Geochemometrics

Surendra P. Verma

Departamento de Sistemas Energéticos, Centro de Investigación en Energía,
Universidad Nacional Autónoma de México,
Priv. Xochicalco s/no., Col. Centro, Temixco, Mor. 62580, Mexico.
spv@cie.unam.mx

ABSTRACT

Analogous to chemometrics, geochemometrics can be defined as the science resulting from the combination of statistics, mathematics and computation with geochemistry. This term in Spanish –geoquimiometría– has already been explicitly used in the literature. Here I elaborate on the numerous basic subjects or areas that the geochemometrics should cover. These include, but are not limited to, the following research topics: data quality, regressions, robust methods, outlier-based methods, significance tests, error or uncertainty propagation in diagrams through Monte Carlo simulation, correlation coefficient, petrogenetic modeling, and geothermometers. Equations for uncertainty propagation in analytical work have also been proposed; similarly, new indications are provided on how to calculate and report the sensitivity and limit of detection of analytical experiments. The conventional linear correlation coefficient, though useful for non-compositional data, is not recommended to be used for interpreting geochemical data. Because compositional data represent a closed unit-sum constrained system and ternary diagrams impose a further unit-sum constraint on any experimental data, these diagrams become statistically unsuitable to handle experimental data, whether compositional or of continuous variable type. Error propagation through Monte Carlo is reported for the first time to illustrate the inconvenience in using such ternary diagrams for compositional data, and an alternative log-transformed bivariate diagram is proposed to replace (or at least complement) ternary diagrams. Topics of further research have been identified, in particular, those applicable to all science and engineering fields.

Kew words: statistics, geochemistry, discrimination diagrams, Monte Carlo simulation, discordancy tests, regression, uncertainty, ternary diagrams.

RESUMEN

En forma análoga a la quimiometría, la geoquimiometría se puede definir como la ciencia que resulta por la combinación de estadística, matemáticas y computación con la geoquímica. El término en español –geoquimiometría– ha sido usado explícitamente con anterioridad en la literatura. En esta reseña, presento numerosos conceptos básicos o áreas que la geoquimiometría debe cubrir. Estos incluyen, pero no son limitados por, los siguientes tópicos de investigación: calidad de datos, regresiones, métodos robustos, métodos basados en valores discordantes, pruebas de significancia, propagación de errores o incertidumbres en diagramas mediante simulación Monte Carlo, coeficiente de correlación, modelado petrogenético, y geotermómetros. Ecuaciones para la propagación de las incertidumbres en el trabajo analítico han sido propuestas; de manera similar, se proveen detalles nuevos sobre cómo calcular y reportar la sensibilidad y el límite de detección de los experimentos analíticos. El coeficiente convencional

de correlación lineal, aunque útil para datos no-composicionales, no debe ser usado para la interpretación de datos geoquímicos. Debido a que los datos composicionales representan un sistema cerrado restringido por suma unitaria constante y los diagramas ternarios imponen la restricción adicional de suma unitaria a cualquier tipo de datos experimentales, estos diagramas vuelven a ser inapropiados estadísticamente para el manejo de los datos experimentales, sean de tipo composición o de cualquier otro tipo de datos continuos. La propagación de errores mediante Monte Carlo se reporta, por vez primera, para ilustrar la inconveniencia de usar estos diagramas ternarios para datos composicionales. Así mismo, se propone un diagrama bivariado alternativo basado en la transformación-log para reemplazar (o al menos, complementar) los diagramas ternarios. Los temas adicionales de investigación han sido identificados, en particular, aquellos aplicables a todos los campos de las ciencias e ingenierías.

Palabras clave: estadística, geoquímica, diagramas de discriminación, simulación Monte Carlo, pruebas de discordancia, regresión, incertidumbre, diagramas ternarios.

INTRODUCTION

According to Wikipedia, the free encyclopedia, the term chemometrics can be defined as "the science of extracting information from chemical systems by data-driven means. It is a highly interfacial discipline, using methods frequently employed in core data-analytic disciplines such as multivariate statistics, applied mathematics, and computer science, but to investigate and address problems in chemistry, biochemistry and chemical engineering. In this way, it mirrors several other interfacial '-metrics' such as psychometrics and econometrics."

The term chemometrics was first coined by Wold almost 40 years ago in 1972. Several reviews have been written on the subject (*e.g.*, Geladi and Esbensen, 1990; Esbensen and Geladi 1990; Lavine and Workman, 2008). Furthermore, numerous books on chemometrics (*e.g.*, Otto, 1999; Miller and Miller, 2005; Bruns *et al.*, 2006) and the journals "Journal of Chemometrics" and "Chemometrics and Intelligent Laboratory Systems" are dedicated to this subject. Other journals, such as "Analytica Chimica Acta", have a section on chemometrics.

Analogous to chemometrics, geochemometrics can be defined as the science resulting from the combination of statistics, mathematics and computation with geochemistry. This term in Spanish –geoquimiometría– has already been explicitly used in the literature (see Verma, 2005).

In the present paper, my aim was to identify the main areas of thrust to illustrate geochemometrics and point out future investigations that could lead to improvements in Earth sciences. The most important topics are as follows: Monte Carlo simulation, data quality, instrumental calibration, sensitivity and limits of detection, error propagation in ternary diagrams, discrimination diagrams, and geothermometers.

MONTE CARLO SIMULATION

The term "Monte Carlo" was first coined in 1940s after the Monte Carlo casino in Monaco. Monte Carlo methods are a class of computational algorithms that rely on repeated random sampling to compute their results. According to Hammersley and Handscomb (1964) the name and systematic development of Monte Carlo methods date back to about 1944. Because of their reliance on repeated computation of random or pseudo-random numbers, these methods are most suited to calculations by a computer. The only quality usually necessary to make good simulations is for the pseudo-random sequence to appear "random enough" as discussed by Law and Kelton (2000) and exemplified by Verma and Quiroz-Ruiz (2006a). Actually, Monte Carlo methods began to be investigated only after the availability of electronic computers from 1945 onwards. In the 1950s the early development took place in relation to the hydrogen bomb and soon became incorporated in the areas of physics, physical chemistry and operational research as well as many other scientific and engineering fields. Because Monte Carlo methods require very long sequences of random numbers, pseudorandom number generators began to be developed, which were quicker to use than the tables of random numbers previously used for statistical sampling. Güell and Holcombe (1990) presented an account of Monte Carlo techniques (experiments on random numbers) that might be useful for analytical applications.

In the area of small size sampling (up to 30) from a normal distribution, Dixon (1950, 1951, 1953) and Grubbs (1950) pioneered the field of Monte Carlo simulation by estimating critical values for their respective discordancy algorithms (Barnett and Lewis, 1994; Verma, 1997, 2005). This initial work (Dixon, 1950, 1951; Grubbs, 1950) was carried out in U.S.A. for military purposes. Nevertheless, those critical values were obviously approximate and were quoted to two to three decimal places. Later, with the availability of faster computers Grubbs and Beck (1972) extended the earlier critical values of Grubbs test to larger sample sizes of up to 147. Other workers (*e.g.*, Rosner, 1975, 1977; Prescott, 1979; Jain, 1981) also simulated critical values for single to multiple-outlier discordancy tests.

More recently in Mexico, Monte Carlo simulation has provided significant advancement for obtaining new critical values of discordancy tests for very large sample sizes up to 30,000. Thus, in a series of papers (Verma and Quiroz-Ruiz, 2006a, 2006b, 2008, 2011; Verma *et al.*, 2008a), Verma

and coworkers reported highly precise and accurate critical values for 33 discordancy test variants.

Monte Carlo methods should be useful for regressions, instrumental calibrations, and other purposes. As an example, Espinosa-Paredes *et al.* (2010) applied this approach for the evaluation of functioning of nuclear reactors. Another novel application of Monte Carlo simulation is presented below in the section of evaluation of errors in "Ternary diagrams".

DATA QUALITY

Data quality in Earth sciences in general and geochemistry in particular should be an important area of research that could be considered an integral part of the science of geochemometrics. When we talk of data quality, we are in fact considering two main aspects —precision and accuracy— of results. The first parameter depends on the quality of instrumental calibration, for which different regression models have been used, as well as of the measurement of the "unknown" material. These aspects will be further discussed in the next section. The second parameter (accuracy), on the other hand, requires adequate reference frame for its quantification. The other important parameters such as uncertainty and traceability are closely related to the precision and accuracy. The uncertainty of a parameter depends on the analytical error expressed in terms of the standard deviation, the number of measurements used for computing the standard deviation, and the Student t value for the appropriate degrees of freedom of the experiment. The section of "REGRESSIONS" below gives more information on this topic. The traceability refers to ongoing validations that the measurement of the final product conforms to the original standard of measurement; therefore, an adequate reference material or materials or synthetic standards are required.

Because geological materials, in general, represent the most complex matrices to be analyzed (theoretically, for example, rocks contain all stable and unstable or long-lived radioactive elements of the Periodic Table), inherent in the data quality are the analytical aspects such as instrumental sensitivities and limits of detection (LODs) for different elements. Sensitivities are seldom reported by researchers, but LOD reports are more common. Unfortunately, there is no consensus regarding how to quantify the LOD values (*e.g.*, see IUPAC, 1978; Ferrús and Egea, 1994; Faber and Kowalski, 1997; Kump, 1997; Mocak *et al.*, 1997; Zorn *et al.*, 1999; del Río Bocio *et al.*, 2003; Miller and Miller, 2005). According to the Vocabulary of Metrology (VIM), LOD can be defined as follows: "Measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a component in a material is β, given a probability α of falsely claiming its presence. IUPAC recommends default values for α and β equal to 0.05". Nevertheless, the estimation of LODs is generally not carried out chemometrically, *i.e.*, a sufficiently

large number n of measurements are not involved (at least n=30 has been recommended – see the following paragraph).

During the first decade of this century, Verma and coworkers, in a series of papers (Verma *et al.*, 2002, 2009a; Santoyo and Verma, 2003; Verma and Santoyo, 2003a, 2003b, 2003c, 2005; Santoyo *et al.*, 2007), demonstrated that the LODs for practically the entire Periodic Table, especially the rare-earth elements, show a systematic behavior likely governed by elemental abundance in the universe and indirectly by the well known odd-even effect of nuclear stability (the Oddo-Harkins rule; see Kaplan, 1963 or Verma *et al.*, 2009a), and that the LOD should be determined from n=30 or more measurements. Thirty measurements as a minimum are justified from the consideration of the Student t value (see table A1 in Verma, 2005), which must be taken into account for obtaining uncertainty estimates – see the section of "Sensitivity and limit of detection (LOD)".

The odd-even systematic behavior of LODs has been independently confirmed by other workers (*e.g.*, Tsakanika *et al.*, 2004; Rodríguez-Ríos *et al.*, 2007) and discussed in a review article by Bacon *et al.* (2006). More work on these lines particularly that leading to a theoretical explanation of such a systematic behavior, is highly desirable. The suggestions on how to calculate LOD are given in the next section.

To obtain accuracy estimates or traceability in the analysis of geological materials, it is mandatory to analyze appropriate, preferably certified, reference materials. Unfortunately, in spite of a large number of reference materials available for geochemical analysis (*e.g.*, see Potts *et al.*, 1992; Govindaraju, 1994; Jochum and Bruckner, 2008; Jochum and Nohl, 2008), none of them is certified for all chemical components, not even for all components of interest in a geochemical study. In fact, this was the objective set forth by one of the pioneers —K. Govindaraju— during more than two decades (1970-1995), but unfortunately, this aim was never achieved. The international community should pay proper attention to this geochemometrics aspect of fundamental research. In my opinion, instead of proposing new reference materials, we should first concentrate on a few selected reference materials to try to certify them for practically the entire Periodic Table, or at least all those chemical elements that would be useful in geochemometrics, for example, petrogenetic modeling, geothermics, or multi-dimensional discrimination diagrams.

To start with, if we could count on at least a few well-certified geochemical reference materials for most elements of geochemometric interest, we might be able to make it mandatory that the Earth science community report their analytical data adjusted to some of these certified reference materials as is customary in isotope geochemistry for reporting, for example, Sr and Nd isotopic compositions. At least, we could reach a consensus on a few geochemical reference materials, which should always be analyzed and reported in geochemical studies for quality control purposes. The very diverse matrices to be analyzed in geochemistry make this simple proposal less viable. Nevertheless, if, for

example, all laboratories analyzing basic and ultrabasic magmas were invited to report data (mean, standard deviation, and number of measurements) on just one reference material out of BIR-1, BHVO-1, or BHVO-2 from the U.S. Geological Survey (U.S.A.), or JB-1a or JB-3 from the Geological Survey of Japan (Japan) and if that particular material were well characterized (certified) for most, if not all, elements of geochemical interest, this would be a very important step forward in geochemometrics for minimizing systematic errors and adequately handling large databases such as those attempted by Agrawal *et al.* (2004) or Verma *et al.* (2006). Thus, the individual data from different laboratories or publications could be adjusted for possible bias before their use in interpretations.

REGRESSIONS

Geochemometrics could also address the topic of regressions related to instrumental calibrations for analysis of geological materials such as rock, ash, soil, mineral, water, and gas, among others. Traditionally, such calibrations have been achieved through an ordinary least-squares linear regression (OLR) model (for more details on regression techniques see the classic book by Draper and Smith, 1998). However, weighted least-squares linear regression (WLR) models should be considered more appropriate for this purpose (*e.g.*, Mahon, 1996; Baumann, 1997; Zorn *et al.*, 1997; Asuero and González, 2007). Alternatively, robust regressions might be more desirable than OLR models (*e.g.*, Hinich and Talwar, 1975; Rousseeuw and Leroy, 1987).

More recently, this adverse situation of calibrating instruments through OLR is changing especially for the calibration stage of data collection, because WLR models are being increasing applied prior to geochemical analyses (*e.g.*, Santoyo and Verma, 2003; Guevara *et al.*, 2005).

Instrumental calibrations for chemical concentration measurements can generally be expressed as the following general linear equation:

$$Resp = a(\pm se_a) + [b(\pm se_b) \cdot Conc] \qquad (1)$$

where a is the intercept term with se_a being its standard error and b is the slope (or sensitivity term) with se_b being its standard error. The concentration term (x-axis, Conc) will also have standard error of se_c, or standard deviation of s_c associated to its estimation for individual calibrator or reference material (RM), which should always be quantified. Similarly, the response term (y-axis, Resp) will have standard error of se_r, or standard deviation of s_r for each calibrator, which can be estimated.

I propose that we should try to switch from the error concept to uncertainty or confidence limits in all such considerations although it is difficult to do so, because error propagation is a very commonly used term; in reality, we are instead dealing with uncertainty propagation.

Weighted least-squares linear regression (WLR)

Our aim is to present the equations useful for calibration in the concentration-response (C-R) space, where C and R are, respectively, the x (independent) and y (dependent) variables of x-y regression line. Let us assume that we have a total number n of calibration reference materials (RMs) with concentration C_i and standard deviation S_{C_i} estimated from m_i measurements (or replicates).

I recall that before the calculation of the central tendency (*e.g.*, here C_i) and dispersion (*e.g.*, S_{C_i}) parameters, it is mandatory to ascertain that all replicate (in this case, m_i) measurements be free from discordant outliers, which can be easily done by computer program DODESSYS (*Di*scordant *Ou*tlier *DE*tection and *S*eparation *SYS*tem; Verma and Díaz-González, 2012). Both normal and log-normal distributions can be handled by the present version of DODESSYS that can handle data arrays of sizes up to 1000; future version of DODESSYS will be able to handle larger sample sizes of up to 30,000. This applies to all such situations described in this paper.

The uncertainty u_{C_i} in the concentration of ith RM can be calculated as follows:

$$u_{C_i} = (S_{C_i} / \sqrt{m_i}) \, t_{(m_i-1)} \qquad (2)$$

where $t_{(m_i-1)}$ is the Student t critical value for (m_i-1) degrees of freedom for the desired confidence level (generally 99% or 95%, two-sided), or significance level of 1% or 5% (α of 0.01 or 0.05). If the t value for the required degrees of freedom were not tabulated, it can be estimated from the interpolation equations put forth by Verma (2009), or obtained from other sources such as R Development Core Team (2009).

Similarly, let the response R_i for the ith RM have standard deviation of S_{R_i} estimated from q_i measurements. Its uncertainty u_{R_i} can then be calculated from:

$$u_{R_i} = \frac{S_{R_i}}{\sqrt{q_i}} \cdot t_{(q_i-1)} \qquad (3)$$

where $t_{(q_i-1)}$ is the Student t critical value for (q_i-1) degrees of freedom and the chosen confidence or significance level (99% or 95%, but this level should be the same as that used for equation (2) and other equations below).

In this way, we have n values of RM concentrations (C_i) and responses (R_i) with their respective uncertainties (u_{C_i} and u_{R_i}). First, we obtain the equation using OLR model as follows:

$$R = a(\pm u_a) + [b(\pm u_b) \cdot C] \qquad (4)$$

where a and b are the intercept and slope (sensitivity) terms, with respective uncertainties u_a and u_b.

Now, our aim is to estimate the total uncertainty u_i of each data point C_i-R_i used in the calibration. The basic idea is to assign the uncertainty in the x-axis variable to the y-axis variable, thus making the x-axis variable as "error-

free" and y-axis variable as having the total uncertainty in the data point under evaluation. Therefore, this can be achieved approximately as follows:

$$u_i = \sqrt{(b \cdot u_{C_i})^2 + (u_{R_i})^2} \qquad (5)$$

Because the OLR (equation 4) is not the statistically appropriate model (for the reasons see *e.g.*, Guevara *et al.*, 2005 or Verma, 2005), the weighing factors for the WLR model can be estimated from the following equation:

$$w_i = \frac{n \cdot (u_i)^{-2}}{\sum\limits_{i=1}^{n} (u_i)^{-2}} \qquad (6)$$

This equation has the following property:

$$\sum\limits_{i=1}^{n} w_i = n \qquad (7)$$

That is, the sum of all weighing factors is equal to the total number of paired data n (C_i-R_i). Thus, the WLR differs from the OLR in such a way that the weighing factors are redistributed as inversely of the total variance of the respective paired data, *i.e.*, the data point with the lowest uncertainty receives the highest weight and vice versa.

The WLR equation is:

$$R = a_w(\pm u_{a_w}) + [b_w(\pm u_{b_w}) \cdot C] \qquad (8)$$

where the slope b_w, its uncertainty u_{b_w}, intercept a_w, and its uncertainty u_{a_w} can be calculated as follows:

$$b_w = \frac{\left(\sum\limits_{i=1}^{n}(w_i \cdot C_i \cdot R_i)\right) - \left(n \cdot \overline{C}_w \cdot \overline{R}_w\right)}{\left(\sum\limits_{i=1}^{n}(w_i \cdot C_i^2)\right) - \left(n \cdot \overline{C}_w^2\right)} \qquad (9)$$

$$u_{b_w} = \sqrt{\frac{\left(\sum\limits_{i=1}^{n}(R_i - \hat{R}_{iw})^2\right)}{(n-2) \cdot \left(\sum\limits_{i=1}^{n}(C_i \cdot \overline{C}_w)^2\right)}} \qquad (10)$$

$$a_w = \overline{R}_w - (b_w \cdot \overline{C}_w) \qquad (11)$$

$$u_{a_w} = \sqrt{\frac{\left(\sum\limits_{i=1}^{n}(R_i - \hat{R}_{iw})^2\right) \cdot \sum\limits_{i=1}^{n} C_i^2}{n \cdot (n-2) \cdot \left(\sum\limits_{i=1}^{n}(C_i \cdot \overline{C}_w)^2\right)}} \qquad (12)$$

For using the above equations, the following three equations are additionally required.

The weighted centroid $\overline{C}_w$ of the concentration variable,

$$\overline{C}_w = \frac{\left(\sum\limits_{i=1}^{n}(w_i \cdot C_i)\right)}{n} \qquad (13)$$

The weighted centroid $\overline{R}_w$ of the response variable,

$$\overline{R}_w = \frac{\left(\sum\limits_{i=1}^{n}(w_i \cdot R_i)\right)}{n} \qquad (14)$$

The response $\hat{R}_w$ of the ith concentration data point (C_i-R_i) used in the calibration, which would actually correspond to the WLR equation,

$$\hat{R}_{iw} = a_w + [b_w \cdot C_i] \qquad (15)$$

The WLR calibration can then be used for measuring the response R_d of an "unknown" sample and its total uncertainty u_{R_d} and calculating its concentration C_d from the regression equation as well as its total uncertainty u_{C_d}, and not just its replication error, as is erroneously customary in chemistry or geochemistry. We recall that if we carry out r measurements of the response R_d and obtain its standard deviation as S_{R_d}, the total uncertainty will have to be first estimated from equation (16).

$$u_{R_d} = \frac{s_{R_d}}{\sqrt{r}} \cdot t_{(r-1)} \qquad (16)$$

where $t_{(r-1)}$ is the Student t critical value for (r-1) degrees of freedom and the chosen confidence level, generally two-sided 99% or 95%. This level should be the same as that used for WLR calibration.

Equation (17) below is postulated for the unknown and rearranged as equation (18) to obtain C_d and equation (19) or (20) for its total uncertainty u_{C_d}.

$$R_d(\pm u_{R_d}) = a_w(\pm u_{a_w}) + [b_w(\pm u_{b_w}) \cdot C_d] \qquad (17)$$

$$C_d(\pm u_{C_d}) = \frac{R_d(\pm u_{R_d}) - a_w(\pm u_{a_w})}{b_w(\pm u_{b_w})} \qquad (18)$$

$$u_{C_d} \approx \sqrt{\left[\frac{(u_{R_d})^2 + (u_{a_w})^2}{(R_d - a_w)^2} + \frac{(u_{b_w})^2}{(b_w)^2}\right] \cdot \frac{(R_d - a_w)}{b_w}} \qquad (19)$$

Equation (19) can also rewritten as:

$$u_{C_d} \approx \sqrt{\left[\left(\frac{u_{R_d}}{b_w}\right)^2 + \left(\frac{u_{a_w}}{b_w}\right)^2 + \left(\left(\frac{u_{b_w}}{b_w}\right)^2 \cdot \left(\frac{R_d - a_w}{b_w}\right)^2\right)\right]} \qquad (20)$$

A better, more appropriate alternative to the uncertainty propagation equation (19) would be to resort to Monte Carlo simulation of equation (18). This is due to the fact that the covariance terms, difficult to determine, are not included in equations (19) or (20) – see the approximate equality sign (≈) in these equations. Thus, once the complete experiment involving instrumental calibration and measurement of the unknown is geochemometrically done, the approximate error propagation equations (Bevington and Robinson, 2003; Verma, 2005) are no longer required in this geochemometric proposal based on Monte Carlo approach.

The statistical significance of the total uncertainty can be expressed in the following inequality, *i.e.*, the population mean μ_{C_d} of the unknown sample would lie within the confidence interval of inequality expression (21) at the chosen confidence level (99% or 95%), as follows:

$$(C_d - u_{c_d}) < \mu_{c_d} < (C_d + u_{c_d}) \qquad (21)$$

For interpreting geochemical data mostly OLR has been used, except in geochronology, for which WLR models, with errors on both variables, have been generally applied (*e.g.*, York, 1966, 1969; McIntyre *et al.*, 1966; Brooks *et al.*, 1972; Mahon, 1996) and are in wide use even today at the data acquisition and calculation stages. If the above equations, for example, equation (8) for WLR calibration and equation (18) for the unknown sample, were routinely used along with Monte Carlo simulations, we should document a significant advancement in geochemometrics, because then we can use appropriate WLR models for data interpretation as well.

On the other hand, in some applications, such as for interpolation or extrapolation purposes, more complex quadratic to higher-order polynomial models might provide better fit to the data. Besides, log-transformation of x-axis (independent or explanatory variable) may be useful in some applications as has been documented for interpolation and extrapolation of Student t critical values, in which the x-axis was the degrees of freedom (Verma, 2009). This kind of transformation could become a more common technique to be used in regressions.

Sensitivity and limit of detection (LOD)

If WLR model is appropriately applied for calibrations, the sensitivity b_w and its uncertainty u_{b_w} can be routinely reported from the regression equation (8).

The estimation of LOD from WLR can now be suggested as the most appropriate geochemometric method, modified after Mocak *et al.* (1997) and Miller and Miller (2005). Let us assume that a "blank" sample (containing no or "very little amount" of analyte) is used for estimating LOD. In most analytical instruments the blank need not contain any pre-established amount of the analyte of interest (*i.e.*, its concentration can be assumed to be zero, $C_0=0$), but in some instruments, such as chromatography, the software does not generally allow the integration of a background or blank signal, therefore, in may be necessary to use a solution containing a small amount of analyte ($C_0>0$); nevertheless, only the smallest amount or concentration that will give a measurable signal should be used. Let the blank be measured k times (where k is recommended to be at least 30 by Verma and Santoyo, 2005; for more explanation see Student t value in table A1 of Verma, 2005 and the uncertainty equations in this work, for example, equation 22; the basic idea is that the uncertainty should be obtained from a large number k

of measurements so that the t value approaches to that of infinity). The response R_0 and standard deviation S_{R_0} are estimated in the instrument under the same conditions as the calibration experiment (remember that R_0 and S_{R_0} should be calculated from discordant outlier-free data array).

The uncertainty u_{R_0} of the blank response R_0 can be calculated as follows:

$$u_{R_0} = \frac{S_{R_0}}{\sqrt{k}} \cdot t_{(k-1)} \qquad (22)$$

Then, the uncertainty u_{C_0} of the blank concentration C_0 can be calculated as follows:

$$u_{C_0} \approx \sqrt{\left[\left(\frac{u_{R_0}}{b_w} \right)^2 + \left(\frac{u_{a_w}}{b_w} \right)^2 + \left(\left(\frac{u_{b_w}}{b_w} \right)^2 \cdot \left(\frac{R_0 - a_w}{b_w} \right)^2 \right) \right]} \qquad (23)$$

Because we are dealing with the uncertainty concept, I propose that the LOD could be defined as follows:

$$LOD = C_0 + U_{C_0} \qquad (24)$$

As stated earlier, for most instruments, C_0 can be assumed as zero, in which case, LOD will be simply estimated from the total uncertainty u_{C_0}. Once again, instead of using the approximate equation (23) Monte Carlo simulation of equation (18) can be undertaken to better determine the LOD. Finally, a comparison of this newly proposed method with the conventional methods already in wide use should be undertaken, which will reinforce the new science of Geochemometrics.

ROBUST METHODS

Robust methods theoretically provide means of handling experimental data in the presence of discordant outliers, because they are considered robust against them (*e.g.*, Barnett and Lewis, 1994; Maronna *et al.*, 2006). Proponents of robust methods always claim their superiority over the outlier-based methods. For central tendency parameter estimates, many different robust statistics have been proposed such as median, mode, mean quartile, Gastwirth mean, trimean, trimmed mean, and Winsorized mean, among others (*e.g.*, Verma, 2005).

It is not clear which robust parameters should be used for a particular application. When these different statistics provide "consistent" estimates for the central tendency parameter, it is immaterial which one is used, but in practice, they may differ significantly from each other, in which case no simple answer can be given as to which statistic is better than the other. Nevertheless, if robust estimators are used for central tendency parameter, adequate robust estimators such as median deviation (also called MAD–median absolute deviation) or interquartile range should also be used for the dispersion parameter as well. The relationship of robust dispersion estimates with the respective "population" standard deviation should also be established.

Thus, what is really required is an objective evaluation of the performance of robust and outlier-based methods, so that the user can independently decide which method or statistic to rely upon and under what circumstances. In other words, is it immaterial or does it matter to use robust or outlier-based methods for handling experimental data of truly or apparently continuous variable type? Not only geochemometrics but also all other scientific and engineering areas will benefit from this proposal. Unpublished preliminary Monte Carlo simulation results by Verma and coworkers point to serious problems in indiscriminately using the median as an unbiased estimate of central tendency of asymmetrically contaminated small-sized statistical samples; these findings will be documented elsewhere.

OUTLIER-BASED METHODS

Such methods have been proposed to handle experimental data as alternative means to robust methods (Barnett and Lewis, 1994). For their statistically correct application, discordancy tests provide an important tool, such as the multiple-test method (MTM) proposed and practiced by Verma (1997). Unfortunately, most people are not even aware of the fact that the most popular statistical parameters of mean and standard deviation belong to this outlier-based category. Therefore, the common practice of calculating these two parameters, without ascertaining that the statistical sample be drawn from a normal population, should be discouraged, and the readers should be made acquainted with the advantages of MTM of Verma (1997) and convenience of a suitable computer program (Verma *et al.*, 1998; Verma and Díaz-González, 2012).

Discordancy tests with new critical values

A large number of statistical tests have been proposed in the literature to detect discordant outliers in univariate data (Barnett and Lewis, 1994). An outlier is an observation that is extreme in an ordered array of a set of univariate observations. For structured data, an outlier is similarly defined as a structure-breaking observation. Nevertheless, an outlier can be a legitimate observation pertaining to the distribution of the rest of the data in the array, or it could be evaluated as discordant from certain statistical criterion. This is the reason why Verma and Díaz-González (2012) have called their computer program as *D*iscordant *O*utlier *DE*tection and *S*eparation *SYS*tem (DODESSYS), which permits the application of 33 discordancy test variants at the strict confidence level of 99%. They also emphasized that DODESSYS should prove an important tool for applying outlier-based methods to experimental data, viz., DODESSYS must always be applied to the experimental data arrays before estimating the mean and standard deviation values. In fact, DODESSYS provides these estimates

both before and after applying the selected discordancy tests, and for this reason alone, it is a highly convenient tool for correctly applying outlier-based methods. It also allows the user to have statistical estimates of discordant outliers, which can therefore be separately interpreted.

Besides masking and swamping effects (Barnett and Lewis, 1994), the detection of discordant outliers may depend on a series of factors such as error of type I when the null hypothesis is inappropriately rejected (*e.g.*, Gawlowski *et al.*, 1998; Efstathiou, 2006) and of type II when the null hypothesis is inappropriately retained (*e.g.*, Miller and Miller, 2005).

What to do with a discordant outlier? In field studies, outliers when detected as discordant can be interpreted separately as manifestation of side events. In laboratory data, they can also arise from faulty instrumentation, inappropriate calibration, systematic errors, and large random errors, among other causes. Therefore, it is advisable to look for the actual cause of such discordancy. When an outlier is evaluated as discordant in a very small sample such as of three observations, I propose that one should carry out more experiments before fully accepting the outlier as discordant. If, by any chance, the new observations are compatible with the discordant outlier, the discordancy is once again evaluated, and either the first apparently discordant observation should be accepted as legitimate and the other observations are consequently declared as suspect values, or the experiment is further repeated to make sure the statistical outcome. If, on the other hand, the new observations are consistent with the initially dominant two observations, which should be more likely, the suspect observation can be definitely declared as discordant and excluded from further considerations. Thus, the cause of discordancy may also become clear. The discordant outliers should be isolated and not actually rejected but interpreted separately from the dominant distribution of data. The latter are thus used separately from the interpretation of the main event under study.

The discordancy procedure includes single as well as multiple outlier tests. The first group is called so, because the tests evaluate one observation at a time for its discordancy. Multiple-outlier tests for which new critical values are available (Verma and Quiroz-Ruiz, 2006a, 2006b, 2008, 2011; Verma *et al.*, 2008a) are designated as k=2, 3, or 4 types (they evaluate for discordancy two, three, or four observations at a time, respectively). Discordancy tests can be used consecutively until no more outliers are detected as discordant. Multiple outlier discordancy tests for more than four observations (k > 4; Barnett and Lewis, 1994) are also known, but new precise critical values have not been simulated.

I describe in detail the use of discordancy tests. Let us assume that we have an array of n univariate data x_i for a parameter x, which can be rearranged in ascending order as an ordered array $x_{(i)}$ where (i) varies from 1 (lowest value) to n (highest value). The observation being tested by a single-outlier test is either $x_{(n)}$ (upper outlier test – the

highest observation or datum is being tested for discordancy) or $x_{(1)}$ (lower outlier test –the lowest observation is being tested for discordancy), or any one of the two extreme observations, $x_{(1)}$ or $x_{(n)}$ depending on which one is considered more distant from the central tendency parameter (extreme outlier test –the extreme observation is being tested for discordancy). The upper or lower outlier tests were also called by Barnett and Lewis (1994) as one-sided and the extreme outlier types as two-sided. These authors, contrary to the current practice in chemistry, also opined that it is the type of test (one-sided or two-sided) that is more important rather than the one-sided or two-sided critical values. For multiple-outlier types, the observations being tested can be on one or both ends of the ordered array, for example, for k=2 type, we can test both $x_{(1)}$ and $x_{(2)}$, or $x_{(n)}$ and $x_{(n-1)}$, or even $x_{(1)}$ and $x_{(n)}$. In the case of geochemistry, however, distinction should be maintained concerning the type of analytical method used, and it is not a good idea to test for $x_{(1)}$ and $x_{(n)}$ together as a group when these two observations were obtained by different analytical methods. I further suggest that in most geochemical applications, it would be safer to use single-outlier tests instead of multiple-outlier types if one wishes to be conservative in declaring outliers as discordant.

For any statistical test, generally two hypotheses are set – null hypothesis H_0 that means that the value(s) being tested was(were) derived from a single normal distribution as the remaining observations in the ordered array and alternate hypothesis H_1 that the observation(s) being tested is(are) discordant, derived from a distribution different from that of the remaining, dominant or more numerous observations. The statistic corresponding to the given test is calculated and compared with the critical value at the chosen confidence level, being 99% according to my suggestion (see also Verma, 1997, 2005; Verma and Quiroz-Ruiz, 2006a, 2006b) or 95% according to most books in chemistry such as Miller and Miller (2005). Most discordancy tests are considered significant for "greater than", *i.e.*, if the calculated statistic is greater than the critical value, H_0 is rejected and, consequently, H_1 is accepted; in other words, the observation(s) being tested is (are) discordant outlier(s). But when the calculated statistic is smaller than the critical value, H_0 is accepted and H_1 is rejected, *i.e.*, the observation(s) tested is (are) declared as legitimate. Unfortunately, "inverse" tests also exist because they are considered significant for "smaller than" (see Barnett and Lewis, 1994; Verma, 2005). The user must therefore be careful in using discordancy tests.

Critical values are required to apply any discordancy test (Barnett and Lewis, 1994; Verma, 1997, 2005; Verma *et al.*, 1998). I recommend the use of new highly precise and accurate values published by Verma and Quiroz-Ruiz (2006a, 2006b, 2008, 2011) and Verma *et al.* (2008a) for 33 discordancy test variants. As an innovation, these new critical values were reported along with individual standard error estimates. These authors also proposed new regression

equations for computing critical values for those sample sizes that were not tabulated, thus permitting the application of these discordancy tests for all sample sizes up to 30,000. Artificial neural network (ANN) was used by Verma *et al.* (2008a) for arriving at best-fitted regression equations. Soon afterwards, Verma and Quiroz-Ruiz (2008) used Statistica software to investigate alternative polynomial fitting in conjunction with natural logarithm transformation of sample sizes. Best-fitted polynomial equations were thus reported and favorably compared with the equations obtained by ANN. More recently, Verma and Quiroz-Ruiz (2011) clarified that the critical values for skewness test N14 published earlier by them were of one-sided type and reported more precise and accurate two-sided critical values for this test.

I present the variation of critical values with sample size for one-sided Grubbs type test N1 and Dixon test N7 in Figure 1 and for powerful two-sided skewness and kurtosis tests N14 and N15 in Figure 2. The dependence of critical value on the sample size for all tests is so high (Figures 1a, 1c, 2a, and 2c) that polynomial regression in these diagrams does not provide any satisfactory fit to the data (see Verma and Quiroz-Ruiz, 2008). Natural logarithm-transformation of the x-axis (sample size) results in significant "smoothing" of the curves (compare the earlier diagrams with Figures 1b, 1d, 2b, and 2d, respectively), which enabled Verma and Quiroz-Ruiz (2008) to propose best-fit polynomial equations for interpolation and extrapolation of critical values from 100 up to 30,000. For all smaller sample sizes up to 100, precise critical values have been simulated, so there is no need for interpolation equations. Note, however, that the complex nature of curves (Figures 2b and 2d) even after log-transformation will not allow best-fit polynomial equations to be proposed for the entire range of sample sizes from 5 to 30,000. These are the reasons why Verma and Quiroz-Ruiz (2008) proposed equations for sample sizes of 100 to 30,000 (and not for 5 to 30,000). Nevertheless, it is now possible to apply any of the 33 discordancy tests to practically any kind of experimental data without having any limitation on the sample sizes.

Another important point to note concerns the horizontal dotted lines denominated as 2s and 3s in Figure 1a. They represent, respectively, the two- and three-standard deviation (2s and 3s) methods also used in the literature as discordancy methods based on "population" criteria, according to which all observations lying outside the range of mean±2s or mean±3s (2s and 3s methods, respectively) are simply rejected as discordant. The correct statistical procedure for small size sampling as carried out in most experiments would be equivalently the Grubbs test (N1) applied, respectively, at 95% and 99% confidence levels. Unfortunately, such statistically erroneous methods (see Barnett and Lewis, 1994), for example, 2s, have been applied in the literature (*e.g.*, Gladney *et al.*, 1992; Imai *et al.* 1995). Their use should, however, be abandoned in favor of Grubbs test N1 as has already been suggested by Verma (1998a), Verma and Quiroz-Ruiz (2006b), and Verma *et al.*
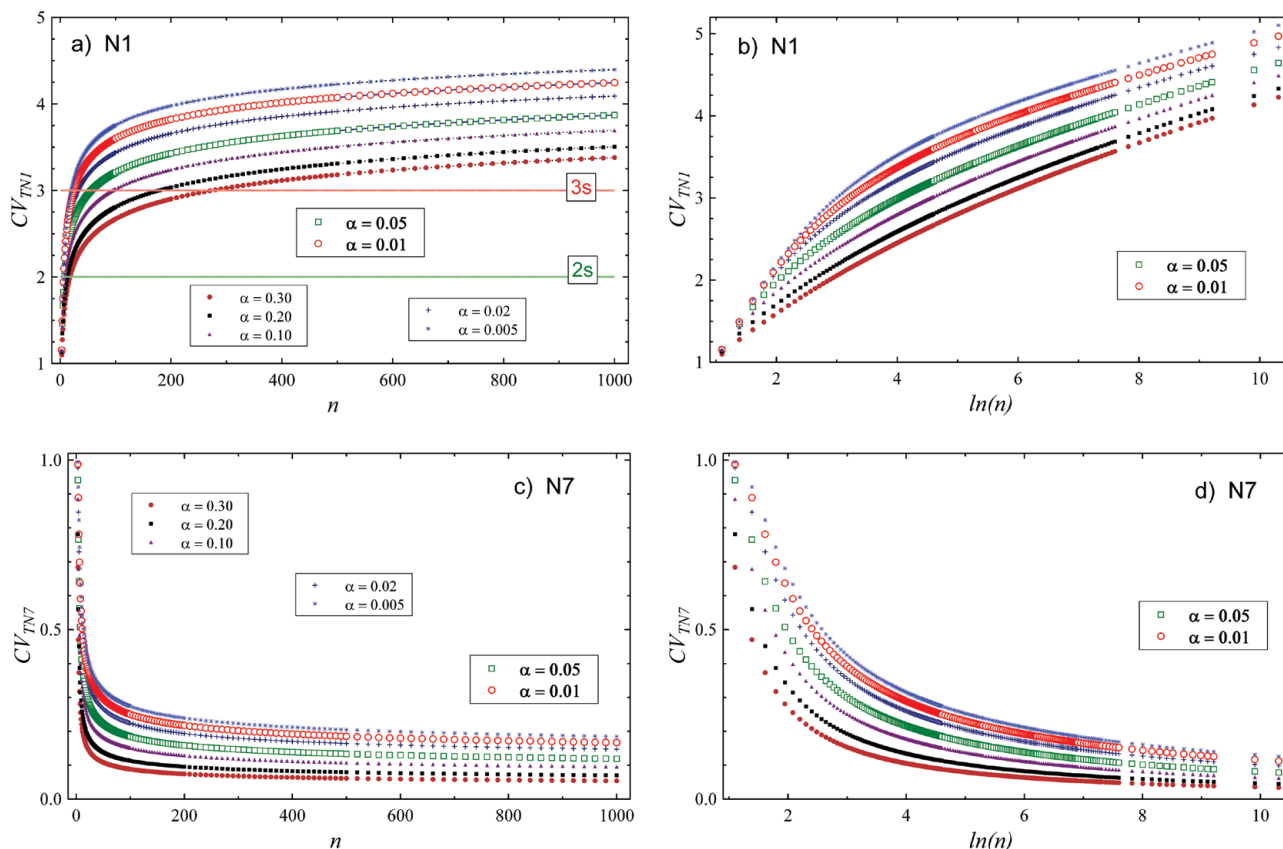
Figure 1. New critical values for tests N1 (Grubbs) and N7 (Dixon) for significance level (α) of 0.30, 0.20, 0.10, **0.05**, 0.02, **0.01**, and 0.005, plotted against sample size of up to 1000, or log-transformed sample size of up to 30,000. Note that the most commonly used α of 0.05 (Miller and Miller, 2005) and 0.01 (Verma, 1997, 2005) are highlighted. All symbols are explained as inset in diagrams (a) and (c). (a) Test N1 with sample size up to 1000 as x-axis, horizontal dotted lines named 2s and 3s represent, respectively, the two- and three-standard deviation methods also used in the literature as discordancy methods based on population criteria; (b) Test N1 with natural logarithm-transformed sample size up to 30,000 as x-axis, note significant reduction in the curvature of the critical value trends in this diagram as compared to (a); (c) Test N7 with sample size as x-axis; (d) Test N7 with natural logarithm-transformed sample size as x-axis, note significant reduction in the curvature of the critical value trends in this diagram as compared to (c).

(2008a). Hayes *et al.* (2007) have also independently criticized and discarded such standard deviation methods based on population criteria. The 1s method sometimes practiced for handling experimental data should be considered worse and statistically erroneous than the 2s or 3s methods.

Discordancy tests without new critical values

A few more discordancy tests, viz., Tietjen and Moore´s statistic (Tietjen and Moore, 1972), Shapiro and Wilk´s statistic (Shapiro and Wilk, 1965; Shapiro *et al.*, 1968), two-sided test for extreme outlier using a robust estimator of standard deviation (Iglewicz and Martinez, 1982; no tabulated critical values are available), and consecutive or recursive test of multiple outliers (Rosner, 1975, 1977; Jain, 1981), have also been proposed, but only old less precise critical values (generally accurate to two decimal places only) are available for their application.

In fact, Tietjen and Moore´s procedure (Tietjen and Moore, 1972) is similar to the Grubbs tests of $S^2_{(n)}/S^2$ to

$S^2_{(n), (n-1), (n-2), (n-3)}/S^2$ types (test N4 k=1 to k=4 types; see Verma, 2005 for more details on test N4 statistics), and the new critical values simulated by Verma and Quiroz-Ruiz (2006b, 2008, 2011) and Verma *et al.* (2008a) are applicable when the outlying observations being tested are on either end of the ordered data array. For k=5-10, only approximate critical values are at present available (Tietjen and Moore, 1972). The multiple or many outlier test RST of Rosner (1975, 1977) is also similar to the Grubbs statistics N2 and N3 (see Verma, 2005), with the difference that RST should be computed from trimmed mean and trimmed standard deviation values.

Therefore, new more precise and accurate critical values are required for completing the MTM of Verma (1997) and significantly improving this line of geochemometric research.

Use of discordancy tests

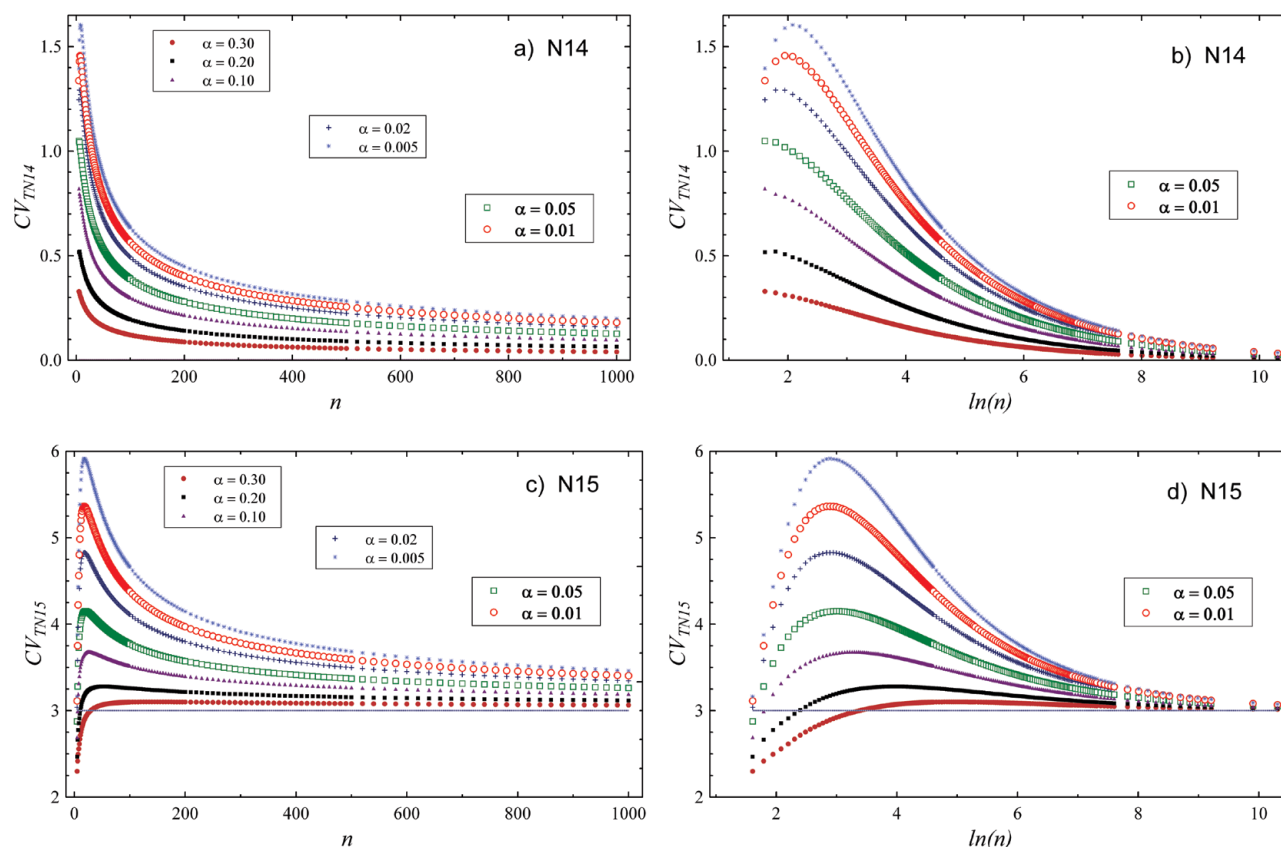The currently available precise critical values of 33

Figure 2. New critical values for tests N14 (skewness) and N15 (kurtosis) for significance level (α) of 0.30, 0.20, 0.10, **0.05**, 0.02, **0.01**, and 0.005, plotted against sample size of up to 1000, or log-transformed sample size of up to 30,000. Note that the most commonly used α of 0.05 (Miller and Miller, 2005) and 0.01 (Verma, 1997, 2005) are highlighted. All symbols are explained as inset in diagrams (a) and (c). (a) Test N14 with sample size up to 1000 as x-axis; (b) Test N14 with natural logarithm-transformed sample size up to 30,000 as x-axis, note significant reduction in the curvature of the critical value trends in this diagram as compared to (a); (c) Test N15 with sample size as x-axis; (d) Test N15 with natural logarithm-transformed sample size as x-axis, note significant reduction in the curvature of the critical value trends in this diagram as compared to (c).

discordancy tests have been used, in conjunction with the MTM (Verma, 1997), by numerous researchers in their respective applications. Just to cite a few recent ones, these are: Armstrong-Altrin (2009); Gómez-Arias *et al.* (2009); Marroquín-Guerra *et al.* (2009); Pandarinath (2009a, 2009b, 2011); Viner *et al.* (2009); Álvarez del Castillo *et al.* (2010); Madhavaraju *et al.* (2010); Najafzadeh *et al.* (2010); Torres-Alvarado *et al.* (2011); Verma *et al.* (2011a); and Zeyrek *et al.* (2010).

It is not clear if we should apply the concept of discordant outliers to raw compositional data without any transformation as done by Verma and coworkers (*e.g.*, Verma, 1997, 1998a, 2005; Velasco and Verma, 1998; Velasco *et al.*, 2000; Guevara *et al.*, 2001; Velasco-Tapia *et al.*, 2001, Verma and Quiroz-Ruiz, 2008; Marroquín-Guerra *et al.*, 2009; Verma *et al.*, 2009a). Verma and Agrawal (2011) and Verma S.K. *et al.* (2012), on the other hand, used discordancy tests to evaluate natural logarithm of element ratios for discordant outliers, prior to the application of linear discriminant analysis to their compiled data.

For evaluation of compositional data, it is possible that some kind of transformation is required prior to the

application of discordancy tests. Although this should be the subject of future research in geochemometrics, for now may I suggest that the log-transformed ratio data, rather than the element concentrations, should be evaluated for discordancy.

Finally, discordancy tests can also be applied during the data acquisition stage of mass spectrometric determinations. Such an application of Dixon tests was reported by Dougherty-Page and Bartlett (1999), although unfortunately it is not a widespread practice explicitly reported in the literature. As an example of unpublished cases, mass spectrometric software in the Geochemistry department of the Max-Planck-Institut für Chemie in Mainz, Germany, allows the application of Dixon tests before the data are printed out from the instrument. More importantly, for correcting inter-laboratory bias in isotopic determinations, all laboratories are supposed to report results of isotopic measurements on established reference materials, such as Eimer & Amend Sr carbonate and more recently, National Bureau of Standards NBS 987 for [87]Sr/[86]Sr (Faure, 2001). Similarly, for Nd isotopic measurements it is customary to report [143]Nd/[144]Nd values obtained on the

La Jolla standard (*e.g.*, Verma, 1992). This practice helps eliminate the systematic errors (inter-laboratory bias) in Sr and Nd isotopic determinations, especially when using data from different laboratories for interpretation of geological processes.

It is also customary in isotopic studies that the analytical errors on isotopic data be individually reported (see, *e.g.*, Verma, 1992). Nevertheless, from the geochemometrics point of view, the shortcoming seems to reside in the fact that the individual analytical errors are reported as two times the standard error of the mean ($2\sigma_E$; note the wrong notation based on population, it should actually be $2s_E$) and not as total within-run uncertainty based on the Student t value. Reporting of simply the standard deviation value without mentioning the total number of measurements is even a more severe mistake. The statistically correct report cannot be easily prepared for the literature data, because the total number of measurements, from which the standard error was calculated, is seldom reported in published literature.

I propose that the geochemometrically correct way to report these individual within-run errors as the uncertainty u_{samp} would be as follows:

$$u_{samp} = \frac{S_{samp}}{\sqrt{p}} \cdot t_{(p-1)} \qquad (25)$$

where s_{samp} is the standard deviation based on p determinations of isotopic ratio in that particular sample and $t_{(p-1)}$ is the two-sided Student t value at 95% or 99% confidence limits. It is needless to say that the isotopic mean ratio and its standard deviation should be calculated only after ascertaining the absence of discordant outliers in the original data array. Only then, this branch of geochemistry will be fully consistent with geochemometrics.

SIGNIFICANCE TESTS

Significance tests (Student t, Fisher F, one-way ANOVA, and two-way ANOVA) are not routinely applied for the interpretation of geochemical data, although some books on geosciences do recommend their use (*e.g.*, Jensen *et al.*, 1997; Verma, 2005). If we were to accept geochemometrics as the emerging science, significance tests should become an integral part of data evaluation in Earth sciences. However, note that these tests require that individual statistical samples be normally distributed. Therefore, DODESSYS (Verma and Díaz-González, 2012) should prove an important tool for the application of significance tests, *i.e.*, for assuring that the basic assumption of normal distribution is complied.

We are attempting to make geochemometrics a reality by reinterpreting some of the published data (*e.g.*, Hernández-Martínez and Verma, 2009) so that the Earth science community could compare and contrast these new geochemometric interpretations with those put forth in the respective original papers. Precise critical values are helpful

in this respect. These can be obtained from freely available software R after proper programming, or can be consulted directly in Verma (2009) for interpolation equations.

Once again, it is not clear if significance tests should be applied to crude compositional data or log-transformed ratios, although it might be desirable to do so to transformed variables.

DIAGRAMS IN GEOCHEMISTRY

Numerous bivariate, ternary or multi-element diagrams are used in geochemistry. However, they should be evaluated from geochemometrics point of view.

Bivariate diagrams

First, I discuss problems with conventional bivariate diagrams (diagrams with two axes) in geochemistry and point out statistical solution to these problems. Such diagrams have been widely used in geochemistry. However, for compositional variables there may be problems when these diagrams are used for geochemical concentrations of chemical elements to draw statistical conclusions. Long ago, Chayes (1960, with more than 210 cites in international journals as judged from the Institute for Scientific Information database) had pointed out difficulties in the use of crude compositional variables. These problems of compositional variables (closure problem and constant sum effect) were later stressed by Aitchison (1982, 1984; these papers with more than 280 cites in international journals). Aitchison, in his pioneering work (Aitchison, 1986; cited more than 940 times in international journals), also proposed solutions to overcome the difficulties of constant sum and closed compositional space of crude variables. He noted that, instead of using crude compositions, one must think in terms of a multivariate approach by calculating compositional ratios having a common denominator and then working in logarithms of these ratios. The division eliminates the compositional units, which may be wt% or %m/m, or μg/g, and renders the compositions as simple numbers opening up the space, *i.e.*, absolute magnitudes are converted in relative magnitudes. The log-transformation of ratios opens up the space theoretically to infinity, in the positive or negative direction, or both, depending on the nature of the common denominator used. The natural logarithm or other kind of logarithm can be used for log-transformation.

More recently, many researchers (*e.g.*, Egozcue *et al.*, 2003; Aitchison and Egozcue, 2005; Buccianti *et al.*, 2006; Verma, 2010; Verma *et al.*, 2010) have stressed the need of abandoning simple bivariate diagrams and using Aitchison's approach in geosciences.

Unfortunately, the most well known among the bivariate diagrams are the so called Harker diagrams based on silica as the x-variable and other major- and trace-elements as the y-axis variables, or Harker type diagrams, in which

a compositional variable other than SiO_2, *e.g.*, MgO, might be used as the x-variable. The geochemical literature is full of such diagrams, which are used to draw statistical inferences. The basic problem with these diagrams is that there is an inherent negative correlation of other chemical variables with SiO_2 because of the constant or unit sum constraints (*e.g.*, Chayes, 1978). In fact, this author additionally showed that even a positive correlation of some variables with SiO_2 is also possible. The existence of negative or positive correlation in these diagrams is routinely used to infer about geological processes, and the inherent statistical correlation from closed sum concept pointed out above is not even mentioned, nor is it taken into account. The reader can, therefore, readily see from the above discussion that the Harker or Harker-type diagrams should not be used any more to draw statistical inferences, or should be used with caution.

Other diagrams obviously unfit for the purpose of drawing statistical inferences are those in which a common variable in both axes is used (see Reyment and Savazzi, 1999, for more discussion), such as A–A/B or A–B/A type diagrams where A and B are two chemical elements.

The new science of geochemometrics should emphasize on this shortcoming of diagrams and popularize the statistically correct solutions.

Correlation coefficient in bivariate diagrams

The simple concept of Pearson´s linear correlation coefficient (r) for compositional data again seems to be irrelevant for interpretation in geochemistry, and should be replaced by the proportionality concept (Aitchison, 1986; Reyment and Savazzi, 1999). For normal, *i.e.*, full-space data there is no problem in using the conventional r. The concept of log-ratio variance is of use in this respect. If we estimate the variance of log-transformed ratios of two elements i and j for a set of samples, we can use this "relative variance" value $(\text{var}\{\log(x_i/x_j)\})$ as an indicator of correlation. If the relative variance value approaches zero, there is a perfect relationship between A and B. Remember here the sizes of the samples or specimens is probably irrelevant. In other words, we can replace the concept of perfect correlation by that of perfect proportionality. Greater values of relative variance express greater departure from the perfect proportionality between parts or components under study (A and B). When the relative variance approaches ∞, the concept of a complete lack of proportionality becomes applicable.

To make this new concept amenable to everyone, Aitchison (1997 cited in Reyment and Savazzi, 1999) introduced a finite scaling transformation as a measure of the relationship between two parts. This scale runs from 0, which signifies a lack of proportional relationship, to 1 that corresponds to a perfect proportional relationship. It requires at least three parts for the computation of the proportionality

measure, but has the disadvantage that association cannot be identified as negative or positive. These concepts have yet to be incorporated in geoscientific research. In the mean time, I suggest that the conventional r can be used for evaluating log-ratio transformed compositional data, and not crude compositions.

Ternary diagrams

Ternary diagrams representing three variables on a plane (two-dimensions, triangular space) are invariably used in many fields of Earth sciences; to cite a few, these are: analytical petrology (Ragland, 1989), environmental chemistry (Andrews *et al.*, 2004), gas geochemistry (Ottonello, 1997), geothermometry of geothermal fluids (Nicholson, 1993; Arnórsson, 2000), granite petrogenesis (Rollinson, 1993; Hall, 1996), groundwater chemistry and classification (Freeze and Cherry, 1979; Appelo and Postma, 1993), igneous rock classification (Rollinson, 1993; Le Maitre *et al.*, 2002), igneous and metamorphic petrology (Spear, 1995; Hall, 1996; Young, 1998), phase diagrams and thermodynamics (Nicholls and Russell, 1990; Tatsumi and Eggins, 1995; Young, 1998; Gasparik, 2003), sedimentary petrography, petrology and provenance (Taylor and McLennan, 1985), tectonomagmatic discrimination (Rollinson, 1993), and even chemometrics (Bruns *et al.*, 2006).

Some workers such as Chayes (1960) and Aitchison (1986) discouraged the use of simple geochemical compositions in bivariate diagrams, but ironically recommended the use of ternary diagrams (Chayes, 1965, 1985; see also Aitchison, 1986). If the proposed ternary diagrams are constructed by the way they are, *i.e.*, by recalculating proportions to 100% sum of the three variables, they are likely to be bound by the problems point out in the present work even if they are based on log-transformed variables. The only application in which these adverse effects are not of much significance will be when the experimental errors or uncertainties in the three ternary variables are exceedingly small or negligible, which is not likely for the compositional variables generally used in such diagrams, particularly trace-elements. Even with the modern analytical techniques, the total propagated uncertainties (combined calibration and measurement uncertainties for the "unknown" samples) for these elements in geological materials are likely to be large enough for serious consequences in ternary diagrams. Effects of analytical errors or uncertainties on individual samples in a ternary diagram are not known. This may be the reason why some workers (*e.g.*, Presnall, 1969; Chayes, 1985) have proposed the use of ternary diagrams as the only choice to visualize and interpret certain kinds of data.

Ternary diagrams are so frequently used that there are tens of thousands of references in published literature. This generalised use is disappointing in view of some existing studies (Butler, 1979; Philip *et al.*, 1987; Howard, 1994), which point out problems with such diagrams. Statistical

summary in ternary diagrams is modified, and genetic inferences may be biased from interactions of other factors that cannot be easily separated from petrogenetic controls (Butler, 1979), for which they are frequently used, for example, the well known AFM (alkalis-iron-magnesium) diagram (Rollinson, 1993) in geochemistry and igneous petrology. The use of ternary diagrams for the comparison of sample sets must also be viewed with caution (Philip *et al.*, 1987). It has also been suggested that instead of error polygons, confidence intervals representing total uncertainty estimates should be used to visualize statistically significant differences between means (Howard, 1994).

To the best of my knowledge, however, no study has yet been reported on correctly propagated errors (or uncertainties) from three individual errors (or uncertainties) of the variables or components used to construct such ternary diagrams. Existing studies (Howard, 1994) on error propagation have been even incorrect because of the unaccounted covariance terms that result from the basic mathematics to construct these ternary diagrams (Bevington and Robinson, 2003; Verma, 2005). Using Monte Carlo simulation, involving very large repetitions of 100,000, I demonstrate, for the first time, the inherent problem of error distortion and visual amplification or reduction in these very frequently-used ternary diagrams.

I report as examples the results of two case studies or error models involving a total of 25 data points with heteroscedastic errors (unequal standard deviations) characterized by an equal relative standard deviation (RSD) simple model and a more realistic, unequal RSD, complex model (see models 1 and 2, respectively, in Table 1). Cases of homoscedastic errors (equal standard deviations independent of the mean values), being unrealistic in Earth sciences and chemistry, are not considered.

Construction of a ternary diagram A-B-C from three measured variables A_m, B_m and C_m, with their respective standard deviation estimates of S_{A_m}, S_{B_m} and S_{C_m}, involves three analogous equations. I present only one such equation (equation 26). The first ternary variable A can be calculated from:

$$A = \frac{A_m}{A_m + B_m + C_m} \times 100 \qquad (26)$$

Even when the initial variables (A_m, B_m and C_m) are not correlated, *i.e.*, their covariance can be neglected, the recalculated ternary variables A, B and C will necessarily be correlated (Bevington and Robinson, 2003), *i.e.*, their covariance terms must be taken into account to estimate final uncertainties of the transformed variables that are constrained within the closed triangular space of ternary diagrams. Unfortunately, the equations to handle the propagated errors are strictly valid only for population standard deviation (σ), provided all terms (quadratic, cubic, etc.) are taken into consideration. For the sample standard deviation (s) being an estimate of σ, the equations are even more approximate (Verma, 2005). One such highly approximate equation for calculating the variance (s_A^2) of the ternary

variable A, which takes into consideration the covariance terms $(S_{(A_m)(B_m)(C_m)})^2$ and $(S_{(A_m)(A_m B_m C_m)})^2$ and, is as follows:

$$s_A^2 \approx A^2 \left[\left(\frac{s_{A_m}}{A_m} \right)^2 \right.$$
$$+ \left(\frac{(s_{A_m})^2 + (s_{B_m})^2 + (s_{C_m})^2 + 2(s_{(A_m)(B_m)(C_m)})^2}{(A_m + B_m + C_m)^2} \right)$$
$$\left. - \left(\frac{2(s_{(A_m)(A_m B_m C_m)})^2}{(A_m)^2(A_m + B_m + C_m)^2} \right) \right] \qquad (27)$$

Use of such complex approximations is not recommended (Verma, 2005).

Therefore, to achieve the objective of correctly and efficiently evaluating ternary diagrams and proposing a statistically viable alternative, I resorted to Monte Carlo simulation, in which, as a first step, a series of independent and identically distributed (IID) random variates –U(0, 1)– uniformly distributed in the space (0, 1), were generated (Law and Kelton, 2000). These were tested for randomness, transformed to normal random variates N(0, 1) (Verma and Quiroz-Ruiz, 2006a), and used for simulation of total propagated errors during the construction of a ternary diagram.

As examples, model 1 (equal RSD) assumes initial errors in the measured variables A_m, B_m and C_m as 5% RSD. Model 2 (unequal RSD) is based on a more realistic case of 1%, 3% and 5% RSD, respectively and additionally, on equation 28 (Thompson, 1988) for the first variable A_m as follows:

$$(s_{A_m})^2 = \left[\left(\frac{LOD_A}{3} \right)^2 + \left(cv_A \times A_m \right)^2 \right] \qquad (28)$$

where the limit of detection (LOD_A) is three times the standard deviation at zero concentration (IUPAC, 1978) of the variable A_m and cv_A is taken to be a constant (being the coefficient of variation for relatively large values of A_m as compared to LOD_A). I note that the use of the new equations for LOD proposed in this work will not change the inferences.

For model 2, the following values (arbitrary units) were set: $LOD_A = 0.015$, $LOD_B = 0.025$, $LOD_C = 0.030$, $cv_A = 0.01$, $cv_B = 0.03$, $cv_C = 0.05$ and These LOD values were assumed so that the quantification limits, being approximately three times the LODs, could be less than the smallest experimentally measured values of A_m, B_m and C_m, used to illustrate these findings. The cv_A, cv_B, cv_C values correspond to the RSD values of 1%, 3%, and 5%, respectively. This allowed me to reasonably model the low concentration values of the variables corresponding to the 25 data, especially those involving 0.1 unit of a measured variable (the last three rows in Table 1).

Values of s_{A_m} (equation 28) and s_{B_m} and s_{C_m} (analogous equations) for each of the 25 data were then calculated and

Table 1. Twenty-five data involving three variables and the synthesis of statistical information for ternary and bivariate natural log ratio-transformed plots.

| Experimental data (assumed cases) | | | Propagated relative standard deviation (%) in ternary plot | | | | | | Transformed experimental data for assumed cases | | Propagated standard deviation in bivariate plot | | | |
| | | | Model 1 (Figure 3a) | | | Model 2 (Figure 3b) | | | | | Model 1 (Figure 3c) | | Model 2 (Figure 3d) | |
A_m	B_m	C_m	A	B	C	A	B	C	$\ln\left(\dfrac{A_m}{C_m}\right)$	$\ln\left(\dfrac{B_m}{C_m}\right)$	$\ln\left(\dfrac{A_m}{C_m}\right)$	$\ln\left(\dfrac{B_m}{C_m}\right)$	$\ln\left(\dfrac{A_m}{C_m}\right)$	$\ln\left(\dfrac{B_m}{C_m}\right)$
47.5	47.5	5	3.56	3.56	5.84	1.54	1.67	4.99	2.2513	2.2513	0.0711	0.0711	0.0511	0.0585
45	45	10	3.60	3.60	5.53	1.54	1.79	4.72	1.5041	1.5041	0.0711	0.0711	0.0515	0.0589
40	40	20	3.76	3.76	4.91	1.67	2.10	4.20	0.6931	0.6931	0.0711	0.0711	0.0512	0.0586
33.33	33.33	33.33	4.10	4.10	4.10	2.05	2.63	3.50	0.0000	0.0000	0.0711	0.0711	0.0511	0.0586
27.5	27.5	45	4.50	4.50	3.38	2.50	3.15	2.89	-0.4925	-0.4925	0.0711	0.0711	0.0511	0.0586
20	20	60	5.12	5.12	2.46	3.17	3.86	2.11	-1.0986	-1.0986	0.0711	0.0711	0.0511	0.0586
12.5	12.5	75	5.82	5.82	1.54	3.88	4.60	1.32	-1.7918	-1.7918	0.0711	0.0711	0.0511	0.0586
7.5	7.5	85	6.32	6.32	0.93	4.37	5.11	0.79	-2.4277	-2.4277	0.0711	0.0711	0.0511	0.0586
2.5	2.5	95	6.85	6.85	0.31	4.88	5.63	0.26	-3.6376	-3.6376	0.0711	0.0711	0.0515	0.0594
75	20	5	1.63	5.51	6.16	0.70	2.53	4.85	2.7081	1.3863	0.0711	0.0711	0.0515	0.0589
70	10	20 *	1.88	5.81	5.36	1.08	2.97	4.07	1.2528	-0.6931	0.0711	0.0711	0.0511	0.0586
20	70	10 *	5.36	1.88	5.81	2.31	1.05	4.98	0.6931	1.9459	0.0711	0.0711	0.0512	0.0586
5	60	35 **	5.90	2.68	4.45	2.69	2.13	3.73	-1.9459	0.5390	0.0711	0.0711	0.0512	0.0586
5	35	60 **	5.90	4.45	2.68	3.32	3.59	2.27	-2.4849	-0.5390	0.0711	0.0711	0.0512	0.0586
5	85	10	6.42	0.94	6.23	2.78	0.68	5.18	-0.6931	2.1401	0.0711	0.0711	0.0513	0.0586
25	10	65	5.00	5.71	2.22	3.36	4.25	1.80	-0.9555	-1.8718	0.0711	0.0711	0.0511	0.0586
50	15	35	3.15	5.25	4.18	1.87	3.14	3.32	0.3567	-0.8473	0.0711	0.0711	0.0511	0.0586
20	55	25	5.03	2.78	4.77	2.22	1.86	4.11	-0.2231	0.7885	0.0711	0.0711	0.0511	0.0586
41	9	50	3.91	5.60	3.28	2.58	3.74	2.55	-0.1985	-1.7148	0.0711	0.0711	0.0511	0.0586
16	42	42	5.16	3.68	3.68	2.59	2.74	3.17	-0.9651	0.0000	0.0711	0.0711	0.0511	0.0586
60	25	15	2.49	4.88	5.37	1.13	2.45	4.36	1.3863	0.5108	0.0711	0.0711	0.0511	0.0586
97	1	2	0.19	6.98	6.93	0.11	3.24	5.01	3.8816	-0.6931	0.0711	0.0711	0.0533	0.0655
12.9	87	0.1	6.19	0.92	6.69	2.76	0.41	11.47	4.8598	6.7685	0.0711	0.0711	0.3607	0.3619
99.8	0.1	0.1	0.01	7.11	7.10	0.01	5.20	8.89	6.9058	0.0000	0.0711	0.0711	0.3607	0.4579
0.1	0.2	99.7	7.10	7.10	0.02	7.17	7.20	0.02	-6.9048	-6.2116	0.0711	0.0711	0.1636	0.1411

The selection of these 25 data was aimed to cover the entire ternary diagram, keeping in mind that these diagrams show certain symmetrical features. Model 1: simple model with 5% RSD assumed for the experimental data; Model 2: more realistic, complex model with 1%, 3%, and 5% RSD assumed for experimental variables A_m, B_m, and C_m, respectively, along with equation (28) for the error dependence on concentration. The data points identified by * and ** confirm the reproducibility of the simulation procedure (see exactly the same results for these pairs of simulations for model 1; note because model 2 simulates more complicated unequal error structure, the results are not expected to be the same).

used in the simulations for model 2. For model 1, however, these calculations were straight forward. Further, to facilitate a visual comparison of the initial data, the three measured variable mean values for all 25 data were assumed to sum up to 100 (see the first three columns in Table 1). The results would not change even if these measured data summed up to values significantly different from 100, because they will necessarily sum up to 100 after the ternary transformations (equation 26 and analogous ones not presented).

The results (Table 1 and Figures 3a and 3b) show significant error distortion in ternary diagrams. The size of the symbols would correspond approximately to 99% confidence limits (total uncertainty estimates) provided each variable were measured about 10 times. If the individual data (mean values of the 25 data under consideration; Table 1) were obtained from a smaller number of measurements (<10), the uncertainty estimates would be greater than those represented in Figures 3a and 3b, depending on the corresponding Student t critical values.

Very large repetitions of 100,000 were used to best represent the size and shape of the data symbols, which would remain practically the same for any smaller or larger repetitions, such as 10,000 or 1,000,000. Only the density of the simulated data symbols will accordingly change.

For model 1 with equal RSD of 5%, the recalculations essential for constructing ternary diagrams (equation 26 and analogous ones not presented) result in totally different RSD values (0.01% to 7.11% in Table 1; see the three columns listed under "Model 1 (Figure 3a)", in which none of the calculated RSD values is equal to the initial 5% RSD), with considerable distortion of symbol shapes (Figure 3a). When two of the three measured variables (A_m, B_m and C_m) are equal for a given ternary datum, the new propagated RSD for these two variables would also be the same, but different from the initial RSD values (*e.g.*, see the rows identified by * and ** and the first 9 rows of results in Table 1, the latter correspond to the data plotted in the vertical direction in Figure 3a). Only for the exceptional case of equal A_m, B_m
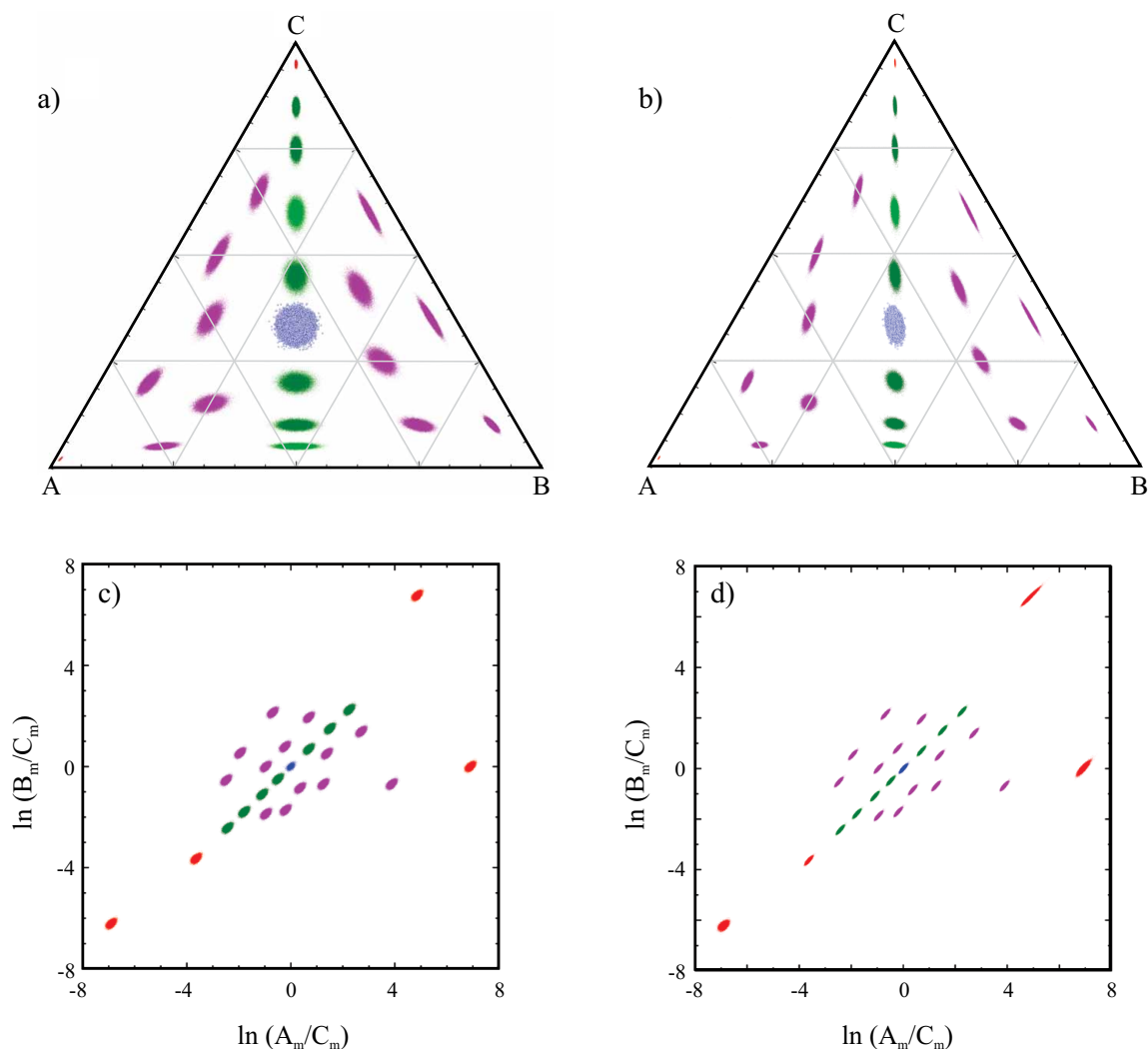
Figure 3. Frequently used ternary (a, model 1; b, model 2) and statistically-correct proposed bivariate (c, model 1; d, model 2) diagrams as illustrated by the use of 25 data involving three variables (initial mean values given in the first three columns of Table 1) and large simulations of 100,000 repetitions. In all diagrams, the symbol sizes would correspond to total uncertainty estimates of 99% confidence limits for sample sizes of about 10. Symbol shapes result from the covariance terms during the construction of ternary diagrams. For more details, see Table 1.

and C_m (each about 33.33% – the fourth row of data in Table 1; the centroid of a ternary diagram – Figure 3a), all three new RSD values are equal (about 4.10%). For all other cases, the new RSD values are totally different from each other and also from the initial RSD of 5%. For data lying close to the apexes or to the boundaries, the shapes of the symbols are even more distorted and their sizes become much smaller (Figure 3a; see the last three data rows in Table 1) than for those lying in the central region of ternary diagrams.

For model 2 with unequal RSD of 1%, 3% and 5% and a realistic error structure (Thompson, 1988), the recalculations inherent in ternary diagrams result in even greater modification of RSD values (0.02% to 11.47% in Table 1; see the three columns listed under "Model 2 (Figure 3b)") and consequent greater distortion of symbol shapes (Figure 3b). For all cases, the new RSD values are totally different from the initial RSD values (Table 1), symbol shapes are

distorted and their sizes become smaller for data lying close to the apexes or to the boundaries.

In ternary diagrams, for a given datum the relative mean value determines the region where it will actually plot, whether central or near the apexes or boundaries, and its total uncertainty estimates containing covariance terms indicate its final shape. Traditionally, small symbols are used to represent the data. However, the symbols occupy a much smaller area near the apexes or boundaries to such an extent that the two data plotting very close to the apexes A and C and one close to the A-B boundary are hardly even visible, whereas the symbols are much larger in the central region (Figures 3a and 3b). Analytical errors of the order of 1% to 5% are reasonable estimates for many applications, but may even be underestimates of total analytical uncertainties in some areas of Earth sciences, such as igneous or sedimentary petrography and trace element geochemistry.

In numerous ternary diagrams proposed in the literature for data interpretation (Rollinson, 1993; Verma, 2010), the variables are generally modified by "suitable" multiplication or dividing factors, *e.g.*, Ti/100–Zr–3Y, so that the "useful" region would lie in the central part of the diagrams, away from the apexes and boundaries. The present Monte Carlo simulation procedure shows that this is precisely the "unwanted" area where the data symbols, when plotted with their respective total uncertainty estimates, would occupy the largest part of the diagram, thus rendering all such proposals of ternary diagrams in Earth sciences and chemistry statistically less powerful and probably even meaningless for the interpretation of experimental data, especially those characterized by large analytical errors.

Furthermore, the existence of constant or unit-sum constraint or closure problem in handling compositional data has long been recognized (Chayes, 1960, 1978). Even if the initial data in ternary diagrams are of truly "continuous" variables (and not of compositions), the methodology to construct such diagrams would result in the closure problem that is very similar to the problem of compositional data in most bivariate diagrams, such as the well known Harker type diagrams frequently used in geochemistry (Rollinson, 1993).

One statistically-correct solution to resolve this problem and open the sample space (theoretically from $-\infty$ to $+\infty$) is the natural logarithm transformation of ratios ("log-ratio") using a common denominator (Aitchison, 1986; Buccianti *et al.*, 2006), although it is not clear to me why Aitchison (1986) used ternary diagrams to illustrate his innovated procedure for compositional data handling. Aitchison and Egozcue (2005) commented that ternary diagrams, complemented with centring and scaling techniques (von Eynatten *et al.*, 2003) are one of the most important and practical tools to represent compositional data. Error propagation in such diagrams was, however, not attempted by any of these authors.

I propose and show that the best statistically-correct alternative to statistically erroneous ternary diagrams is to use bivariate diagrams based on the two log-ratios of the three measured variables A_m, B_m and C_m (Figures 3c and 3d). The equal 5% RSD values (model 1) are projected as a constant standard deviation value (Figure 3c; Table 1). Because the transformed log-ratio variables ln (A_m/C_m) and ln (B_m/C_m), can take any value from $-\infty$ to $+\infty$ (see the two columns of "Transformed experimental data" in Table 1 where both negative and positive mean values result from the 25 data), it would be meaningless to use RSD as a parameter to express the dispersion estimate; instead, standard deviation or total uncertainty estimates (confidence intervals) should directly be used. The standard deviation values for unequal RSD and complex error structure (model 2) are also similarly small (Figure 3d). They increase, as expected from the statistical and chemical principles, when one or two variables in a three-component datum approach the respective LOD values (see the data in the final two columns corresponding to three rows in Table 1; these data plot in Figure 3d as extreme values).

In the present work, although only two models were used to illustrate these critical findings, the use of any other uncertainty values or error structure, or actual error or uncertainty estimates when available, would provide essentially the same conclusion that ternary diagrams are statistically erroneous, because they are characterized by distortion of initial errors and unusually large errors visible in the central region of the closed triangular space, away from the apexes and boundaries. I conclude that ternary diagrams should better be abandoned (Figures 3a and 3b) or, at least, their use minimized. Log-ratio transformed bivariate diagrams (Figures 3c and 3d) could henceforth be adopted to handle three variables in two dimensions. This would not only facilitate correct statistical treatment, but also provide new ways of interpreting data in Earth sciences and chemistry.

DISCRIMINATION DIAGRAMS

Ever since the advent of plate tectonics, this graphic technique came into existence for deciphering the tectonic setting of igneous and sedimentary rocks (Rollinson 1993). A recent comprehensive review by Verma (2010) focused first on the statistical evaluation of traditional bivariate and ternary diagrams, and then presented the advantages of using the more recent (2004-2011) multi-dimensional diagrams. In the section of "Ternary Diagrams" I have provided further evidence against the indiscriminate use of ternary diagrams.

Multi-dimensional linear discriminant function based discrimination diagrams

Aitchison (1986) proposed log-ratio transformation as the solution for correct handling of compositional data. Reyment and Savazzi (1999) presented detailed account of multivariate techniques, including some computer programs, to take Aitchison´s recommendation into account.

As a more recent example, Verma *et al.* (2006) used major-elements in their diagrams by log-ratio transformation with $(SiO_2)_{adj}$ as a common denominator (thus, obtaining ten ratios from eleven major-elements). Because a more abundant component was used as the denominator, all major-element ratios are likely to result in numbers smaller than 1, and therefore their log-ratio transformation will result in negative numbers, thus having opened the space from zero to $-\infty$. Had they chosen a less abundant major-element such as $(MgO)_{adj}$ or $(P_2O_5)_{adj}$, the log-transformed space might have opened in both positive and negative directions.

Agrawal *et al.* (2008), on the other hand, for proposing their diagrams based on immobile trace-elements (La, Sm, Yb, Nb, and Th), chose Th as the common denominator. After log-ratio transformation, they worked in the four-

dimensional space of ln(La/Th), ln(Sm/Th), ln(Yb/Th), and ln(Nb/Th). Here both positive and negative values in the transformed space are likely to occur, opening, thus, the space theoretically in both -∞ and +∞ directions.

Similarly, Verma and Agrawal (2011), in their attempt to propose new discrimination diagrams based on immobile elements $(TiO_2)_{adj}$, Nb, V, Y, and Zr opted for the more abundant $(TiO_2)_{adj}$ as the common denominator. They, therefore, also worked in an open space of log-transformed variables from zero to -∞. They further assured through DODESSYS that the log-transformed ratios be normally distributed.

Finally, Verma S.K. *et al.* (2012) used major-elements after log-ratio transformation with $(SiO_2)_{adj}$ as a common denominator and proposed new multi-dimensional diagrams for acid magmas. They also followed the same methodology of Verma and Agrawal (2011) for ascertaining discordant outlier-free samples. Verma and Díaz-González (2012) have documented additional application cases to confirm the usefulness of DODESSYS and new multi-dimensional discrimination diagrams.

In all of these papers, the authors proposed new discriminant function based diagrams after linear discriminant analysis (LDA) of log-transformed data. Strictly speaking, however, the variables used in LDA should be drawn from a multivariate normal distribution rather than a number of univariate normal distributions. New critical values are required to test for the former, which if achieved will be a significant progress not only for geochemometrics but also for other science and engineering fields.

With the exception of Verma *et al.* (2011b) diagrams for acid magmas, all other multi-dimensional diagrams are meant for tectonic discrimination of basic and ultrabasic magmas. This clearly demonstrates that new multi-dimensional discrimination diagrams are very much needed for intermediate magmas as well as additional diagrams based on immobile elements for acid magmas. The available diagrams were extensively evaluated by the original authors. Four tectonic settings have been successfully discriminated, which are island arc, continental rift, ocean-island, and mid-ocean ridge. More recently, Verma *et al.* (2011b) have evaluated these diagrams from independent datasets and documented high success rates not only for these four tectonic settings, but also for the continental arc of the Andes and the Central American Volcanic Arc interpreted as similar to the island arc setting. Verma *et al.* (2011b) also applied these diagrams to evaluate the dominant tectonic setting of the Mexican Volcanic Belt, which was inferred as continental rift.

PETROGENETIC MODELING

The knowledge of chemical equilibrium constants was incorporated in geochemistry in terms of solid-liquid partition coefficients, and a long history of development exists for modeling magmatic processes of partial melting, fractional crystallization, magma mixing, and assimilation with or without fractional crystallization. Rollinson (1993) is a good source for more detailed information on this topic.

Partial melting of a source region is governed by equations presented by several researchers (*e.g.*, Schilling and Winchester, 1967; Shaw, 1970, 1978; Consolmagno and Drake, 1976; Hertogen and Gijbels, 1976; Langmuir *et al.*, 1977; Wood, 1979). Inversion of partial melting equations was also proposed (Minster and Allègre, 1978; Albarède, 1983; Hofmann and Feigenson, 1983) and used more recently in Mexico by Velasco-Tapia and Verma (2001, in press) for Sierra Chichinautzin, by Verma (2004) for eastern Mexican Volcanic Belt, and by Verma (2006) for Los Tuxtlas volcanic field. For modeling of fractional crystallization, one can resort to the details presented by Schilling and Winchester (1967), Allègre *et al.* (1977), Yanagi and Ishizaka (1978), Villemant *et al.* (1981), and Le Roex and Erlank (1982), among others. Combined or decoupled processes of assimilation and fractional crystallization have been invoked to explain magmatic evolution (DePaolo, 1981; Powell, 1984; Cribb and Barton, 1996).

Other more complex petrogenetic models were presented by O´Hara (1977, 1980, 1993, 1995). Similarly, more complex energy-constrained models have been advocated by Spera and Bohrson (2001, 2002, 2004) and Bohrson and Spera (2001, 2003). "In situ" three-dimensional combined thermal and chemical modeling of magma chambers has also been initiated (*e.g.*, Verma and Andaverde, 2007; Verma *et al.*, 2011c, 2011d), but it is still at its infancy stage.

Uncertainty propagation in these petrogenetic models has not been extensively covered. Verma (1998b, 2000) presented approximate error propagation equations for use in geochemical modeling. There is ample room to carry out Monte Carlo simulation for uncertainty propagation in petrogenetic modeling, because most variables, including solid-liquid partition coefficients (*e.g.*, Torres-Alvarado *et al.*, 2003), in the petrogenetic equations have uncertainties associated to them. Finally, Aitchison´s recommendations (Aitchison, 1986) should be incorporated in this field of geochemistry to reinforce the new science of geochemometrics.

GEOTHERMOMETERS

Solute geothermometers have been widely used in geothermal exploration and exploitation. They have been in use now for nearly forty years. A recent review by Verma *et al.* (2008b) summarized all available equations and reported a new computer program SolGeo for use in such studies. Among solute geothermometers, the use of more recent proposal of silica geothermometers (M.P. Verma, 2008) requires special care. The geothermometers based on Na/K are more commonly used, for which statistically improved equations have been reported by Verma and Santoyo (1997), Díaz-González *et al.* (2008), and Verma and Díaz-González (2012). All these regression equations for the Na/K geother-

mometer report standard errors in the regression coefficients. Therefore, error propagation in such equations can also be correctly handled by Monte Carlo simulations.

Although gas geothermometry has been proposed in geothermics (D´Amore and Panichi, 1980; Giggenbach, 1980; Arnórsson and Gunnalugsson, 1985; Henley *et al.*, 1985), these geothermometers have been less used than solute geothermometers.

In this area of geochemistry, correct statistical handling of compositional data recognizing multivariate nature of fluid composition is yet to take place. Incidentally, most Na/K geothermometers seem to comply with Aitchison´s procedure of log-ratio transformation (see Verma *et al.*, 2008b for review on geothermometers) but only as a bivariate procedure for Na and K (and not a multivariate transformation involving other chemical elements as well). I suggest that additional work should be carried out in the field of geothermometry to make this tool more reliable in the exploration and exploitation of geothermal resources.

FINAL CONSIDERATIONS

Undoubtedly, there are other important areas of research that would reinforce the new science of geochemometrics. Among the currently available fields, I once again cite most of the topics covered in this paper, *i.e.*, data quality, discordancy and significance tests, regressions, error propagation in ternary diagrams through Monte Carlo simulation, and multivariate techniques for correct compositional data handling.

There is some vague notion in the literature that, with the availability of more sophisticated instrumental techniques, the data quality has improved over the years. Is it really true? Is it the precision or the accuracy that has been improved? Or both have been improved? The new science of geochemometrics should answer these crucial questions. My own impression based on unpublished compilations of literature data, without having done yet a systematic research and interpretations, is that it is the precision, and not the accuracy, that has probably improved.

As an integral part of the data quality research in geochemometrics, the systematic behavior of LODs should be further investigated, and a clear theoretical explanation should be put forth. The instrumental sensitivities should always be reported. Total uncertainty estimates would be vital for this line of research. Certified geochemical reference materials for most elements of the Periodic Table are also highly desirable for geochemometric purposes.

Overall performance estimates of discordancy tests in terms of relative efficiency criterion (Verma *et al.*, 2009b; González-Ramírez *et al.*, 2009) should be complemented by the five individual probability calculations as suggested by Barnett and Lewis (1994) and achieved by Hayes and Kinsella (2003) for two discordancy tests. Although new precise and accurate critical values have recently been pro-

posed for 33 test variants, precise values are still required for several other discordancy tests proposed in the literature.

Correct statistical treatment prior to the application of discordancy tests to compositional data has yet to be proposed and its use generalized. Does the application of discordancy tests to log-ratios (Verma and Agrawal, 2011) represent such a correct statistical treatment? Or should we explore discordancy tests for multivariate normal distribution?

Significance tests combined with discordancy tests should be routinely used for interpreting geochemical data. This statistical approach will therefore become an integral part of geochemometrics.

Monte Carlo simulations should be used for an objective comparison of different regression techniques currently available. Are these techniques directly applicable to compositional data? Or is some kind of transformation required to make them suitable for geochemometrics?

Similar to the ternary diagrams evaluated in this work, uncertainty propagation through Monte Carlo simulation in other diagrams such as discriminant function-based multi-element (multi-dimensional bivariate) diagrams should prove useful for future proposals of discrimination diagrams.

An objective comparison of robust and outlier-based methods is urgently needed. This will provide indications of appropriate statistical methods for use in the interpretation of geochemical data. And it will solve the controversies that are central to this important aspect of geochemometrics.

The new science of geochemometrics should warn against the erroneous use of numerous bivariate and ternary diagrams in Earth sciences and facilitate the use of statistically correct methodology for data interpretation. In the light of the simulation results documented in this paper, ternary diagrams should probably be abandoned or at least their use minimized, and bivariate plots involving natural logarithm-ratio transformed variables be adopted as the best, statistically-correct alternative to handle three variables in two dimensions.

Other multivariate techniques, such as principal component analysis and cluster analysis, should also be explored, although to geochemometrics these techniques may not prove more efficient than the linear discriminant analysis. Nevertheless, the statistically correct procedures should be made available to all those interested in using geochemometrics in the interpretation of geological and geochemical data.

Are there suitable methods other than the log-ratio transformation to handle compositional data? Aitchison (1999) seems to have demonstrated that there is none else, other than his extensive discussion on log-ratio transformation using a common denominator (Aitchison, 1986).

Work related to "in situ" thermal and chemical modeling of heat sources, if carried out in combination with Monte Carlo simulations, is likely to provide an important progress in petrogenetic modeling and is therefore highly recommended to reinforce the new science of geochemometrics.

ACKNOWLEDGEMENTS

REFERENCES

Agrawal, S., Guevara, M., Verma, S.P., 2004, Discriminant analysis applied to establish major-element field boundaries for tectonic varieties of basic rocks: International Geology Review, 46(7), 575-594.

Agrawal, S., Guevara, M., Verma, S.P., 2008, Tectonic discrimination of basic and ultrabasic rocks through log-transformed ratios of immobile trace elements: International Geology Review, 50(12), 1057-1079.

Aitchison, J., 1982, The statistical-analysis of compositional data: Journal of the Royal Statistical Society Series B-Methodological, 44(2), 139-177.

Aitchison, J., 1984, Reducing the dimensionality of compositional data set: Mathematical Geology, 16(6), 617-635.

Aitchison, J., 1986, The Statistical Analysis of Compositional Data: Chapman and Hall, London and New York, 416 p.

Aitchison, J., 1999, Logratios and natural laws in compositional data analysis: Mathematical Geology, 31(5), 563-580.

Aitchison, J., Egozcue, J.J., 2005, Compositional data analysis: where are we and where should we be heading?: Mathematical Geology, 37(5), 829-850.

Albarède, F., 1983, Inversion of batch melting equations and the trace element pattern of the mantle: Journal of Geophysical Research, 88(B12), 10573-10583.

Allègre, C.J., Treuil, M., Minster, J.F., Minster, B., Albarède, F., 1977, Systematic use of trace element in igneous process. Part I: Fractional crystallization processes in volcanic suites: Contributions to Mineralogy and Petrology, 60(1), 57-75.

Álvarez del Castillo, A., Santoyo, E., García-Valladares, O., Sánchez-Upton, P., 2010, Evaluación estadística de correlaciones de fracción volumétrica de vapor para la modelación numérica de flujo bifásico en pozos geotérmicos: Revista Mexicana de Ingeniería Química, 9(3), 285-311.

Andrews, J.E., Brimblecombe, P., Jickells, T.D., Liss, P.S., Reid, B., 2004, An Introduction to Environmental Chemistry. Second edition: Oxford, Blackwell Publishing, 296 p.

Appelo, C.A.J., Postma, D., 1993, Geochemistry, Groundwater and Pollution. Second edition: Rotterdam, A.A. Balkema, 649 p.

Armstrong-Altrin, J.S., 2009, Provenance of sands from Cazones, Acapulco, and Bahía Kino beaches, Mexico: Revista Mexicana de Ciencias Geológicas, 26(3), 764-782.

Arnórsson, S. (editor), 2000, Isotopic and chemical techniques in geothermal exploration, development and use. Sampling methods, data handling, interpretation: Vienna, International Atomic Energy Agency, 351 p.

Arnórsson, S., Gunnlaugsson, E., 1985, New gas geothermometers for geothermal exploration–calibration and application: Geochimica et Cosmochimica Acta, 49(6), 1307-1325.

Asuero, A.G., González, G., 2007, Fitting straight lines with replicated observations by linear regression. III weighting data: Critical Reviews in Analytical Chemistry, 37(3), 143-172.

Bacon, J.R., Linge, K.L., Parrish, R.R., Van Vaeck, L., 2006, Atomic spectrometry update. Atomic mass spectrometry: Journal of Analytical Atomic Spectrometry, 21(8), 785-818.

Barnett, V., Lewis, T., 1994, Outliers in Statistical Data. Third edition: Chichester, John Wiley, 584 p.

Baumann, K., 1997, Regression and calibration for analytical separation techniques. Part II: Validation, weighted and robust regression: Process Control and Quality, 10(1), 75-112.

Bevington, P.R., Robinson, D.K., 2003, Data Reduction and Error Analysis for the Physical Sciences. Third edition: McGrawHill, NewYork, 320 p.

Bohrson, W.A., Spera, F.J., 2001, Energy-constrained open-system magmatic processes II: application of energy-constrained assimilation - fractional crystallization (EC-AFC) model to magmatic systems: Journal of Petrology, 42(5), 1019-1041.

Bohrson, W.A., Spera, F.J., 2003, Energy-constrained open-system magmatic processes IV: geochemical, thermal and mass consequences of energy-constrained recharge, assimilation and fractional crystallization (EC-RAFC): Geochemistry Geophysics Geosystems, 4(2), 8002, doi:10.1029/2002GC000316.

Brooks, C., Hart, S.R., Wendt, I., 1972, Realistic use of two-error regression treatments as applied to rubidium-strontium data: Reviews of Geophysics and Space Physics, 10(2), 551-577.

Bruns, R.E., Scarminio, I.S., De Barros Neto, B., 2006, Statistical Design – Chemometrics: Amsterdam, Elsevier, 412 p.

Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V. (editors), 2006, Compositional Data Analysis in the Geosciences: from Theory to Practice: Geological Society Special Publication No. 262, London, 212 p.

Butler, J.C., 1979, Trends in ternary petrologic variation diagrams - fact or fantasy?: American Mineralogist, 64(9-10), 1115-1121.

Chayes, F., 1960, On correlation between variables of constant sum: Journal of Geophysical Research, 65(12), 4185-4193.

Chayes, F., 1965, Classification in a ternary-diagram by means of discriminant functions: The American Mineralogist, 50(10), 1618-1633.

Chayes, F., 1978. Ratio Correlation. A Manual for Students of Petrology and Geochemistry: The University of Chicago Press, Chicago and London, 99 p.

Chayes, F., 1985, Complementary ternaries as a means of characterizing silica saturation in rocks of basaltic composition: Journal of Geology, 93(6), 743-747.

Consolmagno, G.J., Drake, M.J., 1976, Equivalence of equations describing trace element distribution during equilibrium partial melting: Geochimica et Cosmochimica Acta, 40(11), 1421-1422.

Cribb, J.W., Barton, M., 1996, Geochemical effects of decoupled fractional crystallization and crustal assimilation: Lithos, 37(4), 293-307.

D'Amore, F., Panichi, C., 1980, Evaluation of deep temperatures of hydrothermal systems by a new gas geothermometer: Geochimica et Cosmochimica Acta, 44(3), 549-556.

del Río Bocio, J.F., Riu, J., Boqué, R., Rius, F.X., 2003, Limits of detection in linear regression with error in the concentration: Journal of Chemometrics, 17(7), 413-421.

DePaolo, D.J., 1981, Trace element and isotopic effects of combined wallrock assimilation and fractional crystallization: Earth and Planetary Science Letters, 53(2), 189-202.

Díaz-González, L., Santoyo, E., Reyes-Reyes, J., 2008, Tres nuevos geotermómetros mejorados de Na/K usando herramientas computacionales y geoquimiométricas: aplicación a la predicción de temperaturas de sistemas geotérmicos: Revista Mexicana de Ciencias Geológicas, 25(3), 465-482.

Dixon, W.J., 1950, Analysis of extreme values: Annals of Mathematical Statistics, 21(4), 488-506.

Dixon, W.J., 1951, Ratios involving extreme values: Annals of Mathematical Statistics, 22(1), 68-78.

Dixon, W.J., 1953, Processing data for outliers: Biometrics, 9(1), 74-89.

Dougherty-Page, J.S., Bartlett, J.M., 1999, New analytical procedures to increase the resolution of zircon geochronology by the evaporation technique: Chemical Geology, 153(1-4), 227-240.

Draper, N.R., Smith, H., 1998, Applied Regression Analysis, Third edition: New York, John Wiley, 706 p.

Efstathiou, C.E., 2006, Estimation of type I error probability from experimental Dixon's "Q" parameter on testing for outliers within small size data sets: Talanta, 69(5), 1068–1071.

Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G., BArceló-Vidal, C., 2003, Isometric logratio transformations for compositional data analysis: Mathematical Geology, 35(3), 279-300.

Esbensen, K., Geladi, P., 1990, The start and early history of chemometrics - selected interviews. 2. Journal of Chemometrics, 4(6), 389-412.

Espinosa-Paredes, G., Verma, S.P., Vázquez-Rodríguez, A., Núñez-Carrera, A., 2010, Mass flow rate sensitivity and uncertainty analysis in natural circulation boiling water reactor core from Monte Carlo simulations: Nuclear Engineering and Design, 240(5), 1050-1062.

Faber, K., Kowalski, B.R., 1997, Improved estimation of the limit of detection in multivariate calibration: Fresenius Journal of Analytical Chemistry, 357(7), 789-795.

Faure, G., 2001, Origin of Igneous Rocks. The Isotopic Evidence: Springer, Berlin, 496 p.

Ferrús, R., Egea, M.R., 1994, Limit of discrimination, limit of detection and sensitivity in analytical systems: Analytica Chimica Acta, 287(1-2), 119-145.

Freeze, A.R., Cherry, J.A., 1979, Groundwater: New Jersey, Prentice Hall, 604 p.

Gasparik, T., 2003, Phase diagrams for geoscientists: an atlas of the Earth´s interior: Berlin, Springer, 462 p.

Gawlowski, J., Bartulewics, J., Gierczak, T., Niedzielski, J., 1998, Tests for outliers. A Monte Carlo evaluation of the error of first type: Chemical Analysis (Warsaw), 43, 743-753.

Geladi, P., Esbensen, K., 1990, The start and early history of chemometrics. 1. Selected interviews: Journal of Chemometrics, 4(5), 337-354.

Giggenbach, W.F., 1980, Geothermal gas equilibria: Geochimica et Cosmochimica Acta, 44(12), 2021-2032.

Gladney, E.S., Jones, E.A., Nickell, E.J., 1992, 1988 compilation of elemental concentration data for USGS AGV-1, GSP-1 and G-2: Geostandards Newsletter, 16(2), 111–300.

Gómez-Arias, E., Andaverde, J., Santoyo, E., Urquiza, G., 2009, Determination of the viscosity and its uncertainty in drilling fluids used for geothermal well completion: application in the Los Humeros field, Puebla, Mexico: Revista Mexicana de Ciencias Geológicas, 26(2), 516-529.

González-Ramírez, R., Díaz-González, L., Verma, S.P., 2009, Eficiencia relativa de 15 pruebas de discordancia con 33 variantes aplicadas al procesamiento de datos geoquímicos: Revista Mexicana de Ciencias Geológicas, 26(2), 501-515.

Govindaraju, K., 1994, 1994 compilation of working values and sample description for 383 geostandards: Geostandards Newsletter, Special Issue, 1-158.

Grubbs, F.E., 1950, Sample criteria for testing outlying observations: Annals of Mathematical Statistics, 21, 27-58.

Grubbs, F.E., Beck, G., 1972, Extension of sample sizes and percentage points for significance tests of outlying observations: Technometrics, 14(4), 847-854.

Güell, O.A., Holcombe, J.A., 1990, Analytical applications of Monte Carlo techniques: Analytical Chemistry, 62(9), 529A-542A.

Guevara, M., Verma, S.P, Velasco-Tapia, F., 2001, Evaluation of GSJ intrusive rocks JG1, JG2, JG3, JG1a, and JGb1 by an objective outlier rejection statistical procedure: Revista Mexicana de Ciencias Geológicas, 18(1), 74-88.

Guevara, M., Verma, S.P., Velasco-Tapia, F., Lozano-Santa Cruz, R., Girón, P., 2005, Comparison of linear regression models for quantitative geochemical analysis: An example using x-ray fluorescence spectrometry: Geostandards and Geoanalytical Research, 29(3), 271-284.

Hall, A., 1996, Igneous Petrology. Second edition: Essex, England, Longman, 551 p.

Hammersley, J.M., Handscomb, 1964, Monte Carlo Methods: Norwich, England, Fletcher and Son Ltd., 174 p.

Hayes, K., Kinsella, A., 2003, Spurious and non-spurious power in performance criteria for tests of discordancy: The Statistician, 52(1), 69-82.

Hayes, K., Kinsella, A., Coffey, N., 2007, A note on the use of outlier criteria in Ontario laboratory quality control schemes: Clinical Biochemistry, 40(3-4), 147-152.

Henley, R.W., Truesdell, A.H., Barton Jr., P.B., Whitney, J.A., 1985, Fluid-Mineral Equilibria in Hydrothermal Systems: Society of Economic Geologists, 266 p.

Hernández-Martínez, J.L., Verma, S.P., 2009, Reseña sobre las metodologías de campo, analíticas y estadísticas empleadas en la determinación y manejo de datos de los elementos de tierras raras en el sistema suelo-planta: Revista de la Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo, 41(2), 153-189.

Hertogen, J., Gijbels, R., 1976, Calculation of trace element fractionation during partial melting: Geochimica et Cosmochimica Acta, 40(3), 313-322.

Hinich, M.J., Talwar, P.P., 1975, A simple method for robust regression: Journal of the American Statistical Association, 70(349), 113-119.

Hofmann, A.W., Feigenson, M.D., 1983, Case studies on the origin of basalt. I. Theory and reassessment of Grenada basalts: Contributions to Mineralogy and Petrology, 84(4), 382-389.

Howard, J.L., 1994, A note on the use of statistics in reporting detrital clastic compositions. Sedimentology, 41(4), 747-753.

Iglewicz, B., Martínez, J., 1982, Outlier detection using robust measures of scale: Journal of Statistical Computation and Simulation, 15, 285-293.

Imai, N., Terashima, S., Itoh, S., Ando, A., 1995, 1994 compilation values for GSJ reference samples, "Igneous rock series": Geochemical Journal, 29(1), 91-95.

IUPAC, 1978, Nomenclature, symbols, units and their usage in spectrochemical analysis – II data interpretation: Spectrochimica Acta, Part B 33(6), 242-245.

Jain, R.B., 1981, Percentage points of many-outlier detection procedures: Technometrics, 23(1), 71-75.

Jensen, J.L., Lake, L.W., Corbett, P.W.M., Goggin, D.J., 1997, Statistics for Petroleum Engineers and Geoscientists: Prentice-Hall, Upper Saddle River, 390 p.

Jochum, K.P., Bruckner, S.M., 2008, Reference materials in geoanalytical and environmental research - Review for 2006 and 2007: Geostandards and Geoanalytical Research, 32(4), 405-452.

Jochum, K.P., Nohl, U., 2008, Reference materials in geochemistry and environmental research and the GeoReM database: Chemical Geology, 253(1-2), 50-53.

Kaplan, I., 1963, Nuclear Physics: Addition-Wesley, Reading, 770 p.

Kump, P., 1997, Some considerations on the definition of the limit of detection in X-ray fluorescence spectrometry: Spectrochimica Acta Part B, 52(3), 405-408.

Langmuir, C.H., Bender, J.F., Bence, A.E., Hanson, G.N., 1977, Petrogenesis of basalts from the famous area: mid-atlantic ridge: Earth and Planetary Science Letters, 36(1), 133-156.

Lavine, B., Workman, J., 2008, Chemometrics: Analytical Chemistry, 80(12) 4519-4531.

Law, A.M., Kelton, W.D., 2000, Simulation Modeling and Analysis: McGraw Hill, Boston, 760 p.

Le Maitre, R.W., Streckeisen, A., Zanettin, B., Le Bas, M.J., Bonin, B., Bateman, P., Bellieni, G., Dudek, A., Schmid, R., Sorensen, H., 2002, Woolley, A.R., 2002, Igneous rocks. A classification and glossary of terms: recommendations of the International Union of Geological Sciences Subcommission of the Systematics of Igneous Rocks. Second edition: Cambridge, Cambridge University Press, 236 p.

Le Roex, A.P., Erlank, A.J., 1982, Quantitative evaluation of fractional crystallization in Bouvet Island lavas: Journal of Volcanology and Geothermal Research, 13(3-4), 309-338.

Madhavaraju, J., González-León, C.M., Lee, Y.I., Armstrong-Altrin, J.S., Reyes-Campero, and L.M., 2010, Geochemistry of Aptian-Albian Mural Formation of Bisbee Group, Northern Sonora, Mexico: Cretaceous Research, 31(4), 400-414.

Mahon, K.L., 1996, The New "York" regression: application of an improved statistical method to geochemistry: International Geology Review, 38, 293-303.

Maronna, R.A., Martin, R.D., Yohai, V.J., 2006, Robust Statistics: John Wiley, Chichester, 403 p.

Marroquín-Guerra, S.G., Velasco-Tapia, F., Díaz-González, L., 2009, Statistical evaluation of geochemical reference materials from the Centre de Recherches Petrographiques et Geochimiques (France) by applying a schema for the detection and elimination of discordant outlier values: Revista Mexicana de Ciencias Geológicas, 26(2), 530-542.

McIntyre, G.A., Brooks, C.K., Compston, W., Turek, A., 1966, The statistical assessment of Rb-Sr isochrons: Journal of Geophysical Research, 71(22), 5459-5468.

Miller, J.N., Miller, J.C., 2005, Statistics and Chemometrics for Analytical Chemistry. Fifth edition: Pearson Prentice Hall, Essex, England, 271 p.

Minster, J.F., Allègre, C.J., 1978, Systematic use of trace elements in igneous processes Part III. Inverse problem of batch partial melting in volcanic suites: Contributions to Mineralogy and Petrology, 68(1), 37-52.

Mocak, J., Bond, A.M., Mitchell, S., Scollary, G., 1997, A statistical overview of standard (IUPAC and ACS) and new procedures for determining the limits of detection and quantification: application to voltametric and stripping techniques (Technical Report): Pure & Applied Chemistry, 69(2), 297-328.

Najafzadeh, A., Jafarzadeh, M., Musavi-Harami, R., 2010, Provenance and tectonic setting of Upper Devonian silisiclastic rocks of Ilanqareh Formation, NW Iran: Revista Mexicana de Ciencias Geológicas, 27(3), 545-561.

Nicholson, K., 1993, Geothermal Fluids: Chemistry and Exploration Techniques: Berlin, Springer-Verlag, 263 p.

Nicholls, J., Russell, J.K. (editors), 1990, Modern methods of igneous petrology: understanding magmatic processes: Mineralogical Society of America, 314 p.

O´Hara, M.J., 1977, Geochemical evolution during fractional crystallisation of a periodically refilled magma chamber: Nature, 266, 503-507.

O´Hara, M.J., 1980, Nonlinear nature of the unavoidable long-lived isotopic, trace and major element contamination of a developing magma chamber: Philosophical Transactions of the Royal Society of London, 297, 215-227.

O´Hara, M.J., 1993, Trace element geochemical effects of imperfect crystal-liquid separation: The Geological Society of London, Special Publication No. 76, 39-59.

O´Hara, M.J., 1995, Trace element geochemical effects of integrated melt extraction and "shaped" melting regimes: Journal of Petrology, 36(4), 1111-1132.

Otto, M. (1999) Chemometrics. Statistics and Computer Application in Analytical Chemistry: Weinheim, Wiley-VCH. 314 p.

Ottonello, G., 1997, Principles of Geochemistry: New York, Columbia University Press, 894 p.

Pandarinath, K., 2009a, Clay minerals in SW Indian continental shelf sediments cores as indicators of provenance and paleomonsoonal conditions: a statistical approach: International Geology Review, 51(2), 145-165.

Pandarinath, K., 2009b, Evaluation of geochemical sedimentary reference materials of the Geological Survey of Japan (GSJ) by an objective outlier rejection statistical method: Revista Mexicana de Ciencias Geológicas, 26(3), 638-646.

Pandarinath, K., 2011, Solute geothermometry of springs and wells of the Los Azufres and Las Tres Vírgenes geothermal fields, Mexico: International Geology Review, 53(9), 1032-1058.

Philip, G.M., Skilbeck, C.G., Watson, D.F., 1987, Algebraic dispersion fields on ternary diagrams: Mathematical Geology, 19(3), 171-181.

Potts, P.J., Tindle, A.G., Webb, P.C., 1992, Geochemical Reference Material Compositions: Boca raton: Whittles Publishing, CRC Press, 313 p.

Powell, R., 1984, Inversion of the assimilation and fractional crystallization (AFC) equations; characterization of contaminants from isotope and trace element relationships in volcanic suites: Journal of the Geological Society of London, 141, 447-452.

Prescott, P., 1979, Critical values for a sequential test for many outliers: Applied Statistics, 28(1), 36-39.

Presnall, D.C., 1969, The geometrical analysis of partial fusion: American Journal of Science, 267(12), 1178-1194.

R Development Core Team, 2009, R: A language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R/project.org.

Ragland, P.C., 1989, Basic Analytical Petrology: New York, Oxford University Press, 369 p.

Reyment, R.A., Savazzi, E., 1999, Aspects of Multivariate Statistical Analysis in Geology: Amsterdam, Elsevier, 285 p.

Rodríguez-Ríos, R., Aguillón-Robles, A., Leroy, J.L., 2007, Evolución petrológica y geoquímica de un complejo de domos topacíferos en el campo volcánico de San Luis Potosí (México): Revista Mexicana de Ciencias Geológicas, 24(3), 328-343.

Rollinson, H.R., 1993, Using geochemical data: evaluation, presentation, interpretation: Essex, Longman Scientific Technical, 344 p.

Rosner, B., 1975, On the detection of many outliers: Technometrics, 17(2), 221-227.

Rosner, B., 1977, Percentage points for the RST many outlier procedure: Technometrics, 19(3), 307-312.

Rousseeuw, P.J., Leroy, A.M., 1987, Robust Regression and Outlier Detection: John Wiley & Sons, New York, 329 p.

Santoyo, E., Verma, S.P., 2003, Determination of lanthanides in synthetic standards by reversed-phase high performance liquid chromatography with the aid of a weighted least-squares regression model: estimation of method sensitivies and detection limits: Journal of Chromatography A, 997(1-2), 171-182.

Santoyo, E., García, R., Galicia-Alanis, K.A., Verma, S.P., Aparicio, A., Santoyo-Castelazo, A., 2007, Separation and quantification of lanthanides in synthetic standards by capillary electrophoresis: a new experimental evidence of the systematic "odd-even" pattern observed in sensitivities and detection limits: Journal of Chromatography A, 1149(1), 12-19.

Schilling, J.-G., Winchester, J.W., 1967, Rare-earth fractionation and magmatic processes, *in* Runcorn, S.K., (ed.), Mantles of the Earth and Terrestrial Planets. London: Interscience Publishers, pp. 267-283.

Shapiro, S.S., Wilk, M.B., 1965, An analysis of variance test for normality (complete samples): Biometrika, 52(3-4), 591-611.

Shapiro, S.S., Wilk, M.B., Chen, H.J., 1968, A comparative study of various tests for normality: Journal of American Statistical Association, 63(324), 1343-1371.

Shaw, D.M., 1970, Trace element fractionation during anatexis: Geochimica et Cosmochimica Acta, 34(2), 237-243.

Shaw, D.M., 1978, Trace element behaviour during anatexis in the presence of a fluid phase: Geochimica et Cosmochimica Acta, 42(6A), 933-943.

Spear, F.S., 1995, Metamorphic phase equilibria and pressure-temperature-time paths. Second Edition: Washington, D.C., Mineralogical Society of America, 799 p.

Spera, F.J., Bohrson, W.A., 2001, Energy-constrained open-system magmatic processes I: general model and energy-constrained assimilation and fractional crystallization (EC-AFC) formulation: Journal of Petrology, 42(5), 999-1018.

Spera, F.J., Bohrson, W.A., 2002, Energy-constrained open-system magmatic processes 3: energy-constrained recharge, assimilation and fractional crystallization (EC-RAFC): Geochemistry Geophysics Geosystems, 3(12), 8001, doi:10.1029/2002GC000315.

Spera, F.J., Bohrson, W.A., 2004, Open-system magma chamber evolution: an energy-constrained geochemical model incorporating the effects of concurrent eruption, recharge, variable assimilation and fractional crystallization (EC-Ec'RAcFC): Journal of Petrology, 45(12), 2459-2480.

Tatsumi, Y., Eggins, S., 1995, Subduction zone magmatism: Japan, Blackwell Science, 211 p.

Taylor, S.R., McLennan, S.M., 1985, The continental crust: its composition and evolution: Oxford, Blackwell Scientific, 312 p.

Thompson, M., 1988, Variation of precision with concentration in an analytical system: Analyst 113(10), 1579-1587.

Tietjen, G.L., Moore, R.H., 1972, Some Grubbs-type statistics for the detection of several outliers: Technometrics, 14(3), 583-597.

Torres-Alvarado, I.S., Verma, S.P., Palacios-Berruete, H., Guevara, M., González-Castillo, O.Y., 2003, DC_Base: a database system to manage Nernst distribution coefficients and its application to partial melting modeling: Computers & Geosciences, 29(9), 1191-1198.

Torres-Alvarado, I.S., Smith, A.D., Castillo-Román, J., 2011, Sr, Nd, and Pb isotopic and geochemical constraints for the origin of magmas in Popocatépetl volcano (Central Mexico) and their relationship with adjacent volcanic fields: International Geology Review, 53(1), 84-115.

Tsakanika, L.V., Ochsenkühn-Petropoulou, M.T., Mendrinos, L.N., 2004, Investigation of the separation of scandium and rare earth elements from red mud by use of reversed-phase HPLC: Analytical and Bionalytical Chemistry, 379(5-6), 796-802.

Velasco, F., Verma, S.P., 1998, Importance of skewness and kurtosis statistical tests for outlier detection and elimination in evaluation of Geochemical Reference Materials: Mathematical Geology, 30(1), 109-128.

Velasco, F., Verma, S.P., Guevara, M., 2000, Comparison of the performance of fourteen statistical tests for detection of outlying values in geochemical reference material databases: Mathematical Geology, 32(4), 439-464.

Velasco-Tapia, F., Verma, S.P., 2001, First partial melting inversion model for a rift-related origin of the Sierra de Chichinautzin volcanic field, central Mexican Volcanic Belt: International Geology Review, 43(9), 788-817.

Velasco-Tapia, F., Verma, S.P., (in press), Magmatic Processes at the volcanic front of Central Mexican Volcanic Belt: Sierra de Chichinautzin volcanic field (Mexico). Turkish Journal of Earth Sciences.

Velasco-Tapia, F., Guevara, M., Verma, S.P., 2001, Evaluation of concentration data in geochemical reference materials: Chemie der Erde-Geochemistry, 61(1), 69-91.

Verma, M.P., 2008, Qrtzgeotherm: an ActiveX component for the quartz solubility geothermometer: Computers & Geosciences, 34(12), 1918-1925.

Verma, S.K., Pandarinath, K., Verma, S.P., 2012, Statistical evaluation of tectonomagmatic discrimination diagrams for granitic rocks and proposal of new discriminant-function-based multi-dimensional diagrams for acid rocks: International Geology Review, 54(3), 325-347.

Verma, S.P., 1992, Seawater alteration effects on REE, K, Rb, Cs, Sr, U, Th, Pb, and Sr-Nd-Pb isotope systematics of mid-ocean ridge basalt: Geochemical Journal, 26(3), 159-177.

Verma, S.P., 1997, Sixteen statistical tests for outlier detection and rejection in evaluation of international geochemical reference materials: example of microgabbro PM-S: Geostandards Newsletter, Journal of Geostandards and Geoanalysis, 21(1), 59-75.

Verma, S.P., 1998a, Improved concentration data in two international geochemical reference materials (USGS basalt BIR-1 and GSJ peridotite JP-1) by outlier rejection: Geofísica Internacional, 37(3), 215-250.

Verma, S.P., 1998b, Error propagation in geochemical modeling of trace elements in two component mixing: Geofísica Internacional, 37(4), 327-338.

Verma, S.P., 2000, Error propagation in equations for geochemical modelling of radiogenic isotopes in two-component mixing: Proceedings of the Indian Academy of Sciences (Earth and Planetary Sciences), 109(1), 79-88.

Verma, S.P., 2004, Solely extension-related origin of the eastern to west-central Mexican Volcanic Belt (Mexico) from partial melting inversion model: Current Science, 86(5), 713-719.

Verma, S.P., 2005, Estadística Básica para el Manejo de Datos Experimentales: Aplicación en la Geoquímica (Geoquimiometría): México, D. F., Universidad Nacional Autónoma de México, 186 p.

Verma, S.P., 2006, Extension related origin of magmas from a garnet-bearing source in the Los Tuxtlas volcanic field, Mexico: International Journal of Earth Sciences, 95(5), 871-901.

Verma, S.P., 2009, Evaluation of polynomial regression models for the Student t and Fisher F critical values, the best interpolation equations from double and triple natural logarithm transformation of degrees of freedom up to 1000, and their applications to quality control in science and engineering: Revista Mexicana de Ciencias Geológicas, 26(1), 79-92.

Verma, S.P., 2010, Statistical evaluation of bivariate, ternary and discriminant function tectonomagmatic discrimination diagrams: Turkish Journal of Earth Sciences, 19(2), 185-238.

Verma, S.P., Agrawal, S., 2011, New tectonic discrimination diagrams for basic and ultrabasic volcanic rocks through log-transformed ratios of high field strength elements and implications for petrogenetic processes: Revista Mexicana de Ciencias Geológicas, 28(1), 24-44.

Verma, S.P., Andaverde, J., 2007, Coupling of thermal and chemical simulations in a 3-D integrated magma chamber-reservoir model: a new geothermal energy research frontier, *in* Ueckermann, H.I., (ed.) Geothermal Energy Research Trends: Nova Science Publishers, Inc., pp. 149-189.

Verma, S.P., Díaz-González, L., 2012, Application of the discordant outlier detection and separation system in the geosciences: International Geology Review, 54(3), 593-614.

Verma, S.P., Quiroz-Ruiz, A., 2006a, Critical values for six Dixon tests for outliers in normal samples up to sizes 100, and applications in science and engineering: Revista Mexicana de Ciencias Geológicas, 23(2), 133-161.

Verma, S.P., Quiroz-Ruiz, A., 2006b, Critical values for 22 discordancy test variants for outliers in normal samples up to sizes 100, and applications in science and engineering: Revista Mexicana de Ciencias Geológicas, 23(3), 302-319.

Verma, S.P., Quiroz-Ruiz, A., 2008, Critical values for 33 discordancy test variants for outliers in normal samples for very large sizes of 1,000 to 30,000: Revista Mexicana de Ciencias Geológicas, 25(3), 369-381.

Verma, S.P., Quiroz-Ruiz, A., 2011, Corrigendum to Critical values for 22 discordancy test variants for outliers in normal samples up to sizes 100, and applications in science and engineering [Rev. Mex. Cienc. Geol., 23 (2006), 302-319]: Revista Mexicana de Ciencias Geológicas 28 (1): 202.

Verma, S.P., Santoyo, E., 1997, New improved equations for Na/K, Na/Li and SiO_2 geothermometers by outlier detection and rejection: Journal of Volcanology and Geothermal Research, 79(1), 9-23.

Verma, S.P., Santoyo, E., 2003a, An unusual systematic behaviour of detection limits for elements 55Cs to 73Ta: Analytical and Bionalytical Chemistry, 377(1), 82-84.

Verma, S.P., Santoyo, E., 2003b, Evaluation of mass spectrometry and other techniques for the determination of rare-earth elements in geological materials, *in* Aggarwal, S.K. (ed.), ISMAS Silver Jubilee Symposium on Mass Spectrometry 2003, Vol. 1, Dona Paula, Goa: Indian Society for Mass Spectrometry, pp. 471-486.

Verma, S.P., Santoyo, E., 2003c, In search of a systematic behaviour of limits of detection for heavy elements (74W to 92U), *in* Aggarwal, S.K., (ed.), ISMAS Silver Jubilee Symposium on Mass Spectrometry 2003, Vol. 2. Dona Paula, Goa: Indian Society for Mass Spectrometry, pp. 511-516.

Verma, S.P., Santoyo, S., 2005, Is odd-even effect reflected in detection limits?: Accreditation and Quality Assurance, 10(4), 144-148.

Verma, S.P., Orduña-Galván, L.J., Guevara, M., 1998, SIPVADE: A new computer programme with seventeen statistical tests for outlier detection in evaluation of international geochemical reference materials and its application to Whin Sill dolerite WS-E from England and Soil-5 from Peru: Geostandards Newsletter: Journal of Geostandards and Geoanalysis, 22(2), 209–234.

Verma, S.P., Santoyo, E., Velasco-Tapia, F., 2002, Statistical evaluation of analytical methods for the determination or rare-earth elements in geological materials and implications for detection limits: International Geology Review, 44(4), 287-335.

Verma, S.P., Guevara, M., Agrawal, S., 2006, Discriminating four tectonic settings: five new geochemical diagrams for basic and ultrabasic volcanic rocks based on log-ratio transformation of major-element data: Journal of Earth System Science, 115(5), 485-528.

Verma, S.P., Quiroz-Ruiz, A., Díaz-González, L., 2008a, Critical values for 33 discordancy test variants for outliers in normal samples up to sizes 1000, and applications in quality control in Earth sciences: Revista Mexicana de Ciencias Geológicas, 25(1), 82-96, with 209 pages of electronic supplement 25-1-01 Critical values for 33 discordancy tests, available at http://satori.geociencias. unam.mx.

Verma, S.P., Pandarinath, K., Santoyo, E., 2008b, SolGeo: a new computer program for solute geothermometers and its application to Mexican geothermal fields: Geothermics, 37(6), 597-621.

Verma, S.P., Pandarinath, K., Velasco-Tapia, F., Rodríguez-Ríos, R., 2009a, Evaluation of the odd-even effect in limits of detection for electron microprobe analysis of natural minerals: Analytica Chimica Acta, 638(2), 126-132.

Verma, S.P., Díaz-González, L., González-Ramírez, R., 2009b, Relative efficiency of single-outlier discordancy tests for processing geochemical data on reference materials and application to instrumental calibrations by a weighted least-squares linear regression model: Geostandards and Geoanalytical Research, 33(1), 29-49.

Verma, S.P., Pandarinath, K., Verma, S.K., 2010, Statistically correct methodology for compositional data in new discriminant function tectonomagmatic diagrams and application to ophiolite origin: Advances in Geosciences 27, Solid Earth Science, 11-22.

Verma, S.P., González-Ramírez, R., Rodríguez-Ríos, R., 2011a, Comparison of two sample preparation methods in x-ray fluorescence spectrometry for the analysis of Ni and Cr: Geostandards and Geoanalytical Research, 35(2), 183-192.

Verma, S.P., Verma, S.K., Pandarinath, K., Rivera-Gómez, M.A., 2011b, Evaluation of recent tectonomagmatic discrimination diagrams and their application to the origin of basic magmas in Southern Mexico and Central America: Pure and Applied Geophysics, 168(8-9), 1501-1525.

Verma, S.P., Gómez-Arias, E., Andaverde, J., 2011c, Thermal sensitivity analysis of emplacement of the magma chamber in Los Humeros caldera, Puebla, Mexico: International Geology Review, 53(8), 905-925.

Verma, S.P., Arredondo-Parra, U.C., Andaverde, J., Gómez-Arias, E., Guerrero-Martínez, F., 2011d, Three-dimensional temperature field simulation of a cooling of a magma chamber, La Primavera caldera, Jalisco, Mexico, International Geology Review, DOI:1 0.1080/00206814.2011.585036

Villemant, B., Jaffrezic, H., Joron, J.-L., Treuil, M., 1981, Distribution coefficients of major and trace elements; fractional crystallization in the alkali basalt series of Chaine des Puys (Massif Central, France): Geochimica et Cosmochimica Acta, 45(11), 1997-2016.

Viner, R.I., Zhang, T., Second, T., Zabrouskov, V., 2009, Quantification of post-translationally modified peptides of bovine α-crystallin using tandem mass tags and electron transfer dissociation: Journal of Proteomics, 72(5), 874-885.

von Eynatten, H., Barceló-Vidal, C., Pawlowsky-Glahn, V., 2003, Modelling compositional change: the example of chemical weathering of granitoid rocks: Mathematical Geology, 35(3), 231-251.

Wood, D.A., 1979, Dynamic partial melting: its application to the petrogeneses of basalts erupted in Iceland, the Faeroe Islands, the Isle of Skye (Scotland) and the Troodos Massif (Cyprus): Geochimica et Cosmochimica Acta, 43(7), 1031-1046.

Yanagi, T., Ishizaka, K., 1978, Batch fractionation model for the evolution of volcanic rocks in an island-arc: an example from central Japan: Earth and Planetary Science Letters, 40(2), 252-262.

York, D., 1966, Least-squares fitting of a straight line: Canadian Journal of Physics, 44, 1079-1086.

York, D., 1969, Least squares fitting of a straight line with correlated errors: Earth and Planetary Science Letters, 5, 320-324.

Young, D.A., 1998, N.L. Bowen and Crystallization-Differentiation: the Evolution of a Theory: Washington D.C., Mineralogical Society of America, 276 p.

Zeyrek, M., Ertekin, K., Kacmaz, S., Seyis, C., Inan, S., 2010, An ion chromatography method for the determination of major anions in geothermal water samples: Geostandards and Geoanalytical Research, 34(1), 67-77.

Zorn, M.E., Gibbons, R.D., Sonzogni, W.C., 1997, Weighted least-squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration: Analytical Chemistry, 69(15), 3069-3075.

Zorn, M.E., Gibbons, R.D., Sonzogni, W.C., 1999, Evaluation of approximate methods for calculating the limit of detection and limit of quantification: Environmental Science and Technology, 33(13), 2291-2295.