





## Research Article

# Statistically Coherent Calibration of X-Ray Fluorescence Spectrometry for Major Elements in Rocks and Minerals

**Surendra P. Verma** <sup>1</sup>, **Sanjeet K. Verma** <sup>2</sup>, **M. Abdelaly Rivera-Gómez**,<sup>3</sup>  
**Darío Torres-Sánchez** <sup>4</sup>, **Lorena Díaz-González**,<sup>5</sup> **Alejandra Amezcua-Valdez**,<sup>6</sup>  
**Beatriz Adriana Rivera-Escoto**,<sup>7</sup> **Mauricio Rosales-Rivera** <sup>8</sup>, **John S. Armstrong-Altrin**,<sup>3</sup>  
**Héctor López-Loera**,<sup>2</sup> **Fernando Velasco-Tapia**,<sup>9</sup> and **Kailasa Pandarinath**<sup>1</sup>

<sup>1</sup>Instituto de Energías Renovables, Universidad Nacional Autónoma de México, Temixco, Mor 62580, Mexico

<sup>2</sup>División de Geociencias, Instituto Potosino de Investigación en Ciencia y Tecnología, Camino a la Presa San José # 2055, Col. Lomas 4a Sec., San Luis Potosí, SLP 78216, Mexico

<sup>3</sup>Instituto de Ciencias del Mar y Limnología, Unidad de Procesos Oceánicos y Costeros, Universidad Nacional Autónoma de México, Circuito Exterior s/n, 04510 CDMX, Mexico

<sup>4</sup>Posgrado en Geociencias Aplicadas, Instituto Potosino de Investigación en Ciencia y Tecnología, Camino a la Presa San José # 2055, Col. Lomas 4a Sec., San Luis Potosí, SLP 78216, Mexico

<sup>5</sup>Centro de Investigación en Ciencias, Instituto de Investigación en Ciencias Básicas y Aplicadas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Mor 62209, Mexico

<sup>6</sup>Posgrado en Ingeniería, Instituto de Energías Renovables, Universidad Nacional Autónoma de México, Temixco, Mor 62580, Mexico

<sup>7</sup>División de Materiales Avanzados, Instituto Potosino de Investigación en Ciencia y Tecnología, Camino a la Presa San José # 2055, Col. Lomas 4a Sec., San Luis Potosí, SLP 78216, Mexico

<sup>8</sup>Doctorado en Ciencias, Instituto de Investigación en Ciencias Básicas y Aplicadas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Mor 62209, Mexico

<sup>9</sup>Universidad Autónoma de Nuevo León, Facultad de Ciencias de la Tierra, Ex-Hacienda de Guadalupe, Carretera Linares-Cerro Prieto km 8, Linares, N.L. 67700, Mexico

Correspondence should be addressed to Surendra P. Verma; [spv@ier.unam.mx](mailto:spv@ier.unam.mx)

Received 8 August 2018; Revised 8 October 2018; Accepted 19 October 2018; Published 11 December 2018

Academic Editor: Rafal Sitko

Copyright © 2018 Surendra P. Verma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We applied both the ordinary linear regression (OLR) and the new uncertainty weighted linear regression (UWLR) models for the calibration and comparison of a XRF machine through 59 geochemical reference materials (GRMs) and a procedure blank sample. The mean concentration and uncertainty data for the GRMs used for the calibrations (Supplementary Materials) (available here) were achieved from an up-to-date compilation of chemical data and their processing from well-known discordancy and significance tests. The drift-corrected XRF intensity and its uncertainty were determined from mostly duplicate pressed powder pellets. The comparison of the OLR (linear correlation coefficient  $r \sim 0.9523$ – $0.9964$  and  $0.9771$ – $0.9999$ , respectively, for before and after matrix correction) and UWLR models ( $r \sim 0.9772$ – $0.9976$  and  $0.9970$ – $0.9999$ , respectively) clearly showed that the latter with generally higher values of  $r$  is preferable for routine calibrations of analytical procedures. Both calibrations were successfully applied to rock matrices, and the results were generally consistent with those obtained in other laboratories although the UWLR model showed mostly narrower confidence limits of the mean (slope and intercept) or lower uncertainties than the OLR. Similar sensitivity ( $\sim 2.69$ – $46.17$   $\text{kc}\cdot\text{s}^{-1}\cdot\%^{-1}$  for the OLR and  $\sim 2.78$ – $59.69$   $\text{kc}\cdot\text{s}^{-1}\cdot\%^{-1}$  for the UWLR) also indicated that the UWLR could advantageously replace the OLR model. Another novel aspect is that the total uncertainty can be reported for individual chemical data. If the analytical instruments were routinely calibrated from the UWLR model, this action would make the science of geochemistry more quantitative than at present.

## 1. Introduction

All modern analytical instruments require some kind of calibration of the instrumental response ( $y$ -variable) as a function of the concentration ( $x$ -variable) [1–3]. This calibration is generally achieved through an ordinary least-squares linear regression (OLR) model. However, such a procedure is not strictly valid because all requirements for the statistical validity of the OLR model are not fulfilled. Usually, the assumptions “independent concentration variable  $x$  is error-free or less than one-tenth of the error in the dependent response variable  $y$ ” and “error in  $y$  is homoscedastic” (i.e., equal errors for all  $y$  values) are not satisfied and, therefore, more sophisticated and statistically coherent regression procedures, such as weighted least-squares linear regression (WLR) models, should be used [4–18].

X-ray fluorescence (XRF) spectrometry is among the most popular analytical techniques for the determination of all major and some trace elements in rocks [4, 19–27]. Natural geochemical reference materials (GRMs) are commonly used for XRF calibrations and posterior characterization of those and other GRMs as well as of similar rock and mineral matrices [4, 19, 28–30]. As for most other analytical instruments, XRF spectrometers are also calibrated under the statistically incoherent OLR model.

To apply the WLR and compare it with the OLR, both central tendency (e.g., mean) and dispersion (e.g., confidence limits of the mean) estimates on both  $x$ -axis (concentration, generally expressed in the unit of % m/m, i.e., mass/mass unit expressed in percent) and  $y$ -axis (response, in this case XRF intensity, generally reported in the unit of  $\text{kc}\cdot\text{s}^{-1}$ , i.e., kilo counts per second) variables are required. More precise (and accurate) estimates of the central tendency will also be useful for both types of regressions. Therefore, precise concentrations of GRMs with the respective lowest possible “confidence limits of the mean” (referred hereafter as the “uncertainty” of the measured variable) [2, 17, 18] are required to apply the regression procedure. Sometimes, we had to use also the term “error” (instead of the uncertainty) because the use of the error is widespread in the literature.

We report the following five aspects: (a) evaluation of 59 GRMs to achieve the least possible uncertainties in the mean concentrations of all major elements ( $\text{SiO}_2$  to  $\text{P}_2\text{O}_5$ ); (b) the comparison of regression models (OLR and WLR) applied to net drift-corrected XRF intensities before the correction of matrix effects; (c) the second (or final) comparison of both models after achieving the matrix correction as well as for the estimation of sensitivities of the regression models; (d) application of the entire procedure to four GRMs treated as “unknown” samples and their comparison with the previous literature compilations; and (e) development of a computer program to achieve the abovementioned objectives. Thus, the regression equations (intercept and its uncertainty, slope and its uncertainty, and linear correlation coefficient values) for each constituent from  $\text{SiO}_2$  to  $\text{P}_2\text{O}_5$  and their application to similarly complex rock matrices are presented in this work.

## 2. Evaluation of Major Element Data for GRMs

A total of 59 GRMs (listed in alphabetical order in Table S1; this and four other tables are provided in Supplementary Materials), along with a procedure blank, were used in this study. The procedure blank was a pellet prepared in duplicate with only pure  $\text{N,N}'$ -Ethylene bis(stearamide) beads without any sample (Section 3). The individual data reported in earlier compilations [31–47] were first compiled in new databases.

The statistical parameters obtained in these early compilations could not be directly used for instrumental calibrations due to the following reasons: (i) the statistical methods used to achieve the statistical estimates were outdated (see [17, 18, 48, 49] for possible reasons), and the inferred statistical values were of low quality (high values of dispersion); (ii) there are still determinations reported during about 30 or more years (postcompilation years) that were not obviously available to those compilers; (iii) the precision of more recent determinations is likely to have improved due to the availability of online computers on most modern instruments; (iv) newer more reliable statistical techniques are now available for improving both precision and accuracy of the statistical inferences, e.g., the use of discordancy tests with the highest power and lowest swamping and masking effects [18, 48, 50–52]; and (v) importantly, new computer programs have been developed by our group [52–54], available at <http://tlaloc.ier.unam.mx> for download or online processing of data (after previous registration onto our server), which can be advantageously used for efficient processing of experimental databases.

The same kinds of objections are applicable even today for the originator’s websites, such as <https://gbank.gsj.jp/geostandards/welcome.html> for Japanese GRMs or [https://crustal.usgs.gov/geochemical\\_reference\\_standards](https://crustal.usgs.gov/geochemical_reference_standards) for United States GRMs. The statistical information at these websites is based on early compilations (around 30 or more years ago). Furthermore, we were unable to use the recent work [55] because this paper reported significantly larger uncertainty values as compared to those achievable from our new validated statistical procedure [51–54]; besides, updated statistical information on the mean and its uncertainty was not available in [55] for many GRMs used in our work.

The initial databases were complemented by individual data from a large number of posterior publications (~480; Table S1), whose complete listing is available at our server <http://tlaloc.ier.unam.mx> under the heading of “Quality Control.” These major element data were classified according to the analytical method groupings [56]. Data from each method group were considered as a univariate statistical sample. Appropriate discordancy and significance tests were applied from thoroughly automatized software UDASys2 [52] and UDASys3 (unpublished), which, in their “recommended procedure,” apply the most powerful five (two new and three conventional) recursive tests with prior application of respective single-outlier tests having nil swamping and low masking effects [48, 57–60]. Although the application of discordancy tests is identical for both UDASys2 and UDASys3, the difference lies in that the latter applies the

significance (ANOVA,  $F$  and  $t$ ) tests in order to provide the final results automatically.

The resulting statistical information after the application of well-known discordancy tests at the strict 99% confidence level (mean and uncertainty values rounded according to the flexible rules [18]) is listed in Table S2. These GRM compositional data showed by far the lowest 99% uncertainty (Table S2), much lower than any existing compilation [31–47, 55]. We may also stress once again that this was achieved through an objective combination of discordancy tests having the highest performance and lowest swamping and masking effects [17, 18, 48, 53], i.e., from the methodology having the lowest type I and type II errors and the highest power.

Therefore, the population mean of these GRMs is now known within the narrowest possible 99% confidence limits of the mean to best represent the concentration ( $x$ ) axis in the instrumental calibrations as suggested [2, 5, 7, 10, 17, 18, 53]. These data (in units of % m/m; Table S2) will also be useful for those who wish to achieve instrumental calibrations or simply use them for quality control of their results for rock and mineral matrices.

### 3. XRF Instrumentation and Intensity Measurements

A wavelength dispersive X-ray fluorescence (WDXRF) spectrometer Rigaku ZSX Primus II model (rhodium X-ray tube; 4 kW maximum power) was used for this work. We made the effort to best represent the response ( $y$ ) axis ( $x$ -ray intensity in the units of kilo counts per second,  $\text{kc}\cdot\text{s}^{-1}$ ) for the calibrations. For each GRM, duplicate (41 samples) or even triplicate (8 samples) pressed powder pellets were prepared. First, an appropriate amount of each GRM was dried overnight in an oven at about  $105^\circ\text{C}$ . For each pellet, accurately weighed 3.5 g of moisture-free GRM was thoroughly mixed with accurately weighed 3.0 g pure N,N'-ethylene bis(stearamide) beads,  $<840\ \mu\text{m}$  as wax (Sigma-Aldrich), and stored in a desiccator. Pressed powder pellets were prepared at  $20\ \text{tons}\cdot\text{inch}^{-2}$  pressure (about 310 MPa). However, for 10 GRMs, sufficient material was not available; therefore, only a single pellet could be prepared but the measured intensity uncertainty ( $u_{99}$ ) at the 99% confidence level was increased by a factor of 2 to take into account the sample preparation variance. Similarly, accurately weighed 6.5 g of pure N,N'-ethylene bis(stearamide) beads,  $<840\ \mu\text{m}$  as wax, was pressed to prepare a procedure blank sample. This was done in duplicate.

For the intensity measurements, the optimum instrumental conditions were first established through preliminary experiments prior to the routine measurements (Table S2). Each pellet was run at least 8 to 10 times in a random sequence, along with two drift monitors prepared from two volcanic rocks (basalt and rhyolite) from the San Luis Potosí Volcanic Field, San Luis Potosí (central Mexico).

The peak and background measuring conditions and time periods are also listed in Table S3. Appropriate mean drift corrections from two monitors were applied to all intensity measurements. Both monitors were run randomly

8 to 10 times each day. First, the expected monitor intensity was established as an average value of the first two days when the intensities were fairly stable and reproducible. Then, the average drift correction factors were calculated for each chemical element from the two monitors run in the XRF instrument periodically before and after a set of GRMs used for the calibration. These correction factors were then applied to the bracketed GRMs for the entire period of calibration, including the first two days and analysis of “unknown” samples.

Now, although the X-ray counts may obey a Poisson distribution, we are dealing with average values of count rates, which are likely to follow a normal distribution because of the central limit theorem. A normal distribution of measured intensities was also assured for each pellet from the application of discordancy tests as explained above for GRM concentrations. The intensity results for all pellets from a given GRM were then combined, the tests applied again to the combined data, and new mean and 99% uncertainty values were calculated for X-ray intensity of each GRM. This was done to take into account the variance of the sample preparation method, which was significantly higher than the instrumental variance of intensity measurements for individual pellets. The drift-corrected intensity values and their 99% uncertainties ( $\text{kc}\cdot\text{s}^{-1}$ ) for all GRMs, along with the concentration data and their 99% uncertainties (% m/m), are listed in Table S2.

### 4. Regression Models

Two different regression models (OLR and UWLR) were used and compared in this work. The OLR model most frequently used for instrumental calibrations ( $x$ -axis concentration and  $y$ -axis response; GRM concentration and X-ray intensity, respectively, in XRF spectrometry) requires the following assumptions to be fulfilled [4, 7, 10, 12–18]: (i) all errors are in the  $y$ -axis; (ii)  $x$ -axis is either error-free or has at most 10% error of the  $y$ -axis errors; (iii) errors in both axes are normally distributed; and (iv) errors in the  $y$ -axis are homoscedastic. Some or all of these assumptions are violated in most instrumental calibrations through the OLR model.

Thus, from the literature on the GRMs, it has been demonstrated that the concentration axis is not error-free (see non-zero uncertainties for all GRM concentrations in Table S2) [31–47, 51–53]. One can also clearly see that the errors in the intensity axis are not homoscedastic (see unequal, i.e., heteroscedastic uncertainties for any element in different GRMs in Table S2). For a heteroscedastic linear regression system, even if each error or noise term is still Gaussian, the OLR model is no longer the maximum likelihood estimate and consequently, it is no longer efficient [10]. The main advantage that the WLR has over the OLR is the ability to handle regression situations in which the data points are of varying quality as is the case with most instruments including the XRF spectrometers.

However, the major disadvantage of the WLR is that the approach is based on the assumption that the weights are known exactly. They can be estimated using several different equations or algorithms, but when the weights are produced

from small numbers of replicated observations, the regression parameters can be unpredictably affected [10]. In the example of the XRF calibration that we are presenting, the numbers of observations were relatively large for both the  $x$  and  $y$  axes (concentration and X-ray intensity parameters). Besides, instead of the sample variance, we used the uncertainty values (that take into account the number of observations in the formula for uncertainty or confidence limits of the mean calculations) [2, 18] for estimating the weight factors. The problem of the sensitivity to outliers in the regression equations [10] was also appropriately handled by discordancy tests programmed in the UDASys and BiDASys software [53, 54, 61].

Therefore, although frequently used, the OLR model is not statistically correct or coherent. The statistically coherent WLR, especially the uncertainty-based WLR (UWLR [17]) model, should be used. The confidence level, such as 95% or 99% (significance level of 5% or 1%, respectively, or  $\alpha$  of 0.05 and 0.01, respectively), can be explicitly expressed in the confidence limits of the mean or uncertainty used in the UWLR model as well as to estimate the weight factors [17]. We will deal with the 99% uncertainty to have the type I error small (about 1%). Unfortunately, software of most analytical instruments, including XRF spectrometers, allows only the OLR calibration. Therefore, any sophisticated regression model, such as the UWLR, will have to be applied outside the instrumental software. Thus, the probability concept (99% confidence level) can be explicitly used in the UWLR model for weight factors based on the inverse of the squared 99% uncertainty of the mean.

We now present a synthesis of the regression equations for instrumental calibrations [2, 10, 17, 18, 61].

#### 4.1. Ordinary Least-Squares Linear Regression (OLR) Model.

Let us assume that we have a series of  $n$  reference materials or standard calibrators having individual mean concentrations  $x_i$  with respective uncertainties  $u_{x_i}$  where  $i$  varies from 1 to  $n$ . In order to calibrate an instrument, each of these  $n$  calibrators were run several times, obtaining individual mean responses  $y_i$  with respective uncertainties  $u_{y_i}$  where  $i$  varies from 1 to  $n$ . Thus, we have  $n$  bivariate concentration-response data pairs or calibrators  $(x_i, y_i)$  with the respective uncertainties  $(u_{x_i}, u_{y_i})$ .

We can apply the OLR model to these data for obtaining a calibration equation. The OLR fits a least-squares linear equation to the  $n$  pairs  $(x_i, y_i)$  but does not take into account the respective uncertainties  $(u_{x_i}, u_{y_i})$ .

The general regression equation for the OLR is as follows (the subscript O is for the OLR model):

$$y_O(\pm u_{y_O}) = b_O(\pm u_{b_O}) + [m_O(\pm u_{m_O}) \times x], \quad (1)$$

where  $m$  is the slope,  $u_m$  is the resulting uncertainty in the slope,  $b$  is the intercept,  $u_b$  is the resulting uncertainty in the intercept,  $x$  is the independent variable,  $y_O$  is the dependent variable from the OLR model, and  $u_{y_O}$  is the resulting uncertainty in  $y$ . The following equations allow the calculations of these parameters:

$$m_O = \frac{\sum_{i=1}^n \{(x_i - \bar{x}) \times (y_i - \bar{y})\}}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2)$$

where  $\bar{x}$  and  $\bar{y}$  are, respectively, the mean values of the  $x$  and  $y$  variables:

$$u_{m_O} = \left\{ \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n-2) \sum_{i=1}^n (x_i - \bar{x})^2}} \right\} \times t_{(n-2)}^{CL}, \quad (3)$$

where  $\hat{y}_i$  is the value of  $y_O$  for  $x_i$  in equation (1) and  $t$  is the Student's  $t$  test value for  $(n-2)$  degrees of freedom, and the superscript CL is the confidence level, generally 95% or 99%:

$$b_O = \bar{y} - (m_O \times \bar{x}),$$

$$u_{b_O} = \left\{ \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2 \times \sum_{i=1}^n x_i^2}{n(n-2) \sum_{i=1}^n (x_i - \bar{x})^2}} \right\} \times t_{(n-2)}^{CL}. \quad (4)$$

It is a general practice in most instrumental calibrations to ignore all uncertainties in equation (1) and use an OLR equation without any error (or uncertainty) as follows:

$$y_O = b_O + [m_O \times x]. \quad (5)$$

The resulting standard deviation values of repeat measurements of unknown samples are reported as the final errors. However, these are only partial errors because the errors in the calibration equation (1) are not taken into account. In this work, we will use equation (1) to report total errors (in fact, 99% uncertainties) for the OLR model.

**4.2. Uncertainty Weighted Least-Squares Linear Regression (UWLR) Model.** For the UWLR model, the  $n$  pairs  $(x_i, y_i)$  of calibrators as well as the respective uncertainties  $(u_{x_i}, u_{y_i})$  are taken into account in order to achieve the best least-squares linear fit.

The uncertainties  $u_{x_i}$  in the  $x$ -axis are first propagated to the  $y$ -axis, combined with the  $u_{y_i}$ , and the total uncertainty  $u_i$  values on the  $y$ -axis are used for the weighting factors [2, 10, 17, 18, 61]:

$$u_i = \sqrt{(m_O \times u_{x_i})^2 + (u_{y_i})^2}. \quad (6)$$

The weights are calculated from  $u_i$  as follows:

$$w_i = \frac{n \times (u_i)^{-2}}{\sum_{i=1}^n (u_i)^{-2}}, \quad (7)$$

where  $w_i$  values have the following property:

$$\sum_{i=1}^n w_i = n. \quad (8)$$

Thus, the UWLR fits a linear equation to the  $n$  pairs  $(x_i, y_i)$  with the respective weighting factors  $w_i$  as follows (the subscript UW is for the UWLR model):

$$y_{UW}(\pm u_{y_{UW}}) = b_{UW}(\pm u_{b_{UW}}) + [m_{UW}(\pm u_{m_{UW}}) \times x]. \quad (9)$$

Note that this regression line will pass closer to the data with lesser uncertainty  $u_i$ . The intercept and slope variables and their uncertainties are calculated from the following equations:

$$m_{\text{UW}} = \frac{\{\sum_{i=1}^n (w_i \times x_i \times y_i)\} - \{n \times \bar{x}_{\text{UW}} \times \bar{y}_{\text{UW}}\}}{\{\sum_{i=1}^n (w_i \times x_i^2)\} - \{n \times (\bar{x}_{\text{UW}})^2\}}, \quad (10)$$

where  $\bar{x}_{\text{UW}}$  and  $\bar{y}_{\text{UW}}$  are, respectively, the weighted mean values of the  $x$  and  $y$  variables:

$$u_{m_{\text{UW}}} = \left\{ \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_{i\text{UW}})^2}{(n-2)\sum_{i=1}^n (x_i - \bar{x}_{\text{UW}})^2}} \right\} \times t_{(n-2)}^{\text{CL}}, \quad (11)$$

where  $\hat{y}_{i\text{UW}}$  is the value of  $y_{\text{UW}}$  for  $x_i$  in equation (9):

$$b_{\text{UW}} = \bar{y}_{\text{UW}} - (m_{\text{UW}} \times \bar{x}_{\text{UW}}),$$

$$u_{b_{\text{UW}}} = \left\{ \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_{i\text{UW}})^2 \times \sum_{i=1}^n x_i^2}{n(n-2)\sum_{i=1}^n (x_i - \bar{x}_{\text{UW}})^2}} \right\} \times t_{(n-2)}^{\text{CL}}. \quad (12)$$

The best regression equation for a calibration curve should have the following characteristics (without distinguishing the subscripts O and UW): (i) intercept  $b$  small approaching to zero; (ii) slope  $m$  large; and (iii) both  $u_b$  and  $u_m$  small. Further, the quality of the regression, whether a calibration curve or any other bivariate relationship, is also expressed as the linear regression coefficient ( $r$ ;  $r_{\text{O}}$  and  $r_{\text{UW}}$ , respectively, for the OLR and UWLR), which is ideally +1.00000 for a calibration curve [5, 18, 61].

## 5. Application of Regression Models for XRF Calibration

**5.1. Original Drift-Corrected Net Intensities and GRM Concentrations: The First Set of Two Regression Equations for Each Element.** The evaluations for both regression types on the drift-corrected net intensity-concentration (Int-Conc) relationships (Table S2) for all major elements from  $\text{SiO}_2$  to  $\text{P}_2\text{O}_5$  were performed (Table S4), for which the new online software BiDASys was used [61] at <http://tlaloc.ier.unam.mx>. BiDASys allows the application of the conventional OLR as well as the newly proposed UWLR model [17] and provides the output of all regression parameters in an Excel® file. Contrary to the common practice, we will refrain from showing the numerous  $x$ - $y$  (variable  $x$  is drift-corrected net intensity “Int” and variable  $y$  is the GRM concentration “Conc”) plots. This is because Table S4 statistically quantifies the visual interpretation of such diagrams. The quality parameters (standard errors  $se_b$  and  $se_m$ , uncertainty  $u_b$  and  $u_m$ , and linear correlation coefficient  $r$  and its squared value  $R^2$  parameters) are reported in Table S4. Because we are using these several different quality parameters, the concern against the use of solely  $R^2$  parameter [62] is not important for comparison purposes.

We will explain the implications of the statistical results for the first element  $\text{SiO}_2$ ; the statistics for other elements

(Table S4) can be similarly understood. The OLR regression equation from the first row of statistical information in Table S4 is as follows (after the element  $\text{SiO}_2$ , subscript O is for the OLR and p is for provisional concentration; note many decimal places are used for the regression variables in such equations, because these values are not final results, and we should not introduce rounding errors during the calculation stage):

$$C_{\text{SiO}_{2\text{Op}}} \left( \pm u_{C_{\text{SiO}_{2\text{Op}}}} \right) = 11.47071 (\pm 4.90411) + \left[ 0.14325 (\pm 0.01605) \right. \\ \left. \times I_{\text{SiO}_2} \left( \pm u_{I_{\text{SiO}_2}} \right) \right]. \quad (13)$$

Similarly, the UWLR equation from the second row of statistical information in Table S4 is as follows:

$$C_{\text{SiO}_{2\text{UWp}}} \left( \pm u_{C_{\text{SiO}_{2\text{UWp}}}} \right) = -0.01316 (\pm 2.83181) + \left[ 0.18539 (\pm 0.00927) \right. \\ \left. \times I_{\text{SiO}_2} \left( \pm u_{I_{\text{SiO}_2}} \right) \right]. \quad (14)$$

The implications of these regression equations can be understood from the comparison of the uncertainties of the intercept and slope, which are lower for the UWLR (equation (14)) than for the OLR (equation (13)). This means that the uncertainty of the calculated concentration will be lower for the UWLR than for the OLR. Correspondingly, the  $r$  value for the UWLR (0.99004,  $n = 60$ ;  $R^2 = 0.98017$ ) is much higher than that for the OLR (0.95229,  $n = 60$ ;  $R^2 = 0.90687$ ; Table S4). Similar trend in the  $r$  (and  $R^2$ ) values was obtained for all other elements except MnO (Table S4).

**5.2. Matrix-Effect-Corrected Intensities and GRM Concentrations: The Second Set of Two Regression Equations for Each Element.** Matrix correction is certainly required because the abovementioned least-squares linear regression fits are far from “perfect” ( $r \neq +1.00000$ ; in fact,  $r < 1$ ;  $n = 60$ ;  $r = 0.95229$ – $0.99638$  for the OLR and  $r = 0.97715$ – $0.99760$  for the UWLR; Table S4). There is a vast literature on the subject of matrix effects in XRF and their correction procedures [63–75]. In this study, the Lachance-Traill algorithm [73] was used for the matrix effect correction [63, 71]. This was done outside the XRF instrument software. In a review of the existing algorithms, Rousseau [63] showed that the Lachance-Traill algorithm could be considered as one of the most appropriate procedures for the matrix effect correction because other algorithms have limited application range or lack of accuracy. Thus, for each element from  $\text{SiO}_2$  to  $\text{P}_2\text{O}_5$ , a system of overdetermined equations was solved and the resulting alpha coefficients were used to correct all intensities for matrix effects.

From the alpha coefficients, matrix-corrected intensities and improved concentration values for the GRMs and their uncertainties were calculated iteratively under the condition that the convergence parameter (absolute relative difference

of the GRM calculated and input concentrations) for each compositional constituent ( $\text{SiO}_2$  to  $\text{P}_2\text{O}_5$ ) be minimized.

New regression equations for achieving the corrected concentrations were established from the relationship of the calculated GRM concentrations (ConcCalc) and the original GRM concentrations (Conc) given in Table S2, for which the online BiDASys software [61] was used at <http://tlaloc.ier.unam.mx>. These equations can be formulated from the regression coefficient values given in Table S4 (see ConcCalc-Conc rows corresponding to the OLR and UWLR). Again, we will highlight their significance for  $\text{SiO}_2$  only.

The OLR regression equation from the third row of statistical information in Table S4 is as follows:

$$C_{\text{SiO}_{20}}(\pm u_{C_{\text{SiO}_{20}}}) = 1.87016(\pm 4.2048) + \left[ 0.96759(\pm 0.07369) \right] \times C_{\text{SiO}_{20c}}(\pm u_{C_{\text{SiO}_{20c}}}), \quad (15)$$

where the subscripts O and c stand for the OLR model and calculated concentration (ConcCalc), respectively.

Similarly, the UWLR equation from the fourth row of statistical information in Table S4 is as follows:

$$C_{\text{SiO}_{2\text{UW}}}(\pm u_{C_{\text{SiO}_{2\text{UW}}}}) = -0.02185(\pm 1.47356) + \left[ 1.00325(\pm 0.02701) \right] \times C_{\text{SiO}_{2\text{UWc}}}(\pm u_{C_{\text{SiO}_{2\text{UWc}}}}), \quad (16)$$

where the subscripts UW and c stand for the UWLR model and calculated concentration (ConcCalc), respectively.

Equations (15) and (16) show that the concentration values from the UWLR would be more reliable (lesser uncertainty values in both intercept and slope) than the OLR model. The  $r$  value is higher for the UWLR (0.99704,  $n = 60$ ;  $R^2 = 0.99408$ ; Table S4) than the OLR (0.97710,  $n = 60$ ;  $R^2 = 0.95472$ ).

After the matrix correction, in fact most regression equations are better because all  $r$  and  $R^2$  values are higher for both OLR and UWLR than without the correction (Table S4; Figure 1 for  $r$  only). For the OLR, the matrix correction increased the  $r$  values ( $n = 60$ ) from 0.95229–0.99638 ( $R^2 = 0.90687 - 0.99277$ ) to 0.97710–0.99992 ( $R^2 = 0.95472 - 0.99994$ ). Similarly, for the UWLR, this increase was from 0.97715–0.99760 ( $R^2 = 0.95472 - 0.99521$ ) to 0.99704–0.99993 ( $R^2 = 0.99408 - 0.99986$ ). Thus, after matrix correction, all  $r$  values increased for both OLR and UWLR. For the UWLR, the  $r$  values approached the ideal value of +1.00000 (Figure 1). One has to keep in mind that when the  $r$  values are closer to the maximum possible value of 1 (the “ideal” fit), the improvement expressed by the actual (absolute) value of  $r$  will apparently be small. However, as long as the  $r$  value increases for the UWLR as compared to the OLR (Figure 1; Table S4), we can objectively infer that the UWLR is a better regression model than the OLR.

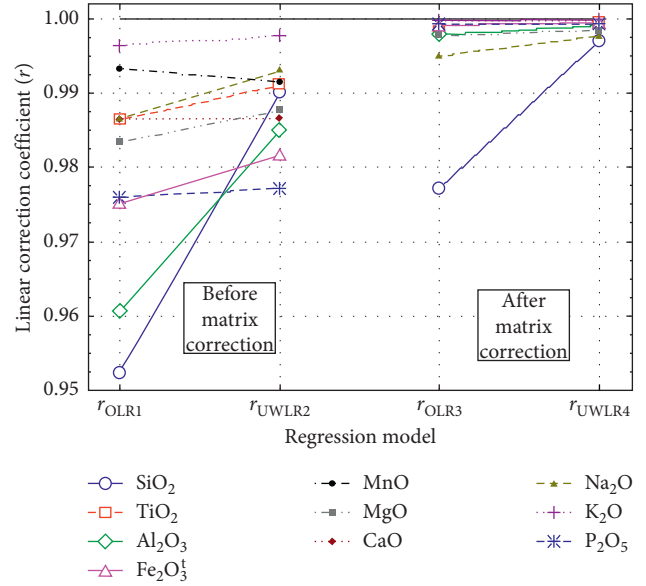


FIGURE 1: Linear correlation coefficient ( $r$ ) values for the ordinary least-squares linear regression (OLR) and uncertainty-based weighted least-squares linear regression (UWLR) models for the XRF calibration of major elements ( $\text{SiO}_2$  to  $\text{P}_2\text{O}_5$ ) in rocks and minerals. OLR1: OLR model 1 for Int-Conc before matrix correction; UWLR2: UWLR model 2 for Int-Conc before matrix correction; OLR3: OLR model 3 for ConcCalc-Conc after matrix correction; and UWLR4: UWLR model 4 for ConcCalc-Conc after matrix correction. Symbols are shown as inset. The horizontal line at the  $r$  value of 1 represents the “ideal” or “perfect” linear fit.

Before the matrix correction, the intercepts of the Int-Conc regression lines were closer to zero for the UWLR (range  $\sim -0.013$  to  $+0.011$ ) than for the OLR (range  $\sim -2.098$  to  $+11.47$ ) model (Table S4; Figure 2). The same is true for the intercept values (ConcCalc-Conc relationship) after the matrix correction ( $\sim -0.025$  to  $+0.021$  for the UWLR and  $\sim -0.110$  to  $+1.87$  for the OLR).

Finally, the uncertainties on both intercept and slope parameters were mostly lower for the UWLR than the OLR (Table S4). We highlight these differences (lower uncertainties for the UWLR) from dimensionless (free of the measurement units) parameters  $\delta u_b$  and  $\delta u_m$  defined as follows:

$$\delta u_b = \frac{u_{b_O} - u_{b_{\text{UW}}}}{u_{b_{\text{UW}}}} \times 100, \quad (17)$$

$$\delta u_m = \frac{u_{m_O} - u_{m_{\text{UW}}}}{u_{m_{\text{UW}}}} \times 100.$$

Plots of these two parameters are presented in Figure 3. If  $u_{b_O} > u_{b_{\text{UW}}}$ , the  $\delta u_b$  will be positive, otherwise it will be negative. The same is true for  $\delta u_m$ . For the comparison of two models OLR and UWLR before the matrix correction, the uncertainty for the UWLR were lower than the OLR for 7 elements (positive  $\delta u_b$  and  $\delta u_m$ ), whereas for after the matrix correction, it was so for 8 elements (out of 10; Figure 3). The exceptions were for 3 elements MnO, CaO, and  $\text{P}_2\text{O}_5$  (negative  $\delta u_b$  and  $\delta u_m$ ) for the uncertainties before matrix

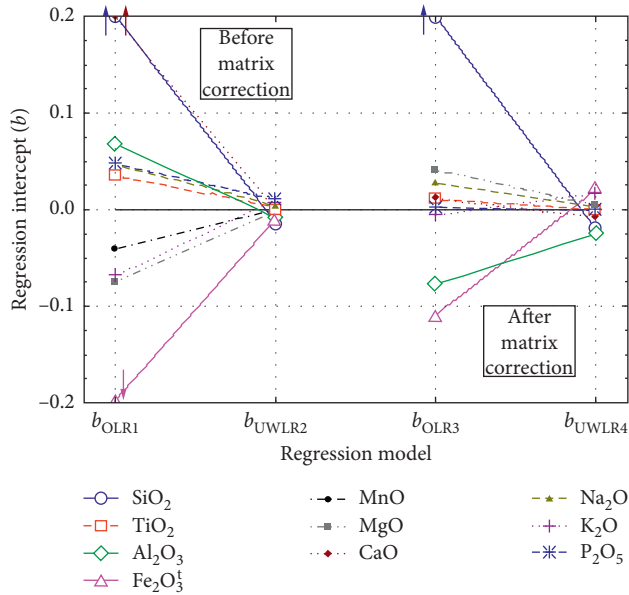


FIGURE 2: Intercept ( $b$ ) values for the ordinary least-squares linear regression (OLR) and uncertainty-based weighted least-squares linear regression (UWLR) models for the XRF calibration of major elements (SiO<sub>2</sub> to P<sub>2</sub>O<sub>5</sub>) in rocks and minerals. Symbols are shown as inset. For abbreviations, see Figure 1. Note some intercept values plotted outside the graph; this is indicated by arrows next to the data point. The horizontal line at the intercept value of zero represents the “ideal” intercept.

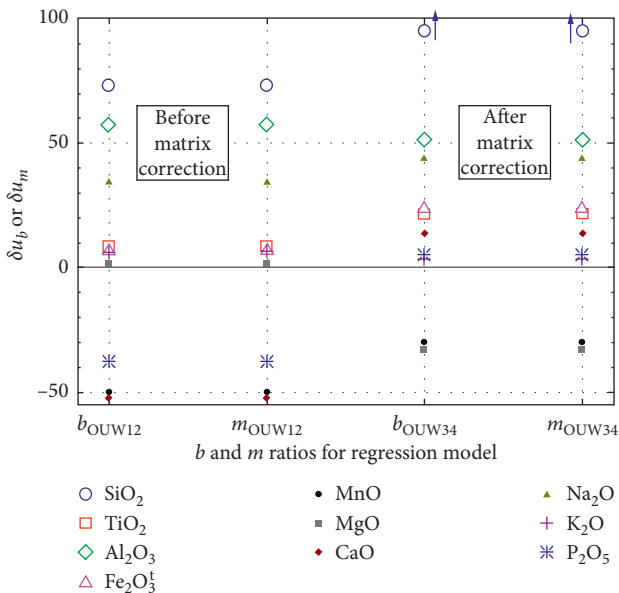


FIGURE 3: New parameters  $\delta u_b$  (for the intercept) and  $\delta u_m$  (for the slope) for the evaluation of intercept ( $b$ ) and slope ( $m$ ) of two regression models (OLR: ordinary least-squares linear regression and UWLR: uncertainty-based weighted least-squares linear regression) before (OUW12) and after (OUW34) the matrix correction. The horizontal solid line at the  $y$  value of zero represents the line with no difference in the uncertainties of the two models. The arrows indicate that these data plotted above the scale are used for the diagram.

correction and for 2 elements MnO and MgO for those after the matrix correction (Figure 3). Even for the exceptions of the elements MnO and MgO, the UWLR values should be usable (Table S4), i.e., it is not actually necessary to resort to the OLR model for these two exceptions (2 out of 10 cases). Thus, we can use the UWLR model for all purposes.

## 6. Sensitivities of Major Elements

We calculated the sensitivities as the slope of the Conc-IntCorr (GRM concentrations of Table S2 and matrix-corrected intensities of Table S5; see Supplementary Materials at <http://tlaloc.ier.unam.mx>) from the regression curve (line) for all 10 elements and for both models (Table 1). Because the  $r$  values are significantly high (all  $>0.961$ ,  $n = 60$ ; Table 1) and the residuals are randomly distributed (graphs not shown), the straight line is the most likely, statistically valid fit for the concentration-matrix-corrected intensity data [5, 17, 18]. Therefore, the slope of the regression line represents an average sensitivity value for a given element under the chosen working conditions (Table S3).

The intercept values were closer to zero (zero being the theoretically ideal intercept) for the UWLR regression ( $\sim -0.113$  to  $+0.104$ ; Table 1) as compared to the OLR ( $\sim -47.8$  to  $+12.3$ ; Table 1). The sensitivity values represented by the slopes of the regression lines (Table 1) were generally similar for both models ( $\sim 2.69$ – $46.17$   $\text{kc}\cdot\text{s}^{-1}\cdot\%^{-1}$  for the OLR and  $\sim 2.78$ – $59.69$   $\text{kc}\cdot\text{s}^{-1}\cdot\%^{-1}$  for the UWLR). The sensitivity actually depends on the measuring conditions (Table S3), which were the same for both models.

For the matrix-corrected intensity-concentration (IntCorr-Conc) regressions, the parameters are listed in Table 2. All intercepts for the UWLR model, without exception, were closer to zero as compared to the OLR model. This confirms the superiority of the UWLR model.

## 7. Application to Rock Matrices

The calibrations achieved in this work (Table S4) were applied to the analysis of four GRMs (attapulgite or Fuller’s earth clay ATT1; bentonite clay CSB1; granite GH; and tonalite TLM1) taken as “unknown” samples. These GRMs, having similarly complex matrices as the calibration samples, were not included in the calibrations because their mean values were available only from early description or compilations (for ATT1 and CSB1 [76]; for GH [77]; and for TLM1 [78]). We were unsuccessful in complementing these “old” concentration values with newer ones for these GRMs. Therefore, these GRMs were used as unknown samples. They were analysed in exactly the same manner as the calibration samples.

All calculations for the unknown samples were done outside the instrumental software. The drift-corrected net intensities and the corresponding uncertainties were processed from the first set of two regression equations (Int-Conc OLR and UWLR models; Table S4) to obtain provisional concentration and uncertainty values. The

TABLE 1: Instrumental sensitivities ( $x$ - $y$ : concentration-matrix-corrected intensity (Conc-IntCorr) regression model; mean and 99% uncertainty) for major elements.

Element	Regression		Number of data pairs (calibrators)	Regression equation parameters						Quality of regression equation	
	Variables $x$ - $y$	Model		Intercept ( $b$ )	$se_b$	$u_b$	Slope ( $m$ )	$se_m$	$u_m$	$r$	$R^2$
SiO <sub>2</sub>	Conc-IntCorr	OLR	60	-47.8441	13.1332	34.9772	6.3592	0.2401	0.6395	0.96105	0.92362
	Conc-IntCorr	UWLR	60	0.1035	5.3870	14.3471	5.3476	0.0985	0.2623	0.99157	0.98321
TiO <sub>2</sub>	Conc-IntCorr	OLR	60	-0.2298	0.3000	0.7991	13.2255	0.2889	0.7694	0.98644	0.97307
	Conc-IntCorr	UWLR	60	-0.0188	0.2637	0.7024	12.3181	0.2539	0.6763	0.99113	0.98235
Al <sub>2</sub> O <sub>3</sub>	Conc-IntCorr	OLR	60	8.1398	4.1673	11.0987	6.1191	0.2269	0.6043	0.96237	0.92615
	Conc-IntCorr	UWLR	60	0.0594	3.0552	8.1369	7.0715	0.1663	0.4430	0.98557	0.97135
Fe <sub>2</sub> O <sub>3</sub> <sup>t</sup>	Conc-IntCorr	OLR	60	12.2628	2.4739	6.5886	4.3943	0.1337	0.3561	0.97419	0.94904
	Conc-IntCorr	UWLR	60	0.0715	2.7723	7.3835	5.4403	0.1498	0.3991	0.98112	0.96259
MnO	Conc-IntCorr	OLR	60	2.0367	0.2963	0.7891	46.1686	0.6958	1.8531	0.99348	0.98700
	Conc-IntCorr	UWLR	60	0.0147	0.7535	2.0069	59.6875	1.7695	4.7127	0.99160	0.98328
MgO	Conc-IntCorr	OLR	60	1.1050	1.0231	2.7246	3.6928	0.0901	0.2400	0.98317	0.96662
	Conc-IntCorr	UWLR	60	0.0108	0.9641	2.5676	3.9863	0.0849	0.2262	0.98745	0.97505
CaO	Conc-IntCorr	OLR	60	-14.3055	3.5661	9.4974	15.4215	0.3371	0.8977	0.98643	0.97304
	Conc-IntCorr	UWLR	60	-0.0121	5.6252	14.9814	11.5666	0.5317	1.4160	0.98657	0.97332
Na <sub>2</sub> O	Conc-IntCorr	OLR	60	0.0166	0.1553	0.4137	2.6888	0.0571	0.1521	0.98716	0.97449
	Conc-IntCorr	UWLR	60	-0.0111	0.1154	0.3073	2.7843	0.0424	0.1130	0.99344	0.98692
K <sub>2</sub> O	Conc-IntCorr	OLR	60	1.2200	0.5896	1.5702	13.8296	0.1431	0.3811	0.99691	0.99383
	Conc-IntCorr	UWLR	60	-0.1130	0.5491	1.4623	14.4829	0.1333	0.3549	0.99793	0.99587
P <sub>2</sub> O <sub>5</sub>	Conc-IntCorr	OLR	60	-1.7376	0.5433	1.4469	39.1055	1.2932	3.4442	0.96972	0.94035
	Conc-IntCorr	UWLR	60	0.0626	0.7039	1.8747	29.4822	1.6756	4.4624	0.96926	0.93946

$b$ , intercept;  $se$ , standard error;  $u$ , uncertainty at 99%;  $m$ , slope;  $r$ , linear correlation coefficient;  $R^2$ , squared linear correlation coefficient.

provisional concentrations were then used to obtain matrix corrections for each sample. The method was iteratively applied with the newer concentrations to obtain the final calculated concentration values (Table 3). These calculated concentration values were used to compute the final mean concentrations ( $\bar{x}$ ) and 99% uncertainties of the mean ( $u_{99}$ ) for each sample from the second sets of regression equations (ConcCalc-Conc, OLR and UWLR models; Table S4). The loss on ignition (LOI) was required to optimise the final results.

The results are listed in Table 3 and compared with the literature compilations [75–77]. On the other hand, because 99% uncertainties were not reported in the original compilations, they were computed for the comparison from the standard deviation, number of determinations, and appropriate two-sided  $t$  values at 99% confidence level [2, 18].

Firstly, although the mean concentration values determined by the OLR and UWLR models showed a general agreement, the 99% uncertainty values ( $u_{99}$ ; Table 3) were generally lower for the UWLR models, which clearly



TABLE 2: Regression ( $x$ - $y$ ): matrix-corrected intensity-concentration (IntCorr-Conc) parameters.

Element	Regression		Number of data pairs (calibrators)	Regression equation parameters						Quality of regression equation	
	Variables $x$ - $y$	Model		Intercept ( $b$ )	$se_b$	$u_b$	Slope ( $m$ )	$se_m$	$u_m$	$r$	$R^2$
SiO <sub>2</sub>	IntCorr-Conc	OLR	60	10.8541	1.6762	4.4641	0.1452	0.0055	0.0146	0.96105	0.92362
	IntCorr-Conc	UWLR	60	-0.0148	0.9765	2.6006	0.1848	0.0032	0.0085	0.99157	0.98321
TiO <sub>2</sub>	IntCorr-Conc	OLR	60	0.0344	0.0220	0.0587	0.0736	0.0016	0.0043	0.98644	0.97307
	IntCorr-Conc	UWLR	60	0.0020	0.0204	0.0542	0.0801	0.0015	0.0040	0.99114	0.98235
Al <sub>2</sub> O <sub>3</sub>	IntCorr-Conc	OLR	60	-0.1675	0.6763	1.8011	0.1514	0.0056	0.0150	0.96237	0.92615
	IntCorr-Conc	UWLR	60	-0.0070	0.4419	1.1769	0.1397	0.0037	0.0098	0.98557	0.97135
Fe <sub>2</sub> O <sub>3</sub> <sup>t</sup>	IntCorr-Conc	OLR	60	-2.0576	0.5960	1.5873	0.2160	0.0066	0.0175	0.97419	0.94904
	IntCorr-Conc	UWLR	60	-0.0092	0.5458	1.4536	0.1772	0.0060	0.0160	0.98112	0.96259
MnO	IntCorr-Conc	OLR	60	-0.0412	0.0067	0.0178	0.0214	0.0003	0.0009	0.99348	0.98700
	IntCorr-Conc	UWLR	60	0.0001	0.0134	0.0356	0.0165	0.0007	0.0017	0.99160	0.98328
MgO	IntCorr-Conc	OLR	60	-0.0977	0.2748	0.7319	0.2618	0.0064	0.0170	0.98317	0.96662
	IntCorr-Conc	UWLR	60	-0.0014	0.2692	0.7168	0.2376	0.0063	0.0167	0.98745	0.97505
CaO	IntCorr-Conc	OLR	60	1.0561	0.2173	0.5789	0.0631	0.0014	0.0037	0.98643	0.97304
	IntCorr-Conc	UWLR	60	0.0064	0.4616	1.2293	0.0853	0.0029	0.0078	0.98657	0.97331
Na <sub>2</sub> O	IntCorr-Conc	OLR	60	0.0447	0.0567	0.1511	0.3624	0.0077	0.0205	0.98716	0.97449
	IntCorr-Conc	UWLR	60	0.0054	0.0421	0.1122	0.3558	0.0057	0.0152	0.99345	0.98693
K <sub>2</sub> O	IntCorr-Conc	OLR	60	-0.0720	0.0430	0.1146	0.0719	0.0007	0.0020	0.99691	0.99383
	IntCorr-Conc	UWLR	60	0.0082	0.0391	0.1042	0.0687	0.0007	0.0018	0.99793	0.99587
P <sub>2</sub> O <sub>5</sub>	IntCorr-Conc	OLR	60	0.0537	0.0128	0.0341	0.0241	0.0008	0.0021	0.96972	0.94035
	IntCorr-Conc	UWLR	60	0.0010	0.0202	0.0538	0.0320	0.0013	0.0033	0.96929	0.93952

$b$ , intercept;  $se$ , standard error;  $u$ , uncertainty at 99%;  $m$ , slope;  $r$ , linear correlation coefficient;  $R^2$ , squared linear correlation coefficient.

indicates that this model should be used routinely, instead of the conventional OLR model. Secondly, there is also a general agreement among all mean values, especially for granite GH and tonalite TLM1. The two clay samples (ATT1 and CSB1) showed some differences with the preliminary values obtained by the originators of these GRMs [75]. These values for comparison were obtained in only one laboratory. The errors (uncertainties) reported in the literature were underestimated, because they did not include those resulting from the calibrations. Furthermore, the accuracy data of the originator's laboratory were not reported [75], such as the

results for established GRMs and their comparison to other laboratories.

## 8. Computer Program XRFCalcUnknown

An online computer program JSpectrum\_XRFCalcUnknown will be available at our server <https://tlaloc.ier.unam.mx> for use for unknown samples, which will guide other users to achieve the UWLR calibration outside of the instrumental software and its routine application to unknown samples. This program incorporates the iteration process to achieve reliable

TABLE 3: Application of the XRF calibrations for the determination of major elements (mean concentration in % m/m; mass/mass unit expressed in percentage, habitually called wt.% and 99% uncertainty  $u_{99}$  or simply  $u$ ) in four rock samples (data reported as rounded value following flexible criteria proposed by [18]) and their comparison with the literature data.

Rock	Method (ref.)	SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	Fe <sub>2</sub> O <sub>3</sub> <sup>†</sup>	MnO	MgO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	P <sub>2</sub> O <sub>5</sub>										
		$\bar{x}$	$u_{99}$	$\bar{x}$	$u_{99}$	$\bar{x}$	$u_{99}$	$\bar{x}$	$u_{99}$	$\bar{x}$	$u_{99}$										
ATTI	OLR	61.7	6.1	0.4961	0.0173	10.00	0.49	4.246	0.292	0.02394	0.00188	5.373	0.377	2.139	0.086	0.083	0.094	0.7777	0.0378	0.8074	0.0134
	UWLR	62.30	2.23	0.4951	0.0143	9.666	0.321	4.402	0.235	0.02416	0.00269	6.17	0.52	2.144	0.076	0.058	0.065	0.7781	0.0366	0.7828	0.0125
	[76]	59.6	0.8	0.49	0.015	9.50	0.29	3.31	0.06	0.025	0.018	9.14	0.06	1.87	0.03	0.1	0.03	0.86	0.06	0.76	0.01
CSBI	OLR	52.7	5.5	0.1448	0.0157	24.60	0.71	3.493	0.289	0.01408	0.00188	1.523	0.270	1.117	0.085	2.126	0.119	0.4169	0.0374	0.0383	0.0060
	UWLR	52.76	2.04	0.1343	0.0130	24.40	0.48	3.564	0.233	0.01398	0.00268	1.604	0.398	1.105	0.075	2.123	0.083	0.4337	0.0361	0.0357	0.0057
	[76]	55.3	0.8	0.16	0.03	21.75	0.18	3.88	0.03	0.020	0.003	1.94	0.08	1.28	0.06	2.07	0.20	0.54	0.03	—	—
GH	OLR	73.3	6.5	0.0968	0.0157	13.98	0.53	1.336	0.285	0.05084	0.00190	0.074	0.262	0.843	0.084	3.618	0.151	5.302	0.059	0.0119	0.0059
	UWLR	73.62	2.38	0.0861	0.0129	13.714	0.349	1.466	0.229	0.05064	0.00271	0.038	0.390	0.828	0.074	3.625	0.105	5.216	0.057	0.0091	0.0057
	[77]	75.8	0.39	0.08	0.07	12.50	0.16	1.34	0.2	0.05	0.009	0.03	0.18	0.69	0.13	3.85	0.11	4.76	0.05	0.08	0.03
TLM1	OLR	57.7	5.7	0.8497	0.0200	17.38	0.58	7.637	0.306	0.11614	0.00199	3.972	0.319	6.807	0.100	2.675	0.130	1.7585	0.0403	0.1378	0.0063
	UWLR	57.67	2.11	0.8402	0.0165	17.046	0.386	7.682	0.247	0.11592	0.00281	4.281	0.447	6.821	0.088	2.667	0.090	1.742	0.0390	0.1336	0.0060
	[78]	58.85	0.35	0.820	0.013	17.48	0.25	7.67	0.15	0.105	0.013	3.3	0.08	6.67	0.08	2.98	0.07	1.67	0.08	0.15	0.013

concentrations as demonstrated in this work. Example input data files and a ReadMe document are provided to facilitate the application of JSpectrom\_XRFCalcUnknown. One important aspect of the program is that for a sample to be identified as an “unknown” sample, the value of LOI (loss on ignition in percent) should be input in the first sheet of the measured intensity file.

A novel aspect of the present work is that total 99% uncertainty can be calculated for individual datum in a given sample (treated as unknown; Table 3). This innovation if put into practice can entirely change the geochemical literature, and in fact make geochemistry a more quantitative science. Further, if an appropriate GRM is analysed as unknown and the analytical data (both mean and total uncertainty) are reported along with the field samples, the data accuracy can be statistically judged from such reports.

## 9. Conclusions

The XRF spectrometer calibrated under both the OLR and UWLR models clearly showed that the UWLR provides more reliable results (lower uncertainty estimates) than the OLR model commonly practiced for most XRF instruments. The sensitivity and LOD values presented for both models also supported the use of the UWLR model. The UWLR model should therefore be used routinely in such calibrations. The use of a large number of well-characterized GRMs is also recommended for this purpose as illustrated in the present work. The application of our procedure was well documented for the analysis of similarly complex rock matrices. The reporting of total uncertainty values for individual datum is highly recommended for all future geochemical research. This work for the XRF shows that such a practice is easy to achieve in any other analytical calibration procedures. As the major conclusion, we can confirm that the statistically coherent WLR model was shown to perform better than the frequently used conventional statistically incoherent OLR model.

## Data Availability

The list of all compiled references (Table S1) will be available at <http://tlaloc.ier.unam.mx>. These references are not included with the manuscript because they are too many (~480). Similarly, as stated, the online program JSpectrom\_XRFCalcUnknown will also be added onto this web portal <http://tlaloc.ier.unam.mx>. This program needs to be available online for future use; it cannot be submitted to the journal.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported through the Newton Advanced Fellowship Award (grant NA160116) of the Royal Society, U.K., to the second author (SKV) and from the sabbatical stay of SPV at IPICYT. We are grateful to the Nanoscience

and Nanotechnology National Research Laboratory (LINAN), Carbon Nanostructures and Two-Dimensional Systems Laboratory at IPICYT, and Dr. Emilio Muñoz-Sandoval for providing access to the required facilities. M. Abdelaly Rivera-Gómez is grateful to CTIC and DGAPA for a postdoctoral fellowship at the ICML-UNAM. Darío Torres-Sánchez and Mauricio Rosales-Rivera thank CONACYT for the doctoral fellowship. The GRM compilation was initiated long ago in our group by the participation of R. González-Ramírez although the bulk of the work was carried out by the present authors. We are grateful to the IER-UNAM library personnel for efficiently providing some of the literature materials for compilation and to Alfredo Quiroz-Ruiz for the maintenance of the computing facility at IER-UNAM. Diego Villanueva-López helped us during the pressed powder pellet preparation and for checking the correctness of the information in GRM databases.

## Supplementary Materials

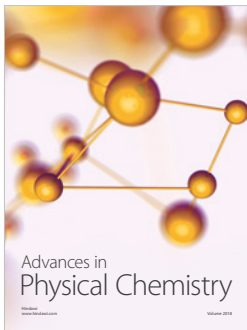
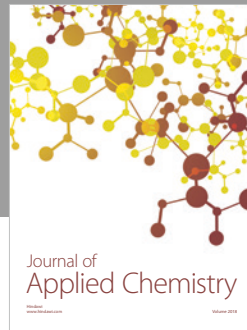
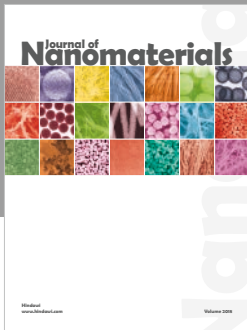
Five tables (Tables S1–S5) are provided. (*Supplementary Materials*)

## References

- [1] H. R. Rollinson, *Using Geochemical Data: Evaluation, Presentation, Interpretation*, Longman Scientific Technical, Essex, UK, 1993.
- [2] J. N. Miller and J. C. Miller, *Statistics and Chemometrics for Analytical Chemistry*, Pearson Prentice Hall, Essex, UK, 2010.
- [3] R. G. Brereton, “Chemometrics in analytical chemistry. A review,” *Analyst*, vol. 112, pp. 1635–1657, 1987.
- [4] M. Guevara, S. P. Verma, F. Velasco-Tapia, R. Lozano-Santa Cruz, and P. Girón, “Comparison of linear regression models for quantitative geochemical analysis: an example using x-ray fluorescence spectrometry,” *Geostandards and Geoanalytical Research*, vol. 29, no. 3, pp. 271–284, 2005.
- [5] P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*, Mc-Graw Hill Book Company, New York, NY, USA, 1969.
- [6] D. York, “Least squares fitting of a straight line with correlated errors,” *Earth and Planetary Science Letters*, vol. 5, pp. 320–324, 1969.
- [7] A. H. Kalantar, “Weighted least squares evaluation of slope from data having errors in both axes,” *Trends in Analytical Chemistry*, vol. 9, no. 5, pp. 149–151, 1990.
- [8] K. L. Mahon, “The New “York” regression: application of an improved statistical method to geochemistry,” *International Geology Review*, vol. 38, no. 4, pp. 293–303, 1996.
- [9] M. E. Zorn, R. D. Gibbons, and W. C. Sonzogni, “Weighted least-squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration,” *Analytical Chemistry*, vol. 69, no. 15, pp. 3069–3075, 1997.
- [10] N. R. Draper and H. Smith, *Applied Regression Analysis*, John Wiley & Sons, New York, NY, USA, 1998.
- [11] A. Schick, “Improving weighted least-squares estimates in heteroscedastic linear regression when the variance is a function of the mean response,” *Journal of Statistical Planning and Inference*, vol. 76, no. 1-2, pp. 127–144, 1999.
- [12] A. Sayago, M. Boccio, and A. G. Asuero, “Fitting straight lines with replicated observations by linear regression: the least

- squares postulates," *Critical Review of Analytical Chemistry*, vol. 34, no. 1, pp. 39–50, 2004.
- [13] A. Sayago and A. G. Asuero, "Fitting straight lines with replicated observations by linear regression: Part II. Testing for homogeneity of variances," *Critical Review of Analytical Chemistry*, vol. 34, no. 3–4, pp. 133–146, 2004.
- [14] A. G. Asuero, A. Sayago, and A. G. González, "The correlation coefficient: an overview," *Critical Review of Analytical Chemistry*, vol. 36, no. 1, pp. 41–59, 2006.
- [15] J. Tellinghuisen, "Weighted least-squares in calibration: what difference does it make?," *Analyst*, vol. 132, no. 6, pp. 536–543, 2007.
- [16] S. P. Verma, L. Díaz-González, and R. González-Ramírez, "Relative efficiency of single-outlier discordancy tests for processing geochemical data on reference materials and application to instrumental calibration by a weighted least-squares linear regression model," *Geostandards and Geoanalytical Research*, vol. 33, no. 1, pp. 29–49, 2009.
- [17] S. P. Verma, "Geochemometrics," *Revista Mexicana de Ciencias Geológicas*, vol. 29, no. 1, pp. 276–298, 2012.
- [18] S. P. Verma, *Análisis Estadístico de Datos Composicionales*, Universidad Nacional Autónoma de México, CDMX, Mexico, 2016.
- [19] P. J. Potts, *A Handbook of Silicate Rock Analysis*, Blackie, Glasgow, UK, 1987.
- [20] P. J. Potts and P. C. Webb, "X-ray fluorescence spectrometry," *Journal of Geochemical Exploration*, vol. 44, no. 1–3, pp. 251–296, 1992.
- [21] M. El Maghraoui, J.-L. Joron, J. Etoubleau, P. Cambon, and M. Treuil, "Determination of forty four major and trace elements in GPMA magmatic rock reference materials using x-ray fluorescence spectrometry (XRF) and instrumental neutron activation analysis (INAA)," *Geostandards Newsletter: Journal of Geostandards and Geoanalysis*, vol. 23, no. 1, pp. 59–68, 1999.
- [22] K. Tani, H. Kawabata, Q. Chang, K. Sato, and Y. Tatsumi, "Quantitative analyses of silicate rock major and trace elements by X-ray fluorescence spectrometer: evaluation of analytical precision and sample preparation," *Frontiers in Research of Earth Evolution*, vol. 2, pp. 1–8, 2004.
- [23] J. Enzweiler and M. A. Vendemiato, "Analysis of sediments and soils by x-ray fluorescence spectrometry using matrix corrections based on fundamental parameters," *Geostandards and Geoanalytical Research*, vol. 28, no. 1, pp. 103–112, 2005.
- [24] L. P. Bédard, "Neutron activation analysis, atomic absorption and x-ray fluorescence spectrometry review for 2004–2005," *Geostandards and Geoanalytical Research*, vol. 30, no. 3, pp. 183–186, 2006.
- [25] K. Nakayama, Y. Shibata, and T. Nakamura, "Glass beads/x-ray fluorescence analysis of 42 components in felsic rocks," *X-Ray Spectrometry*, vol. 36, no. 2, pp. 130–140, 2007.
- [26] W. Wu, T. Xu, Q. Hao, Q. Wang, S. Zhang, and C. Zhao, "Applications of x-ray fluorescence analysis of rare earths in China," *Journal of the Rare Earths*, vol. 28, pp. 30–36, 2010.
- [27] H. Mashima, "XRF analyses of major and trace elements in silicate rocks calibrated with synthetic standard samples," *Natural Resources of Environment and Humans*, vol. 6, pp. 39–50, 2016.
- [28] D. Robinson and M. C. Bennett, "XRF determination of 19 trace elements in international geochemical reference samples," *Geostandards Newsletter*, vol. 5, no. 2, pp. 175–181, 1981.
- [29] S. P. Verma, T. Besch, M. Guevara, and B. Schulz-Dobrich, "Determination of twelve trace elements in twenty-seven and ten major elements in twenty-three geochemical reference samples by X-Ray fluorescence spectrometry," *Geostandards Newsletter*, vol. 16, no. 2, pp. 301–309, 1992.
- [30] X. Wang, G. Li, Q. Zhang, and Y. Wang, "Determination of major/minor and trace elements in seamount phosphorite by XRF spectrometry," *Geostandards and Geoanalytical Research*, vol. 28, no. 1, pp. 81–88, 2004.
- [31] K. Govindaraju, "1987 compilation report on Ailsa Craig Granite AC-E with the participation of 128 GIT-IWG laboratories," *Geostandards Newsletter*, vol. 11, no. 2, pp. 203–255, 1987.
- [32] K. Govindaraju, "Report (1980) on three GIT-IWG rock reference samples: anorthosite from Greenland, AN-G; basalte d' Essey-la-Côte, BE-N; granite de Beauvoir, MA-N," *Geostandards Newsletter*, vol. 4, no. 1, pp. 49–138, 1980.
- [33] E. S. Gladney and I. Roelandts, "1987 compilation of elemental concentration data for USGS BHVO-1, MAG-1, QLO-1, RGM-1, SCo-1, SDC-1, SGR-1, and STM-1," *Geostandards Newsletter*, vol. 12, no. 2, pp. 253–262, 1988.
- [34] K. Govindaraju, "Report (1967–1981) on four ANRT rock reference samples: diorite DR-N, serpentine UB-N, bauxite BX-N, disthene DT-N," *Geostandards Newsletter*, vol. 6, no. 1, pp. 91–159, 1982.
- [35] E. S. Gladney and I. Roelandts, "1987 compilation of elemental concentration data for USGS BIR-1, DNC-1 and W-2," *Geostandards Newsletter*, vol. 12, no. 1, pp. 63–118, 1988.
- [36] S. Abbey, C. R. McLeod, and W. Liang-Guo, "FeR-1, FeR-2, FeR-3 and FeR-4 Four Canadian iron-formation samples prepared for use as reference materials," Geological Survey of Canada Report, Geological Survey of Canada, Ottawa, ON, Canada, 1983.
- [37] K. Govindaraju, "Report (1973–1984) in two ANRT geochemical reference samples: granite GS-N and Potash Feldspar FK-N," *Geostandards Newsletter*, vol. 8, no. 2, pp. 173–206, 1984.
- [38] E. S. Gladney, C. E. Burns, and I. Roelandts, "1982 compilation of elemental concentration data for the United States Geological Survey's geochemical exploration reference samples GXR-1 to GXR-6," *Geostandards Newsletter*, vol. 8, no. 2, pp. 119–154, 1984.
- [39] E. S. Gladney and I. Roelandts, "1988 compilation of elemental concentration data for USGS geochemical exploration reference materials GXR-1 to GXR-6," *Geostandards Newsletter*, vol. 14, no. 1, pp. 21–118, 1990.
- [40] K. Govindaraju, "Report (1984) on two GIT-IWG geochemical reference samples: albite from Italy, AL-I and Iron Formation sample from Greenland, IF-G," *Geostandards Newsletter*, vol. 8, no. 1, pp. 63–113, 1984.
- [41] N. Imai, S. Terashima, S. Itoh, and A. Ando, "1994 compilation of analytical data for minor and trace elements in seventeen GSJ geochemical reference samples, "igneous rock series"," *Geostandards and Geoanalytical Research*, vol. 19, no. 2, pp. 135–213, 1995.
- [42] S. Terashima, "Elemental concentrations in nine new GSJ rock reference samples," *Geostandards Newsletter*, vol. 14, no. 1, pp. 1–5, 1990.
- [43] S. Terashima, S. Itoh, M. Ujiie, H. Kamioka, T. Tanaka, and H. Hattori, "Three new GSJ rock reference samples: rhyolite JR-3, gabbro JGb-2 and hornblende JH-1," *Geostandards Newsletter*, vol. 17, no. 1, pp. 1–4, 1993.
- [44] S. Abbey, "Reference materials: rock samples SY-2, SY-3, MRG-1," Energy, Mines and Resources Canada Report, Natural Resources Canada, Ottawa, ON, Canada, 1979.

- [45] K. Govindaraju, "Report (1968-1978) on two mica reference samples: biotite Mica-Fe and phlogopite Mica-Mg," *Geostandards Newsletter*, vol. 3, no. 1, pp. 3–24, 1979.
- [46] T. W. Steele and R. G. Hansen, "Major element data (1966-1978) for the six "Nimroc" reference samples," *Geostandards Newsletter*, vol. 3, no. 2, pp. 135–172, 1979.
- [47] E. S. Gladney, E. A. Jones, E. J. Nickell, and I. Roelandts, "1988 compilation of elemental concentration data for USGS DTS-1, G-1, PCC-1, and W-1," *Geostandards Newsletter*, vol. 15, no. 2, pp. 199–396, 1991.
- [48] V. Barnett and T. Lewis, *Outliers in Statistical Data*, John Wiley & Sons, Chichester, UK, 1994.
- [49] K. Hayes, A. Kinsella, and N. Coffey, "A note on the use of outlier criteria in Ontario laboratory quality control schemes," *Clinical Biochemistry*, vol. 40, no. 3-4, pp. 147–152, 2007.
- [50] S. P. Verma, L. Díaz-González, J. A. Pérez-Garza, and M. Rosales-Rivera, "Quality control in geochemistry from a comparison of four central tendency and five dispersion estimators and example of a geochemical reference material," *Arabian Journal of Geosciences*, vol. 9, p. 740, 2016.
- [51] S. P. Verma, L. Díaz-González, J. A. Pérez-Garza, and M. Rosales-Rivera, "Erratum to: quality control in geochemistry from a comparison of four central tendency and five dispersion estimators and example of a geochemical reference material," *Arabian Journal of Geosciences*, vol. 10, p. 24, 2017.
- [52] S. P. Verma, M. Rosales-Rivera, L. Díaz-González, and A. Quiroz-Ruiz, "Improved composition of Hawaiian basalt BHVO-1 from the application of two new and three conventional recursive discordancy tests," *Turkish Journal of Earth Science*, vol. 26, no. 5, pp. 331–353, 2017.
- [53] S. P. Verma, R. Cruz-Huicochea, and L. Díaz-González, "Univariate data analysis system: deciphering mean compositions of island and continental arc magmas, and influence of underlying crust," *International Geology Review*, vol. 55, no. 15, pp. 1922–1940, 2013.
- [54] M. Rosales Rivera, *Desarrollo de herramientas estadísticas computacionales con nuevos valores críticos generados por simulación computacional*, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, Mexico, 2018.
- [55] K. P. Jochum, U. Weis, B. Schwager et al., "Reference values following ISO guidelines for frequently requested rock reference materials," *Geostandards and Geoanalytical Research*, vol. 40, no. 3, pp. 333–350, 2016.
- [56] F. Velasco-Tapia, M. Guevara, and S. P. Verma, "Evaluation of concentration data in geochemical reference materials," *Chemie der Erde*, vol. 61, no. 1, pp. 69–91, 2001.
- [57] B. Rosner, "On the detection of many outliers," *Technometrics*, vol. 17, no. 2, pp. 221–227, 1975.
- [58] F. E. Grubbs and G. Beck, "Extension of sample sizes and percentage points for significance tests of outlying observations," *Technometrics*, vol. 14, no. 4, pp. 847–854, 1972.
- [59] R. B. Jain and L. A. Pingel, "On the robustness of recursive outlier detection procedures to nonnormality," *Communications in Statistics - Theory and Methods*, vol. 10, no. 13, pp. 1323–1334, 1981.
- [60] R. B. Jain, "Detecting outliers: power and some other considerations," *Communications in Statistics - Theory and Methods*, vol. 10, no. 22, pp. 2299–2314, 1981.
- [61] M. Rosales-Rivera, L. Díaz-González, and S. P. Verma, "A new online computer program (BiDASys) for ordinary and uncertainty weighted least-squares linear regressions: case studies from food chemistry," *Revista Mexicana de Ingeniería Química*, vol. 17, no. 2, pp. 507–522, 2018.
- [62] J. B. Willet and J. D. Singer, "Another cautionary note about R2: its use in weighted least-squares regression analysis," *American Statistician*, vol. 42, no. 3, pp. 236–238, 1988.
- [63] R. M. Rousseau, "Corrections for matrix effects in X-ray fluorescence analysis—a tutorial," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 61, no. 7, pp. 759–777, 2006.
- [64] T. Shiraiwa and N. Fujino, "Theoretical calculation of fluorescent x-ray intensities in fluorescent x-ray spectrochemical analysis," *Japanese Journal of Applied Physics*, vol. 5, no. 10, pp. 886–899, 1966.
- [65] R. Rousseau and F. Claisse, "Theoretical alpha coefficients for the Claisse-Quintin relation for x-ray spectrochemical analysis," *X-Ray Spectrometry*, vol. 3, no. 1, pp. 31–36, 1974.
- [66] R. M. Rousseau, "Fundamental algorithm between concentration and intensity in XRF analysis 2-Practical application," *X-Ray Spectrometry*, vol. 13, no. 3, pp. 121–125, 1984.
- [67] R. M. Rousseau, J. P. Willis, and A. R. Duncan, "Practical XRF calibration procedures for major and trace elements," *X-Ray Spectrometry*, vol. 25, no. 4, pp. 179–189, 1996.
- [68] A. J. Klimasara, "XRF analysis—theory, experiment, and regression," in *Proceedings of X-Ray Conference (DXC) on Applications of X-Ray Analysis*, Denver, CO, USA, August 1997.
- [69] B. Tan and W. Sun, "Correction method for the matrix effect in x-ray fluorescence spectrometric analysis," *X-Ray Spectrometry*, vol. 27, no. 2, pp. 95–104, 1998.
- [70] B. I. Kitov, "Calculation features of the fundamental parameter method in XRF," *X-Ray Spectrometry*, vol. 29, no. 4, pp. 285–290, 2000.
- [71] A. J. Klimasara, "Logical steps in the automated Lachamp-Triall XRF matrix correction method utilizing an electronic spreadsheet," in *Advances in XRF Analysis*, vol. 42, pp. 53–83, JCPDS-International Centre for Diffraction Data, Newtown Square, PA, USA, 2000.
- [72] R. M. Rousseau, "Correction for long-term instrumental drift," *X-Ray Spectrometry*, vol. 31, no. 6, pp. 401–407, 2002.
- [73] G. R. Lachance and R. J. Traill, "A practical solution to the matrix problem in X-ray analysis," *Canadian Journal of Spectroscopy*, vol. 11, pp. 43–48, 1966.
- [74] J. P. Willis and G. R. Lachance, "Comparison between some common influence coefficient algorithms," *X-Ray Spectrometry*, vol. 33, no. 3, pp. 181–188, 2004.
- [75] J. P. Willis and G. R. Lachance, "A new approach to correcting theoretical emitted intensities for absorption and enhancement effects," *X-Ray Spectrometry*, vol. 33, no. 3, pp. 204–211, 2004.
- [76] J. W. Hosterman and F. J. Flanagan, "USGS reference samples Attapulgitte ATT-1 and Bentonite CSB-1," *Geostandards and Geoanalytical Research*, vol. 11, no. 1, pp. 1–9, 1987.
- [77] K. Govindaraju, "1995 working values with confidence limits for twenty-six CRPG, ANRT and IWG-GIT geostandards," *Geostandards and Geoanalytical Research*, vol. 19, pp. 1–32, 1995.
- [78] F. J. Flanagan, "Rock reference samples, san marcos gabbro, GSM-1 and lakeview mountain tonalite, TLM-1," *Geostandards and Geoanalytical Research*, vol. 10, no. 2, pp. 111–119, 1986.



Hindawi

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

