New critical values for F and their use in the ANOVA and Fisher’s F tests for evaluating geochemical reference material granite G-2 (U.S.A.) and igneous rocks from the Eastern Alkaline Province (Mexico)

Nuevos valores críticos de F y su uso en las pruebas ANOVA y F de Fisher para evaluar el material de referencia geoquímica del granito G-2 (U.S.A.) y rocas ígneas de la Provincia Alcalina Oriental (México)

R. Cruz-Huicochea¹, S.P. Verma²*

¹Posgrado en Ingeniería (Energía), Instituto de Energías Renovables, Universidad Nacional Autónoma de México, Priv. Xochicalco s/n, Col. Centro, Apartado Postal 34, Temixco 62580, Morelos, México. recrh@live.com
²Departamento de Sistemas Energéticos, Instituto de Energías Renovables, Universidad Nacional Autónoma de México, Priv. Xochicalco s/n, Col. Centro, Apartado Postal 34, Temixco 62580, Morelos, México. spv@ier.unam.mx

*corresponding author

Received: 16/05/2011 / Accepted: 12/04/2013

Abstract

Monte Carlo method was used to simulate precise and accurate critical values (with 2 to 7 decimal places) of Fisher’s F test for degrees of freedom (ν₁ and ν₂) of up to 2000 and confidence levels of 20% to 99.9% for two-sided or two-tailed test (or equivalently to 60% to 99.95% for one-sided or one-tailed test). Unlike the existing literature, we present 15 new critical value tables along with 15 additional tables of the error estimates of individual values. For the two degrees of freedom – horizontal (ν₁) as well as vertical (ν₂) – the critical values correspond to 1(1)30(5)100(10)160(20)200(50)400(100)1000 and 6(1)30(5)100(10)160(20)200(50)400 (100)1000(200)2000, respectively, where the numbers before and after the parenthesis are the initial and final values for which critical values were simulated and the numbers in parenthesis indicate the step size how the initial value of degrees of freedom approached to the final one. Thus these critical values consist of 62x62 values for each of the eleven confidence levels. Even though a large number (3844 for each confidence level, amounting to a total of 42284 values for all confidence levels) of new values of F were simulated, values do not exist for many different degrees of freedom in the range of 1-2000. This problem of unavailable critical values was resolved by evaluating regression models based on simple polynomial functions as well as those involving log-transformation of the independent variable. New best-fit equations were thus proposed to estimate the not-tabulated critical values as well as for the estimation of probability of calculated F value. This methodology of log transformations should be useful also for other kinds of applications involving polynomial fitting. We compared these simulated results with commercial as well as freely available software. The new precise and accurate critical values were used to illustrate the application of analysis of variance (ANOVA) and Fisher’s F tests to geochemical data for international geochemical reference material granite G-2 from U.S.A. We also present the methodology to statistically compare geochemical data for ultrabasic, basic and intermediate magmas from eight regions of the Eastern Alkaline Province of Mexico and U.S.A. Finally, we use these as well as the discordant outlier-free data for ultrabasic and basic rocks in multi-dimensional discrimination diagrams to explore the tectonic setting for this volcanic province, which was inferred to be mainly an extensional zone or a continental rift.
**Keywords:** Monte Carlo simulation, significance test, polynomial regressions, natural logarithm-transformation, discrimination diagrams, discordant outliers.

**1. Introduction**

Any experimental measurement should be characterised by at least two statistical parameters - central tendency or location and dispersion or scale (e.g., Ebdon, 1988; Barnett and Lewis, 1994; Jensen et al., 1997; Bevington and Robinson, 2003; Miller and Miller, 2005; Verma, 2005). The use of sample mean value for the former and sample standard deviation for the latter is common in geological literature. In geological or geochemical discussion, it is frequently necessary to compare measurements of a given geochemical parameter from two geological areas or rock formations, for which two mean values are simply compared visually, and inference in made about one mean value is higher, lower, or different from the other, without any reference to the corresponding dispersion estimates. This is a statistically erroneous practice and should be replaced by the proper so called significance tests (e.g., Verma, 2005, 2012a, 2013a), although these tests should be strictly applied to log-transformed compositional data (e.g., Aitchison, 1986). These significance tests are as follows: Fisher’s F test, Student’s t test, and ANOVA (ANalysis Of VAriance). The simulation of new critical values for the t statistic and their applications are extensively discussed in the companion paper by Verma and Cruz-Huicochea (2013) in this Monograph.

The Fisher’s F test is generally applied for the comparison of variances of two statistical samples drawn from a normal (or Gaussian) distribution, which allows us to choose the appropriate equation for the application of the Student’s t test (e.g., Ebdon, 1988; Verma, 2005). More importantly, the ANOVA test is a powerful procedure to evaluate if three or more statistical normal samples are drawn from one or more populations (e.g., Jensen et al., 1997; Bevington and Robinson, 2003; Miller and Miller, 2005; Verma, 2005).

Evaluation of hypotheses by these tests (F and ANOVA) requires that the calculated statistic be compared with the corresponding critical value at the chosen confidence or significance level. At present, critical value tables are available for confidence levels of 95% and 99% (or equivalently, significance levels of 5% and 1%, or a of 0.05 and 0.01, respectively, and for horizontal degrees of freedom (ν1 for the numerator of the statistic) of 1(1)12(3)15(5)30(10)50, 100, 1000 and ∞, and for vertical degrees of freedom (ν2 for the denominator of the statistic) of 1(1)30, 35, 40, 50, 60, 80, 100, 200, 500, 1000 and ∞, and have the precision of two or three decimal
places (e.g., Neave, 1981; Morrison, 1990; Kanji, 1993; Miller and Miller, 2005; Verma, 2005; Walker and Madden, 2005). Note in the above sequences that the numbers within the brackets “($)” represent the step size and those outside them are the initial and final values of the degrees of freedom. Thus, critical values are available for all degrees of freedom $v_1$ from 1 to 12, but for higher $v_1$ only for $v_1=15, 20, 25, 30, 40, 50, 100, 1000, \text{ and } \infty$. For critical values not tabulated, interpolations are required such as those proposed by Verma (2009). Nevertheless, software systems have been developed to generate more precise critical values from the density function of the F distribution; this is the case of the R Development Core Team (2009).

Multivariate techniques, such as linear discriminant analysis, require the data to be drawn from a multivariate normal distribution (e.g., Morrison, 1990). Barnett and Lewis (1994) and Rencher (2002) provided details on the Wilks’ statistic (Wilks, 1963) for detecting single or more outliers in a multivariate normal distribution. They also indicated that the critical values by Wilks (1963) are only approximate. Jennings and Young (1988) simulated more precise critical values for one or more outliers in such multivariate distributions. However, these critical value tables are limited to sample sizes of $5(1)10(2)20(5)50, 100, 200, 500$ (where the numbers before and after the parenthesis are the initial and final values and the numbers in parenthesis indicate the step size). Thus, the expression “$5(1)10(2)20$” means that the critical values are available for multivariate sample sizes of $5, 6, 7, 8, 9, 10, 12, 14, 16, 18,$ and 20. In the absence of critical values for sample sizes not tabulated in any existing work, for example, for sample size of 11, 13, or 15, the Wilks’ test cannot be easily applied, because precise interpolations are required. Rencher (2002) also presented the equations to convert the Wilks’ statistic to the F test statistic. Therefore, in the absence of a complete set of critical values for all sample sizes, in spite of the fact that Jennings and Young (1988) had simulated more tables in addition to the approximate values by Wilks (1963), the new critical values for the F test can be advantageously used for identifying and separating discordant outliers in multivariate normal distributions.

In this work, we developed an alternative Monte Carlo method to simulate precise and accurate critical values for ANOVA and F tests and natural logarithm-transformation based polynomial regressions for the interpolation of simulated values and calculation of sample probabilities. Our results are favourably compared with the commercially and freely available software systems. We illustrate the use of these critical values for the evaluation of the geochemical reference material granite G-2 from U.S.A. and of geochemical data for volcanic rocks from the Eastern Alkaline Province (EAP) of Mexico and U.S.A. Finally, the geochemical data for ultrabasic and basic rocks from the EAP were also used in multi-dimensional tectonomagmatic discrimination diagrams to infer its dominant tectonic setting of extension or continental rift.

2. Alternate procedure

2.1. Simulation of critical values for the ANOVA and F tests

Monte Carlo method has been used in Mexico to simulate precise and accurate critical values of discordancy tests (Verma and Quiroz-Ruiz, 2006a, 2006b, 2008, 2011; Verma et al., 2008) and Student’ t test (also known as t-statistic; Verma and Cruz-Huicochea, 2013) as well as for other applications such as the evaluation of nuclear reactor performance (Espinosa-Paredes et al., 2010) and error propagation in ternary diagrams and alternative proposal of bivariate diagrams based on log-ratio transformations (for advantages and requirements of such transformations, see Verma, 2012a and references therein). Our procedure is similar to that described by Verma and Quiroz-Ruiz (2006a) and Verma and Cruz-Huicochea (2013); therefore, only the differences of the present approach will be specifically mentioned.

2.1.1. Optimum simulation size or repetitions

The mean critical values for 95% and 99% confidence levels and associated standard errors were calculated for degrees of freedom $v_1$ of 5 and 10 and $v_2$ of 10, 20, and 30, using seven different simulation or repetition sizes of 10,000 to 5’000,000 and 100 independent streams of normal random variates IID N(0,1). The standard error of the mean significantly decreased with increase in the repetition sizes from 10,000 to 1’000,000, but remained practically constant for higher repetitions (from 1’000,000 to 5’000,000). Therefore, for routine simulations we used the repetitions of 1’000,000 and 100 independent streams of random normal variates.

2.1.2. Calculation of the statistic

The F statistic was calculated for each set of degrees of freedom $v_1$ and $v_2$ where $v_1=(n_y-1)$ and $v_2=(n_y-1)$, for which 1’000,000 samples of sizes $n_y$ and $n_x$ were simulated from each of the 100 streams. The ANOVA statistic was calculated for each set of $v_1$ and $v_2$ where $v_1=[k-1]$ and $v_2=[N-k]$, $k$—number of samples and N—total number of data in all $k$ samples, for which $k$ samples for appropriate combination of sizes were simulated, each of them
1,000,000 times. For example, for \( \nu_1=4 \) and \( \nu_2=20 \), \( k=5 \) samples had to be simulated with the total number of data \( N \) as \( \nu_2+k=25 \) and the number of data in each sample as \( N/k=5 \).

### 2.1.3. Calculation of the critical values

The simulated values correspond to the following degrees of freedom: horizontal \( \nu_1 \) of 1(1)30(5)100(10)160(20)200(50)400(100)1000 and vertical \( \nu_2 \) of 6(1)30(5)100(10)160(20)200(50)400(100)1000(200)2000. The confidence levels (two-tailed or two-sided) were as follows: 20%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 98%, 99%, 99.5%, 99.6%, 99.8% and 99.9%, which correspond to the significance levels \( \alpha = 0.80, 0.60, 0.50, 0.40, 0.30, 0.20, 0.10, 0.05, 0.04, 0.02, 0.01, 0.005, 0.004, 0.002, \) and 0.001, respectively (two-tailed or two-sided). These values correspond respectively to the following confidence levels of one-tailed or one-sided test: 60%, 70%, 75%, 80%, 85%, 90%, 95%, 97.5%, 98%, 99%, 99.5%, 99.75%, 99.8%, 99.9% and 99.95%, which correspond to the significance levels \( \alpha = 0.40, 0.30, 0.25, 0.20, 0.15, 0.10, 0.05, 0.025, 0.02, 0.01, 0.005, 0.0025, 0.002, 0.001, \) and 0.0005, respectively (one-tailed or one-sided).

Finally, the mean and standard error of the mean of the 100 values for each \( \nu_1 \) and \( \nu_2 \) and each confidence level, were estimated and reported.

### 2.2. Interpolation of critical values for the ANOVA and F tests

As in Verma (2009) and Verma and Cruz-Huicochea (2013), for estimating the interpolated critical values (not tabulated) the natural logarithm-transformation of both \( \nu_1 \) and \( \nu_2 \) was a necessary step before the polynomial regressions. A computer program was written in Java that enabled us to evaluate 28 different regression models based on simple polynomial fits as well as single, double and triple natural logarithms of \( \nu_1 \) and \( \nu_2 \), all up to 8th order polynomial terms. The best-fit equation was chosen from the criteria of the multiple-correlation coefficient (R²) and the averaged sum of the squared residuals (SSR/N where N is the total number of residuals).

### 2.3. Polynomial fits for the estimation of confidence levels for the ANOVA and F tests

Computational packages, both commercial and freely available, do calculate the confidence level that corresponds to a given set of statistical samples, but the actual method used is not clear. Therefore, we developed a similar polynomial fitting method as for the interpolation of critical values. This consisted of the natural logarithm transformations of degrees of freedom as well as polynomial terms up to 8th order.

### 3. Results of Monte Carlo simulations

#### 3.1. Critical value tables

We simulated 100 sets of critical values for 62 values of \( \nu_1 \) as 1(1)30(5)100(10)160(20)200(50)400(100)1000 and also 62 values of \( \nu_2 \) as 6(1)30(5)100(10)160(20)200(50)400(100)1000(200)2000 and for confidence levels (two-tailed or two-sided) of 20%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 98%, 99%, 99.5%, 99.6%, 99.8% and 99.9%.

The standard error of the mean calculated from these 100 sets of values was used as the criterion for reporting rounded critical values (Bevington and Robinson, 2003; Verma, 2005). The newly simulated critical values, although more precise than the tabulated critical values in the literature, are consistent with the literature values as well as with those generated from the language R (R Development Core Team, 2009). Abridged form of the mean values for the most frequently used 95% and 99% confidence levels (two-tailed and one-tailed) are presented in Tables 1-4 (Appendix A - Supplementary material), respectively, and the respective standard error values in Tables 5-8 (Appendix A - Supplementary material). These values are for the horizontal \( \nu_1 \) of 1-11 only, but for all vertical degrees of freedom \( \nu_2 \) simulated in this work, i.e., 6(1)30(5)100(10)160(20)200(50)400(100)1000(200)2000.

The complete simulated critical values for all confidence levels and all \( \nu_1 \) and \( \nu_2 \) (360 pages of tables; Tables ES1-ES15 and ES16-ES30, respectively, for mean values and their standard errors) are available on request to any of the authors as a supplementary file in pdf format. The complete critical value tables for all simulated \( \nu_1 \) and \( \nu_2 \) corresponding to partial Tables 1-4 are, respectively, Tables ES8, ES7, ES11, and ES10. Similarly, the complete standard error tables corresponding to Tables 5-8 are, respectively, Tables ES23, ES22, ES26, and ES25.

#### 3.2. Critical value equations

For the critical values not tabulated, the regression equations based on 28 different models were tested for the most frequently used confidence levels of 95% and 99%. The best models were obtained in this work (i.e., with the lowest R² and SSR/N; for more details on these criteria, the reader is referred to a companion paper by Verma and Cruz-Huicochea, 2013, in this Monograph). A few regression equations are presented in Tables 9 and 10 (Appendix A - Supplementary material) for one-tailed
95%, to estimate, respectively, the missing values of \( \nu_1 \) for given \( \nu_2 \) of 1-11 only and the missing values of \( \nu_1 \) for given \( \nu_2 \) of 6-17. Similar equations for 99% confidence levels and similar \( \nu_1 \) and \( \nu_2 \) are summarized in Tables 11 and 12 (Appendix A - Supplementary material). A large number of equations are summarised in Tables ES31-ES35 (now 25 pages of tables; more tables for other degrees of freedom are being prepared; all will be available on request from any of the authors).

3.3. Equations for probability estimates

In an analogous manner, 21 regression models were evaluated to propose equations (as examples in Table ES35) for estimating probability of statistical samples for the ANOVA and F tests. In fact, here we converted the probabilities to confidence levels in order to make the probability estimates more understandable to scientists working in different disciplines (other than mathematics or statistics). In the teaching experience of the second author of this paper, most students and researchers, if not all, do not fully understand the use of commercial or freely available software, being the reason why he clarified these concepts in Verma (2005), and we converted the estimates to confidence levels. For using these equations, the parameter “cv” should be replaced by the calculated value of ANOVA or F statistic (\( F_{\text{calc}} \)) for the set of statistical samples, for which the probability estimate or confidence level is desired.

4. Applications

It is important to use statistical tests for processing geochemical data to evaluate similarities and differences among several compositional variables. The erroneous practice of simply comparing only the mean values without any reference to the standard deviation estimates should be abandoned (Verma, 2012a, 2013).

We exemplify the application of the ANOVA (for three or more statistical samples) and Fisher’s F and Student’s t (both for two statistical samples) tests, using the new critical values for F (this work) and t (Verma and Cruz-Huicochea, 2013). These significance tests also require that the individual data arrays (statistical samples under evaluation) be drawn from a normal distribution (Morrison, 1990). To ascertain this assumption, we used only the multiple-outlier type tests (Barnett and Lewis, 1994; González-Ramírez et al., 2009; Verma et al., 2009) programmed in DODESSYS (Verma and Díaz-González, 2012) for identifying and separating discordant outliers at the strict 99% confidence level. We applied the ANOVA and F tests both before and after the application of DODESSYS. We also note that our application examples were processed in an earlier version of a computer program for the application of significance tests, which used our earlier less precise critical values for F. This is the reason why the critical values in our examples may not totally match with those in Tables 1-4 (Appendix A - Supplementary material).

Two application examples are presented as follows: (i) geochemical data for international geochemical reference material granite G-2 from U.S.A. and (ii) geochemical data for volcanic rocks from the Eastern Alkaline Province (EAP) of Mexico and U.S.A. For the application of significance tests, strictly speaking only logratios should be used (e.g. Aitchison, 1986; Verma, 2012a, 2013). However, we present the examples of chemical compositions, as is still customary in the geological literature.

4.1. Evaluation of geochemical data for international reference material granite G-2

We used an unpublished compilation of geochemical data for G-2 as done by Verma and Cruz-Huicochea (2013) in a companion paper of this Monograph. This application therefore will allow the user to compare the results of application of ANOVA (this work) with the combined application of F and t tests (Verma and Cruz-Huicochea, 2013). Although it is customary to present major-element data as oxides, we have shown them as elements, because the compilation by Gladney et al. (1992), which lists the data as elements, was used initially. Thus, because these authors had compiled the major-elements as elements (%), we continued their “unconventional” approach in this work.

The method grouping was the same as that proposed by Velasco-Tapia et al. (2001) as follows: Gr1–classical methods; Gr2–atomic absorption methods; Gr3–x-ray fluorescence methods; Gr4–emission spectrometry methods; Gr5–nuclear methods; Gr6–mass spectrometry methods; Gr7–chromatography methods; and Gr8–miscellaneous methods.

In the processing of reference materials, the geochemical data from different analytical methods should only be combined after ascertaining that these data were drawn from a single or equivalent normal population (Verma, 1998). Therefore, we maintained the identity of different analytical method groups (Gr1 to Gr8) and applied the significance tests of ANOVA or the combination of F and t, to first evaluate the similarities and differences among these data arrays. When these tests suggested that a certain number of groups were drawn from the same population, i.e., they did not show statistically significant differences, they were combined to propose final mean values for the granite G-2 under evaluation.
Table 13 (Appendix A - Supplementary material) presents ANOVA results for 10 major-elements (Si to P, all in %), 13 rare-earth elements (La to Lu), and 33 other trace elements (B to W). The calculated F value ($F_{calc}$) and critical F value ($F_{crit}$) for the corresponding degrees of freedom ($v_1$, and $v_2$) and the total number of groups for a given element can be used to test the null hypothesis ($H_0$) that the data from all method groups were drawn from a single population at the 99% confidence level, as compared to the alternate hypothesis ($H_1$), according to which the data from all method groups did not arise from the same population. Although in some books such as Miller and Miller (2005) ANOVA is applied in conjunction with the one-sided critical values, we opted to apply this test using two-sided values, because for the alternate hypothesis to be true or accepted the mean or variance, or both for one or more samples under evaluation can be either smaller or greater, i.e., different from the remaining samples. It is only for the other significance tests, namely F and t, that a suitable hypothesis for one-sided values can be proposed, for example, for F test the alternate hypothesis could be that one sample has a smaller or greater variance than the other sample, and for t test, one sample has a smaller or greater mean than the other.

We illustrate this methodology for three elements (Si, Ti, and Al; Table 13- Appendix A - Supplementary material) as examples.

For Si (%), $F_{calc} = 5.64635769$ was compared with the critical value (two-sided 99% confidence level, 99%ts; where ts means two-sided or two-tailed) for $v_1=6$ and $v_2=143$ (obtained from interpolation of critical values in Table 3) $cv_{99\%ts, (v_1=6, v_2=143)} = 3.254998$. Because $F_{calc} > cv_{99\%ts, (v_1=6, v_2=143)}$, we concluded that $H_0$ is false (F), i.e., all method groups for Si (%) did not come from a single population, i.e., some sample(s) originated from a population having different mean or variance, or both.

Another case would be Ti (%), for which $F_{calc} (3.1966244) < cv_{99\%ts, (v_1=6, v_2=162)} = 3.234068$, corresponding to $v_1=6$, $v_2=162$, making $H_0$ to be true (T), and we would conclude that all method groups were drawn from the same population. Note this inference is different from that for Si (%).

The third example is for Al (%), with $F_{calc} (1.8238030) < cv_{99\%ts, (v_1=6, v_2=150)} (3.24457 from Table 3), corresponding to $v_1=6$, $v_2=150$, making $H_0$ to be true (T), and we would conclude that all method groups were drawn from the same population. Note that this example does not require the use of interpolation equations, because the critical value can be found in Table ES11 (available from any of the authors).

Thus, for 45 elements $H_0$ is true (identified by T–true– in the last column of Table 13), i.e., data from different method groups can be combined to process them further. However, because significant differences were inferred in the data from different method groups for 11 elements (identified by F–false– in the last column of Table 13; see also Table 14, Appendix A - Supplementary material), the results of only those method groups that do not show significant differences among them, can be combined (Table 14). For example (see Table 14), for Si (%), the data from all method groups (Gr1, Gr2, Gr3, Gr4, Gr5, and Gr8), except Gr6, can be combined to calculate the central tendency and dispersion parameters.

Nevertheless, because the significance tests (ANOVA, F and t) require that the data in individual samples be normally distributed, all samples were first processed by DODESSYS and after that the ANOVA test was applied to discordant-outlier-free data (Tables 15 and 16, Appendix A - Supplementary material). Twenty elements (when ANOVA was applied after the use of DODESSYS) instead of eleven (when ANOVA was applied without the use of DODESSYS) showed significant differences among method groups (compare Tables 15 and 16 with Tables 13 and 14, respectively). This implies that the application of ANOVA does depend on the presence of outlying observations, and it is recommended to assure that the samples under evaluation are free from such data.

Further, it appears that for most elements the mass spectrometry methods (Gr6) provided results different from the other methods (Table 16).

### 4.1.1. Comparison of the ANOVA with F-t tests and the two standard deviation method (2s) for G-2

After the application of DODESSYS and ANOVA as explained above, the combined geochemical data for G-2 from those analytical methods that showed no significant differences, were processed once again in DODESSYS (using multiple-outlier tests at 99% confidence level) to separate any other discordant outliers and obtain the final statistics (see the three columns marked ANOVA test in Table 17, Appendix A - Supplementary material). Note that the identify of the geochemical data from the analytical method or methods that showed significant differences from the remaining methods was maintained, and their statistics were presented separately in Table 17, for example, note Ti (%) is listed two times in Table 17. Thus, for seven cases (Ti (%), Ca (%), Na (%), La, Sm, Sc and Th; Table 17), the ANOVA test had shown that the data from one method group (e.g., Gr3, x-ray fluorescence spectrometry for Ti) should not be combined with the other methods.

These results were compared with those reported by Verma and Cruz-Huicochea (2013) from the t test and by Gladney et al. (1992) from the 2s method. We note that the comparison of the ANOVA, F-t, and 2s methods
should, however, be better or more objectively performed if the ANOVA and F-t tests were applied at a less strict 95% confidence level.

Although we have not applied significance tests for this comparison, the mean values obtained from ANOVA showed some differences from the F-t tests for 42 cases and from the literature for 57 cases (Table 17). Out of 60 major and trace elements (Table 17), the standard deviation values obtained from the ANOVA test were lower than the F-t tests for 37 elements, equal for 10 elements, and higher for the remaining 13 elements. Similarly, the standard deviation values from ANOVA were lower than the literature data (2s method) for 28 elements, equal for 2 elements, and higher for 30 elements. Such differences in the mean and standard deviation values, irrespective of whether statistically significant or not, may affect the calibration of instruments or the evaluation of analytical methods (e.g., Santoyo and Verma, 2003; Guevara et al., 2005; Verma et al., 2009; Verma, 2012a).

Although as done in our companion paper (Verma and Cruz-Huicochea, 2013) we have presented only one example (granite G-2), this kind of application will be useful for evaluating geochemical data for all other rock reference materials (e.g., Gladney and Roelandt, 1988; Imai et al., 1995; Verma, 1997, 1998; Velasco-Tapia et al., 2001; Marroquín-Guerrera et al., 2009; Pandarinath, 2009a) or water reference materials (e.g., M.P. Verma, 2013).

4.2. Evaluation of geochemical data from the Eastern Alkaline Province (EAP) of Mexico and U.S.A.

We compiled data for 575 rock samples of all igneous rock types from eight regions in the Eastern Alkaline Province (EAP; Rg1 to Rg8; see Figure 1 for more details on regions). The region numbering is arranged approximately from north to south as follows: (1) Rg1–U.S.A.;
Rg2–Sierra Picacho (Nuevo León); Rg3–Sierra San Carlos (Tamaulipas); Rg4–Sierra Tamaulipas (Tamaulipas); Rg5–various locations in the Hidalgo and Veracruz states; Rg6–various locations in the Veracruz state; Rg7–two volcanic regions in the Veracruz state; and Rg8–Los Tuxtlas (Veracruz). The magma and rock types were automatically determined from the computer program SINCLAS (Verma et al., 2002).


For comparing compositional data, although it is not customary to apply the ANOVA test, this test for three or more statistical samples or the F-t combination for two statistical samples should always be applied for statistically correct comparisons. However, ANOVA requires that the data for each region be normally distributed (Jensen et al., 1997; Verma, 2005). Therefore, in order to evaluate the effects of the fulfillment of this condition of normal distribution, we compared the data from these eight regions both before and after the application of DODESSYS (Verma and Díaz-González, 2012), which helps identify possible discordant outliers for each parameter. The results of application of ANOVA are summarised in Tables 18-23 for ultrabasic, basic, and intermediate magmas (Appendix A - Supplementary material). The symbols used in these tables are presented in Table 24 (Appendix A - Supplementary material).

After the application of DODESSYS, the statistics of compositional data (number of samples, mean, and standard deviation) from each region are summarized in Tables 25, 26, and 27 for ultrabasic, basic, and intermediate rocks, respectively (Appendix A - Supplementary material). If we were interested to know the average compositions of this entire area (Eastern Alkaline Province of Mexico and U.S.A.), the results of ANOVA will have to be taken into account (Tables 19, 21, and 23, respectively) to combine the regions with no significant differences for a given parameter. This has not been specifically done in this work; we have simply summarized the statistical data individually for each region.

4.2.1. Ultrabasic rocks

Ultrabasic rocks are present in five of the eight regions compiled in this study. The ANOVA test was applied to these data to determine if these rocks showed statistically significant differences (Table 18). When DODESSYS was applied to these data before ANOVA, the results are summarized in Table 19. The statistics of compositions of these ultrabasic rocks are presented in Table 20.

For the major element \((SiO_2)_{adj}\) ANOVA did not identify any three regions (out of five) that have similar concentrations. Therefore, all five regions are listed in the last column of Table 18. On the other hand, for two major elements \((Al_2O_3)_{adj}\) and \((MnO)_{adj}\) all regions could be grouped together because none of them showed significant differences. For other seven major elements (adjusted TiO\textsubscript{2,2}, FeO\textsubscript{2,3}, MgO, CaO, Na\textsubscript{2}O, K\textsubscript{2}O and P\textsubscript{2}O\textsubscript{5}; Table 18) one or two regions showed significant differences with respect to the remaining regions. For example, for \((TiO_2)_{adj}\) two regions (Rg4 and Rg8) showed significantly different concentrations as compared to the remaining three regions (Rg3, Rg5, and Rg6). For two parameters (sialic and femic minerals) based on major elements, all regions could be grouped together, because no region showed significant differences. For other such parameters, one or two regions showed differences with respect to the remaining ones.

One trace element (Sc) showed significant differences among the three regions under study. For seventeen trace elements (Pr, Dy, Ho, Er, Tm, Lu, Co, Cs, Cu, Ga, Pb, Sr, Th, U, V, Y, and Zn), none of the regions showed significant difference. For the remaining trace elements (Table 18) at least three regions did not show any significant difference.

Thirteen ratio parameters, such as \((LREE4)_{NOE}\) and \((HREE3)_{NOE}\), were statistically similar for all regions listed in Table 18 (for the explanation of these parameters, see Table 24 in Appendix A - Supplementary material). For most other parameters, one region (Rg4) showed significant differences as compared to the remaining regions (Table 18).

The application of DODESSYS showed that for \((Na_2O)_{adj}\) in addition to Rg3 (Table 6), Rg4 also showed significant differences in comparison with the remaining regions (Table 19). Similarly, total iron values expressed as \((FeO)_{adj}\) for only one region (Rg8) were different from the other regions (Table 18), but the prior application of DODESSYS rendered that for this parameter two regions (Rg6 and Rg8) were different from the remaining regions (Table 19). The results for all other major-elements remained unchanged.
Two trace elements (Rb and Zn) showed additional significant differences for one more region (Rg3 and Rg4, respectively). For Sc, significant differences were observed in all groups. Solidification index (S.I.) showed significant difference for only one region (Rg4). Finally, five ratio parameters – \((\text{LILE4})_{\text{NOE1}}\), \((\text{LILE3})_{\text{NOE2}}\), \((\text{LILE3})_{\text{NOE3}}\), \((\text{Ba/Nb})_{\text{Rg4}}\), and \((\text{Nb/Nb}^*)_{\text{Rg5}}\) – also showed differences for one or more additional region–Rg3, Rg3, Rg4, and Rg4-Rg5, respectively. Note that some chemical parameters, such as \((\text{K}_2\text{O})_{\text{adj}}\) in Rg3 or \((\text{P}_2\text{O}_5)_{\text{adj}}\) in Rg8, are represented by a lesser number of samples (Table 25), because a few data were identified by DODESSYS as discordant. For some trace elements in some regions, for example, REE (La to Lu) in Rg3, the data are absent because these samples were not analyzed for these elements by the original authors.

Further, as is customary in geochemical studies, not all samples analyzed for major elements were studied for trace elements, being the reason that the number of samples for trace elements is equal or smaller than that for major elements.

Table 25 (Appendix A - Supplementary material) presents the statistical synthesis for all major elements, 12 lanthanides and 20 other trace elements in ultrabasic rock samples for only five regions, because only for them information was available in our database.

It is important to note that the number of samples listed along with the central tendency (mean) and dispersion (standard deviation) parameters can, in future, be statistically evaluated, for example, estimation of total uncertainty or comparison with other geological areas involving significance tests (ANOVA or F-t). Student t values would be required for the estimation of total uncertainty (Verma and Cruz-Huicochea, 2013). Therefore, we stress that in all studies, the statistical synthesis should consist of at least these three statistical parameters (n, mean, and standard deviation; Table 25); a fuller statistical report would present the total uncertainty as the upper and lower confidence limits of the mean.

4.2.2. Basic rocks

Basic rocks are present in six of the eight regions compiled in this study. The ANOVA test was applied to these data to determine if these rocks showed statistically significant differences (Table 20). When DODESSYS was applied to these data before ANOVA, the results are summarised in Table 21. The statistics of compositions of these basic rocks are presented in Table 20.

For \((\text{MnO})_{\text{adj}}\), all regions could be grouped together because ANOVA did not identify any significant difference for this parameter. For other nine major elements (adjusted \(\text{SiO}_2\), \(\text{TiO}_2\), \(\text{Al}_2\text{O}_3\), \(\text{Fe}_2\text{O}_3\), \(\text{MgO}\), \(\text{CaO}\), \(\text{Na}_2\text{O}\), \(\text{K}_2\text{O}\) and \(\text{P}_2\text{O}_5\); Table 20) one, two, or even three regions showed significant differences with respect to the remaining regions (Table 20). For example, for \((\text{SiO}_2)_{\text{adj}}\) two regions (Rg8 and Rg6) showed significantly different concentrations as compared to the remaining four regions (Rg2, Rg3, Rg4, and Rg5). For all parameters based on major elements, one, two or three regions showed differences with respect to the remaining ones. For example, Mg# values for Rg8, Rg3 and Rg6 were different from Rg2, Rg4, and Rg5 (Table 20).

One trace element (V) showed significant differences among all four regions for which data were available. For ten trace elements (Ce, Pr, Nd, Sm, Eu, Gd, Er, Tm, Ga, and Zn), none of the regions showed significant differences. For the remaining trace elements (Table 20), at least three regions did not show any significant differences.

Three ratio parameters, \((\text{LREE4})_{\text{NOE}}\), \((\text{LREE3})_{\text{NOE1}}\) and \((\text{LILE4/LREE4})_{\text{NOE}}\), were statistically similar for all regions listed in Table 20. For other parameters, one, two, or three regions showed significant differences as compared to the remaining regions (Table 20).

The application of DODESSYS showed significant differences for \((\text{MnO})_{\text{adj}}\) from one region (Rg6) in comparison with the remaining regions (Table 21); note that if DODESSYS was not applied before ANOVA, no differences existed (Table 20). For \((\text{K}_2\text{O})_{\text{adj}}\), in addition to Rg2 and Rg8 (Table 20), Rg3 also showed significant differences in comparison with the remaining regions (Table 21). The application of DODESSYS prior to ANOVA also resulted in different results for \((\text{MgO})_{\text{adj}}\), \((\text{Na}_2\text{O})_{\text{adj}}\) and \((\text{Fe}_2\text{O}_3)_{\text{adj}}\) as well as for \(\text{T}_{\text{alk}},\ \text{FeO}^*/\text{Mg}, \text{Salic}, \text{Femic}\) and A.R. (compare Tables 20 and 21). For C.I., significant differences were observed in all groups.

As a result of DODESSYS, seven trace elements (Ho, Er, Ba, Cu, Nb, Rb and Th) showed additional significant differences for one more region (Table 21). In four trace elements (Lu, Cs, Hf, and Pb), the results showed different regions with significant differences in comparison with the previous results (compare Tables 20 and 21). For V, significant differences were observed for all groups. Finally, for one ratio parameter, significant differences were observed for all groups and the other eight ratio parameters, also showed differences for one or more additional regions (compare Tables 20 and 21). For ten parameters, the results (Table 21) showed different regions with significant differences in comparison with the previous results (Table 20).

Table 26 (Appendix A - Supplementary material) presents statistical parameters for basic rocks from the individual regions of the EAP in the same way as done for ultrabasic rocks (Table 25), and all earlier comments for the subsection of ultrabasic rocks are valid for basic rocks.
Fig. 2. - Application of the set of five discriminant function multi-dimensional tectonic diagrams based on major-elements (Verma et al., 2006) to basic and ultrabasic rocks from the different regions of the Eastern Alkaline Province (EAP) of Mexico. The symbols used for the different regions of the EAP (for more details see Figure 1) are explained as inset in Figure 2 (a); filled symbols are used for discordant outliers as inferred from DODESSYS (Verma and Díaz-González, 2012). The names of x-y axes in the different diagrams explicitly show the tectonic field discriminated in that particular diagram and the subscript m2 stands for the second set of such multi-dimensional diagrams based on major elements. (a) Four-field diagram IAB-CRB-OIB-MORB; (b) Three-field diagram IAB-CRB-OIB; (c) Three-field diagram IAB-CRB-MORB; (d) Three-field diagram IAB-OIB-MORB; and (e) Three-field diagram CRB-OIB-MORB. The abbreviations used are as follows: IAB—island arc basic (and ultrabasic) rocks; CRB—continental rift basic (and ultrabasic) rocks; OIB—ocean island basic (and ultrabasic) rocks; and MORB—mid-ocean ridge basic (and ultrabasic) rocks.

4.2.3. Intermediate rocks

Intermediate rocks are present in all regions compiled in this study. In order to determine if these rocks showed statistically significant differences, the ANOVA test was applied to these data (Table 22). When DODESSYS was applied to these data before ANOVA, the results are summarized in Table 23. The statistics of compositional data for these intermediate rocks are presented in Table 27 (Appendix A - Supplementary material).

For major elements, in all cases (adjusted SiO$_2$, TiO$_2$, Fe$_2$O$_3$, Al$_2$O$_3$, MnO, MgO, CaO, Na$_2$O, K$_2$O and P$_2$O$_5$; Table 22) ANOVA identified one or more regions that showed significant differences with respect to the remaining regions. For all parameters based on major elements, at least one or more regions showed differences with respect to the remaining regions.

For six trace elements (Sm, Tb, Ho, Tm, Co and U), none of the regions showed significant differences. For the remaining trace elements (Table 22) one or more re-
regions showed significant differences compared to the remaining regions.

For ratio parameters, five (LILE5/HFSE6, LILE5/HFSE5a, LILE5/HFSE5b, LILE4/HFSE4b, and LREE3/HFSE4a) were statistically similar for all regions listed in Table 22. For most other parameters at least three regions did not show any significant difference.

For three major elements (adjusted SiO₂, MnO and MgO), the application of DODESSYS showed different results than those obtained without DODESSYS (compare Tables 22 and 23). In the same way, Tₘ, A.R. and salic parameters showed additional significant differences for one and two more regions (Rg3, Rg3 and Rg3, respectively). Six trace elements (La, Sm, Ho, Ga, U, and Zn) showed additional significant differences for one or more regions. For three trace elements (Ba, Cs and Sr), the results showed different regions with significant differences (Table 23) in comparison with the previous results (Table 22). Finally, six ratio parameters (LILE5/HFSE6, LILE5/HFSE5a, LILE5/HFSE5b, LILE4/HFSE4a, LILE4/HFSE4b, and LREE3/HFSE4a) were statistically similar for all regions listed in Table 22. For most other parameters at least three regions did not show any significant difference.
SE4b, and LILE4/HFSE4c) also showed differences for one additional region (Rg1) and for six parameters the results showed different regions with significant differences in comparison with the previous result (Table 22).

Table 27 presents the statistical parameters for intermediate rocks from the individual regions of the EAP, and the comments made in the subsection of ultrabasic rocks are also valid.

5. Application of multi-dimensional discrimination diagrams

New multi-dimensional discrimination diagrams have been recently proposed for tectonic discrimination of basic and ultrabasic magmas (Agrawal et al., 2004, 2008; Verma et al., 2006; Verma and Agrawal, 2011) and for acid magmas (S.K. Verma et al., 2012). New diagrams based on all major elements, selected immobile major and trace elements, and immobile trace elements have also recently put forth for intermediate magmas (Verma and Verma, 2013) and for acid magmas (Verma et al., 2013a), which can be used in future to better understand the tectonic setting and their relationship to magma compositions.

These multi-dimensional diagrams have been positively evaluated and used by the original authors as well as several workers (e.g., Rajesh, 2007; Sheth, 2008; Aparicio and García, 2009; Polat et al., 2009a, 2009b, 2011; Rama-akrishnan, 2009; Verma, 2009, 2010, 2012; Bai- lie et al., 2010, 2012; Slovenec et al., 2010; Wang et al., 2010; Zhang et al., 2010; Pandarinath and Verma, 2013; Verma and Verma, 2013). Importantly, the proposal of these newer diagrams (Verma et al., 2006; Agrawal et al., 2008; Verma and Agrawal, 2011) and their use fully complies with the statistical requirements for handling compositional data (Aitchison, 1986; Agrawal and Verma, 2007; Verma, 2012a). Computer program TecD (Verma and Rivera-Gómez, 2013) greatly facilitates the use of these diagrams. More significantly, the log-transformed ratio variables and not the crude compositions should be processed by DODESSYS to identify discordant outliers (Verma and Díaz-González, 2012), as also suggested by Verma (2012a).

5.1. Application of ANOVA and DODESSYS to log-transformed ratios

As a statistically coherent procedure, we computed natural logarithm of element ratios and compared the transformed compositional data for ultrabasic and basic rocks (taken together) from different regions of the EAP (Tables 28, Appendix A - Supplementary material). The combined ultrabasic and basic rock data from a given area are required because the new diagrams are meant to be used for both types of magma. The first set of log-transformed ratios (Table 28) are used in the discrimination diagrams of Verma et al. (2006) for ultrabasic and basic rocks, whereas the other two sets are employed in the diagrams by Agrawal et al. (2008) and Verma and Agrawal (2011), respectively, both of which also proposed for ultrabasic and basic rocks. Then, we applied the ANOVA test to statistically compare these parameters at the strict 99% confidence level. Even in the log-transformed ratios, significant differences exist among the regions of EAP under study. For example, for $\ln\left(\frac{(TiO_2)_{adj}}{(SiO_2)_{adj}}\right)$, two regions (Rg6 and Rg8) are different from the remaining five regions (Table 28).

After the application of DODESSYS and separating the discordant outliers from the main data, ANOVA provided the comparisons summarized in Table 29 (Appendix A - Supplementary material). Although the conclusions of ANOVA about most ratio parameters remained practically the same as without the prior application of DODESSYS, they were different for two parameters (compare Tables 28 and 29), which are the following: (i) for $\ln\left(\frac{(Al_2O_3)_{adj}}{(SiO_2)_{adj}}\right)$ two regions (Rg2 and Rg8) were different from the other regions without the application of DODESSYS, but with this prior application, ANOVA showed that only one region (Rg8) is different from the remaining regions; and (ii) for $\ln\left(\frac{(K_2O)_{adj}}{(SiO_2)_{adj}}\right)$ two regions (Rg3 and Rg4) were different from the other regions without the application of DODESSYS, but with this application, ANOVA showed that three regions (Rg3, Rg4, and Rg5) are different from the other regions.

5.2. Application of new multi-dimensional discrimination diagrams for tectonic inferences

These diagrams therefore can be used for inferring the tectonic setting of this area. We plotted the samples from our study in Figure 2 – the set of five diagrams based on log-transformed ratios of major elements (Verma et al., 2006) and in Figure 3 – the set of five diagrams based on log-transformed ratios of immobile trace elements (Agrawal et al., 2008). The results of application of these diagrams are summarized in Tables 30 and 31, respectively (Appendix A - Supplementary material).

Finally, instead of plotting the samples in the set of five diagrams (Verma and Agrawal, 2011), we decided to compute probabilities for individual samples as suggested by these authors. Note that in a given diagram, a particular sample will plot in the field for which the probability is the highest. Therefore, it is unnecessary to plot the samples in diagrams, and it is better to identify
for any given sample the highest probability for a tectonic field. The samples along with their respective probabilities are summarized in Table 32 (Appendix A - Supplementary material).

As stated earlier, the log-transformed data used in the three sets of multi-dimensional diagrams (Verma et al., 2006; Agrawal et al., 2008; Verma and Agrawal, 2011), were also processed in DODESSYS (Verma and Díaz-González, 2012), which separated the discordant outliers and rendered the bulk of data as discordant outlier-free. We also indicate the usefulness of this additional approach in the use of new discrimination diagrams.

We discuss first how these different diagrams are interpreted using the example of Rg2 (Sierra Picacho, Nuevo León, Tertiary rocks; Fig. 2a-e and Table 30). These diagrams can discriminate four tectonic settings of IAB (island arc basic rock; field or group no. 1), CRB (continental rift basic rocks; field or group no. 2), OIB (ocean island basic rocks; field or group no. 3), and MORB (mid-ocean ridge basic rocks; field or group no. 4).

For the first diagram (1-2-3-4) of Verma et al. (2006) for basic and ultrabasic rocks, 14 out of 16 samples from Rg2 plot in the continental rift (CRB) field and the remaining two samples in the MORB field (Fig. 2a; see the first row of data in Table 30). Thus, a large proportion or percentage of these samples indicates a continental rift setting for Rg2.

The other diagrams (1-2-3, Fig. 2b; 1-2-4, Fig. 2c; 2-3-4, Fig. 2e; see also Table 30) also show that 14 of these 16 samples plot in the continental rift field. Given the consistency of these diagrams (Fig. 2a-c, e), the remaining diagram (1-3-4, Fig. 2d, in which the field no. 2 is absent) should be considered as the inapplicable diagram (see the asterisk mark after 1-3-4 in Rg2 of Table 30) for this case study. In such inapplicable diagrams, samples will plot in the tectonic field that has characteristics similar to the inferred tectonic setting from other diagrams; in this case, most (10+1) samples plot in the ocean island field, which is very similar to the continental rift tectonic setting. Thus, the success rate for Rg2 discrimination as a continental rift is about 88% (14 out 16 samples).

It is interesting to note that the multiple-outlier type discordancy tests applied at the strict 99% confidence level identified two samples as having discordant data and incidentally both samples plotted in a field different from the continental rift, that is, they represent mis-discrimination (see solid diamond symbol in figure 2a-e and number enclosed in [ ] in Table 30). Thus, if we consider only the fourteen discordant outlier-free samples, the success rate for the continental rift setting would then be 100% (14 out of 14 samples were correctly discriminated)!

The other regions (Rg4, Rg5, Rg6, and Rg8) were also discriminated as an extensional or a continental rift setting in these diagrams (Fig. 2 and Supplementary material: Table 30), whereas one region (Rg3) likely indicated an ocean island setting. The discordant outliers for these regions were also a few in number and generally plotted in a tectonic field different from the inferred setting (Supplementary material: table 30).

The application of Agrawal et al. (2008) diagrams (Fig. 3a-e; Table 31, in Appendix A - Supplementary material) also indicated a continental rift setting for Rg2, Rg5, Rg6, and Rg8, whereas the remaining region (Rg4) is likely an ocean island setting. No discordant outlier was observed for four regions (Rg2, Rg4, Rg5, and Rg8), and only one discordant observation was present in Rg6 (Table 31).

The third set of diagrams (Verma and Agrawal, 2011; Table 32, in Appendix A - Supplementary material) indicated a continental rift setting for Rg3, transitional from continental rift to ocean island for Rg4, transitional from continental rift to MORB for Rg6, and inconclusive results for Rg8. No discordant outlier was observed for two regions (Rg3 and Rg4). One discordant observation was present in Rg6 and two in Rg8 (Table 32).

In summary, from all three sets of diagrams the results of an extensional or a continental rift setting for this area can be inferred, which is fully consistent with earlier conclusions by Verma (2006) for Rg8 and S.K. Verma et al. (2012) for Rg8 and some other regions of the EAP from a limited compilation of data.

6. Additional considerations

The ANOVA test is a powerful method to statistically decipher similarities and differences among three or more statistical samples drawn from normal populations. The computer program DODESSYS (Verma and Díaz-González, 2012) is useful to identify and separate discordant outliers in such samples. This procedure should be applied prior to the ANOVA or F-t tests as well as for the use of new multi-dimensional tectonomagmatic diagrams. The F test has an additional application to know which version of the Student t test should be applied to evaluate the mean values of two statistical samples drawn from normal populations. Sometimes, the t test is applied probably without the prior explicitly stated application of the F test. For example, Wani and Mondal (2011), although presenting a rare example of statistical application in the geological literature, compared the compositions of calcareous and non-calcareous shales, through the t test at 95% confidence level. We suggest that the t test should always be applied in combination with the F test, and preferably at the strict 99% confidence level. Furthermore, it is very important that prior to the application of ANOVA or F-t tests, DODESSYS (Verma and Díaz-González, 2012) should be advantageously used at
the strict 99% confidence level. Finally, instead of applying these statistical tests to crude compositional data as done in section 4 above, log-transformed ratios should be evaluated as done in section 5 (for more details, see Aitchison, 1986; Verma, 2012a).

Comparative studies, such as those reported by several workers (Balaram et al., 1995a, 1995b; Dai Kin et al., 1999; Tamisier-Korolak et al., 1999; Almeida et al., 2002; Kubala-Kukus et al., 2003; M.P. Verma, 2004, 2008; Bansal et al., 2008; and Kumar et al., 2008), can be correctly carried out by the use of the new critical values and the statistical procedures outlined in this work.

Finally, in future using the present new critical values for F and those for t (Verma and Cruz-Huicochea, 2013) numerous studies (e.g., Díaz-González et al., 2008; Armstrong-Altrin, 2009; Hernández-Martínez and Verma, 2009; Gómez-Arias et al., 2009; Madhavaraju and Lee, 2009; Marroquin-Guerra et al., 2009; Pandarinath, 2009a, 2009b; Álvarez del Castillo et al., 2010; Zeyrek et al., 2010; Torres-Alvarado et al., 2011), can be improved through the application of significance tests (ANOVA and F-t tests), provided the geochemometrics techniques (Verma, 2012a) are followed for handling compositional data.

6.1. Relevant software for the application of discordancy and significance tests

All critical values and equations were programmed in a new software UDASYS (Univariate Data Analysis SYStem), which is documented recently by Verma et al. (2013b). UDASYS facilitates the application of both categories of statistical methods for univariate data (robust and outlier-based; Verma, 2005), 33 discordancy tests (Verma and Díaz-González, 2012), and significance tests (Fisher’s F, Student’s t and ANOVA). All critical values simulated in this work and in the companion paper by Verma and Cruz-Huicochea (2013) and the relevant polynomial equations are programmed in UDASYS.

UDASYS helps apply the ANOVA, F and t tests more powerfully than the available commercial software systems. Anyone interested in using UDASYS for his or her work is invited to obtain it free of cost from any of the authors.

7. Conclusions

New precise and accurate critical values along with the corresponding standard error estimates as well as best-fit equations were developed for a better application of the ANOVA and F tests. Application examples have been presented to illustrate the statistically correct procedure for the comparison and evaluation of geochemical data. The geochemical data for ultrabasic and basic rock samples from the Eastern Alkaline Province of Mexico used in discriminant-function based multi-dimensional diagrams suggest an extensional or a continental rift setting for this volcanic province.

Acknowledgements

We are grateful to Alfredo Quiroz-Ruiz for help and guidance in the procurement and use of modern computing facilities. The Sistema Nacional de Investigadores (SNI-Conacyt) provided the necessary financial support for purchasing the special computing facilities and for assigning an assistantship to the first author (RCH); both these actions enabled the development of the research work leading to this paper on application of critical values and other reports as well as for the culmination of Bachelor level thesis of RCH in Computer Engineering. We also thank the two reviewers Dr. V. Balaram and Dr. M. E. A. Mondal for kindly reading our manuscript; we are glad to see that both of them highly appreciated our work.

References

Baillie, R., Gutzmer, J., Rajesh, H.M. (2010): Lithogeochemistry as a tracer of the tectonic setting, lateral integrity and mineralization of a highly metamorphosed Mesoproterozoic volcanic arc sequence on
Torres-Alvarado, I.S., Smith, A.D., Castillo-Roman, J. (2011): Sr, Nd and Pb isotopic and geochemical constraints for the origin of mag-
mas in Popocatepetl volcano (central Mexico) and their relationship with the adjacent volcanic fields. *International Geology Review* 53, 84-115. doi: 10.1080/00206810902906738


**Appendix A - Supplementary material**

Tables 1 to 32 cited in this paper can be downloaded from http://revistas.ucm.es/index.php/JIGE/