for inorganic soil analysis. With homogeneity verified, the material is ready to proceed to subsequent stages of characterization and certification, as outlined by current regulations. The procedures applied in this study provide a solid and practical foundation for the preparation and evaluation of reference materials in similar contexts.

Chapter 3: Interlaboratory program and stability of the candidate for soil reference material.

1. Introdução

Certified reference materials are essential inputs in analytical chemistry laboratories, which exhibit adequate homogeneity and stability for certain defined properties and has been proven suitable for its intended use in measurement processes (12). These materials are required in the following situations: (i) use in the validation process confirming the accuracy of analytical methods (60, 62); (ii) use in the preparation of calibration curves for direct methods based on solid sampling (63, 64); (iii) procedures established using slurry sampling (63, 65); (iv) validation of analytical equipment and sensors (66-68).

The development of a reference material involves several critical steps and must adhere to strict quality criteria to ensure its suitability for analytical use. The process begins with the selection of a bulk sample that accurately represents the physical and chemical characteristics of the intended matrix. This material then undergoes preparation procedures—such as drying, milling, sieving, and homogenization—to reduce variability and achieve a consistent particle size distribution across the entire batch. To confirm the uniformity of the material, homogeneity testing is carried out, aiming to determine whether different portions of the sample yield comparable results for specific analytes. This step is conducted in accordance with the guidelines provided in ISO Guide 35:2017. Stability assessments are also performed to verify that the chemical composition of the material remains unchanged over time and under varying storage conditions. After these preliminary evaluations, the material is characterized through an interlaboratory comparison study. In this stage, several independent laboratories analyze the material using validated and traceable methods. The goal is to assign reliable reference values and assess the overall measurement uncertainty. The interlaboratory program plays a crucial role in confirming the accuracy, reproducibility, and broader applicability of the reference material across different analytical settings (10, 11, 43).

A study was developed to prepare a human hair reference material for determining mercury, methylmercury, and other elements. The identification of outliers in the data of the interlaboratory program was performed using the Grubbs test and the Z-score (69). A Chinese study prepared a CRM of tea containing five residues of five pesticides. The interlaboratory study involved 73 laboratories, and the proficiency assessment of the results was performed using Z-score (70). The preparation of pumpkin seed flour reference material for inorganic analysis involved an interlaboratory trial with 12 participants, and the evaluation of the data received was assessed using the Grubbs, Cochran, and robust Z-score tests (49). The interlaboratory test data for characterizing the powdered milk reference material to determine inorganic nutrients were evaluated using the Grubbs, Cochran, Z-score, robust Z-score tests, and confidence ellipses (71). A laboratory reference material of corn flour was prepared to determine eight inorganic nutrients. Of the 13 participant laboratories, eight presented data for all chemical elements. The robust Z-score test was used to detect outliers among the data received by the assay. In addition, the Data-Driven Soft Independent Modeling of Class Analog (DD-SIMCA) chemometric technique was used to evaluate the compatibility among the assay data (72).

2. Objectives

2.1. Main Objective

To develop a soil reference material for inorganic analysis, characterized and validated through an interlaboratory study, and further evaluated for stability.

2.2. Specific Aims

- To prepare the soil reference material through homogenization and sterilization processes.
- To perform the characterization of the material by means of an interlaboratory program involving 28 laboratories.
- To apply Grubbs and Cochran statistical tests to detect outliers in the obtained data.

- To use Z-score and robust Z-score techniques for the selection and evaluation of the results in the certification process.
- To certify the concentration of eight inorganic chemical elements in the produced reference material.

3. Experimental

3.1. Collection and preparation of the soil

Soil samples were collected near a former lead mining company located in *Santo Amaro da Purificação,* Brazil — a region historically known for its high concentrations of toxic metals in the environment (73). Due to the elevated levels of these contaminants, it was also necessary to collect samples from areas farther from the site to dilute the material and better control concentration levels. In total, approximately 4.0 kg of soil was gathered for the preparation stage.

The material was brought to the laboratory in plastic bags. To begin the preparation, the soil was sieved using a 50 µm mesh to eliminate coarse debris such as stones and plant matter. It was then dried at 40 °C for approximately two weeks, or until it reached a stable weight (74). Once dry, the soil was further refined using a finer 10 µm mesh sieve and thoroughly homogenized with the aid of a liquid mixing system typically used for wall paints — specifically, a gyroscopic mixer (model LA250-MJ) manufactured by Vitória & Cia, kindly provided by a commercial paint company.

The homogenized soil was then portioned into 100 amber glass bottles, each containing around 30 grams. To ensure the preservation of the material’s original properties for future certification, sterilization was performed using gamma irradiation at a dose of 15 kGy, carried out at the IPEN laboratory in São Paulo, Brazil.

3.2. Methodology

The methodology employed varied among the participating laboratories. Table 10 shows the extraction procedures and analytical methods used by each of them.

Table 10 - Decomposition methods and analytical instrumental techniques used by participants in the interlaboratory comparison for soil candidate reference material.

Laboratory	Extraction method	analytical method
1	Microwave	ICP OES
2	Digester block	FAAS
3	Microwave	ICP OES
4	Microwave	ICP OES
5	Microwave	ICP OES
6	Microwave	ICP OES
7	Digester block	MIP OES
8	Microwave	ICP OES e ICP-MS
9	Microwave	MP-AES
10	Digester block	GF AAS e DMA
11	Microwave	ICP OES
12	Digester block	ICP OES
13	Digester block	ICP OES
14	Heating plate	ICP MS
15	Microwave	ICP OES
16	Microwave	ICP OES
17	Microwave	ICP OES
18	Digester block	MP OES
19	Microwave	ICP OES
20	Microwave	ICP OES
21	Digester block	ICP OES
22	Digester block	ICP OES
23	Digester block	ICP OES e DMA
24	Microwave	FAAS
25	MAE-PDC	ICP OES, CVG-ICP-MS e ICP-MS
26	Microwave	ICP OES
27	Microwave	ICP MS
28	Microwave	ICP OES, ICP-MS e DMA
29	Microwave	FAAS e GFAAS

3.3. Statistical tests for evaluating interlaboratory trial data

3.3.1. Grubbs test.

The Grubbs test, widely used in evaluating interlaboratory program results, is a tool used to identify outliers in a data series. The value $\boldsymbol{G}$ is established by equation 11, where $\boldsymbol{x_i}$ is the suspect value, and $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are the mean and standard deviation calculated with all series data, including the suspect value (75-77).

$$\boldsymbol{G = \frac{Xi - \mu}{\sigma}} \quad \textit{\textbf{eq. 11}}$$

When the G value obtained is greater than the critical G value, it denotes that the suspect value investigated is an outlier of the data set. The critical G value is tabulated with (P = 0.05) (75, 76).

Another application of the Grubbs test involves evaluating the extreme values of the data series. The numbering of the equation 1 is changed to ($\boldsymbol{x_{max} - \mu}$) or ($\boldsymbol{\mu - x_{min}}$), where $\boldsymbol{x_{max}}$, $\boldsymbol{x_{min}}$, and $\boldsymbol{\mu}$, respectively, are the largest value, the mean, and the smallest value of the data set (75).

This test must be applied involving an outlier from time to time. After removing an outlier, the test should be used again until no outlier is detected (75, 77).

3.3.2. Cochran test.

In interlaboratory programs, the Cochran test is used to assess the repeatability of data sent by participants by evaluating the variance of the suspect data. Thus, $\boldsymbol{C}$ is calculated by equation 12, where $\boldsymbol{s_s^2}$ is the variance of the suspect data set, and $\boldsymbol{s_n^2}$ is the variance of all participating laboratories (75-77).

$$\boldsymbol{C = \frac{S_s^2}{\sum S_n^2}} \quad \textit{\textbf{eq. 12}}$$

When the ***C*** value obtained is greater than the tabulated critical value, it indicates that the results sent by the laboratory in question must be rejected (76).

3.3.3. The Z score test

The Z score also called the standard score, evaluates the results of each laboratory using a standard normal distribution. This test is calculated by equation 9 (can be seen in chapter 1 section 3.2), where x_i is the mean value presented by a single participant, μ and σ are the mean value and the standard deviation of the data series (75, 77).

The Z score is a dimensionless quantity, which has a positive or negative sign when the result of a participant is greater or less than the mean of the data (75).

The Z score classifies the performance of the laboratories according to the absolute values of $Z \leq 2$ as satisfactory, questionable for $2 < |Z| \leq 3$, and unsatisfactory for $|Z| \geq 3$ (75, 77, 78).

3.3.4. The robust Z-score

The robust Z-score (***R-Z-score***) is the most used and recommended test in interlaboratory programs. The calculation is performed by the median and the difference between the third (***Q3***) and first (***Q1***) quartiles of the data set. This fact removes the influence of the extreme values of the data series on the results found (51, 78, 79).

Equation 10 (can be seen in chapter 1 section 3.2) presents the calculation of this test, where x_i is the value indicated by the participant, ***med*** is the median of the data set, and the ***IQN*** is the normalized interquartile range. Normalization is obtained by the correction factor 0.7413 (51, 52, 77, 79).

3.4. Interlaboratory test uncertainty

For the calculation of uncertainty, Equation 8 was applied (as presented in Chapter 1, Section 2.5.3). In this context, s* represents the standard deviation obtained from the laboratories that were not excluded after the application of the robust Z-score test, and p corresponds to the number of laboratories used in the calculation of s*.

3.5. Evaluation of Material Stability

The stability of the candidate reference material was evaluated by monitoring the analyte concentrations at different storage times. Measurements were performed at three time points: at the beginning of the study ($t = 0$), after 3 months, and after 6 months. For each time point, three replicate determinations were carried out. The data obtained were then used to verify whether significant variations in the concentrations occurred during the evaluated period.

The assessment followed the statistical approach recommended in ISO Guide 35 (11), based on simple linear regression. In this model, the analyte concentration (y) is expressed as a function of time (x). The slope of the regression line (b_1) was tested against its uncertainty $s(b_1)$, using the criteria previously described in Chapter one Section 2.4.2. Briefly, the regression parameters were calculated according to Equations 3–5, where b_1 represents the rate of change of the concentration with time, and $s(b_1)$ reflects the variability of this estimate (80).

This procedure allows verifying whether the slope differs significantly from zero. If no significant trend is observed, the material can be considered stable for the evaluated period, and the associated uncertainty can be estimated from the standard deviation of the slope (80).

4. Results and discussions

4.1. interlaboratory program

The interlaboratory program received results from 29 laboratories of chemistry teaching and research institutions, and the chemical elements were determined according to the instrumentation available in each laboratory. At least one professor with a PhD in chemistry submitted the reports. The chemical elements determined were arsenic, cadmium, copper, manganese, magnesium, sodium, nickel, lead, antimony, selenium, strontium, and zinc.

All results processed by the interlaboratory program are composed of two averages obtained from triplicates of chemical analyses.

The results sent were summarized by each chemical element and evaluated using the following statistical tests: Grubbs test, Cochran test, Z-score test, robust Z-score, and confidence ellipse.

4.1.1. Results for lead

Of the 29 participating laboratories, 21 results were sent for the chemical element lead. The Grubs test, which can identify outliers, was applied to the forty-two data received. The highest G values found considering the duplicates were for laboratories 12, 28, and 21, which were (1.82), (2.29) and (1.76). Considering the critical G value of 2.58 for 21 laboratories with 95% confidence, it is concluded that there are no outliers among the forty averages received for lead (75, 76).

The Cochran test, which identifies outliers about the repeatability of the results of the averages presented by each laboratory, was also applied to the lead data. Laboratory 20 was investigated because it presented the highest variance value. Then, the relationship between the highest variance (2940) and the sum of the variances (8266) results in a C value of 0.356, which is a value lower than the critical C value of 0.389, indicating that the laboratory data have no outliers (75, 76).

The Z-score test makes it possible to classify the quality of the two averages presented by the laboratories. All forty-two means were evaluated, and only the two from laboratory 28 had Z values greater than 2.0, which classifies these data as questionable. All the other forty means had Z values less than 2.0, indicating that the results presented by the other twenty laboratories are acceptable.

The robust Z-score test was also applied to the 42 averages sent by the 21 laboratories. Of these, thirty-nine had Z values lower than 2.0, indicating that all these results are satisfactory. The two averages from Laboratory 28 and one from Laboratory 12 had Z values higher than 2.0, suggesting that these data are questionable. This result is partially compatible with the results found by the Z-score test, which indicated only the data from Laboratory 28 as questionable. This difference in the results of these two tests highlights the efficiency of the robust Z-score, which performs the evaluation using the median and quartiles, eliminating the effect of extreme values in the test calculation (51, 52).

Table 11 - Interlaboratory program - assessment of the results from lead

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z - Robust A	Z - Robust B	Test Z-A	Test Z-B
1	426.70	421.20	15.14	-0.61	-0.61	-0.79	-0.74	reliable	reliable
2	426.80	419.80	24.50	-0.61	-0.62	-0.79	-0.75	reliable	reliable
4	415.72	420.20	10.05	-0.70	-0.62	-0.88	-0.74	reliable	reliable
5	478.00	459.23	176.17	-0.19	-0.31	-0.33	-0.42	reliable	reliable
7	621.26	609.92	64.30	0.96	0.91	0.94	0.82	reliable	reliable
10	575.00	525.00	1250	0.59	0.22	0.53	0.12	reliable	reliable
11	366.00	374.00	32.00	-1.10	-0.99	-1.33	-1.13	reliable	reliable
12	269.10	274.60	15.13	-1.88	-1.79	-2.19	-1.95	questionable	reliable
13	515.30	510.40	12.00	0.11	0.10	0.00	0.00	reliable	reliable
16	456.00	441.00	112.50	-0.37	-0.45	-0.53	-0.57	reliable	reliable
17	560.17	561.82	1.36	0.47	0.52	0.40	0.42	reliable	reliable
18	621.76	595.52	344.27	0.97	0.79	0.95	0.70	reliable	reliable
20	644.65	721.33	2939.91	1.15	1.80	1.15	1.74	reliable	reliable
21	721.70	677.13	993.10	1.77	1.45	1.83	1.38	reliable	reliable
22	569.99	614.75	1001.73	0.55	0.95	0.49	0.86	reliable	reliable
23	550.00	565.00	112.50	0.39	0.54	0.31	0.45	reliable	reliable
25	468.92	465.85	4.71	-0.27	-0.25	-0.41	-0.37	reliable	reliable
26	578.98	541.17	714.93	0.62	0.35	0.57	0.25	reliable	reliable
27	479.72	446.82	541.28	-0.18	-0.41	-0.32	-0.52	reliable	reliable

28	210.00	215.00	12.50	-2.36	-2.27	-2.71	-2.44	questionable	questionable
29	584.00	585.00	0.50	0.66	0.71	0.61	0.62	reliable	reliable

4.1.2. Results for calcium

Seventeen laboratories of the twenty-nine participants submitted results of the calcium content of the soil sample of the interlaboratory program. The thirty-four averages were evaluated using the Grubbs test, and the highest G values were obtained for the samples from Laboratories 13 and 4, which were (1.61) and (1.87). This test indicated the absence of outliers among the analyzed data, considering that the critical G value for 17 laboratories is 2.47 (75, 76).

The Cochran test assesses the precision between the means of the participating laboratory. Laboratory 20 had the largest variance (128028) and the sum of all variances (298167), resulting in a C value of 0.420, lower than the critical C value of 0.434 for 17 laboratories. Therefore, the Cochran test did not exclude the data from Laboratory 20 as outliers (75, 76).

Seventeen laboratories were selected to estimate the certified calcium value. The Z-score test was applied to the thirty-four means, and all the results obtained were less than 2.0, indicating that the data were satisfactory for the final analysis. The highest Z-scores for laboratories 4 and 9 were (-1.72) and (-1.73), respectively.

Applying the robust Z-score to the calcium element data indicated that all two means from laboratories 4 and 9 showed questionable results, all with negative signs. Considering: (i) the robust Z-score test is established to minimize the effect of the data series's extreme values (maximum or minimum values). (ii) All four means in question have minimum values. It was decided not to consider the means of these two laboratories in calculating the certified calcium value (51, 52).

Table 12 - Interlaboratory program - assessment of the results from calcium

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z – Robust A	Z - Robust B	Test Z-A	Test Z-B
4	1822.57	1818.2	9.55	-1.72	-1.87	-2.21	-3.35	questionable	questionable
5	2045.87	1995.72	1257.64	-1.17	-1.42	-1.61	-2.69	reliable	questionable
6	2647.30	2698.30	1300.50	0.33	0.33	0.00	-0.07	reliable	reliable
8	2716.51	3051.84	56223.10	0.50	1.21	0.19	1.25	reliable	reliable
9	1876.55	1872.85	6.85	-1.59	-1.73	-2.06	-3.15	questionable	questionable
11	2335.00	2482.00	10804.50	-0.45	-0.21	-0.84	-0.88	reliable	reliable
13	3165.30	2855.70	47926.08	1.61	0.72	1.39	0.52	reliable	reliable
16	2966.00	2827.00	9660.50	1.12	0.65	0.85	0.41	reliable	reliable
17	2911.06	2917.88	23.26	0.98	0.88	0.71	0.75	reliable	reliable
20	2311.98	2828.06	133169.28	-0.51	0.65	-0.90	0.42	reliable	reliable
21	2871.84	2716.45	12073.07	0.88	0.37	0.60	0.00	reliable	reliable
22	2446.34	2483.70	697.88	-0.17	-0.21	-0.54	-0.87	reliable	reliable
24	2040.12	2049.35	42.60	-1.18	-1.29	-1.62	-2.49	reliable	questionable
25	2717.65	2773.10	1537.35	0.50	0.52	0.19	0.21	reliable	reliable
26	2817.00	3018.00	20200.50	0.75	1.13	0.45	1.13	reliable	reliable
27	2340.57	2465.98	7864.28	-0.44	-0.25	-0.82	-0.93	reliable	reliable
29	2744.00	2776.00	512.00	0.57	0.52	0.26	0.22	reliable	reliable

4.1.3. Results for manganese

The data analysis for the manganese element certification involved 16 laboratories among the 29 participants. Applying the Grubbs test indicated that the highest G values were for laboratory 1 (1.77) and laboratory 4 (1.72), which were lower than the critical G, which is 2.44. Thus, all the thirty-two averages were validated for certification by this test (75, 76).

The Cochran test assessed the repeatability between duplicates of manganese content means. The test was applied with the variance of Laboratory 27 (410) and the sum of the variances of (1197), resulting in a C value of (0.343), which is less than the critical C value of (0.452). This way, the 16 laboratories met the requirements in this test to be considered eligible for the certification of manganese. (75, 76).

The Z-score test was applied for all 16 laboratories, and the highest Z-score values were obtained for laboratory 1 and 4 with values of 1.76 and -1.72, all values were greater than -2 and less than 2, indicating that there were no outliers in the Z-score test.

The robust Z-score test was applied for the thirty-two means, and similarly to the standard Z-score analysis, all values fell within the acceptable range, indicating the absence of outliers. The highest robust Z-score values were observed for laboratory 1 (1.96) and laboratory 3 (1.91). (51, 52).

Table 13 - Interlaboratory program - assessment of the results from manganese

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z - Robust A	Z - Robust B	Test Z-A	Test Z-B
1	171.10	181.12	50.20	1.14	1.77	1.44	1.96	reliable	reliable
2	114.40	125.70	63.85	-1.60	-1.07	-1.69	-0.86	reliable	reliable
3	179.59	173.87	16.36	1.55	1.39	1.91	1.59	reliable	reliable
4	111.73	111.60	0.01	-1.73	-1.79	-1.84	-1.58	reliable	reliable
5	132.14	139.54	27.40	-0.74	-0.36	-0.71	-0.16	reliable	reliable
11	145.00	145.00	0.00	-0.12	-0.08	0.00	0.12	reliable	reliable
12	120.50	142.70	246.42	-1.30	-0.20	-1.35	0.01	reliable	reliable
13	140.30	136.30	8.00	-0.35	-0.53	-0.26	-0.32	reliable	reliable
14	144.97	140.86	8.45	-0.12	-0.30	0.00	-0.09	reliable	reliable
15	157.50	149.80	29.64	0.48	0.16	0.69	0.37	reliable	reliable
16	165.00	161.00	8.00	0.84	0.74	1.11	0.94	reliable	reliable
17	173.11	175.4	2.62	1.23	1.47	1.55	1.67	reliable	reliable
18	142.67	128.15	105.42	-0.24	-0.95	-0.13	-0.74	reliable	reliable
19	141.79	142.48	0.24	-0.28	-0.21	-0.18	-0.01	reliable	reliable
25	161.97	162.28	0.05	0.70	0.80	0.94	1.00	reliable	reliable
27	158.93	130.29	410.34	0.55	-0.84	0.77	-0.63	reliable	reliable

4.1.4. Results for cadmium

The first analysis of the cadmium data was performed with 18 participants using the Grubbs test. This test indicated that the two means sent by Laboratory 26 presented the highest G values: (3.39) and (1.64). The critical G value in this condition is 2.47, indicating that one of the means of this participant is an outlier and should therefore be excluded from the analysis set. Then, this test was reapplied to the remaining data, and no outliers were detected. The evaluation continued with 17 laboratories (75, 76).

Following, the Cochran test was applied to the variances obtained by the replicates of the means sent. The highest C value was attributed to Laboratory 28, which had a value of 0.858. Considering the critical C value of (0.434) for n equal to 17, it was concluded that the data from this participant are outliers and should be deleted from the data series. The reapplication of this test showed the replicates of laboratory 6 with a C value of (0.290) being lower than the critical C value for n of 16, which is 0.452. Therefore, the analysis continued with 16 laboratories (75, 76).

The application of the Z-score test showed that all thirty-two averages have Z values lower than 2.0, which guarantees the acceptability of the data, despite both results presented by Laboratory 15 having the highest Z-scores, which are 1.96 and 1.59.

The robust Z-score test results demonstrated that all thirty-two analyzed averages have values lower than 2.00. So, in agreement with the standard Z-score test, no outliers were identified. The highest robust Z-score values were 1.83 and 1.44, both from laboratory 15.

Table 14 - Interlaboratory program - assessment of the results from cadmium

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z - Robust A	Z - Robust B	Test Z-A	Test Z-B
1	34.56	35.72	0.68	0.48	1.01	0.62	0.73	reliable	reliable
4	25.00	24.90	0.01	-0.72	-0.87	-0.57	-0.67	reliable	reliable
5	24.74	24.89	0.01	-0.75	-0.87	-0.60	-0.68	reliable	reliable
6	22.30	25.60	5.45	-1.05	-0.75	-0.90	-0.58	reliable	reliable
7	24.82	24.89	0.00	-0.74	-0.87	-0.59	-0.68	reliable	reliable
8	31.10	33.8	3.64	0.05	0.68	0.19	0.48	reliable	reliable
10	30.20	30.10	0.00	-0.07	0.03	0.08	0.00	reliable	reliable
11	24.10	23.70	0.08	-0.83	-1.08	-0.68	-0.83	reliable	reliable
13	24.60	22.60	2.00	-0.77	-1.27	-0.62	-0.97	reliable	reliable
15	41.22	38.43	3.89	1.31	1.48	1.44	1.08	reliable	reliable
18	24.24	24.33	0.00	-0.81	-0.97	-0.66	-0.75	reliable	reliable
19	37.76	37.73	0.00	0.88	1.36	1.01	0.99	reliable	reliable
21	36.64	34.50	2.30	0.74	0.80	0.88	0.57	reliable	reliable
24	29.57	30.38	0.33	-0.15	0.08	0.00	0.04	reliable	reliable
25	35.50	35.31	0.02	0.59	0.94	0.74	0.68	reliable	reliable
27	24.13	24.45	0.05	-0.82	-0.95	-0.67	-0.73	reliable	reliable
28	52.00	37.00	112.50	2.65	1.23	2.78	0.89	questionable	reliable

4.1.5. Results for arsenic

The arsenic results were evaluated using data from sixteen laboratories. The presence of outliers among the thirty-two averages received was checked using the Grubbs test. Laboratories 20 and 21 presented the highest G values, (1.96) and (1.70), both values being lower than the critical G value (2.44 for n = 16), indicating the absence of outliers in the data provided for certification (75, 76).

The agreement between the 32 means presented by the sixteen participants was investigated using the Cochran test. The C value calculated from the largest variance between the data was 0.389 for laboratory 18, being lower than the critical C value (0.452) for n equal to 16, indicating the absence of outliers among the data (75, 76).

The results from the Z-score test showed Z values lower than 2.0 for all thirty-two averages presented by the laboratories, confirming the data's acceptability for the certification process.

The application of the robust Z-score test to the arsenic analytical data revealed values ranging from (-0.99) to (1.71) for replicates-A, and from (-1.33) to (1.34) for replicates-B. These results indicate that the data can be reliably utilized to calculate the certified arsenic value in the analyzed soil matrix (52).

Table 15 - Interlaboratory program - assessment of the results from arsenic

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z - Robust A	Z - Robust B	Test Z-A	Test Z-B
1	52.85	52.32	0.14	0.08	0.04	0.06	0.00	reliable	reliable
4	37.58	36.40	0.69	-0.76	-0.85	-0.69	-0.72	reliable	reliable
5	31.54	31.40	0.01	-1.10	-1.12	-0.99	-0.94	reliable	reliable
6	43.19	52.30	41.50	-0.45	0.04	-0.42	0.00	reliable	reliable
8	71.40	70.20	0.72	1.11	1.03	0.97	0.81	reliable	reliable
11	25.70	28.30	3.38	-1.42	-1.30	-1.28	-1.08	reliable	reliable
12	27.40	22.67	11.19	-1.33	-1.61	-1.20	-1.34	reliable	reliable
13	52.30	50.06	2.50	0.05	-0.09	0.03	-0.10	reliable	reliable
14	32.13	33.37	0.77	-1.07	-1.02	-0.96	-0.86	reliable	reliable
16	63.60	53.80	48.02	0.68	0.12	0.58	0.07	reliable	reliable
18	50.41	64.43	98.28	-0.05	0.71	-0.06	0.55	reliable	reliable
20	86.68	78.65	32.24	1.96	1.50	1.72	1.19	reliable	reliable
21	76.99	82.12	13.21	1.42	1.69	1.24	1.35	reliable	reliable
24	51.12	54.45	5.54	-0.01	0.16	-0.03	0.10	reliable	reliable
28	64.00	69.00	12.5	0.70	0.96	0.60	0.75	reliable	reliable
29	55.00	47.00	32.00	0.20	-0.26	0.16	-0.24	reliable	reliable

4.1.6. Results for cooper

The Cochran test evaluated the variances found for the copper replicates and revealed that the data from Laboratory 26 with C (0.51) is an outlier in the data set, because this value is greater than the critical C (0.452) for an n value of 16. This way, the data from this participant was deleted and a new test was applied, generating the highest C value of 0.340 (laboratory 25) which is less than the critical C (0.471) for n equal to 15. Thus, all data from the 15 participants will be used to validate the copper content in the soil matrix (75, 76).

The Grubbs test was used to identify outliers in the individual averages presented by the laboratories. The highest G values were found for Laboratory 4 (1.92) and (1.72), which are lower than the critical G of 2.42 for n equal to 15, indicating that the entire set of data evaluated is satisfactory for copper certification (75, 76).

The Z-score test demonstrated that, except for the results from Laboratory 26, all data submitted to validate the copper element are acceptable. Laboratory 4 presented the highest values (-1.93) and (-1.72), with values lower than 2.0.

The results obtained in applying the robust Z-score were entirely compatible with the results of the Z-score. All 30 averages evaluated are acceptable for determining the certified value of copper in the laboratory reference material of soil (52).

Table 16 - Interlaboratory program - assessment of the results from copper

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z - Robust A	Z - Robust B	Test Z-A	Test Z-B
1	29.04	29.03	0.00	-0.79	-0.83	-0.77	-0.86	reliable	reliable
2	28.90	28.40	0.13	-0.82	-1.01	-0.81	-1.02	reliable	reliable
3	32.92	34.21	0.83	0.17	0.66	0.33	0.46	reliable	reliable
4	26.17	26.10	0.00	-1.50	-1.67	-1.59	-1.61	reliable	reliable
5	30.40	29.05	0.92	-0.45	-0.82	-0.38	-0.86	reliable	reliable
8	35.3	36.00	0.25	0.76	1.18	1.01	0.92	reliable	reliable
9	32.53	32.88	0.06	0.08	0.28	0.22	0.12	reliable	reliable
11	35.00	35.30	0.04	0.69	0.98	0.93	0.74	reliable	reliable
12	34.15	31.92	2.49	0.48	0.00	0.68	-0.12	reliable	reliable
13	35.1	34.7	0.08	0.71	0.80	0.95	0.59	reliable	reliable
14	29.81	29.4	0.08	-0.60	-0.72	-0.55	-0.77	reliable	reliable
24	32.5	31.6	0.40	0.07	-0.09	0.21	-0.20	reliable	reliable
25	30.56	33.50	4.32	-0.41	0.46	-0.34	0.28	reliable	reliable
26	43.6	38.40	13.52	2.82	1.87	3.38	1.53	questionable	reliable
27	28.59	26.98	1.29	-0.90	-1.42	-0.90	-1.39	reliable	reliable
28	31.00	33.00	2.00	-0.30	0.32	-0.21	0.15	reliable	reliable

4.1.7. Results for Zinc

The evaluation of the analytical data for the zinc element using the Cochran test showed that all results are acceptable, considering the highest C value for Laboratory 27 (0.416) and the critical C value of 0.452 for an n value equal to 16. However, the high median values observed in Laboratories 26, 27, and 20 stand out (75, 76).

The Grubbs test was applied to the thirty-two means to detect possible outliers. The highest G values found were (1,78) and (1.65) both from Laboratory 9. Considering the critical G value of 2.44 for n equal to 16, it is concluded that there are no outliers among the zinc means analyzed (75, 76).

The Z-score test was applied to the thirty-two means received to determine the zinc content in the material. The results showed that the Z-score values ranged from -1.89 to 1.77, all below the threshold of 2.0, indicating that no questionable values were identified in the study.

Applying the robust Z-score to the analytical data indicated the acceptability of all 16 Laboratories evaluated, with robust Z-score variations from -1.68 to 1.36 for one set of means and from -1.94 to 1.75 for the other (52).

Table 17 - : Interlaboratory program - assessment of the results from zinc

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z - Robust A	Z - Robust B	Test Z-A	Test Z-B
1	264.68	254.79	48.92	-0.71	-0.92	-0.71	-0.78	reliable	reliable
3	251.26	278.97	383.92	-1.03	-0.32	-0.99	-0.28	reliable	reliable
8	332.39	335.71	5.51	0.89	1.10	0.69	0.89	reliable	reliable
9	364.72	362.89	1.67	1.65	1.78	1.36	1.45	reliable	reliable
11	265.00	258.00	24.5	-0.71	-0.84	-0.70	-0.71	reliable	reliable
13	318.1	322.3	8.82	0.55	0.77	0.40	0.61	reliable	reliable
14	221.72	220.71	0.51	-1.73	-1.77	-1.60	-1.48	reliable	reliable
15	300.00	307.7	29.64	0.12	0.40	0.02	0.31	reliable	reliable
16	298.00	288.00	50.00	0.08	-0.09	-0.02	-0.09	reliable	reliable
17	284.25	275.23	40.68	-0.25	-0.41	-0.31	-0.36	reliable	reliable
18	321.98	321.95	0.00	0.64	0.76	0.48	0.60	reliable	reliable
20	263.12	297.08	576.64	-0.75	0.14	-0.74	0.09	reliable	reliable
21	229.11	236.7	28.82	-1.55	-1.37	-1.45	-1.15	reliable	reliable
22	328.56	324.69	7.49	0.80	0.83	0.61	0.66	reliable	reliable
23	338.00	328.00	50.00	1.02	0.91	0.81	0.73	reliable	reliable
27	336.21	253.85	3391.49	0.98	-0.95	0.77	-0.80	reliable	reliable

4.1.8. Results for nickel

The Cochran technique was applied to the variances found for nickel levels in the soil. The highest C value (0.334) was observed for Laboratory 11. Considering the C critical of 0.515 for n equal to 13, it was concluded that there are no outliers among the data presented by the laboratories (75, 76).

The Grubbs test evaluated the averages with nickel contents sent by the participants. Laboratories 26 (1.81) and 25 (2.29) presented the highest G values for each data series, indicating that no data should be rejected, considering the critical G value of 2.33 for 13 laboratories.

Applying the Z-score test to the nickel data revealed that the means from Laboratory 25 and another from Laboratory 26, with Z values of (2.29) and (2.03), should be deleted from the data set because they were greater than 2.0. Therefore, the data from these participants were removed, and the validation process continued with 11 laboratories.

With the data from laboratories 25 and 26 removed, the robust Z-score test was applied to the remaining 11 participants. The results showed a robust Z-score value of 2.30 for one of the replicates from Laboratory 5, indicating the rejection of the data from this participant. A new analysis using this test revealed that the results from the remaining ten laboratories are satisfactory for determining the certified value of nickel in the soil matrix (52).

Table 18 - Interlaboratory program - assessment of the results from nickel

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z - Robust A	Z - Robust B	Test Z-A	Test Z-B
1	10.28	10.68	0.08	-0.89	-0.61	-0.70	-1.13	reliable	reliable
4	11.8	11.2	0.18	-0.02	-0.29	0.00	0.00	reliable	reliable
5	13.29	12.11	0.69	0.83	0.28	0.69	1.98	reliable	reliable
6	12.20	10.60	1.28	0.21	-0.66	0.19	-1.30	reliable	reliable
8	10.63	11.27	0.20	-0.68	-0.24	-0.54	0.15	reliable	reliable
11	13.20	11.30	1.81	0.79	-0.23	0.65	0.22	reliable	reliable
12	10.06	11.01	0.45	-1.01	-0.41	-0.80	-0.41	reliable	reliable
13	10.9	11.3	0.08	-0.53	-0.23	-0.41	0.22	reliable	reliable
16	10.1	10.4	0.05	-0.99	-0.79	-0.78	-1.74	reliable	reliable
20	9.76	10.49	0.27	-1.18	-0.73	-0.94	-1.54	reliable	reliable
25	14.54	15.33	0.31	1.55	2.29	1.27	8.98	reliable	questionable
26	15.00	14.9	0.00	1.81	2.03	1.48	8.05	reliable	questionable
28	12.00	11.00	0.50	0.10	-0.41	0.09	-0.43	reliable	reliable

4.1.9. Results for magnesium

Of the twenty-eight participating laboratories, only thirteen submitted results for the chemical element magnesium, yielding twenty-six mean values (considering duplicate measurements). The Grubbs test was applied to assess the presence of potential outliers. The highest G values were observed for laboratories 25 and 26, with values of 2.16 and 2.01, respectively. Given that the critical G value for thirteen laboratories at a 95% confidence level is 2.33, it was concluded that none of the results could be classified as outliers (75, 76).

The Cochran test was also performed to evaluate the consistency of repeatability across laboratories. Laboratory 25 exhibited the highest variance (208.41). When this value was divided by the total sum of variances (414.172), the resulting C value was 0.503, below the critical threshold of 0.515. This result indicates that the variance reported by laboratory 25 does not represent a significant deviation and should not be considered an outlier (75, 76).

The Z-score test was applied to evaluate the quality of the submitted results further. Among the 26 values analyzed, only one from laboratory 25 and one from laboratory 26 had Z-scores greater than 2.0, classifying them as questionable. The remaining 24 results had Z-scores below this threshold, suggesting that the outcomes from the other 11 laboratories are within acceptable limits.

After excluding laboratories 25 and 26 from the analysis, a new evaluation of the robust Z-scores was conducted for the remaining 11 laboratories. In this second assessment, laboratory 16 also exhibited questionable robust Z-score values and was subsequently excluded. As a result, the robust Z-scores of the remaining 10 laboratories ranged from -2 to 2 and were considered acceptable for calculating the certified magnesium value (51, 52).

Table 19 - Interlaboratory program - assessment of the results from magnesium

Laboratory	Replicate A	Replicate B	Variance	Score Z - A	Score Z - B	Z - Robust A	Z - Robust B	Test Z-A	Test Z-B
1	987.23	1005.09	159.51	-0.49	-0.41	-0.24	-0.21	reliable	reliable
3	1330.00	1280.00	1250.00	0.76	0.47	1.11	1.07	reliable	reliable
4	923.37	920.10	5.34	-0.72	-0.68	-0.49	-0.61	reliable	reliable
5	1026.11	990.59	630.92	-0.35	-0.46	-0.09	-0.28	reliable	reliable
11	1154.78	1150.49	9.20	0.12	0.06	0.42	0.47	reliable	reliable
14	792.33	767.29	313.50	-1.20	-1.17	-1.01	-1.32	reliable	reliable
15	994.60	1050.80	1579.22	-0.46	-0.26	-0.21	0.00	reliable	reliable
16	1472.00	1405.00	2244.5	1.27	0.87	1.67	1.65	reliable	reliable
22	1072.72	1040.67	513.60	-0.18	-0.30	0.09	-0.05	reliable	reliable
23	1391.00	1191.00	20000.00	0.98	0.19	1.35	0.65	reliable	reliable
25	1048.70	1808.19	288412.53	-0.27	2.16	0.00	3.54	reliable	questionable
26	1675.00	1493.00	16562.00	2.01	1.15	2.47	2.06	questionable	questionable
27	720.90	629.20	4204.16	-1.46	-1.61	-1.29	-1.97	reliable	reliable

4.2. Assessment of interlaboratory mean and uncertainty

Table 20 presents the mean concentration values reported by the laboratories participating in the interlaboratory program proposed for the production of the soil reference material. In this table, s* represents the robust standard deviation, while u_{char} corresponds to the characterization uncertainty associated with the mean.

Table 20 - Mean concentrations and characterization uncertainty of the proposed soil reference material

Element	Mean	S*	u_{char}
Pb	526,73	92,84	26,62
Ca	2541,35	390,53	118,40
Mn	147,09	19,32	6,04
Cd	30,32	6,69	2,03
As	51,51	17,75	5,55
Cu	32,06	3,65	1,14
Zn	293,24	39,25	12,27
Ni	11,74	1,60	0,56
Mg	1127,70	269,25	93,34

u_{char} = standard uncertainty due to the characterization

S* = Robust standard deviation

4.2. Stability study results

Table 21 presents the regression parameters and uncertainty values obtained for the stability evaluation of the soil reference material. For each analyte, the intercept (b_0), slope, residual standard deviation (s), and the standard error of the slope [$s(b_1)$] were calculated. The term ($t_{0.95,n-2}$) × $s(b_1)$ corresponds to the critical value used to test whether the slope is significantly different from zero. The parameter u_{lts} represents the uncertainty associated with the long-term stability study. Equations 3, 4, 5, and 7 were used to calculate the parameters presented in Table 21.

$$b_0 = \bar{y} - \overline{b_1}\bar{x} \quad \textbf{\textit{eq. 3}}$$

$$s(b_1) = \frac{s}{\sqrt{(\Sigma^{\mathrm{n}}{}_{\mathrm{i}=1}\,(x_{\mathrm{i}} - \bar{x})^2\,)}} \qquad \textbf{\textit{eq. 4}}$$

$$s^2 = \frac{\Sigma^{\mathrm{n}}{}_{\mathrm{i}=1}\,(y_{\mathrm{i}} - b_0 - b_1 x_{\mathrm{i}})^2}{(n-2)} \qquad \textbf{\textit{eq. 5}}$$

$$u_{lts} = s(b_1)\,.\,t_{(0.95,n-1)} \qquad \textbf{\textit{eq. 7}}$$

As shown in the “Response” column, all analytes were classified as *stable*, since the calculated slopes were not significantly different from zero within the studied period. This indicates that the material maintained its chemical stability under the evaluated storage conditions, ensuring its suitability as a reference material.

Table 21 - Stability evaluation parameters and long-term stability uncertainty (u_{lts})

	b_0	**Slope**	**s**	**s(b1)**	**($t_{0.95,n-2}$) * s(b1)**	**Response**	**u_{lts}**
As	y = -0,1239x + 52,671	0,12	1,09	0,26	3,25	stable	1,54
Ca	y = 32,627x + 1143,4	32,63	59,50	14,03	178,20	stable	84,15
Cd	y = -0,7678x + 36,646	0,77	1,18	0,28	3,54	stable	1,67
Cu	y = 0,2805x + 27,501	0,28	0,78	0,18	2,33	stable	1,10
Mg	y = 25,503x + 862,54	25,50	80,80	19,04	241,98	stable	114,27
Mn	y = 4,4233x + 165,34	4,42	1,62	0,38	4,85	stable	2,29
Ni	y = 0,2592x + 9,0388	0,26	1,38	0,33	4,14	stable	1,95
Pb	y = 4,1325x + 420,56	4,13	7,74	1,82	23,17	stable	10,94
Zn	y = 3,6321x + 252,95	3,63	8,01	1,89	23,99	stable	11,33

($t_{0.95,n-2}$) = 12,706; $\sqrt{\sum(xi - x^{-})^2}$ = 4.243

5. Conclusion

The interlaboratory study and stability evaluation allowed for a comprehensive assessment of the candidate soil reference material for inorganic analysis. Data from 29 participating laboratories confirmed its applicability to different analytical techniques, with results carefully examined through statistical approaches such as Grubbs, Cochran, Z-score, and robust Z-score tests. These methods ensured the identification of outliers and reinforced the reliability of the final dataset, supporting the accuracy and comparability of the certified values.

Certified concentrations were assigned for eight environmentally and toxicologically relevant elements - Pb, Ca, Mn, Cd, As, Cu, Zn, and Ni - together with associated uncertainties derived from homogeneity, characterization, and stability studies. The stability assessment, in particular, confirmed that the material preserved its chemical integrity throughout the evaluated storage period, demonstrating its suitability for long-term use as a reference material.

Taken together, the results show that the developed soil candidate successfully meets the essential requirements of homogeneity, stability, and traceability, in accordance with international recommendations for reference material production.

References

1. Adams F, Adriaens M. The metamorphosis of analytical chemistry. Analytical and Bioanalytical Chemistry. 2020;412(15):3525-37.
2. Rosenberg E, Krska R. Analytical chemistry in front of the curtain! Analytical and bioanalytical chemistry. 2024;416(8):1787-95.
3. Guide I. 30: 1992. Terms and definitions used in connection with reference materials. 2008.
4. Quesada C, Lloyd J, Anderson L, Fyllas N, Schwarz M, Czimczik C. Soils of Amazonia with particular reference to the RAINFOR sites. Biogeosciences. 2011;8(6):1415-40.
5. Kuselman I, Pavlichenko M. Designs of experiment for proficiency testing with a limited number of participants. Accreditation and quality assurance. 2004;9(7):387-90.
6. Smith P, Cotrufo M, Rumpel C, Paustian K, Kuikman P, Elliott J, et al. Biogeochemical cycles and biodiversity as key drivers of ecosystem services provided by soils. Soil. 2015;1(2):665-85.
7. Piekut A, Baranowska R, Marchwińska-Wyrwał E, Ćwieląg-Drabek M, Hajok I, Dziubanek G, et al. Is the soil quality monitoring an effective tool in consumers' protection of agricultural crops from cadmium soil contamination?—a case of the Silesia region (Poland). Environmental Monitoring and Assessment. 2018;190(1):25.
8. Ayamba BE, Abaidoo RC, Opoku A, Ewusi-Mensah N. Mechanisms for nutrient interactions from organic amendments and mineral fertilizer inputs under cropping systems: a review. PeerJ. 2023;11:e15135.
9. Jaiswal A, Verma A, Jaiswal P. Detrimental effects of heavy metals in soil, plants, and aquatic ecosystems and in humans. Journal of Environmental Pathology, Toxicology and Oncology. 2018;37(3).
10. Guide I. General requirements for the competence of reference material producers. ISO: Geneva, Switzerland. 2009.
11. Guide I. 35: 2017 Reference materials—Guidance for characterization and assessment of homogeneity and stability. ISO [website] URL: https://www iso org/ru/standard/60281 html. 2017.
12. Emons H. The 'RM family'—Identification of all of its members. Accreditation and Quality Assurance. 2006;10:690-1.
13. Azevedo RS, Dos Santos LO, Chagas AV, Felix CS, de Andrade JB, Melo VS, et al. Preparation and characterization of new reference material for inorganic analysis of pumpkin seed flour–An interlaboratory program. Food Chemistry. 2025;477:143501.
14. Rutkowska M, Kochańska K, Kandel L, Bajger-Nowak G, Słomińska M, Marć M, et al. Homogeneity study of candidate reference material (contaminated soil) based on determination of selected metals, PCBs and PAHs. Measurement. 2018;128:1-12.
15. Sainio P, Mäkinen I, Hutter J-W, den Ouden T, Krysell M. Feasibility study for the preparation of certified reference materials part I–mineral oil contaminated soils. Accreditation and quality assurance. 2006;11(3):116-21.
16. Pershina MV, Asnin LD, Tselishchev YG. Short-range spatial variations in chemico-mineralogical composition, morphological and adsorption properties of clayey soil, Middle Urals. Environmental Earth Sciences. 2025;84(16):1-12.
17. Lucena ICd, Amorim RSS, Lobo FdA, Baldoni RN, Matos DMdS. Spatial heterogeneity of soils of the Cerrado-Pantanal ecotone. Revista Ciência Agronômica. 2014;45:673-82.
18. Guerreiro QLdM, Oliveira Júnior RCd, Santos GRd, Ruivo MdLP, Beldini TP, Carvalho EJM, et al. Spatial variability of soil physical and chemical aspects in a Brazil nut tree stand in the Brazilian Amazon. African Journal of Agricultural Research. 2017.

19. Barnekow U, Fesenko S, Kashparov V, Kis-Benedek G, Matisoff G, Onda Y, et al. Guidelines on soil and vegetation sampling for radiological monitoring. 2019.
20. Barth DS, Mason BJ. Soil sampling quality assurance user's guide. Nevada Univ., Las Vegas (USA); 1984.
21. Forster J. Soil sampling, handling, storage and analysis. Methods in applied soil microbiology and biochemistry: Elsevier; 1995. p. 49-121.
22. Tholen DW. ISO/IEC 17043: the new International Standard for proficiency testing. Accreditation and quality assurance. 2008;13(12):727-30.
23. Guide I. 31 (2000) Reference materials—contents of certificates and labels. ISO, Geneva. 2000.
24. Standardization IOf. ISO Guide 32:* calibration in Analytical Chemistry and Use of Certified Reference Materials: International organization for standardization; 1997.
25. Normalización OId. ISO Guide 33: Reference Materials--Good Practice in Using Reference Materials: ISO; 2015.
26. Standardization IOf. ISO 13528: Statistical methods for use in proficiency testing by interlaboratory comparisons: ISO; 2005.
27. Metrology JCfGi. Evaluation of measurement data—Guide to the expression of uncertainty in measurement. JCGM. 2008;100(2008):1-116.
28. Chagas AV, Felix CS, Dantas SV, Lima AS, Cunha FA, Ferreira MM, et al. Use of multi-element robust Z-score and DD-SIMCA methods in the homogeneity study for preparation of a soil laboratory reference material. Microchemical Journal. 2025:114783.
29. Castro L, Moreira EG, Vasconcellos MB, Lange CN, Pedron T, Batista BL. Stability assessment of a bovine kidney reference-material candidate. Journal of radioanalytical and nuclear chemistry. 2018;317(2):1133-9.
30. Thompson M, Ellison SL, Wood R. The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report). Pure and Applied Chemistry. 2006;78(1):145-96.
31. ISO I. 13528: 2005 Statistical Methods for Use in Proficiency Testing by Inter-laboratory Comparison. International Organization for Standardization (ISO), Geneva. 2005.
32. Pommé S, Spasova Y. A practical procedure for assigning a reference value and uncertainty in the frame of an interlaboratory comparison. Accreditation and quality assurance. 2008;13(2):83-9.
33. Leys C. Taking parametric assumptions seriously: Arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. International Review of Social Psychology. 2019.
34. Linsinger TP, Pauwels J, van der Veen AM, Schimmel H, Lamberty A. Homogeneity and stability of reference materials. Accreditation and Quality Assurance. 2001;6(1):20-5.
35. Sarembaud J, Pinto R, Rutledge DN, Feinberg M. Application of the ANOVA-PCA method to stability studies of reference materials. Analytica chimica acta. 2007;603(2):147-54.
36. Dos Santos LO, Dos Santos AM, Ferreira MM, Ferreira SL, Nepomuceno AFS. The use of ANOVA-PCA and DD-SIMCA in the development of corn flour laboratory reference materials. Food Chemistry. 2022;367:130748.
37. Gomes CA, da Silva CJM, Carneiro MTW, de Souza JR, de Almeida CMS. DD-SIMCA as an alternative tool to assess the short-term stability of a marine sediment reference material candidate. Chemometrics and Intelligent Laboratory Systems. 2025;257:105312.
38. No A, Committee AM. z-Scores and other scores in chemical proficiency testing—their meanings, and some common misconceptions. Analytical Methods. 2016;8(28):5553-5.
39. Kucheryavskiy S, Rodionova O, Pomerantsev A. A comprehensive tutorial on Data-Driven SIMCA: Theory and implementation in web. Journal of Chemometrics. 2024;38(7):e3556.
40. Grassi S, Tarapoulouzi M, D'Alessandro A, Agriopoulou S, Strani L, Varzakas T. How chemometrics can fight milk adulteration. Foods. 2022;12(1):139.
41. De Angelis D, Summo C, Pasqualone A, Faccia M, Squeo G. Advancements in food authentication using soft independent modelling of class analogy (SIMCA): a review. Food Quality and Safety. 2024;8:fyae032.

42. Padariya C, Rutkowska M, Konieczka P. The importance and availability of marine certified reference materials. Critical Reviews in Environmental Science and Technology. 2022;52(18):3322-73.

43. Ellison SL, Botha A. Principles for the assessment of homogeneity and stability in the new ISO Guide 35: 2017. Accreditation and Quality Assurance. 2018;23(1):47-51.

44. Wold S. Pattern recognition by means of disjoint principal components models. Pattern recognition. 1976;8(3):127-39.

45. Pomerantsev AL, Rodionova OY. Popular decision rules in SIMCA: Critical review. Journal of Chemometrics. 2020;34(8):e3250.

46. Zontov Y, Rodionova OY, Kucheryavskiy SV, Pomerantsev A. DD-SIMCA–a MATLAB GUI tool for data driven SIMCA approach. Chemometrics and Intelligent Laboratory Systems. 2017;167:23-8.

47. Azevedo RSDA, Neto JHS, Dos Santos LO, Gomes AD, Santana DD, Dos Santos AMP. Homogeneity and stability assessment of a candidate to pumpkin seed flour reference material by means of computer vision based chemometrics assisted approach. Food Chemistry. 2022;368:130842.

48. dos Santos AMP, dos Santos LO, Brandao GC, Araujo RGO, de Santana FA, Lemos VA, et al. Preparation and characterization of a new reference material for the inorganic analysis of corn flour. Accreditation and Quality Assurance. 2017;22(1):37-43.

49. Azevedo RSDA, dos Santos LO, Chagas AVB, Felix CSA, de Andrade JB, Melo VSC, et al. Preparation and characterization of new reference material for inorganic analysis of pumpkin seed flour - An interlaboratory program. Food Chemistry. 2025;477:143501.

50. Tripathy SS. Comparison of statistical methods for outlier detection in proficiency testing data on analysis of lead in aqueous solution. American Journal of Theoretical and Applied Statistics. 2013;2(6):233.

51. Gupta KK, Bhattacharjee S, Kar S, Chakrabarty S, Thakur P, Bhattacharyya G, et al. Mineral compositions of eight common spices. Communications in Soil Science and Plant Analysis. 2003;34(5-6):681-93.

52. Zhou Q, Hu JH, Li XP, Li SN, Gao ZX, Xu JP, et al. Construction and application of the robust Youden plot in an EQA program. Accreditation and Quality Assurance. 2015;20(3):195-201.

53. Portugal LA, Ferreira HS, dos Santos WNL, Ferreira SLC. Simultaneous pre-concentration procedure for the determination of cadmium and lead in drinking water employing sequential multi-element flame atomic absorption spectrometry. Microchemical Journal. 2007;87(1):77-80.

54. Novaes CG, Ferreira SLC, Neto JHS, de Santana FA, Portugal LA, Goicoechea HC. A Multiple Response Function for Optimization of Analytical Strategies Involving Multi-elemental Determination. Current Analytical Chemistry. 2016;12(2):94-101.

55. Ferreira SLC, Lima CC, Garcia RJL, da Silva JB, Jr., Coutinho JJ, Garcia KS, et al. Risk Assessment of Toxic Elements in Surface Soils Collected Near and Far from a Deactivated Lead Smelter in Bahia, Brazil. Analytical Letters. 2025;58(2):285-98.

56. Santos AS, dos Santos LO, dos Santos IF, dos Santos KR, Ferreira SLC. Application of chemometric tools for homogeneity and stability evaluation during the preparation of a powdered milk laboratory reference material for inorganic analysis. Analytical Methods. 2020;12(8):1055-63.

57. Kennard RW, Stone LA. Computer aided design of experiments. Technometrics. 1969;11(1):137-48.

58. Rodionova OY, Pomerantsev AL. Subset selection strategy. Journal of Chemometrics: A Journal of the Chemometrics Society. 2008;22(11-12):674-85.

59. Kuselman I, Fajgelj A. IUPAC/CITAC Guide: Selection and use of proficiency testing schemes for a limited number of participants chemical analytical laboratories (IUPAC Technical Report). Pure and Applied Chemistry. 2010;82(5):1099-135.

60. Thompson M, Ellison SL, Wood R. Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC Technical Report). Pure and applied chemistry. 2002;74(5):835-55.

61. Miller J, Miller JC. Statistics and chemometrics for analytical chemistry: Pearson education; 2018.
62. de la Guardia M, Garrigues S. The social responsibility of environmental analysis. Trends in Environmental Analytical Chemistry. 2014;3:7-13.
63. Butcher DJ. Recent advances in graphite furnace atomic absorption spectrometry: a review of fundamentals and applications. Applied Spectroscopy Reviews. 2024;59(2):247-75.
64. Zhang Z, Dong JH, Feng L, Chen LF, Xing PJ, Du YY, et al. Rapid and sensitive determination of trace Cd in carbonate samples using direct solid sampling electrothermal vaporization atomic absorption spectrometry. Microchemical Journal. 2024;201.
65. Silva AS, Brandao GC, Ferreira SLC, dos Santos AMP. Sequential and Simultaneous Determination of Cd, Fe and Ni in Toothpastes Employing Slurry Sampling High-Resolution Continuum Source Graphite Furnace Atomic Absorption Spectrometry. Analytical Letters. 2022;55(8):1192-206.
66. Kunpatee K, Kusonpan P, Chaiyo S, Chailapakul O, Jitcharoen J, Kalcher K, et al. An electrochemical sensor utilizing zeolite synergistically combined with copper(I) oxide and bismuth(III) oxide for the sensitive and selective determination of paraquat. Microchemical Journal. 2025;212.
67. Korolczuk M, Ochab M, Geca I. Anodic Stripping Voltammetric Procedure of Thallium(I) Determination by Means of a Bismuth-Plated Gold-Based Microelectrode Array. Sensors. 2024;24(4).
68. Yildiz D, Atasoy M, Demirak A. Ultra-Trace Determination of Arsenic by Platinum-Coated Tungsten Coil Hydride Generation (HG) Atomic Absorption Spectrometry (AAS). Analytical Letters. 2022;55(17):2776-88.
69. Yamakawa A, Nagano K, Onishi K, Ukachi M, Ieda T, Kobayashi M, et al. Human hair certified reference material for total mercury, methylmercury, and trace element analyses. Accreditation and Quality Assurance. 2025;30(3):277-90.
70. Chen YH, Qiu XL, Feng C, Xu Q, Lin YJ, Le SY, et al. Preparation of a reference material for tea containing five pesticide residues and its evaluation in an interlaboratory comparison study in China. Accreditation and Quality Assurance. 2022;27(2):93-101.
71. Santos AS, dos Santos LO, Pereira JB, de Oliveira FM, Ferreira SLC. Development of reference material from powdered milk: Uncertainties and interlaboratory evaluation through confidence ellipses. Microchemical Journal. 2020;159.
72. dos Santos LO, dos Santos AMP, Ferreira MMC, Ferreira SLC, Nepomuceno AFSF. The use of ANOVA-PCA and DD-SIMCA in the development of corn flour laboratory reference materials. Food Chemistry. 2022;367:130748.
73. Ferreira SLC, Lima CC, Garcia RJL, da Silva Júnior JB, Coutinho JJ, Garcia KS, et al. Risk Assessment of Toxic Elements in Surface Soils Collected Near and Far from a Deactivated Lead Smelter in Bahia, Brazil. Analytical Letters. 2025;58(2):285-98.
74. Ferreira SLC, Teixeira LSG, Ferreira WQ, Dantas SVA, Lima AS, Korn MGA, et al. Determination and assessment of contamination of toxic chemical elements in soils–review. TrAC Trends in Analytical Chemistry. 2024:118055.
75. Dias RD, de Abreu CA, de Abreu MF, González AP. Statistical Methods for Evaluating Results from Soil Micronutrient Analyses in Interlaboratory Programs. Communications in Soil Science and Plant Analysis. 2015;46:57-71.
76. Miller JN. Using the Grubbs and Cochran tests to identify outliers. Analytical Methods. 2015;7(19):7948-50.
77. Miller JN, Miller JC. Statistics and chemometrics for analytical chemistry. Sixth ed2010. 278 p.
78. Tripathy SS, Saxena RK, Gupta PK. Comparison of statistical methods for outlier detection in proficiency testing data on analysis of lead in aqueous solution. American Journal of Theoretical and Applied Statistics. 2013;2(6):233.
79. Bhattacharjee S, Chakravarty S, Maity S, Dureja V, Gupta KK. Metal contents in the groundwater of Sahebgunj district, Jharkhand, India, with special reference to arsenic. Chemosphere. 2005;58(9):1203-17.

80. SANTOS ADS. Produção e caracterização de um material de referência de laboratório a partir do leite em pó: requisito de qualidade para análise inorgânica. Salvador: Federal University of Bahia; 2020.