Critical values for six Dixon tests for outliers in normal samples up to sizes 100, and applications in science and engineering

Surendra P. Verma[1,2,*] and Alfredo Quiroz-Ruiz[1]

[1] *Centro de Investigación en Energía, Universidad Nacional Autónoma de México, Priv. Xochicalco s/no., Col Centro, Apartado Postal 34, Temixco 62580, Mexico.*
[2] *Centro de Investigación en Ingeniería y Ciencias Aplicadas, Universidad Autónoma del Estado de Morelos, Av. Universidad No. 1001, Col. Chamilpa, Cuernavaca 62210, Mexico*
* *spv@cie.unam.mx*

ABSTRACT

In this paper we report the simulation procedure along with new, precise, and accurate critical values or percentage points (with 4 decimal places; standard error of the mean ≤0.0001) for six Dixon discordance tests with significance levels α = 0.30, 0.20, 0.10, 0.05, 0.02, 0.01, 0.005 and for normal samples of sizes n up to 100. Prior to our work, critical values (with 3 decimal places) were available only for n up to 30, which limited the application of Dixon tests in many scientific and engineering fields. With these new tables of more precise and accurate critical values, the applicability of these discordance tests (N7 and N9-N13) is now extended to 100 observations of a particular variable in a statistical sample. We give examples of applications in many diverse fields of science and engineering including geosciences, which illustrate the advantage of the availability of these new critical values for a wider application of these six discordance tests. Statistically more reliable applications in science and engineering to a greater number of cases can now be achieved with our new tables than was possible earlier. Thus, we envision that these new critical values will result in wider applications of the Dixon tests in a variety of scientific and engineering fields such as agriculture, astronomy, biology, biomedicine, biotechnology, chemistry, environmental and pollution research, food science and technology, geochemistry, geochronology, isotope geology, meteorology, nuclear science, paleontology, petroleum research, quality assurance and assessment programs, soil science, structural geology, water research, and zoology.

Key Words: Outlier methods, normal sample, Monte Carlo simulations, reference materials, earth sciences.

RESUMEN

En este trabajo se presenta el procedimiento para la simulación junto con valores críticos o puntos porcentuales nuevos y más precisos y exactos (con 4 puntos decimales; el error estándar de la media ≤0.0001) de las seis pruebas de discordancia de Dixon y para los niveles de significancia α = 0.30, 0.20, 0.10, 0.05, 0.02, 0.01, 0.005 y para tamaños n de las muestras normales de hasta 100. Antes de nuestro trabajo, se disponía de valores críticos (con 3 puntos decimales) solamente para n hasta 30, lo cual limitaba seriamente la aplicación de las pruebas de Dixon en muchos campos de las ciencias e ingenierías. Con las nuevas tablas de valores críticos más precisos y exactos obtenidos en el presente trabajo, la aplicabilidad de las pruebas de Dixon (N7 y N9-N13) se ha extendido a 100 observaciones de una variable en una muestra estadística. Presentamos ejemplos de aplicaciones en muchos campos de ciencias e ingenierías incluyendo las geociencias. Estos ejemplos demuestran la ventaja de la disponibilidad de estos nuevos *valores críticos para una aplicación muy amplia de esas seis pruebas de discordancia. Se esperan aplicaciones a un mayor número de casos en ciencias e ingenierías, estadísticamente más confiables que como era posible anteriormente. De esta manera, prevemos que los nuevos valores críticos resulten en aplicaciones de las pruebas de Dixon mucho más amplias en una variedad de campos de ciencias e ingenierías tales como agronomía, astronomía, biología, biomedicina, biotecnología, ciencia del suelo, ciencia nuclear, ciencia y tecnología de los alimentos, contaminación ambiental, geocronología, geología estructural, geología isotópica, geoquímica, investigación del agua y del petróleo, programas de aseguramiento y evaluación de calidad, paleontología, química, meteorología y zoología.*

Palabras clave: Métodos de valores desviados, muestra normal, simulaciones Monte Carlo, materiales de referencia, pruebas de discordancia de Dixon, Ciencias de la Tierra.

INTRODUCTION

Two main sets of methods (Outlier methods and Robust methods; Barnett and Lewis, 1994) exist for correctly estimating location (central tendency) and scale (dispersion) parameters for a set of experimental data likely to be drawn, in most cases in science and engineering, from a normal or Gaussian distribution (Verma, 2005). The outlier scheme is based on a set of tests for normality (or detection of outliers) such as Dixon tests described here. However, caution is required when applying such outlier tests for samples that are not normally distributed. The alternative scheme for arriving at these parameters consists of a series of robust or accommodation approach methods (for location parameter: *e.g.*, median, mode, Winsorized mean, trimmed mean, and mean quartile; and for scale parameter: *e.g.*, interquartile range and median deviation; see Barnett and Lewis, 1994; Verma, 2005, or any standard text book on statistics), all of which rely on not "taking into account" the outlying and other peripheral observations in a set of experimental data. These methods, although in use in many branches of science and engineering, will not be considered here any further because the main objective of this paper is to comment on and improve the applicability of six discordance tests, proposed by Dixon more than 50 years ago, which are still widely used as explained below.

Dixon (1950, 1951, 1953) proposed six discordance tests for normal univariate samples and estimated critical values or percentage points for these tests for sizes up to 30 and reported them to 3 decimal places. These tests were designated N7 and N9-N13 by Barnett and Lewis (1994). Dixon (1951) also stated that the estimated critical values for tests N7 (test statistic r_{10} in this paper), N9 (statistic r_{11}), and N10 (statistic r_{12}) were "in error by not more than one or two units in the third (decimal) place", whereas those for tests N11 (statistic r_{20}), N12 (statistic r_{21}), and N13 (statistic r_{22}) were "believed to be accurate to within three or four units in the third (decimal) place".

These tests have been widely used –and are still in use– in the outlier-based scheme for correctly estimating the location and scale parameters (*e.g.*, Thomulka and Lange, 1996; Freeman *et al.*, 1997; Hanson *et al.*, 1998; Verma *et al.*, 1998; Woitge *et al.*, 1998; Muranaka, 1999; Tigges *et al.*, 1999; Taylor, 2000; Hofer and Murphy, 2000; Buckley and Georgianna, 2001; Langton *et al.*, 2002; Reed *et al.*, 2002; Stancak *et al.*, 2002; Yurewicz, 2004; Kern *et al.*, 2005). However, these tests are applicable to only samples of sizes up to 30, which severely limits their application in many scientific and engineering fields, because, today, the number of individual data in a statistical sample has considerably increased (to much greater than 30) than was customary a few decades ago. Furthermore, Gawlowski *et al.* (1998) considered the Dixon tests for normal univariate samples as inferior to the Grubbs tests because the critical values for the former (quoted to only three significant digits, or 3 decimal places; Dixon, 1951) are less accurate than for the latter (quoted to four significant digits, or 3 or 4 decimal places depending on the critical values being >1 or <1; Grubbs and Beck, 1972). In fact, other reasons (see pp. 121-125 and p. 222 in Barnett and Lewis, 1994) might account for the relative efficiency of discordance tests than the one stated by Gawlowski *et al.* (1998).

The computation of new critical values for Dixon discordance tests through Monte Carlo simulations was motivated from multiple reasons: (1) The still wide use of these tests by researchers in many scientific and engineering fields (see selected references for the past ten years 1996-2005 cited above); (2) the availability of critical values for Dixon tests with 3 decimal places as compared to Grubbs tests with critical values with 3 or 4 decimal places; and most importantly (3) the inapplicability of these discordance tests to the actual data for numerous chemical elements in reference materials (RMs) in the field of (a) alloy industry (*e.g.*, Roelandts, 1994); (b) biology (Ihnat, 2000); (c) biomedicine (Patriarca *et al.*, 2005); (d) cement industry (Sieber *et al.*, 2002); (e) food industry (In't Veld, 1998, Langton *et al.*, 2002); (f) environmental research (Dybczyński *et al.*, 1998; Gill *et al.*, 2004; Holcombe *et al.*, 2004); (g) rock geochemistry (*e.g.*, Guevara *et al.*, 2001); and (h) soil science (Dybczyński *et al.*, 1979; Hanson *et al.*, 1998; Verma *et al.*, 1998), as well as to experimental data in numerous other scientific and engineering applications as will be explained later in this paper.

We included all six discordance tests (N7 and N9-N13; see pp. 218-236 of Barnett and Lewis, 1994), initially proposed by Dixon (1950, 1951, 1953), for simulating new, precise, and accurate critical values for n up to 100 (number of data in a given statistical sample, $n = 3(1)100$ for test N7, *i.e.*, for *all* values of n between 3 and 100; $n = 4(1)100$ for tests N9 and N11; $n = 5(1)100$ for tests N10 and N12; and $n = 6(1)100$ for test N13). The minimum number of data to be tested in a given sample (*i.e.*, the minimum sample size) varies from 3 to 6 depending on the type of statistics to be computed (Table 1).

In this paper, we outline the simulation procedure and present new critical values for all six discordance tests and their comparison with the available literature critical values for n up to 30. We also highlight applications to evaluate experimental data in different science or engineering fields, including many branches of earth sciences.

SIX DIXON DISCORDANCE TESTS (N7 AND N9-N13)

Assume a univariate data set (a random sample from a normal population) of n observations represented by an array: $x_1, x_2, x_3,..., x_{n-2}, x_{n-1}, x_n$. If we arrange these data in ascending order, from the lowest to the highest observations, we may call the new array as: $x_1, x_2, x_3,...,x_{n-2}, x_{n-1}, x_n$ where $x_{(1)}$ is the lowest observation and $x_{(n)}$ is the highest one.

Tests N7, N9, and N10 are discordance tests for an extreme outlier ($x_{(n)}$ or $x_{(1)}$) in a normal sample with population variance (σ^2) unknown, whereas tests N11-N13 are for two extreme observations (either the upper-pair $x_{(n)}$, $x_{(n-1)}$ or the lower-pair $x_{(1)}$, $x_{(2)}$) in a similar normal sample. The corresponding test statistics are given in Table 1.

As an example, the test statistic for test N7 is:

$$TN7 = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}} \qquad (1)$$

Suppose $x_{(n)}$ is an outlier, *i.e.*, it appears unusually far from the rest of the sample. The procedure for testing $x_{(n)}$ includes first the computation of the statistic *TN7* (equation 1) for an actual data set under evaluation. It is said that the value $x_{(n)}$ is under evaluation, *i.e.*, tested to see if it was drawn from the same normal population as the rest of the sample (null hypothesis H_0), or it came from a different normal sample (with a different mean or a different variance or both), *i.e.*, if it happens to be a discordant outlier (alternate hypothesis H_1).

The computed value of test statistic *TN7* is then compared with the critical value (percentage point) for a given number of observations n and at a given confidence level (CL) or significance level (SL or α), generally recommended to be 99% CL or 1% SL (or 0.01 α) or even more strict; for most applications in science and engineering (*e.g.*, Verma, 1997, 1998; Gawlowski *et al.*, 1998), although less strict CL of 95% or 5% SL (or 0.05 α) (*e.g.*, Dybczyński *et al.*, 1979; Dybczyński, 1980; Rorabacher, 1991) or even 90% or 10% SL (or 0.10 α) (*e.g.*, Ebdon, 1988 suggested 10% SL for some other statistical tests) have also been used. If computed *TN7* is less than the critical value at a given confidence level, H_0 is said to be true at that particular confidence level, *i.e.*, there is no outlier at the chosen confidence level. But if computed *TN7* is greater than the respective critical value at a given confidence level, H_0 is said to be false and, consequently, H_1 is said to be true at that particular confidence level, *i.e.*, the observation tested ($x_{(n)}$) by *TN7* is detected as a discordant outlier which can

Table 1. Dixon discordance tests for univariate normal samples (modified after Dixon, 1951; Barnett and Lewis, 1994)

Test code *	Value(s) Tested		Test statistic	Test significance	Applicability of test n_{min} - n_{max} Literature	Present work
N7 (r_{10})	Upper	$x_{(n)}$	$TN7 = (x_{(n)}-x_{(n-1)})/(x_{(n)}-x_{(1)})$	Greater	3 – 30	3 – 100
N9 (r_{11})	Upper	$x_{(n)}$	$TN9_u = (x_{(n)}-x_{(n-1)})/(x_{(n)}-x_{(2)})$	Greater	4 – 30	4 – 100
	Lower	$x_{(1)}$	$TN9_l = (x_{(2)}-x_{(1)})/(x_{(n-1)}-x_{(1)})$	Greater	4 – 30	4 – 100
N10 (r_{12})	Upper	$x_{(n)}$	$TN10_u = (x_{(n)}-x_{(n-1)})/(x_{(n)}-x_{(3)})$	Greater	5 – 30	5 – 100
	Lower	$x_{(1)}$	$TN10_l = (x_{(2)}-x_{(1)})/(x_{(n-2)}-x_{(1)})$	Greater	5 – 30	5 – 100
N11 (r_{20})	Upper pair	$x_{(n)}, x_{(n-1)}$	$TN11_{up} = (x_{(n)}-x_{(n-2)})/(x_{(n)}-x_{(1)})$	Greater	4 – 30	4 – 100
	Lower pair	$x_{(1)}, x_{(2)}$	$TN11_{lp} = (x_{(3)}-x_{(1)})/(x_{(n)}-x_{(1)})$	Greater	4 – 30	4 – 100
N12 (r_{21})	Upper pair	$x_{(n)}, x_{(n-1)}$	$TN12_{up} = (x_{(n)}-x_{(n-2)})/(x_{(n)}-x_{(2)})$	Greater	5 – 30	5 – 100
	Lower pair	$x_{(1)}, x_{(2)}$	$TN12_{lp} = (x_{(3)}-x_{(1)})/(x_{(n-1)}-x_{(1)})$	Greater	5 – 30	5 – 100
N13 (r_{22})	Upper pair	$x_{(n)}, x_{(n-1)}$	$TN13_{up} = (x_{(n)}-x_{(n-2)})/(x_{(n)}-x_{(3)})$	Greater	6 – 30	6 – 100
	Lower pair	$x_{(1)}, x_{(2)}$	$TN13_{lp} = (x_{(3)}-x_{(1)})/(x_{(n-2)}-x_{(1)})$	Greater	6 – 30	6 – 100

* Test code (N series) is from Barnett and Lewis (1994), whereas test code (r series) is from Dixon (1951). The symbols for test statistics *TN7*, $TN9_u$, etc. are proposed in the present work. The subscripts u, l, up and lp are, respectively, upper (the highest), lower (the lowest), upper pair, and lower pair observations. Finally, note that, in the present work, n_{max} has been increased to 100 for all tests (see Tables 2-7).

then be discarded, and the test applied consecutively for other extreme values until H_0 is true.

Similar reasoning is valid for other single-value outlier tests N9 and N10 as well as for an upper- or lower-pair outlier tests N11-N13.

Some critical values for these tests were estimated by Dixon (1951), and are available only for n up to 30. Different kinds of interpolations of these values for n up to 30 have also been reported (Bugner and Rutledge, 1990; Rorabacher, 1991). Thus, because of the unavailability of critical values for $n > 30$, this test could not be applied for such data sets (with $n > 30$).

SIMULATION PROCEDURE

Our Monte Carlo type simulation procedure for new, precise, and accurate critical values for six Dixon discordance tests N7 and N9-N13 can be summarized in the following five steps:

(1) *Generating random numbers uniformly distributed in the space (0, 1), i.e., samples from a uniform U(0, 1) distribution*: After exploring a number of different generators for their properties, the Marsenne Twister algorithm of Matsumoto and Nishimura (1998) was employed because this seems to be a widely used generator with a very long (2^{19937}-1) period – a highly desirable property for such applications (Law and Kelton, 2000). Thus, a total of 20 different and independent streams were generated, each one consisting of at least 5,000,000 or more random numbers (IID U(0, 1)). In this way, more than 100,000,000 random numbers of 64 bits were generated.

(2) *Testing of the random numbers if they resemble independent and identically distributed IID U(0, 1) random variates*: Each stream was tested for randomness using Marsaglia (1968) two- and three-dimensional plot method (see also Law and Kelton, 2000, for more details). Two- and three-dimensional typical plots are shown in Figures 1 and 2, respectively. The simulated data clearly fill the (0, 1) space as required by this randomness test in both two- (Figure 1 a-c for 10,000, 100,000, and 10,000,000 random numbers, respectively) and three-dimensions (Figure 2 a-c). Another test for randomness was also applied, which checks how many individual numbers are actually repeated in a given stream of random numbers, and if such repeat-numbers are few, the simulated random numbers can be safely used for further applications. On the average, only around 1 number out of 100,000 numbers in individual streams of IID U(0, 1) was repeated. Between two streams the repeat-numbers were on the average around 3 in 200,000 combined numbers, amounting to about 150 in the combined total of 10,000,000 numbers for two streams. Thus, because the repeat-numbers were so few, all 20 streams were considered appropriate for further work.

(3) *Converting the random numbers to continuous random variates for a normal distribution N(0, 1)*: The polar method of Marsaglia and Bray (1964) was employed instead of the somewhat slower trigonometric method of Box and Muller (1958). Further, this polar method was found to be sufficiently fast for our simulations and, therefore, we did not investigate any other faster scheme such as the algorithm proposed by Kinderman and Ramage (1976). Two parallel streams of random numbers (R_1 and R_2) were used for generating one set of IID N(0, 1) normal random variates. Thus, from 20 different streams of IID U(0, 1), 10 sets of N(0, 1) were obtained, each one of the size ~10,000,000. These simulated data were graphically examined for normality. A typical density function graph (by converting the simulated data N(0, 1) to this function using the well-known conversion equation; see Law and Kelton, 2000 or Verma, 2005 for details) is shown in Figure 3, where the data seem to approximate a normal distribution (Gaussian) curve extremely well. The "imperfections" towards the beginning $[(x\text{-}\mu) < -5]$ and end $[(x\text{-}\mu) > +5]$ of the density curve (Figure 3) are expected (see Verma, 2005 for more details) for the sample size of 10,000,000 to represent the density function ($-\infty$ to $+\infty$). Practically no repeat-numbers were found in tests with 100,000 numbers in these sets of random normal variates. Therefore, the data were considered of a high quality to represent a normal distribution and could, therefore, be safely used for further applications.

(4) *Computing test statistics from random normal samples of sizes up to 100*: From each of these 10 sets of IID N(0, 1), random samples of sizes from 3 to 100 were drawn sequentially from each set of N(0, 1) and test statistics for all 6 tests (*TN7* and *TN9-TN13*; Table 1) were computed for each of these 10 sets of random normal variates. Thus, 10 sets of 100,000 test statistics for each value of n from 3 (for test N7) or 4 (for tests N9 and N11) or 5 (for tests N10 and N12) or 6 (for N13) to 100 were generated.

(5) *Inferring critical values and evaluating their reliability*: Critical values (percentage points) were computed for each of the 10 sets of 100,000 simulated test statistic values for sample sizes from 3, or 4, or 5, or 6 (depending on the type of test statistic) to 100 for all six discordance tests (N7 and N9-N13) and for different values of α from 0.30 to 0.005. By maintaining the number of individual test statistic values constant (100,000) irrespective of the size of the samples (from 3 to 100), we wanted to simulate critical values with similar standard errors of the mean for the entire set of n up to 100. This procedure was accomplished for each formula (Table 1), obtaining 100,000 test statistics for each of the 10 independent simulated sets of univariate normal samples. Thus, for test N7 (with only one test statistic formula; Table 1) only 10 such sets were generated, whereas for all other discordance tests (N9-N13), 20 sets of critical values were generated (10 sets for each formula given in Table 1). The final overall mean and median (central tendency) as well as standard deviation and standard error of the mean (dispersion) parameters were computed from 10 sets of values for test N7 and from 20 sets for tests N9-N13.

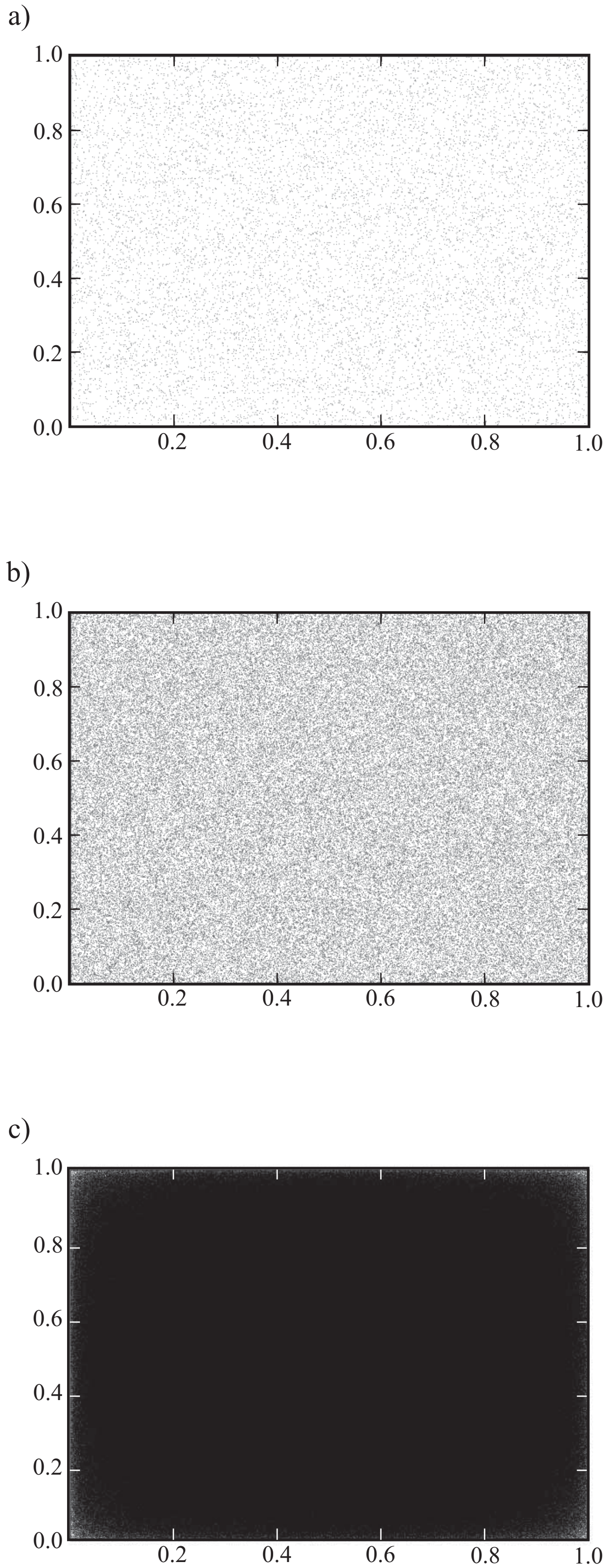

Figure 1. A typical two-dimensional lattice structure plot of the simulated IID U(0, 1) data for testing them for randomness. The x- and y-axes are U_i and U_{i+1}, respectively, where U_1, U_2,… is the sequence of random numbers generated in this work. Note that the simulated data clearly fill the two-dimensional space as required by this randomness test. (a) plot of only one set of 10,000 sequential data; (b) a set of 100,000 sequential data; and (c) the entire 10,000,000 sequential data for a given set.

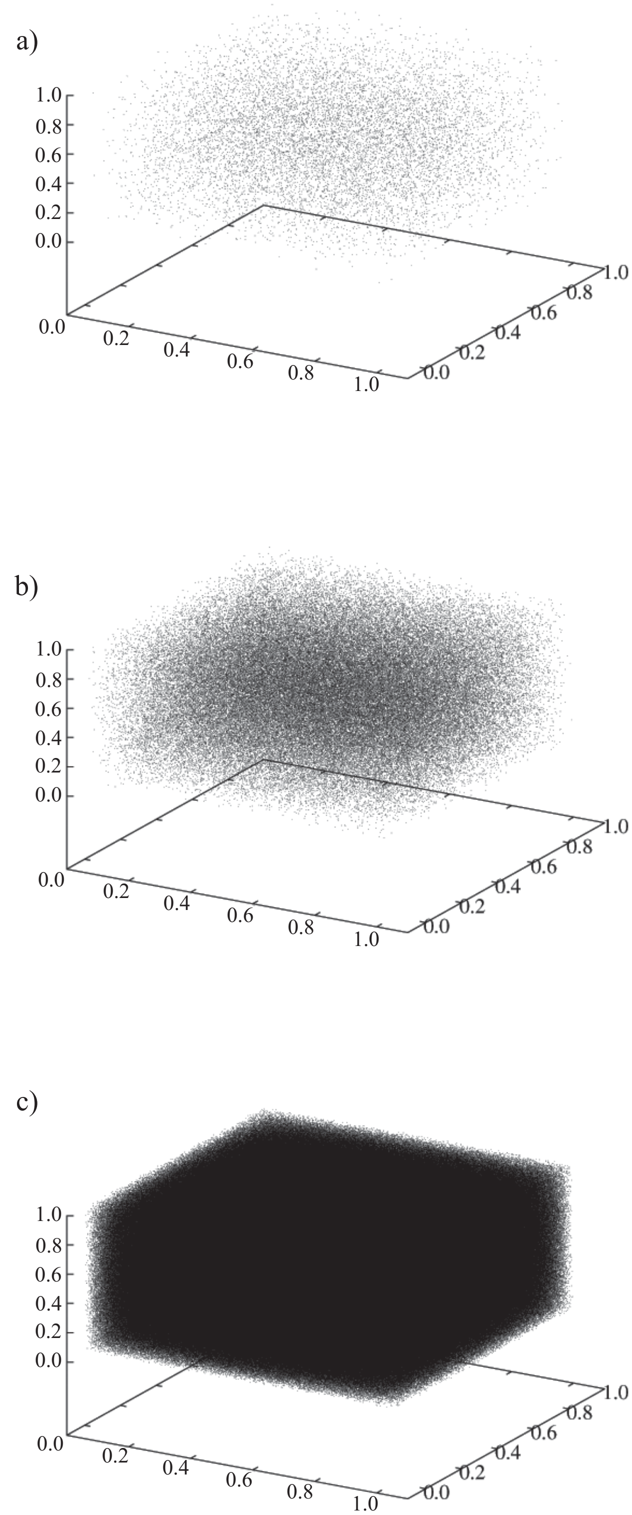

Figure 2. A typical three-dimensional lattice structure plot of the simulated IID U(0, 1) data for testing them for randomness. The x-, y-, and z-axes are U_i, U_{i+1}, and U_{i+2}, respectively, where U_1, U_2, U_3,… is the sequence of random numbers generated in this work. Note that the simulated data clearly fill the space as required by this randomness test. The data plotted in (a), (b) and (c) are the same as in Figure 1.

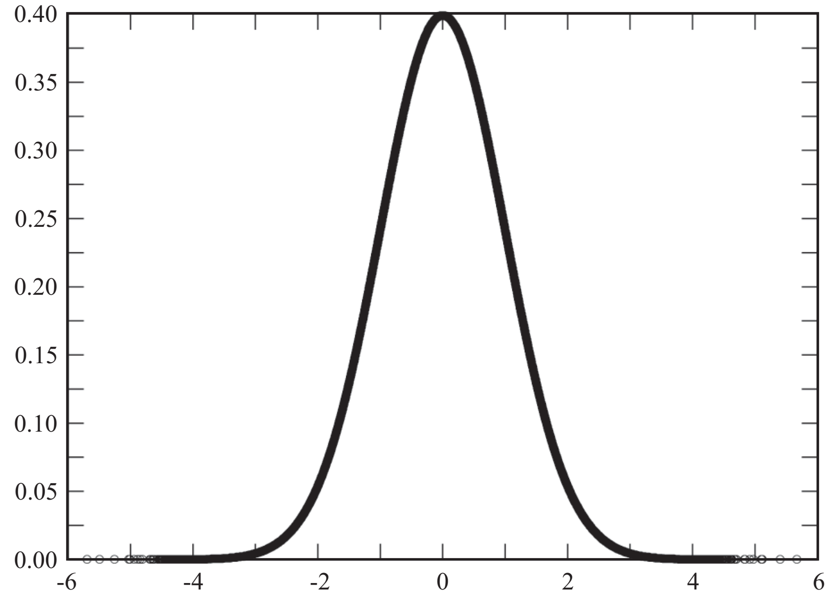

Figure 3. A typical density function f(x) [y-axis] versus (x-μ) [x-axis] plot. The density function *f*(*x*) (or probability PG(*x*, μ, σ)) was estimated from one set of 10,000,000 normal random variates generated in this work through Monte Carlo simulations (generation of random numbers IID U(0, 1) and their conversion to a normal distribution IID N(0, 1) variates).

RESULTS OF NEW CRITICAL VALUES FOR TESTS N7 AND N9-N13

The new critical values for discordance tests N7 and N9-N13, for n from 3 up to 100 and $\alpha = 0.30$ to 0.005 (corresponding to confidence levels of 70% to 99.5% or significance level of 30% to 0.5%) are summarized in Tables 2-7. These new critical value data (Tables 2-7), along with their uncertainty estimates, are available in other formats such as *txt* or *Excel* or *Statistica*, on request from any of the authors (S.P. Verma spv@cie.unam.mx, or A. Quiroz-Ruiz aqr@cie.unam.mx).

For each n (up to 100) and α (= 0.30, 0.20, 0.10, 0.05, 0.02, 0.01, 0.005), these critical values (Tables 2-7) are the mean (x) of 10 individual simulation sets of results for test N7 and of 20 sets for tests N9-N13. It should be noted that the two formulae for a given test (N9-N13) gave very similar critical values and, therefore, these values could be combined to report single tables with more precise results (20 sets of simulations) than for test N7 (based on only 10 sets). The median critical values were found to be in close agreement with these mean values, ascertaining the simulated critical values of N7 and N9-N13 to be also normally distributed. The median values are not tabulated to limit the length of this paper and also because the sample mean for a "normal" sample appears to provide a better approximation for the central tendency than does the sample median (see *e.g.*, Rorabacher, 1991). The standard error of the mean ($\overline{x}_{se}$) was also computed and generally found to be ≤ 0.0001, irrespective of the actual value of n from 3 (or 4, or 5, or 6) to 100. The mean values of these standard errors are summarized in the footnotes of Tables 2-7 to give a clear idea of the reliability (precision and accuracy) of our new critical values (or percentage points). Thus, our present critical values are much more reliable (error is on the fourth or even on the fifth decimal place; see footnotes of Tables 2-7) than the earlier literature values quoted to only 3 decimal places (with their approximate errors ~0.001–0.004 being on the third decimal place, *i.e.*, ≥ 0.001) for n only up to 30 (Dixon, 1951; Rorabacher, 1991). In fact, this is the first time that the reliability of any set of critical values is being explicitly estimated and clearly reported; however, to limit the length of this paper, the errors available for the simulated critical values were not individually tabulated for each of the data. We also consider our new values highly accurate (accuracy being similar to the precision reported in Tables 2-7) because our simulation procedure was highly elaborated and well-tested at different stages of our work (see evaluations of simulated data in Figures 1 to 3) and our simulated critical values agreed to the exact values of Dixon (1951) for $n = 3$ and 4 to about 0.08% (see below).

Dixon (1951) presented exact solutions only for the cases where $n = 3$ and 4, whereas critical values for $n = 5$, 7, 10, 15, 20, 25, and 30 were calculated by using numerical methods; all other values for n up to 30 were estimated by interpolation with a presumed uncertainty of ± 0.001–0.004. Later, Bugner and Rutledge (1990) used a non-linear model (multi-exponential approximations of the data) to the Dixon's statistical tables. Similarly, Rorabacher (1991) used a cubic regression interpolation of Dixon's critical value data to obtain new critical values for only the so called two-sided test at the 95% confidence level; all other values were simply reproduced from Dixon (1951). Erroneous entries for test N11 corresponding to $\alpha = 0.05$ in

Table 2. Critical values for Dixon-type discordance test N7 of an upper outlier in a normal sample.

n	CL SL α	70% 30% 0.30	80% 20% 0.20	90% 10% 0.10	**95%** **5%** **0.05**	98% 2% 0.02	**99%** **1%** **0.01**	99.5% 0.5% 0.005
3		0.6836	0.7808	0.8850	0.9411	0.9763	0.9881	0.9940
4		0.4704	0.5603	0.6789	0.7651	0.8457	0.8886	0.9201
5		0.3730	0.4508	0.5578	0.6423	0.7291	0.7819	0.8234
6		0.3173	0.3868	0.4840	0.5624	0.6458	0.6987	0.7437
7		0.2811	0.3444	0.4340	0.5077	0.5864	0.6371	0.6809
8		0.2550	0.3138	0.3979	0.4673	0.5432	0.5914	0.6336
9		0.2361	0.2915	0.3704	0.4363	0.5091	0.5554	0.5952
10		0.2208	0.2735	0.3492	0.4122	0.4813	0.5260	0.5658
11		0.2086	0.2586	0.3312	0.3922	0.4591	0.5028	0.5416
12		0.1983	0.2467	0.3170	0.3755	0.4405	0.4831	0.5208
13		0.1898	0.2366	0.3045	0.3615	0.4250	0.4664	0.5034
14		0.1826	0.2280	0.2938	0.3496	0.4118	0.4517	0.4869
15		0.1764	0.2202	0.2848	0.3389	0.3991	0.4385	0.4739
16		0.1707	0.2137	0.2765	0.3293	0.3883	0.4268	0.4614
17		0.1656	0.2077	0.2691	0.3208	0.3792	0.4166	0.4504
18		0.1613	0.2023	0.2626	0.3135	0.3711	0.4081	0.4423
19		0.1572	0.1973	0.2564	0.3068	0.3630	0.4002	0.4333
20		0.1535	0.1929	0.2511	0.3005	0.3562	0.3922	0.4247
21		0.1504	0.1890	0.2460	0.2947	0.3495	0.3854	0.4173
22		0.1474	0.1854	0.2415	0.2895	0.3439	0.3789	0.4109
23		0.1446	0.1820	0.2377	0.2851	0.3384	0.3740	0.4051
24		0.1420	0.1790	0.2337	0.2804	0.3328	0.3674	0.3986
25		0.1397	0.1761	0.2303	0.2763	0.3287	0.3625	0.3935
26		0.1376	0.1735	0.2269	0.2725	0.3242	0.3583	0.3889
27		0.1355	0.1710	0.2237	0.2686	0.3202	0.3543	0.3843
28		0.1335	0.1687	0.2208	0.2655	0.3163	0.3499	0.3801
29		0.1318	0.1664	0.2182	0.2622	0.3127	0.3460	0.3762
30		0.1300	0.1645	0.2155	0.2594	0.3093	0.3425	0.3718
31		0.1283	0.1624	0.2132	0.2567	0.3060	0.3390	0.3685
32		0.1268	0.1604	0.2110	0.2541	0.3036	0.3357	0.3646
33		0.1255	0.1590	0.2088	0.2513	0.2999	0.3323	0.3610
34		0.1240	0.1571	0.2066	0.2488	0.2973	0.3294	0.3583
35		0.1227	0.1555	0.2045	0.2467	0.2948	0.3266	0.3548
36		0.1215	0.1540	0.2026	0.2445	0.2921	0.3238	0.3522
37		0.1202	0.1525	0.2008	0.2423	0.2898	0.3213	0.3498
38		0.1192	0.1512	0.1993	0.2408	0.2879	0.3187	0.3465
39		0.1181	0.1499	0.1974	0.2383	0.2853	0.3163	0.3443
40		0.1169	0.1484	0.1958	0.2366	0.2836	0.3141	0.3415
41		0.1160	0.1472	0.1944	0.2350	0.2815	0.3124	0.3400
42		0.1153	0.1462	0.1930	0.2334	0.2794	0.3102	0.3377
43		0.1141	0.1449	0.1915	0.2319	0.2778	0.3081	0.3353
44		0.1134	0.1441	0.1902	0.2302	0.2758	0.3061	0.3332
45		0.1124	0.1430	0.1890	0.2288	0.2744	0.3050	0.3325
46		0.1116	0.1418	0.1875	0.2273	0.2726	0.3028	0.3298
47		0.1108	0.1408	0.1865	0.2257	0.2711	0.3009	0.3279
48		0.1102	0.1400	0.1850	0.2241	0.2690	0.2991	0.3256
49		0.1093	0.1390	0.1839	0.2228	0.2676	0.2972	0.3235
50		0.1087	0.1381	0.1829	0.2216	0.2662	0.2960	0.3225

Table 2. Critical values for Dixon-type discordance test N7 of an upper outlier in a normal sample (continued).

n	CL / SL / *α*	70% / 30% / 0.30	80% / 20% / 0.20	90% / 10% / 0.10	**95% / 5% / 0.05**	98% / 2% / 0.02	**99% / 1% / 0.01**	99.5% / 0.5% / 0.005
51		0.1079	0.1374	0.1819	0.2206	0.2651	0.2941	0.3204
52		0.1071	0.1365	0.1808	0.2191	0.2632	0.2927	0.3191
53		0.1067	0.1357	0.1797	0.2182	0.2620	0.2920	0.3177
54		0.1060	0.1349	0.1788	0.2169	0.2606	0.2899	0.3163
55		0.1052	0.1340	0.1777	0.2160	0.2595	0.2880	0.3140
56		0.1047	0.1334	0.1768	0.2145	0.2582	0.2873	0.3136
57		0.1041	0.1326	0.1759	0.2135	0.2570	0.2859	0.3118
58		0.1036	0.1320	0.1752	0.2126	0.2555	0.2845	0.3098
59		0.1030	0.1312	0.1741	0.2116	0.2545	0.2828	0.3089
60		0.1024	0.1304	0.1733	0.2106	0.2531	0.2816	0.3075
61		0.1019	0.1299	0.1726	0.2095	0.2522	0.2812	0.3071
62		0.1014	0.1294	0.1717	0.2085	0.2510	0.2792	0.3061
63		0.1009	0.1286	0.1707	0.2075	0.2500	0.2784	0.3041
64		0.1004	0.1281	0.1703	0.2070	0.2493	0.2775	0.3031
65		0.1000	0.1275	0.1694	0.2057	0.2480	0.2766	0.3025
66		0.0997	0.1272	0.1689	0.2053	0.2472	0.2754	0.3006
67		0.0991	0.1264	0.1679	0.2045	0.2466	0.2742	0.2996
68		0.0987	0.1260	0.1674	0.2037	0.2457	0.2735	0.2990
69		0.0982	0.1254	0.1667	0.2030	0.2445	0.2724	0.2983
70		0.0979	0.1249	0.1660	0.2020	0.2436	0.2714	0.2968
71		0.0974	0.1243	0.1652	0.2013	0.2429	0.2709	0.2959
72		0.0970	0.1238	0.1648	0.2005	0.2420	0.2696	0.2946
73		0.0967	0.1234	0.1641	0.1996	0.2409	0.2682	0.2934
74		0.0961	0.1228	0.1635	0.1990	0.2402	0.2677	0.2932
75		0.0960	0.1225	0.1631	0.1984	0.2398	0.2667	0.2922
76		0.0955	0.1221	0.1626	0.1980	0.2387	0.2662	0.2912
77		0.0952	0.1217	0.1620	0.1973	0.2382	0.2656	0.2905
78		0.0948	0.1212	0.1613	0.1964	0.2372	0.2646	0.2897
79		0.0943	0.1205	0.1605	0.1955	0.2365	0.2637	0.2885
80		0.0939	0.1201	0.1601	0.1950	0.2360	0.2633	0.2876
81		0.0937	0.1198	0.1596	0.1943	0.2349	0.2621	0.2870
82		0.0935	0.1195	0.1594	0.1940	0.2345	0.2614	0.2859
83		0.0930	0.1189	0.1586	0.1934	0.2337	0.2608	0.2852
84		0.0928	0.1187	0.1583	0.1927	0.2330	0.2599	0.2844
85		0.0925	0.1182	0.1576	0.1922	0.2322	0.2588	0.2836
86		0.0921	0.1178	0.1573	0.1918	0.2319	0.2584	0.2832
87		0.0918	0.1174	0.1567	0.1909	0.2309	0.2573	0.2818
88		0.0915	0.1171	0.1563	0.1906	0.2304	0.2568	0.2811
89		0.0913	0.1167	0.1557	0.1899	0.2298	0.2566	0.2808
90		0.0910	0.1165	0.1554	0.1896	0.2294	0.2558	0.2798
91		0.0906	0.1160	0.1547	0.1887	0.2285	0.2548	0.2790
92		0.0903	0.1156	0.1544	0.1885	0.2279	0.2543	0.2788
93		0.0902	0.1154	0.1540	0.1881	0.2272	0.2539	0.2784
94		0.0899	0.1151	0.1537	0.1876	0.2272	0.2535	0.2775
95		0.0896	0.1147	0.1532	0.1869	0.2259	0.2524	0.2766
96		0.0894	0.1144	0.1528	0.1865	0.2257	0.2521	0.2764
97		0.0892	0.1141	0.1524	0.1860	0.2251	0.2512	0.2755
98		0.0890	0.1138	0.1521	0.1856	0.2247	0.2513	0.2751
99		0.0887	0.1134	0.1516	0.1851	0.2240	0.2499	0.2738
100		0.0885	0.1131	0.1512	0.1846	0.2234	0.2498	0.2737

CL: Confidence level (%); SL: Significance level (%); α: Significance level. Headers for commonly used CL or SL or α are given in bold face (*e.g.*, for RM applications). The mean values of the standard error of the mean ($\overline{x}_{se}$) for these critical values ($\overline{x}$) are (respective % errors are also reported in parentheses): ~0.00011 (for α = 0.30, 0.09%); ~0.00011 (for α = 0.20, 0.07%); ~0.00009 (for α = 0.10, 0.041%); ~0.00008 (for α = 0.05, 0.029%); ~0.00007 (for α = 0.02, 0.020%); ~0.000043 (for α = 0.01, 0.012%); and ~0.000028 (for α= 0.005, 0.007%).

Table 3. Critical values for Dixon-type discordance test N9 of an upper or lower outlier in a normal sample.

n	CL SL *α*	70% 30% 0.30	80% 20% 0.20	90% 10% 0.10	**95%** **5%** **0.05**	98% 2% 0.02	**99%** **1%** **0.01**	99.5% 0.5% 0.005
4		0.7351	0.8226	0.9105	0.9550	0.9820	0.9910	0.9955
5		0.5244	0.6147	0.7281	0.8064	0.8762	0.9120	0.9376
6		0.4208	0.5020	0.6098	0.6916	0.7722	0.8185	0.8544
7		0.3589	0.4328	0.5332	0.6111	0.6918	0.7399	0.7812
8		0.3177	0.3858	0.4793	0.5539	0.6321	0.6808	0.7226
9		0.2883	0.3519	0.4404	0.5114	0.5876	0.6346	0.6757
10		0.2660	0.3260	0.4102	0.4778	0.5509	0.5972	0.6375
11		0.2483	0.3054	0.3857	0.4510	0.5215	0.5663	0.6055
12		0.2341	0.2886	0.3660	0.4291	0.4977	0.5412	0.5796
13		0.2223	0.2747	0.3496	0.4111	0.4780	0.5208	0.5581
14		0.2125	0.2632	0.3357	0.3955	0.4605	0.5026	0.5396
15		0.2039	0.2530	0.3236	0.3819	0.4456	0.4868	0.5229
16		0.1965	0.2443	0.3129	0.3698	0.4322	0.4723	0.5080
17		0.1900	0.2365	0.3037	0.3594	0.4204	0.4595	0.4945
18		0.1843	0.2297	0.2952	0.3500	0.4102	0.4495	0.4845
19		0.1791	0.2235	0.2878	0.3418	0.4010	0.4395	0.4734
20		0.1744	0.2178	0.2810	0.3340	0.3926	0.4303	0.4639
21		0.1703	0.2128	0.2747	0.3271	0.3847	0.4220	0.4551
22		0.1664	0.2083	0.2692	0.3207	0.3776	0.4143	0.4472
23		0.1630	0.2041	0.2644	0.3151	0.3714	0.4081	0.4406
24		0.1596	0.2002	0.2593	0.3092	0.3646	0.4006	0.4325
25		0.1567	0.1965	0.2550	0.3043	0.3591	0.3949	0.4267
26		0.1540	0.1933	0.2509	0.2995	0.3537	0.3893	0.4208
27		0.1514	0.1901	0.2472	0.2954	0.3493	0.3848	0.4158
28		0.1489	0.1872	0.2435	0.2912	0.3444	0.3795	0.4107
29		0.1466	0.1845	0.2404	0.2874	0.3403	0.3748	0.4053
30		0.1446	0.1821	0.2371	0.2837	0.3362	0.3702	0.4007
31		0.1425	0.1795	0.2341	0.2804	0.3324	0.3662	0.3972
32		0.1406	0.1773	0.2313	0.2773	0.3289	0.3625	0.3929
33		0.1390	0.1753	0.2289	0.2744	0.3254	0.3590	0.3891
34		0.1372	0.1731	0.2262	0.2714	0.3222	0.3555	0.3855
35		0.1357	0.1712	0.2238	0.2686	0.3193	0.3522	0.3816
36		0.1342	0.1693	0.2215	0.2662	0.3164	0.3491	0.3788
37		0.1327	0.1676	0.2194	0.2636	0.3134	0.3462	0.3753
38		0.1314	0.1660	0.2175	0.2613	0.3108	0.3432	0.3720
39		0.1301	0.1644	0.2153	0.2588	0.3078	0.3400	0.3689
40		0.1287	0.1627	0.2133	0.2567	0.3057	0.3378	0.3666
41		0.1275	0.1613	0.2117	0.2546	0.3035	0.3353	0.3642
42		0.1265	0.1600	0.2099	0.2527	0.3011	0.3329	0.3615
43		0.1253	0.1587	0.2081	0.2507	0.2988	0.3305	0.3589
44		0.1244	0.1573	0.2068	0.2489	0.2967	0.3282	0.3562
45		0.1232	0.1561	0.2051	0.2473	0.2951	0.3265	0.3547
46		0.1223	0.1549	0.2036	0.2456	0.2930	0.3242	0.3520
47		0.1213	0.1536	0.2021	0.2438	0.2908	0.3219	0.3498
48		0.1204	0.1526	0.2008	0.2422	0.2890	0.3202	0.3482
49		0.1194	0.1514	0.1993	0.2405	0.2873	0.3183	0.3460
50		0.1186	0.1504	0.1981	0.2391	0.2858	0.3167	0.3441

Table 3. Critical values for Dixon-type discordance test N9 of an upper or lower outlier in a normal sample (continued).

n	CL SL α	70% 30% 0.30	80% 20% 0.20	90% 10% 0.10	**95%** **5%** **0.05**	98% 2% 0.02	**99%** **1%** **0.01**	99.5% 0.5% 0.005
51		0.1178	0.1495	0.1968	0.2377	0.2843	0.3150	0.3423
52		0.1169	0.1484	0.1957	0.2361	0.2821	0.3129	0.3403
53		0.1162	0.1476	0.1946	0.2351	0.2809	0.3116	0.3388
54		0.1155	0.1466	0.1932	0.2336	0.2792	0.3095	0.3371
55		0.1146	0.1456	0.1920	0.2322	0.2778	0.3078	0.3346
56		0.1140	0.1448	0.1911	0.2309	0.2764	0.3069	0.3337
57		0.1133	0.1439	0.1901	0.2297	0.2751	0.3052	0.3319
58		0.1126	0.1431	0.1891	0.2286	0.2737	0.3040	0.3304
59		0.1119	0.1422	0.1878	0.2274	0.2724	0.3021	0.3287
60		0.1113	0.1415	0.1869	0.2261	0.2708	0.3002	0.3269
61		0.1107	0.1407	0.1860	0.2251	0.2697	0.2995	0.3257
62		0.1101	0.1400	0.1851	0.2239	0.2685	0.2982	0.3249
63		0.1095	0.1392	0.1840	0.2227	0.2673	0.2969	0.3236
64		0.1090	0.1386	0.1834	0.2220	0.2665	0.2957	0.3222
65		0.1085	0.1380	0.1825	0.2209	0.2652	0.2946	0.3208
66		0.1079	0.1374	0.1817	0.2200	0.2641	0.2934	0.3194
67		0.1074	0.1366	0.1808	0.2193	0.2632	0.2924	0.3185
68		0.1069	0.1360	0.1800	0.2182	0.2620	0.2913	0.3175
69		0.1062	0.1353	0.1791	0.2173	0.2609	0.2899	0.3164
70		0.1059	0.1348	0.1786	0.2164	0.2600	0.2890	0.3148
71		0.1054	0.1341	0.1777	0.2156	0.2594	0.2884	0.3139
72		0.1049	0.1336	0.1770	0.2145	0.2578	0.2867	0.3124
73		0.1045	0.1330	0.1764	0.2139	0.2569	0.2855	0.3113
74		0.1040	0.1325	0.1756	0.2129	0.2562	0.2851	0.3110
75		0.1037	0.1320	0.1751	0.2124	0.2554	0.2838	0.3096
76		0.1031	0.1316	0.1745	0.2117	0.2546	0.2828	0.3087
77		0.1028	0.1311	0.1738	0.2109	0.2537	0.2821	0.3078
78		0.1023	0.1305	0.1730	0.2101	0.2525	0.2811	0.3069
79		0.1020	0.1300	0.1723	0.2092	0.2518	0.2802	0.3057
80		0.1015	0.1295	0.1718	0.2087	0.2510	0.2793	0.3046
81		0.1012	0.1290	0.1712	0.2077	0.2503	0.2785	0.3040
82		0.1008	0.1285	0.1707	0.2074	0.2496	0.2777	0.3030
83		0.1004	0.1280	0.1700	0.2065	0.2487	0.2769	0.3022
84		0.1001	0.1277	0.1696	0.2059	0.2479	0.2760	0.3009
85		0.0997	0.1271	0.1689	0.2052	0.2470	0.2748	0.3005
86		0.0992	0.1267	0.1684	0.2047	0.2465	0.2742	0.2996
87		0.0989	0.1262	0.1678	0.2041	0.2457	0.2736	0.2986
88		0.0986	0.1259	0.1673	0.2035	0.2449	0.2724	0.2979
89		0.0983	0.1254	0.1668	0.2028	0.2445	0.2724	0.2973
90		0.0980	0.1251	0.1663	0.2021	0.2437	0.2711	0.2959
91		0.0976	0.1247	0.1657	0.2016	0.2433	0.2706	0.2958
92		0.0972	0.1242	0.1652	0.2011	0.2423	0.2699	0.2947
93		0.0970	0.1239	0.1647	0.2005	0.2418	0.2690	0.2941
94		0.0968	0.1235	0.1643	0.2000	0.2412	0.2686	0.2936
95		0.0963	0.1231	0.1638	0.1993	0.2403	0.2678	0.2926
96		0.0961	0.1227	0.1634	0.1989	0.2399	0.2671	0.2920
97		0.0958	0.1224	0.1629	0.1983	0.2390	0.2665	0.2910
98		0.0956	0.1221	0.1624	0.1978	0.2388	0.2660	0.2905
99		0.0952	0.1216	0.1619	0.1972	0.2379	0.2649	0.2895
100		0.0950	0.1213	0.1616	0.1968	0.2374	0.2646	0.2893

For the explanation of abbreviations see footnote of Table 2. The mean values of the standard error of the mean ($\overline{x}_{se}$) for these critical values ($\overline{x}$) are (respective % errors are also reported in parentheses): ~0.00008 (for α = 0.30, 0.06%); ~0.00008 (for α = 0.20, 0.044%); ~0.00007 (for α = 0.10, 0.029%); ~0.00006 (for α = 0.05, 0.021%); ~0.00005 (for α = 0.02, 0.015%); ~0.000034 (for α = 0.01, 0.009%); and ~0.000023 (for α = 0.005, 0.005%).

Table 4. Critical values for Dixon-type discordance test N10 of an upper or lower outlier in a normal sample.

n	CL SL α	70% 30% 0.30	80% 20% 0.20	90% 10% 0.10	**95%** **5%** **0.05**	98% 2% 0.02	**99%** **1%** **0.01**	99.5% 0.5% 0.005
5		0.7551	0.8381	0.9195	0.9597	0.9839	0.9920	0.9960
6		0.5503	0.6399	0.7503	0.8248	0.8895	0.9220	0.9448
7		0.4450	0.5281	0.6356	0.7152	0.7918	0.8347	0.8683
8		0.3814	0.4572	0.5588	0.6365	0.7151	0.7618	0.8005
9		0.3385	0.4091	0.5048	0.5798	0.6575	0.7047	0.7447
10		0.3073	0.3735	0.4645	0.5361	0.6116	0.6587	0.6984
11		0.2836	0.3461	0.4329	0.5018	0.5752	0.6207	0.6603
12		0.2646	0.3242	0.4076	0.4740	0.5455	0.5903	0.6290
13		0.2496	0.3065	0.3868	0.4519	0.5215	0.5651	0.6031
14		0.2370	0.2921	0.3697	0.4327	0.5006	0.5431	0.5803
15		0.2263	0.2793	0.3547	0.4162	0.4827	0.5243	0.5613
16		0.2170	0.2685	0.3418	0.4015	0.4665	0.5076	0.5435
17		0.2090	0.2590	0.3304	0.3889	0.4523	0.4929	0.5285
18		0.2019	0.2507	0.3203	0.3779	0.4406	0.4807	0.5161
19		0.1958	0.2433	0.3116	0.3682	0.4298	0.4692	0.5041
20		0.1901	0.2365	0.3035	0.3590	0.4199	0.4588	0.4931
21		0.1851	0.2306	0.2962	0.3511	0.4106	0.4493	0.4833
22		0.1804	0.2252	0.2897	0.3436	0.4025	0.4404	0.4742
23		0.1765	0.2202	0.2839	0.3369	0.3952	0.4329	0.4664
24		0.1725	0.2156	0.2781	0.3302	0.3874	0.4244	0.4573
25		0.1690	0.2114	0.2731	0.3245	0.3814	0.4179	0.4507
26		0.1659	0.2076	0.2683	0.3191	0.3752	0.4117	0.4437
27		0.1628	0.2039	0.2641	0.3143	0.3702	0.4062	0.4383
28		0.1599	0.2005	0.2597	0.3095	0.3645	0.4007	0.4325
29		0.1573	0.1973	0.2560	0.3052	0.3600	0.3952	0.4266
30		0.1550	0.1946	0.2524	0.3012	0.3552	0.3905	0.4219
31		0.1526	0.1917	0.2491	0.2972	0.3507	0.3857	0.4169
32		0.1503	0.1890	0.2458	0.2937	0.3472	0.3818	0.4125
33		0.1485	0.1868	0.2430	0.2904	0.3432	0.3776	0.4081
34		0.1465	0.1844	0.2400	0.2870	0.3394	0.3736	0.4039
35		0.1447	0.1821	0.2373	0.2838	0.3360	0.3701	0.4001
36		0.1430	0.1800	0.2347	0.2810	0.3331	0.3669	0.3967
37		0.1413	0.1780	0.2323	0.2783	0.3296	0.3632	0.3928
38		0.1398	0.1762	0.2300	0.2757	0.3267	0.3599	0.3895
39		0.1382	0.1743	0.2277	0.2728	0.3235	0.3564	0.3867
40		0.1367	0.1725	0.2254	0.2705	0.3210	0.3542	0.3839
41		0.1354	0.1709	0.2235	0.2682	0.3186	0.3514	0.3810
42		0.1342	0.1695	0.2215	0.2660	0.3160	0.3485	0.3779
43		0.1329	0.1679	0.2196	0.2638	0.3134	0.3457	0.3750
44		0.1318	0.1665	0.2181	0.2619	0.3110	0.3431	0.3723
45		0.1305	0.1651	0.2162	0.2599	0.3093	0.3415	0.3704
46		0.1295	0.1637	0.2146	0.2582	0.3070	0.3390	0.3674
47		0.1284	0.1623	0.2129	0.2560	0.3045	0.3364	0.3649
48		0.1273	0.1612	0.2114	0.2542	0.3026	0.3345	0.3634
49		0.1263	0.1598	0.2096	0.2525	0.3006	0.3324	0.3606
50		0.1253	0.1587	0.2083	0.2508	0.2989	0.3305	0.3584

Table 4. Critical values for Dixon-type discordance test N10 of an upper or lower outlier in a normal sample (continued).

n CL	70%	80%	90%	**95%**	98%	**99%**	99.5%
SL	30%	20%	10%	**5%**	2%	**1%**	0.5%
α	0.30	0.20	0.10	**0.05**	0.02	**0.01**	0.005
51	0.1245	0.1575	0.2070	0.2494	0.2972	0.3287	0.3565
52	0.1235	0.1564	0.2056	0.2476	0.2951	0.3266	0.3545
53	0.1226	0.1555	0.2044	0.2464	0.2938	0.3250	0.3525
54	0.1219	0.1545	0.2030	0.2446	0.2918	0.3229	0.3511
55	0.1209	0.1533	0.2016	0.2433	0.2901	0.3208	0.3485
56	0.1202	0.1525	0.2006	0.2418	0.2887	0.3198	0.3474
57	0.1194	0.1515	0.1994	0.2405	0.2872	0.3180	0.3452
58	0.1186	0.1505	0.1982	0.2392	0.2856	0.3165	0.3436
59	0.1179	0.1495	0.1970	0.2379	0.2842	0.3148	0.3420
60	0.1172	0.1487	0.1959	0.2365	0.2823	0.3126	0.3398
61	0.1165	0.1478	0.1949	0.2354	0.2814	0.3119	0.3389
62	0.1159	0.1470	0.1938	0.2341	0.2800	0.3104	0.3377
63	0.1152	0.1463	0.1928	0.2329	0.2786	0.3091	0.3361
64	0.1146	0.1455	0.1921	0.2320	0.2776	0.3077	0.3346
65	0.1140	0.1448	0.1911	0.2308	0.2763	0.3065	0.3332
66	0.1134	0.1442	0.1902	0.2299	0.2752	0.3051	0.3316
67	0.1128	0.1433	0.1893	0.2290	0.2742	0.3041	0.3307
68	0.1122	0.1426	0.1883	0.2279	0.2730	0.3028	0.3297
69	0.1116	0.1418	0.1873	0.2268	0.2718	0.3014	0.3284
70	0.1112	0.1413	0.1868	0.2260	0.2706	0.3004	0.3267
71	0.1106	0.1406	0.1858	0.2249	0.2701	0.2997	0.3255
72	0.1101	0.1400	0.1850	0.2238	0.2684	0.2979	0.3241
73	0.1096	0.1394	0.1843	0.2231	0.2673	0.2967	0.3231
74	0.1091	0.1387	0.1834	0.2221	0.2667	0.2961	0.3227
75	0.1088	0.1383	0.1829	0.2215	0.2657	0.2947	0.3213
76	0.1082	0.1378	0.1822	0.2206	0.2647	0.2938	0.3199
77	0.1078	0.1372	0.1815	0.2199	0.2638	0.2928	0.3190
78	0.1072	0.1365	0.1807	0.2189	0.2626	0.2918	0.3181
79	0.1068	0.1360	0.1799	0.2180	0.2618	0.2909	0.3169
80	0.1064	0.1354	0.1793	0.2174	0.2609	0.2900	0.3159
81	0.1060	0.1349	0.1786	0.2164	0.2602	0.2889	0.3150
82	0.1056	0.1344	0.1781	0.2160	0.2593	0.2881	0.3140
83	0.1051	0.1338	0.1773	0.2151	0.2583	0.2873	0.3132
84	0.1047	0.1334	0.1769	0.2143	0.2574	0.2860	0.3120
85	0.1043	0.1328	0.1762	0.2137	0.2565	0.2852	0.3113
86	0.1039	0.1324	0.1755	0.2131	0.2561	0.2844	0.3100
87	0.1035	0.1319	0.1749	0.2123	0.2550	0.2838	0.3094
88	0.1031	0.1315	0.1745	0.2117	0.2543	0.2826	0.3085
89	0.1028	0.1309	0.1739	0.2110	0.2539	0.2824	0.3078
90	0.1025	0.1306	0.1733	0.2102	0.2528	0.2809	0.3063
91	0.1020	0.1302	0.1726	0.2097	0.2522	0.2806	0.3064
92	0.1016	0.1297	0.1721	0.2091	0.2514	0.2796	0.3050
93	0.1014	0.1293	0.1715	0.2084	0.2508	0.2788	0.3043
94	0.1011	0.1289	0.1712	0.2079	0.2503	0.2784	0.3038
95	0.1006	0.1285	0.1706	0.2072	0.2492	0.2774	0.3027
96	0.1004	0.1280	0.1701	0.2068	0.2487	0.2768	0.3020
97	0.1001	0.1276	0.1696	0.2061	0.2479	0.2762	0.3013
98	0.0998	0.1274	0.1691	0.2056	0.2475	0.2755	0.3005
99	0.0994	0.1268	0.1685	0.2049	0.2467	0.2744	0.2994
100	0.0992	0.1266	0.1682	0.2044	0.2461	0.2740	0.2995

For the explanation of abbreviations see footnote of Table 2. The mean values of the standard error of the mean ($\bar{x}_{se}$) for these critical values ($\bar{x}$) are (respective % errors are also reported in parentheses): ~0.00009 (for α = 0.30, 0.06%); ~0.00008 (for α = 0.20, 0.044%); ~0.00007 (for α = 0.10, 0.029%); ~0.00006 (for α = 0.05, 0.021%); ~0.00006 (for α = 0.02, 0.015%); ~0.000037 (for α = 0.01, 0.009%); and ~0.000025 (for α = 0.005, 0.006%).

Table 5. Critical values for Dixon-type discordance test N11 of an upper or lower outlier-pair in a normal sample.

n	CL / SL / α	70% / 30% / 0.30	80% / 20% / 0.20	90% / 10% / 0.10	**95% / 5% / 0.05**	98% / 2% / 0.02	**99% / 1% / 0.01**	99.5% / 0.5% / 0.005
4		0.8069	0.8703	0.9345	0.9669	0.9867	0.9934	0.9967
5		0.6230	0.6934	0.7822	0.8445	0.9005	0.9294	0.9496
6		0.5205	0.5863	0.6734	0.7400	0.8070	0.8458	0.8770
7		0.4556	0.5164	0.5990	0.6640	0.7324	0.7743	0.8091
8		0.4102	0.4671	0.5451	0.6077	0.6741	0.7156	0.7514
9		0.3770	0.4305	0.5048	0.5644	0.6289	0.6694	0.7055
10		0.3511	0.4021	0.4731	0.5305	0.5929	0.6325	0.6674
11		0.3304	0.3793	0.4472	0.5028	0.5635	0.6024	0.6367
12		0.3135	0.3607	0.4264	0.4798	0.5391	0.5769	0.6106
13		0.2994	0.3449	0.4090	0.4613	0.5188	0.5562	0.5888
14		0.2873	0.3316	0.3937	0.4448	0.5012	0.5379	0.5702
15		0.2768	0.3199	0.3803	0.4300	0.4857	0.5213	0.5536
16		0.2676	0.3098	0.3690	0.4178	0.4714	0.5066	0.5383
17		0.2596	0.3008	0.3587	0.4066	0.4598	0.4944	0.5256
18		0.2524	0.2927	0.3498	0.3967	0.4492	0.4831	0.5133
19		0.2458	0.2855	0.3414	0.3878	0.4392	0.4724	0.5029
20		0.2400	0.2788	0.3340	0.3797	0.4301	0.4636	0.4931
21		0.2347	0.2728	0.3272	0.3721	0.4221	0.4547	0.4840
22		0.2297	0.2673	0.3210	0.3655	0.4151	0.4473	0.4763
23		0.2253	0.2623	0.3152	0.3591	0.4082	0.4401	0.4687
24		0.2211	0.2577	0.3100	0.3534	0.4018	0.4331	0.4614
25		0.2172	0.2533	0.3048	0.3478	0.3956	0.4273	0.4555
26		0.2137	0.2494	0.3005	0.3429	0.3902	0.4215	0.4495
27		0.2104	0.2455	0.2961	0.3382	0.3853	0.4160	0.4433
28		0.2074	0.2420	0.2919	0.3336	0.3803	0.4108	0.4381
29		0.2043	0.2388	0.2882	0.3296	0.3757	0.4057	0.4331
30		0.2015	0.2356	0.2845	0.3256	0.3714	0.4014	0.4285
31		0.1989	0.2327	0.2813	0.3219	0.3675	0.3972	0.4237
32		0.1965	0.2300	0.2781	0.3185	0.3636	0.3931	0.4198
33		0.1944	0.2274	0.2749	0.3152	0.3602	0.3896	0.4162
34		0.1921	0.2249	0.2722	0.3120	0.3567	0.3864	0.4130
35		0.1900	0.2225	0.2696	0.3091	0.3535	0.3829	0.4088
36		0.1880	0.2203	0.2669	0.3064	0.3504	0.3795	0.4053
37		0.1861	0.2182	0.2645	0.3035	0.3473	0.3763	0.4023
38		0.1842	0.2160	0.2620	0.3010	0.3444	0.3730	0.3989
39		0.1826	0.2140	0.2597	0.2982	0.3416	0.3704	0.3960
40		0.1808	0.2122	0.2575	0.2960	0.3392	0.3679	0.3931
41		0.1792	0.2103	0.2555	0.2936	0.3365	0.3651	0.3903
42		0.1777	0.2087	0.2536	0.2914	0.3342	0.3624	0.3876
43		0.1763	0.2070	0.2516	0.2894	0.3320	0.3599	0.3854
44		0.1750	0.2055	0.2498	0.2873	0.3298	0.3578	0.3827
45		0.1736	0.2040	0.2481	0.2854	0.3277	0.3555	0.3804
46		0.1723	0.2025	0.2464	0.2836	0.3254	0.3533	0.3784
47		0.1710	0.2010	0.2448	0.2817	0.3234	0.3512	0.3763
48		0.1699	0.1997	0.2431	0.2798	0.3215	0.3489	0.3735
49		0.1686	0.1982	0.2415	0.2782	0.3197	0.3469	0.3720
50		0.1674	0.1969	0.2399	0.2764	0.3178	0.3449	0.3695

Table 5. Critical values for Dixon-type discordance test N11 of an upper or lower outlier-pair in a normal sample (continued).

n	CL SL *α*	70% 30% 0.30	80% 20% 0.20	90% 10% 0.10	**95%** **5%** **0.05**	98% 2% 0.02	**99%** **1%** **0.01**	99.5% 0.5% 0.005
51		0.1664	0.1957	0.2384	0.2749	0.3162	0.3434	0.3678
52		0.1653	0.1945	0.2371	0.2732	0.3144	0.3415	0.3664
53		0.1643	0.1934	0.2358	0.2717	0.3127	0.3399	0.3645
54		0.1633	0.1923	0.2345	0.2702	0.3110	0.3379	0.3626
55		0.1623	0.1911	0.2332	0.2688	0.3094	0.3361	0.3605
56		0.1614	0.1900	0.2319	0.2676	0.3080	0.3350	0.3588
57		0.1604	0.1890	0.2306	0.2660	0.3065	0.3335	0.3574
58		0.1595	0.1880	0.2296	0.2649	0.3051	0.3317	0.3557
59		0.1587	0.1869	0.2282	0.2635	0.3037	0.3304	0.3544
60		0.1578	0.1860	0.2273	0.2623	0.3021	0.3286	0.3527
61		0.1570	0.1851	0.2262	0.2611	0.3010	0.3276	0.3515
62		0.1561	0.1841	0.2249	0.2599	0.2997	0.3263	0.3498
63		0.1554	0.1832	0.2240	0.2588	0.2983	0.3246	0.3481
64		0.1547	0.1824	0.2231	0.2579	0.2971	0.3235	0.3474
65		0.1540	0.1816	0.2221	0.2567	0.2960	0.3220	0.3457
66		0.1532	0.1808	0.2211	0.2554	0.2948	0.3209	0.3445
67		0.1525	0.1799	0.2202	0.2546	0.2937	0.3198	0.3434
68		0.1518	0.1791	0.2193	0.2535	0.2924	0.3186	0.3423
69		0.1511	0.1782	0.2182	0.2525	0.2915	0.3176	0.3411
70		0.1504	0.1775	0.2175	0.2516	0.2904	0.3162	0.3394
71		0.1498	0.1768	0.2166	0.2506	0.2896	0.3153	0.3388
72		0.1492	0.1761	0.2158	0.2497	0.2885	0.3141	0.3372
73		0.1486	0.1755	0.2150	0.2487	0.2874	0.3127	0.3363
74		0.1480	0.1748	0.2141	0.2478	0.2865	0.3121	0.3354
75		0.1474	0.1742	0.2134	0.2471	0.2853	0.3109	0.3341
76		0.1468	0.1735	0.2128	0.2462	0.2846	0.3097	0.3332
77		0.1463	0.1729	0.2119	0.2454	0.2837	0.3087	0.3320
78		0.1457	0.1722	0.2111	0.2444	0.2827	0.3081	0.3310
79		0.1452	0.1715	0.2104	0.2438	0.2818	0.3072	0.3302
80		0.1446	0.1709	0.2096	0.2428	0.2810	0.3062	0.3292
81		0.1440	0.1703	0.2088	0.2421	0.2800	0.3053	0.3283
82		0.1436	0.1697	0.2083	0.2414	0.2792	0.3044	0.3270
83		0.1431	0.1692	0.2077	0.2407	0.2784	0.3033	0.3258
84		0.1426	0.1686	0.2070	0.2398	0.2774	0.3024	0.3257
85		0.1421	0.1680	0.2063	0.2392	0.2768	0.3021	0.3249
86		0.1416	0.1674	0.2058	0.2384	0.2758	0.3009	0.3236
87		0.1411	0.1669	0.2050	0.2377	0.2751	0.3003	0.3228
88		0.1407	0.1664	0.2044	0.2370	0.2743	0.2994	0.3220
89		0.1403	0.1659	0.2039	0.2364	0.2740	0.2992	0.3216
90		0.1398	0.1654	0.2032	0.2357	0.2730	0.2979	0.3204
91		0.1393	0.1649	0.2026	0.2351	0.2725	0.2974	0.3198
92		0.1389	0.1644	0.2021	0.2345	0.2714	0.2966	0.3189
93		0.1385	0.1639	0.2014	0.2338	0.2708	0.2957	0.3181
94		0.1380	0.1634	0.2009	0.2332	0.2702	0.2949	0.3174
95		0.1377	0.1630	0.2004	0.2326	0.2695	0.2942	0.3166
96		0.1372	0.1625	0.1998	0.2320	0.2690	0.2936	0.3160
97		0.1369	0.1620	0.1993	0.2313	0.2682	0.2933	0.3156
98		0.1365	0.1616	0.1988	0.2308	0.2676	0.2925	0.3148
99		0.1360	0.1611	0.1982	0.2302	0.2669	0.2915	0.3138
100		0.1357	0.1607	0.1978	0.2296	0.2663	0.2910	0.3131

For the explanation of abbreviations see footnote of Table 2. The mean values of the standard error of the mean ($\overline{x}_{se}$) for these critical values ($\overline{x}$) are (respective % errors are also reported in parentheses): ~0.00008 (for $\alpha = 0.30$, 0.043%); ~0.00008 (for $\alpha = 0.20$, 0.033%); ~0.00008 (for $\alpha = 0.10$, 0.027%); ~0.00008 (for $\alpha = 0.05$, 0.024%); ~0.00008 (for $\alpha = 0.02$, 0.020%); ~0.00007 (for $\alpha = 0.01$, 0.016%); and ~0.00006 (for $\alpha = 0.005$, 0.014%).

Table 6. Critical values for Dixon-type discordance test N12 of an upper or lower outlier-pair in a normal sample.

n	CL SL *α*	70% 30% 0.30	80% 20% 0.20	90% 10% 0.10	**95%** **5%** **0.05**	98% 2% 0.02	**99%** **1%** **0.01**	99.5% 0.5% 0.005
5		0.8501	0.9019	0.9519	0.9760	0.9905	0.9952	0.9976
6		0.6802	0.7461	0.8248	0.8778	0.9236	0.9462	0.9620
7		0.5763	0.6408	0.7236	0.7842	0.8427	0.8756	0.9012
8		0.5068	0.5682	0.6493	0.7110	0.7735	0.8106	0.8411
9		0.4579	0.5162	0.5942	0.6545	0.7176	0.7561	0.7886
10		0.4209	0.4763	0.5515	0.6102	0.6722	0.7111	0.7444
11		0.3915	0.4451	0.5172	0.5745	0.6354	0.6736	0.7068
12		0.3684	0.4197	0.4893	0.5454	0.6050	0.6425	0.6754
13		0.3491	0.3988	0.4668	0.5213	0.5801	0.6174	0.6497
14		0.3330	0.3810	0.4472	0.5007	0.5583	0.5945	0.6268
15		0.3192	0.3659	0.4302	0.4823	0.5392	0.5754	0.6073
16		0.3072	0.3528	0.4158	0.4667	0.5221	0.5573	0.5887
17		0.2967	0.3413	0.4028	0.4530	0.5076	0.5425	0.5734
18		0.2874	0.3310	0.3917	0.4409	0.4948	0.5292	0.5594
19		0.2791	0.3218	0.3814	0.4300	0.4829	0.5168	0.5468
20		0.2716	0.3135	0.3721	0.4198	0.4721	0.5059	0.5360
21		0.2649	0.3061	0.3639	0.4109	0.4624	0.4955	0.5247
22		0.2586	0.2992	0.3561	0.4026	0.4538	0.4866	0.5160
23		0.2530	0.2930	0.3492	0.3952	0.4456	0.4784	0.5075
24		0.2479	0.2872	0.3426	0.3879	0.4377	0.4701	0.4982
25		0.2431	0.2818	0.3364	0.3814	0.4311	0.4633	0.4918
26		0.2386	0.2769	0.3310	0.3751	0.4244	0.4564	0.4845
27		0.2346	0.2723	0.3259	0.3700	0.4184	0.4498	0.4778
28		0.2306	0.2680	0.3209	0.3644	0.4127	0.4438	0.4718
29		0.2271	0.2641	0.3164	0.3595	0.4076	0.4381	0.4657
30		0.2237	0.2602	0.3122	0.3550	0.4024	0.4330	0.4605
31		0.2205	0.2567	0.3082	0.3508	0.3978	0.4281	0.4552
32		0.2175	0.2533	0.3043	0.3465	0.3933	0.4234	0.4508
33		0.2148	0.2503	0.3008	0.3426	0.3893	0.4195	0.4464
34		0.2122	0.2474	0.2974	0.3389	0.3851	0.4154	0.4423
35		0.2096	0.2444	0.2941	0.3356	0.3815	0.4113	0.4376
36		0.2072	0.2418	0.2911	0.3324	0.3778	0.4076	0.4339
37		0.2050	0.2392	0.2882	0.3290	0.3744	0.4039	0.4301
38		0.2027	0.2367	0.2853	0.3259	0.3710	0.4005	0.4263
39		0.2007	0.2344	0.2826	0.3227	0.3678	0.3967	0.4232
40		0.1986	0.2321	0.2800	0.3201	0.3648	0.3942	0.4198
41		0.1967	0.2300	0.2776	0.3174	0.3619	0.3911	0.4170
42		0.1949	0.2279	0.2753	0.3148	0.3590	0.3878	0.4135
43		0.1932	0.2259	0.2730	0.3125	0.3565	0.3851	0.4112
44		0.1916	0.2242	0.2710	0.3101	0.3539	0.3824	0.4080
45		0.1900	0.2223	0.2689	0.3078	0.3516	0.3801	0.4055
46		0.1884	0.2205	0.2669	0.3059	0.3491	0.3775	0.4030
47		0.1869	0.2188	0.2649	0.3035	0.3467	0.3750	0.4005
48		0.1855	0.2173	0.2630	0.3014	0.3445	0.3727	0.3973
49		0.1840	0.2156	0.2611	0.2993	0.3424	0.3704	0.3955
50		0.1826	0.2141	0.2594	0.2974	0.3402	0.3682	0.3929

Table 6. Critical values for Dixon-type discordance test N12 of an upper or lower outlier-pair in a normal sample (continued).

n	CL / SL / α	70% / 30% / 0.30	80% / 20% / 0.20	90% / 10% / 0.10	**95% / 5% / 0.05**	98% / 2% / 0.02	**99% / 1% / 0.01**	99.5% / 0.5% / 0.005
51		0.1815	0.2126	0.2576	0.2956	0.3383	0.3662	0.3914
52		0.1801	0.2111	0.2560	0.2938	0.3363	0.3640	0.3890
53		0.1789	0.2098	0.2546	0.2920	0.3342	0.3621	0.3870
54		0.1778	0.2086	0.2530	0.2905	0.3324	0.3601	0.3851
55		0.1765	0.2071	0.2514	0.2887	0.3305	0.3577	0.3826
56		0.1755	0.2059	0.2500	0.2872	0.3290	0.3563	0.3809
57		0.1744	0.2047	0.2485	0.2854	0.3270	0.3545	0.3795
58		0.1733	0.2035	0.2472	0.2840	0.3258	0.3530	0.3772
59		0.1722	0.2023	0.2457	0.2826	0.3241	0.3511	0.3762
60		0.1713	0.2013	0.2446	0.2812	0.3222	0.3493	0.3740
61		0.1703	0.2001	0.2433	0.2799	0.3210	0.3482	0.3724
62		0.1694	0.1991	0.2420	0.2784	0.3195	0.3466	0.3709
63		0.1685	0.1980	0.2408	0.2770	0.3180	0.3448	0.3694
64		0.1676	0.1971	0.2399	0.2762	0.3166	0.3438	0.3678
65		0.1668	0.1960	0.2386	0.2747	0.3151	0.3422	0.3661
66		0.1660	0.1952	0.2375	0.2734	0.3141	0.3407	0.3650
67		0.1651	0.1941	0.2365	0.2725	0.3129	0.3394	0.3636
68		0.1642	0.1931	0.2354	0.2711	0.3114	0.3382	0.3625
69		0.1634	0.1922	0.2342	0.2698	0.3104	0.3368	0.3610
70		0.1627	0.1914	0.2334	0.2690	0.3091	0.3352	0.3589
71		0.1619	0.1905	0.2324	0.2678	0.3079	0.3342	0.3582
72		0.1612	0.1898	0.2314	0.2667	0.3068	0.3329	0.3568
73		0.1605	0.1889	0.2305	0.2657	0.3056	0.3317	0.3553
74		0.1598	0.1881	0.2296	0.2646	0.3047	0.3310	0.3546
75		0.1592	0.1874	0.2287	0.2638	0.3033	0.3293	0.3530
76		0.1585	0.1867	0.2278	0.2628	0.3025	0.3283	0.3520
77		0.1579	0.1860	0.2271	0.2618	0.3013	0.3273	0.3510
78		0.1571	0.1852	0.2262	0.2608	0.3004	0.3264	0.3501
79		0.1566	0.1845	0.2253	0.2600	0.2994	0.3253	0.3487
80		0.1559	0.1837	0.2243	0.2591	0.2984	0.3241	0.3476
81		0.1553	0.1829	0.2235	0.2582	0.2975	0.3232	0.3467
82		0.1547	0.1824	0.2229	0.2573	0.2965	0.3223	0.3454
83		0.1541	0.1817	0.2221	0.2565	0.2955	0.3208	0.3440
84		0.1536	0.1811	0.2213	0.2556	0.2945	0.3202	0.3437
85		0.1530	0.1804	0.2206	0.2548	0.2939	0.3196	0.3429
86		0.1524	0.1797	0.2199	0.2540	0.2926	0.3183	0.3414
87		0.1519	0.1792	0.2192	0.2532	0.2920	0.3174	0.3407
88		0.1514	0.1786	0.2185	0.2524	0.2910	0.3166	0.3398
89		0.1509	0.1780	0.2178	0.2518	0.2905	0.3161	0.3393
90		0.1503	0.1774	0.2171	0.2509	0.2894	0.3147	0.3378
91		0.1498	0.1768	0.2164	0.2503	0.2888	0.3142	0.3374
92		0.1493	0.1762	0.2158	0.2497	0.2879	0.3134	0.3363
93		0.1489	0.1756	0.2151	0.2489	0.2870	0.3122	0.3354
94		0.1484	0.1751	0.2145	0.2481	0.2863	0.3116	0.3344
95		0.1480	0.1746	0.2138	0.2474	0.2855	0.3108	0.3335
96		0.1475	0.1741	0.2132	0.2468	0.2848	0.3101	0.3328
97		0.1470	0.1736	0.2127	0.2460	0.2841	0.3095	0.3325
98		0.1466	0.1731	0.2122	0.2454	0.2835	0.3088	0.3315
99		0.1460	0.1725	0.2114	0.2447	0.2826	0.3077	0.3304
100		0.1457	0.1721	0.2109	0.2441	0.2819	0.3070	0.3301

For the explanation of abbreviations see footnote of Table 2. The mean values of the standard error of the mean ($\overline{x}_{se}$) for these critical values ($\overline{x}$) are (respective % errors are also reported in parentheses): ~0.00008 (for α = 0.30, 0.038%); ~0.00008 (for α = 0.20, 0.032%); ~0.00008 (for α = 0.10, 0.026%); ~0.00008 (for α = 0.05, 0.024%); ~0.00008 (for α = 0.02, 0.020%); ~0.00008 (for α = 0.01, 0.017%); and ~0.00007 (for α = 0.005, 0.014%).

Table 7. Critical values for Dixon-type discordance test N13 of an upper or lower outlier-pair in a normal sample.

n	CL / SL / α	70% / 30% / 0.30	80% / 20% / 0.20	90% / 10% / 0.10	**95% / 5% / 0.05**	98% / 2% / 0.02	**99% / 1% / 0.01**	99.5% / 0.5% / 0.005
6		0.8668	0.9140	0.9582	0.9794	0.9918	0.9959	0.9980
7		0.7068	0.7697	0.8438	0.8919	0.9327	0.9526	0.9665
8		0.6042	0.6679	0.7477	0.8051	0.8591	0.8890	0.9122
9		0.5346	0.5962	0.6756	0.7351	0.7942	0.8289	0.8573
10		0.4842	0.5431	0.6207	0.6802	0.7405	0.7771	0.8084
11		0.4457	0.5023	0.5775	0.6359	0.6965	0.7336	0.7654
12		0.4153	0.4697	0.5425	0.5998	0.6599	0.6974	0.7296
13		0.3908	0.4436	0.5148	0.5707	0.6299	0.6671	0.6990
14		0.3704	0.4214	0.4907	0.5457	0.6039	0.6403	0.6722
15		0.3534	0.4029	0.4702	0.5240	0.5817	0.6177	0.6494
16		0.3386	0.3868	0.4528	0.5054	0.5617	0.5973	0.6286
17		0.3258	0.3728	0.4372	0.4891	0.5447	0.5798	0.6108
18		0.3145	0.3606	0.4238	0.4746	0.5299	0.5643	0.5951
19		0.3046	0.3495	0.4118	0.4621	0.5162	0.5503	0.5804
20		0.2956	0.3397	0.4010	0.4501	0.5036	0.5381	0.5682
21		0.2876	0.3309	0.3913	0.4398	0.4925	0.5263	0.5562
22		0.2802	0.3228	0.3821	0.4305	0.4825	0.5162	0.5455
23		0.2736	0.3156	0.3742	0.4218	0.4735	0.5065	0.5359
24		0.2675	0.3089	0.3666	0.4133	0.4643	0.4969	0.5260
25		0.2619	0.3026	0.3596	0.4058	0.4566	0.4893	0.5186
26		0.2568	0.2970	0.3533	0.3989	0.4492	0.4819	0.5102
27		0.2520	0.2916	0.3474	0.3927	0.4425	0.4744	0.5028
28		0.2475	0.2867	0.3415	0.3866	0.4362	0.4676	0.4956
29		0.2434	0.2821	0.3365	0.3812	0.4301	0.4617	0.4891
30		0.2394	0.2777	0.3317	0.3759	0.4245	0.4556	0.4835
31		0.2359	0.2737	0.3271	0.3710	0.4192	0.4502	0.4776
32		0.2324	0.2699	0.3226	0.3665	0.4142	0.4448	0.4725
33		0.2293	0.2664	0.3188	0.3621	0.4098	0.4405	0.4675
34		0.2262	0.2631	0.3150	0.3577	0.4052	0.4359	0.4629
35		0.2232	0.2597	0.3114	0.3540	0.4012	0.4318	0.4582
36		0.2205	0.2566	0.3079	0.3504	0.3971	0.4273	0.4539
37		0.2180	0.2538	0.3047	0.3467	0.3932	0.4232	0.4502
38		0.2154	0.2510	0.3013	0.3432	0.3895	0.4194	0.4460
39		0.2132	0.2483	0.2983	0.3398	0.3860	0.4157	0.4421
40		0.2108	0.2458	0.2955	0.3367	0.3826	0.4126	0.4385
41		0.2087	0.2433	0.2928	0.3339	0.3792	0.4089	0.4351
42		0.2066	0.2411	0.2901	0.3309	0.3760	0.4057	0.4316
43		0.2047	0.2388	0.2876	0.3282	0.3730	0.4027	0.4288
44		0.2029	0.2369	0.2854	0.3256	0.3706	0.3996	0.4253
45		0.2010	0.2348	0.2830	0.3232	0.3679	0.3970	0.4225
46		0.1993	0.2329	0.2808	0.3209	0.3651	0.3940	0.4201
47		0.1975	0.2308	0.2786	0.3183	0.3625	0.3914	0.4171
48		0.1960	0.2292	0.2765	0.3159	0.3601	0.3889	0.4141
49		0.1943	0.2272	0.2745	0.3138	0.3577	0.3864	0.4120
50		0.1928	0.2256	0.2725	0.3116	0.3553	0.3838	0.4091

Table 7. Critical values for Dixon-type discordance test N13 of an upper or lower outlier-pair in a normal sample (continued).

n	CL SL *α*	70% 30% 0.30	80% 20% 0.20	90% 10% 0.10	**95%** **5%** **0.05**	98% 2% 0.02	**99%** **1%** **0.01**	99.5% 0.5% 0.005
51		0.1915	0.2239	0.2705	0.3095	0.3533	0.3817	0.4072
52		0.1900	0.2223	0.2688	0.3076	0.3511	0.3793	0.4049
53		0.1886	0.2208	0.2671	0.3057	0.3490	0.3771	0.4022
54		0.1874	0.2194	0.2654	0.3039	0.3470	0.3752	0.4003
55		0.1861	0.2178	0.2636	0.3021	0.3448	0.3726	0.3979
56		0.1848	0.2165	0.2621	0.3004	0.3431	0.3712	0.3961
57		0.1836	0.2152	0.2605	0.2984	0.3409	0.3689	0.3943
58		0.1825	0.2138	0.2591	0.2969	0.3395	0.3672	0.3921
59		0.1813	0.2125	0.2576	0.2953	0.3378	0.3655	0.3911
60		0.1802	0.2113	0.2561	0.2937	0.3357	0.3633	0.3884
61		0.1791	0.2101	0.2548	0.2922	0.3345	0.3622	0.3867
62		0.1781	0.2088	0.2534	0.2907	0.3328	0.3605	0.3849
63		0.1771	0.2077	0.2520	0.2892	0.3310	0.3585	0.3833
64		0.1761	0.2067	0.2510	0.2881	0.3297	0.3570	0.3820
65		0.1752	0.2056	0.2497	0.2866	0.3280	0.3554	0.3800
66		0.1743	0.2046	0.2485	0.2852	0.3267	0.3539	0.3785
67		0.1733	0.2035	0.2473	0.2842	0.3255	0.3527	0.3770
68		0.1724	0.2024	0.2461	0.2828	0.3240	0.3514	0.3758
69		0.1715	0.2013	0.2448	0.2815	0.3226	0.3499	0.3742
70		0.1707	0.2004	0.2439	0.2804	0.3214	0.3480	0.3724
71		0.1698	0.1995	0.2427	0.2791	0.3201	0.3469	0.3714
72		0.1690	0.1987	0.2417	0.2779	0.3188	0.3455	0.3700
73		0.1683	0.1978	0.2407	0.2769	0.3176	0.3442	0.3684
74		0.1675	0.1969	0.2396	0.2756	0.3166	0.3435	0.3676
75		0.1668	0.1961	0.2388	0.2747	0.3150	0.3418	0.3662
76		0.1661	0.1954	0.2378	0.2737	0.3142	0.3407	0.3646
77		0.1654	0.1945	0.2369	0.2727	0.3128	0.3395	0.3637
78		0.1646	0.1936	0.2359	0.2715	0.3119	0.3384	0.3625
79		0.1639	0.1929	0.2350	0.2707	0.3110	0.3373	0.3611
80		0.1632	0.1920	0.2340	0.2697	0.3099	0.3361	0.3597
81		0.1625	0.1912	0.2331	0.2686	0.3086	0.3350	0.3589
82		0.1619	0.1906	0.2323	0.2677	0.3078	0.3341	0.3577
83		0.1613	0.1899	0.2316	0.2668	0.3067	0.3327	0.3561
84		0.1607	0.1892	0.2307	0.2659	0.3056	0.3318	0.3557
85		0.1601	0.1884	0.2299	0.2650	0.3049	0.3310	0.3548
86		0.1594	0.1876	0.2291	0.2640	0.3037	0.3298	0.3533
87		0.1588	0.1871	0.2283	0.2632	0.3031	0.3289	0.3523
88		0.1582	0.1864	0.2276	0.2624	0.3018	0.3280	0.3516
89		0.1577	0.1857	0.2269	0.2617	0.3013	0.3274	0.3511
90		0.1571	0.1851	0.2260	0.2609	0.3000	0.3260	0.3495
91		0.1565	0.1845	0.2253	0.2600	0.2994	0.3254	0.3488
92		0.1560	0.1838	0.2247	0.2594	0.2986	0.3243	0.3478
93		0.1555	0.1832	0.2239	0.2585	0.2975	0.3233	0.3470
94		0.1550	0.1826	0.2232	0.2577	0.2968	0.3226	0.3456
95		0.1545	0.1821	0.2225	0.2569	0.2959	0.3217	0.3448
96		0.1539	0.1815	0.2218	0.2563	0.2952	0.3210	0.3441
97		0.1535	0.1809	0.2213	0.2554	0.2943	0.3204	0.3438
98		0.1530	0.1804	0.2207	0.2548	0.2937	0.3195	0.3427
99		0.1524	0.1797	0.2198	0.2541	0.2927	0.3185	0.3414
100		0.1520	0.1793	0.2193	0.2535	0.2920	0.3178	0.3413

For the explanation of abbreviations see footnote of Table 2. The mean values of the standard error of the mean ($\overline{x}_{se}$) for these critical values ($\overline{x}$) are (respective % errors are also reported in parentheses): ~0.00008 (for α = 0.30, 0.037%); ~0.00008 (for α = 0.20, 0.032%); ~0.00008 (for α = 0.10, 0.025%); ~0.00008 (for α = 0.05, 0.024%); ~0.00009 (for α = 0.02, 0.021%); ~0.00008 (for α = 0.01, 0.018%); and ~0.00008 (for α = 0.005, 0.015%).

Dixon's original paper due to an upward shift of one row were, however, corrected by Rorabacher (1991).

In spite of this important observation (*i.e.*, our simulation results being much more precise and accurate than the earlier literature values by Dixon, 1951 and Rorabacher, 1991), we decided to compare our results with the literature critical values (Figure 4) for $\alpha = 0.05$ (5% SL) and 0.01 (1% SL) to find out the similarities and differences between them.

First, our present simulated critical values are characterized by very small standard errors for all values of n up to 100 (see footnotes of Tables 2-7; see also graphically the dashed and dotted near-horizontal line pairs in Figure 4a-f for n up to 100). These % errors are: 0.09% – 0.007% for test N7 (Table 2); 0.06% – 0.005% for N9 (Table 3); 0.06% – 0.006% for N10 (Table 4); 0.043% – 0.014% for N11 (Table 5); 0.038% – 0.014% for N12 (Table 6); and 0.037% – 0.015% for N13 (Table 7). The errors of the literature critical values, on the other hand, were not really estimated by Dixon (1951) nor were stated by Barnett and Lewis (1994), but the indications are (Dixon, 1951) that they are much larger than those obtained in our present simulations (see dashed-dotted and dashed curves in Figure 4a-f, for n up to 30). Equally large errors also apply for Rorabacher (1991) critical values. Finally, due to the inaccuracies of literature critical values they differ from our present simulated values; for example, for 5% and 1% SL (Figure 4a-f), these (absolute) % differences may be as large (or even larger than) ~0.4% for tests N7-N10, up to about 1.0% for test N11 (ignoring the shifted $\alpha = 0.05$ values; see large circles at the lower part of the diagram in Fig. 4d), and up to about 2.0% for tests N12 and N13.

It is noteworthy that a comparison of our simulated values for $n = 3$ and 4 with the "exact" solutions by Dixon (1951) shows extremely small differences (mean absolute % differences of ~0.08%, ~0.08%, and ~0.06% for tests N7, N9, and N11, respectively), assuring thus the high accuracy of our simulated results. A part of the differences might be due to the fact that, although our values were rounded to 4 decimal places, Dixon (1951) reported these "exact" critical values to only 3 decimal places (and we do not know whether they were rounded or truncated; in any case, the report to only 3 decimal places should have caused unknown deviations from the "exact" nature of Dixon's critical values).

At first sight, from the truly statistical point of view it may appear that the more precise and accurate critical values (to 4 decimal places) such as those obtained in the present work may not represent a major advantage against the earlier less precise and accurate literature values (to 3 decimal places). Suppose we have a statistic TN (*e.g.*, $TN7$) for an observation (*e.g.*, $x_{(n)}$), and we wish to weigh the evidence for disbelieving H_0, *i.e.*, for judging this observation to be a discordant outlier. We assess the weight of evidence by seeing how unusual the value of T_{sample} is in the distribution of TN given H_0, *i.e.*, by $P(TN > T_{sample} \vdots H_0)$ – this is the significance probability SP (T_{sample}), and it has a continuum of values over the range of values of T_{sample} – not just the familiar 0.05, 0.02, 0.01, etc. If SP (T_{sample}) is very small, say 1/500 (*i.e.*, 0.002), the weight of evidence for disbelieving H_0 is very strong. If SP (T_{sample}) is quite large, say 1/5 (*i.e.*, 0.20), the evidence for disbelieving H_0 is very weak because a 1 to 5 chance is nothing unusual. If SP (T_{sample}) is 0.06 or, for that matter, 0.04, there is some evidence for regarding the outlier as discordant, but it is not conclusive. If it were practicable for us to calculate the value of SP (T_{sample}) for our observed value (T_{sample}), we could assess directly the weight of evidence for rejecting H_0. But for most discordance tests, this determination of the value of SP (T_{sample}) is not practicable or convenient, and we make the judgment with a few "milestones": the value of (T_{sample}) that gives SP (T_{sample}) = 0.05, 0.01, or whatever. These are the tabulated critical values of TN (Tables 2-7) such as $TN7$ (Table 2). Reference to these gives some indication of the weight of evidence for judging the outlier (value tested such as $x_{(n)}$) to be discordant. These are basically statistical arguments.

In practice, however, one decides *a priori* what value of SP (T_{sample}) whether 0.05 (Dybczyński, 1980; Rorabacher, 1991) or 0.01 (Verma, 1997, 2005), is to be used for routine operation of discordance tests. Suppose we have a value of TN that is accurate and precise to 4 decimal places, and for which the value of SP (T_{sample}) for an extreme observation is 0.05025, and if we are using the criterion of 0.05 to detect discordant outliers, this observation *will be* detected as a discordant outlier and eliminated from the initial data set. However, suppose if we are using less precise and accurate critical values that are out by ~1% as mentioned above (see Figure 4), the SP (T_{sample}) for the same observation happens to be 0.04975 (*i.e.*, off by ~1% with respect to 0.05025). This observation will *not* be detected as a discordant outlier (*i.e.*, retained in the data set) because of our initial assumption that the value of SP (T_{sample}) should be greater than 0.05 set for this application, although statistically speaking both SP (T_{sample}) (0.05025 and 0.04975) are very similar. In other cases, the opposite action might result from the application of different sets of critical values. We, therefore, conclude that more precise and accurate critical values are preferable, if available, for all routine application of discordance tests.

APPLICATIONS IN SCIENCE AND ENGINEERING

In this section we present a number of examples of published data sets in science and engineering where application of these extended Dixon tests (for n up to 100) could be useful. In only some cases, the original authors reported individual experimental data. For these cases, the tests can be actually applied to provide examples of applications in science and engineering as explained below. For other cases, we can only point out how these tests will be

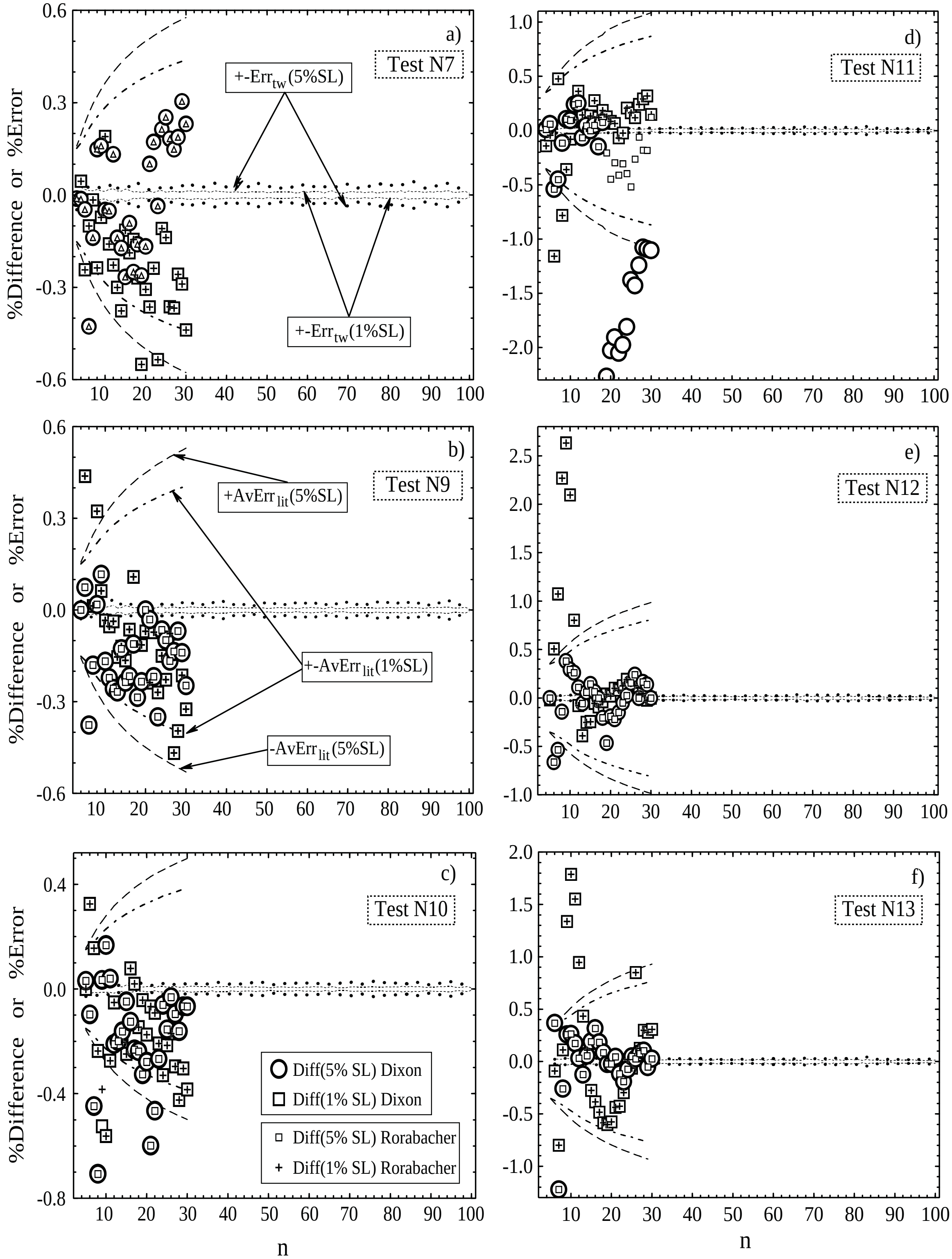

Figure 4. Comparison of the literature critical values with the new simulation results for *n* up to 30 and $\alpha = 0.05$ and 0.01 (most frequently used α values); the y-axis either represents (%Difference), *i.e.*, $[100\cdot(CV_{lit}-CV_{tw})/CV_{tw}]$, being the % difference of the literature values (CV_{lit}: critical value from Dixon, 1951; Barnett and Lewis, 1994; or from Rorabacher, 1991) from the present critical values (CV_{tw}: critical value obtained in this work), or %Error (in literature values as reported by Dixon, 1951, or Rorabacher, 1991, for *n* up to 30, or in the present values). The symbols used are shown in Figures 4a-c. +-Err_{tw}(5%SL) and +-Err_{tw}(1%SL): ±%Standard error of the mean (this work) for 5% and 1% significance level, respectively; +-$AvErr_{lit}$(5%SL) and +-$AvErr_{lit}$(1%SL): ±%Average error of critical value data reported (assumed) by Dixon (1951) for 5% and 1% significance level, respectively; and Diff(5% SL) and Diff(1% SL): % difference between the literature critical value data reported by Dixon (1951) or Rorabacher, 1991, with respect to the present simulated data for 5% and 1% significance level, respectively. (a) Single outlier test N7; (b) Single outlier test N9; (c) Single outlier test N10; (d) Upper- or lower-pair outlier test N11; (e) Upper- or lower-pair outlier test N12; and (f) Upper- or lower-pair outlier test N13.

useful. Nevertheless, we have included numerous examples from different fields of earth sciences to highlight the use of Dixon tests to actual data sets.

A spreadsheet (in *Statistica* commercial software) for applying these tests is available from the first author (SPV) of this paper. We may also point out that an updated version of the existing SIPVADE software (Verma *et al.*, 1998), which will include critical values from these new tables (Tables 2-7) as well as others (under preparation), will also be made available to the scientific and industrial community for applying all discordance tests, including the six Dixon tests presented in this paper.

Rorabacher (1991), who reported new interpolated critical values for Dixon tests, argued in favor of a "two-tailed" (instead of the conventional "one-tailed") test at the 95% confidence level for such applications. We instead follow the recommendations by Verma (1997, 1998, 2005) to apply these tests "one-tailed" as is customary in such applications (see *e.g.*, Dybczyński *et al.*, 1979; Dybczyński, 1980; Barnett and Lewis, 1994; Verma *et al.*, 1998; Velasco *et al.*, 2000; Guevara *et al.*, 2001) at the strict 99% confidence level (significance level α of 0.99) to limit the associated significance probability SP (T_{sample}) to 0.01 and thus have a more conclusive evidence that the outlier is discordant. Rorabacher (1991) recommendations of a "two-tailed" test would correspond to a less strict value of SP (T_{sample}) of 0.025 (the half of 0.05), and for this statistical reason, Verma (1997, 1998, 2005) proposal of using the 99% confidence level (*i.e.*, SP (T_{sample}) of 0.01) would be preferable for such applications.

Dixon tests are generally used for detecting a small number of outliers, because the power of these tests decreases as the number of outliers increases in a given data set (*e.g.*, Gibbons, 1994). Further, different kinds of masking effects make the detection of discordant outliers by a given test statistic difficult (Barnett and Lewis, 1994; Velasco *et al.*, 2000; Buckley and Georgianna, 2001). Therefore, although in this paper we illustrate the application of Dixon tests only, the joint concurrent use of several discordance tests, such as Dixon and Grubbs tests as well as skewness and kurtosis tests and appropriate variants of multiple-outlier tests, is highly recommended and, in fact, considered essential (Dybczyński, 1980; Verma, 1997, 1998, 2005; Verma *et al.*, 1998; Velasco *et al.*, 2000; Guevara *et al.*, 2001).

Further work is in progress to simulate new, precise, and accurate critical values for numerous other tests summarized by Barnett and Lewis (1994) and used by Verma and collaborators in the study of geochemical reference materials. New results of this ongoing investigation we plan to publish soon in an international journal.

Agricultural and Soil Sciences

Stevens *et al.* (1995) and Lugo-Ospina *et al.* (2005) studied nutrients in animal manures – valuable inputs for agronomic crop production. Dixon tests can be applied to find outliers in studentized residuals (see pp. 320-323 in Barnett and Lewis, 1994, for details on such residuals) of several linear relationships used by these authors to interpret their data. In a different study, Batjes (2005) presented organic carbon data in major soil groups of Brazil, to which Dixon tests can be used to detect outliers for 13 major soil groups with the number of representative profiles from 6 to 53 (see table 2 of Batjes paper). Similarly, Luedeling *et al.* (2005) studied drainage, salt leaching, and physico-chemical properties of irrigated man-made terrace soils in a mountain oasis of northern Oman and used one of the Dixon tests for the evaluation of outliers. All six Dixon tests can now be better applied in such studies.

Aquatic Environmental Research

Thomulka and Lange (1996) studied the impact of various chemicals to aquatic environments and used a Dixon test for the evaluation of outliers. With the availability of new critical values, we suggest that all six Dixon tests be applied to such data. Similarly, Buckley and Georgianna (2001) used Dixon (1950, 1951) tests for handling statistical outliers in whole effluent toxicity data, for which new critical values of Dixon tests will be of great use.

Astronomy

As an example of outlier-based applications, statistical analysis of the metallicities of superclusters and moving groups by Taylor (2000) is worthy of comment. Taylor presented, in Appendix B of this paper, the use of one Dixon (1951) test for outlier detection. The new critical values derived here for all six Dixon tests would certainly facilitate the use of outlier-methods in such studies. Furthermore, for evaluating the data (in table 8 of Taylor, 2000) for the number of stars varying from 21 to 71, the author had to use a non-parametric x^2 test instead of the more powerful parametric tests F and Student's t (or ANOVA). With the availability of these new critical values for n up to 100, discordance tests N7 and N9-N13 can first be applied to remove any outliers and then the parametric tests (F and Student's t) can be safely applied (see Verma, 2005 for more details).

Biology

Linkosalo *et al.* (1996) and Schaber and Badeck (2002) studied tree physiology and used King (1953) test (listed as test N8 by Barnett and Lewis, 1994) for outlier detection. To the data in both papers, all Dixon tests can now be better applied than was possible earlier. For example, Schaber and Badeck (2002) presented phenological data for 9 stations (see their table 2), in which the number of observations varied from 7 to 44. Four of these stations have $n > 30$, for which Dixon tests, with these new critical values for n up to 100, will be now readily applicable.

Biomedicine and Biotechnology

Freeman *et al.* (1997) investigated the effects of recombinant granulocyte colony-stimulating factor during canine bacteria pneumonia. Similarly, Sevransky *et al.* (2005) investigated nitric oxide as a possible cause of the cardiac dysfunction associated with high, lethal doses of tumor necrosis factor-α in dogs. The number of animals in each treatment group varied from 4 to 18 in the first study and 3 to 12 in the second one. In both studies, although the authors used one of the Dixon (1950) tests, the new, precise, and accurate critical values for all six Dixon tests will render these tests to be better applicable to these data than the single Dixon test with less precise critical values. Further, even for experiments with a greater number of subjects (up to 100 in each group), Dixon tests can be now applied. Woitge *et al.* (1998), on the other hand, evaluated biochemical markers of bone turnover to provide information for the diagnosis and monitoring of metabolic bone disease and applied a Dixon test. All Dixon tests can now be better applied to such data.

Chemistry

Zaric and Niketic (1997) compiled data on Co–NO_2 bond lengths in the crystal structures of ammine-nitro complexes of cobalt(III) and applied only one outlier test (N8 by King, 1953) to these data. Neither this test nor any of the Dixon tests could have been applied to their total compiled data because the total number of data ($n = 54$) was much greater than 30 (being the highest n for which critical values were available). With the availability of the new critical values, all Dixon tests can now be applied to these data. The initial statistical information for Co–NO_2 bond lengths was: mean ± standard deviation, 1.946 ± 0.039 Å ($n = 54$). After the application of Dixon tests (applied at the 95% confidence level for illustration purposes), the final statistics for the bond lengths can be summarized as: 1.938 ± 0.024 Å ($n = 50$).

Geochronology

Both Wang *et al.* (1998) and Dougherty-Page and Bartlett (1999) used only one test (Dixon, 1950) for outlier detection in their geochronology data. All Dixon tests with more precise and accurate critical values will be better applicable for such studies, especially when n is greater than 30. For example, Wang *et al.* (1998) presented Pb-Pb evaporation data (their table 3) with number of blocks up to 94, for which all Dixon tests can be applied to identify outliers. Dougherty-Page and Bartlett (1999), on the other hand, programmed a Dixon test during the data acquisition stage, for which the combination of all tests will now prove a better and more effective choice.

From the field of geochronology, we present an example from Bartlett *et al.* (1998) to highlight the use of Dixon tests. Single crystal zircon Pb isotopic compositions and the inferred ages were presented in their table 1 (to limit the length of this paper, the data under evaluation are not reproduced here because the reader can easily consult them in the original paper). The authors grouped the data for seven zircon grains (#1, 3, 4, 6, 7, 8, and 9 in their table 1) from south India to discuss the consistency of the relevant ages (2436±11 Ma for grain #1 and 2438±12 Ma for all 7 grains; the quoted error is one standard deviation throughout this subsection). Dixon tests can now be first applied to the isotopic data of individual grains (*e.g.*, $^{207}Pb/^{204}Pb$ data) because these experimental data should be normally distributed according to the Gauss theorem. We applied these tests to grain #1 $^{207}Pb/^{204}Pb$ data ($n = 20$) at the strict confidence level of 99% and found two lower and two upper outliers; the resulting ages can, therefore, be stated as 2435±4 Ma ($n = 16$). Similarly, using all ($n = 33$) data for the seven zircon grains and Dixon tests (at the strict confidence level of 99%), the respective age was estimated to be 2427±18 Ma ($n = 31$). The significantly larger dispersion of this age (±18 Ma) as compared to a single zircon grain (±4 Ma) probably reflects some age heterogeneity of the zircon grains – a characteristic not inferred by the original authors of this paper.

Meteorology

Graybeal *et al.* (2004) applied two Dixon tests (Dixon, 1950, 1951) to the seasonal and station-based analyses of hourly meteorological (temperature) data. All Dixon tests will be better applicable to this work because of more precise and accurate critical values estimated in the present study. Similarly, studentized residuals for the regressions presented by these authors can also be evaluated by these tests to detect discordant outliers.

Zoology

Harcourt *et al.* (2005) studied distribution-abundance (density) relationship of tropical mammals at the level of species, genera, and families/subfamilies. These authors eliminated outliers but their statistical method of outlier detection was not clear. Here, all Dixon tests can be easily applied to objectively detect outliers at the genera and family levels. Furthermore, outliers may be useful for additional geographical analysis of the data.

Quality assurance and assessment programs

RMs are widely used for the purpose of traceability, precision, accuracy, and sensitivity of routine analysis as well as in calibrations of analytical methods (*e.g.*, Verma, 1997). Weighted least-squares linear regression models, instead of the conventional ordinary least-squares linear regression, are now becoming a requirement for such instrumental calibrations (*e.g.*, Santoyo and Verma, 2003; Guevara *et al.*, 2005). Reliable concentration (central tendency or location parameter) as well as standard deviation, standard error of the mean, or confidence limit (dispersion or scale parameter) data for each chemical constituent in the RMs are, therefore, required. However, RMs, being highly complex natural materials, are not easily prone to this type of characterization, and proper statistical methods must be applied (*e.g.*, Barnett and Lewis, 1994; Verma, 1997, 1998, 2005; Verma *et al.*, 1998, and references therein).

Biology and Biomedicine

Ihnat (2000) evaluated the performance of neutron activation and other methods in an international reference material characterization campaign, in which the author summarized such data for large *n* (see table 2 of this paper). Dixon tests can now be applied to 26 sets of data, 16 of them are with $n > 30$; for the latter cases, Dixon tests were not earlier applicable. Similarly, Patriarca *et al.* (2005), in an inter-laboratory study related to their toxic metals project, used Grubbs test, among others, for the identification of outliers. Because a large number of participants (74) were involved in this study, Dixon tests can now be successfully applied to their trace element data in serum, blood, and urine samples, particularly because the new critical values are more precise and accurate than the earlier literature values.

Cement industry

Sieber *et al.* (2002) evaluated new cement and concrete reference materials. As an example, for Fe_2O_3 data in one reference material (SRM 1880a; see their table 5 for the inter-laboratory data; to limit the length of our paper we have not reproduced here the raw data), they reported 37 individual values, for which all Dixon tests can now be applied because of the availability of critical values for *n* up to 100. When Dixon tests are applied to these data at the strict confidence level of 99%, two outliers are detected in XRF data from "Construction Technology Laboratories", with the resulting statistics of mean ± standard deviation values being 2.799 ± 0.018 ($n = 35$). In their inter-laboratory data, Sieber *et al.* (2002) did not detected these outliers. The usefulness of Dixon tests is, therefore, clear from this case study, in which the application of these tests under the assumption that the data are normally distributed showed that there were two discordant outliers and, consequently, the resulting location and scale parameters will be more reliable after the application of these statistical tests.

Food science and technology

Morabito *et al.* (2004) and Villeneuve *et al.* (2004) evaluated data on organochlorinated compounds and petroleum hydrocarbons in a fish RM and methylmercury and arsenobetaine in an oyster tissue RM, respectively. Dixon tests will be readily applicable in such studies. Similarly, these tests will be useful for the food microbiology data presented by In't Veld (1998) and Langton *et al.* (2002).

Environmental and pollution research

Dybczyński *et al.* (1998), Gill *et al.* (2004), and Holcombe *et al.* (2004) analyzed inter-laboratory data for tobacco leaves, human hair, and sewage sludge RMs, respectively; it is obvious that Dixon tests with new critical values can be applied for the evaluation of these inter-laboratory data.

Nuclear science

Lin *et al.* (2001) evaluated radio-nuclide inter-laboratory data for the certification of RMs; in this work, the number of data varied from 18 to 84 (see table 2 of this paper). All Dixon tests with new critical values can now be applied to the data summarized by these authors.

Rock chemistry

Advantages of the availability of new critical values for Dixon tests are readily seen for several rock RMs summarized in Table 8. As an example of andesite AGV-1 from the U.S. Geological Survey, tests N7 and N9-N13, earlier applied to 0 major and 15 trace elements (Velasco-Tapia *et al.*, 2001), can now be applied to 4 major and 39 trace elements (Table 8). In a similar way, these tests can be applied to the recently available single-laboratory raw data for Mexican RMs (Lozano and Bernal, 2005); in fact, this practice is highly recommended before estimating the location and scale parameters (see Verma, 2005 for details).

Soil science

An example of a soil RM from Peru is also listed in Table 8, for which Dixon tests, earlier applied to 5 major and 28 trace elements (Verma *et al.*, 1998), can now be used for testing the data of 7 major and 34 trace elements.

Water research

M.P. Verma (2004) compiled results of several inter-laboratory studies related to the International Association of Geochemistry and Cosmochemistry (IAGC) and International Atomic Energy Agency (IAEA), in which the number of laboratories varied from 15 to 38. Although a statistically incorrect 2s method (two standard deviation method; for more details on this method see Gladney and Roelandts, 1988a; Gladney *et al.*, 1991; Imai *et al.*, 1996; note that this method has been shown to be statistically incorrect by Verma, 1997, 1998) was used for outlier detection and elimination, all six Dixon tests can be readily and correctly applied to such data. Similarly, Holcombe *et* *al.* (2004) evaluated chemical data on river water, drinking water, and estuary water RMs, for which Dixon tests can now be recommended.

As a further example, we present the results of application of Dixon tests to one set of inter-laboratory HCO_3^- data in water samples compiled by M.P. Verma (2004; see sample IAEA1, n = 21, in Table 2 of the original paper), for which a mean value of 295.3 µg/ml, with a standard deviation of 18.7 µg/ml was reported. Application of all Dixon tests to these data (at the 95% confidence level, which will be the statistically correct confidence level for the erroneous "2s method") detected two upper- and two lower-outliers, obtaining the final statistics of 296 ± 9 µg/ml (n = 17).

Other Applications in Geosciences

We have already presented application of Dixon tests in geochronology and quality assurance and assessment programs in different areas of geosciences. Here, we include more areas of earth sciences to further illustrate the application of the Dixon tests using the new critical values (Tables 2-7).

Petroleum hydrocarbons and organic compounds in sediment samples

Villeneuve *et al.* (2002) presented such data on a sediment sample and used the Box-and-Whisker plot to

Table 8. Applicability of all six Dixon discordance test for some RM databases.

Reference Material (RM)		Earlier application possible for # of		Present application possible for # of		Literature Reference(s)
Code	Description	Major elements	Trace elements	Major elements	Trace elements	
AGV-1	Andesite from U.S.G.S., U.S.A.	0	15	4	39	Gladney *et al.* (1992); Velasco-Tapia *et al.* (2001)
BIR-1	Basalt from U.S.G.S., U.S.A.	6	28	13	32	Gladney and Roelandts (1988a); Verma (1998)
BHVO-1	Basalt from U.S.G.S., U.S.A.	5	25	13	46	Gladney and Roelandts (1988b); Velasco-Tapia *et al.* (2001)
JG-1	Granodiorite from G.S.J., Japan	2	25	15	46	Imai *et al.* (1996); Guevara *et al.* (2001)
JG-2	Granite from G.S.J., Japan	6	28	15	39	Imai *et al.* (1996); Guevara *et al.* (2001)
JG-3	Granodiorite from G.S.J., Japan	6	36	15	40	Imai *et al.* (1996); Guevara *et al.* (2001)
JG-1a	Granodiorite from G.S.J., Japan	5	23	15	43	Imai *et al.* (1996); Guevara *et al.* (2001)
JGb-1	Gabbro from G.S.J., Japan	5	23	15	42	Imai *et al.* (1996); Guevara *et al.* (2001)
JP-1	Peridotite from G.S.J., Japan	8	28	15	32	Imai *et al.* (1996); Verma (1998)
PM-S	Microgabbro (Scotland) from G.I.T.-I.W.G.	5	19	8	47	Govindaraju *et al.* (1994, 1995); Verma (1997)
RGM-1	Rhyolite from U.S.G.S., U.S.A.	11	41	13	41	Gladney and Roelandts (1988b); Velasco-Tapia *et al.* (2001)
Soil-5	Soil (Preu) from I.A.E.A., Vienna	5	28	7	34	Dybczyński *et al.* (1979); Verma *et al.* (1998)
W-1	Diabase from U.S.G.S., U.S.A.	0	21	3	47	Gladney *et al.* (1991); Velasco-Tapia *et al.* (2001)
W-2	Diabase from U.S.G.S., U.S.A.	5	32	14	38	Gladney and Roelandts (1988a); Velasco-Tapia *et al.* (2001)
WS-E	Whin Sill dolerite (England) from G.I.T.-I.W.G.	5	15	8	44	Govindaraju *et al.* (1994, 1995); Verma *et al.* (1998)

U.S.G.S.: United States Geological Survey; G.S.J.: Geological Survey of Japan; I.A.E.A.: International Atomic Energy Agency; G.I.T.-I.W.G.: Groupe International de Travail "Etalons analytiques des mineraux, minerais et roches or International Working Group "Analytical Standards of Minerals, Ores, and Rocks".

detect outliers. Our Tables 2-7 enable us to apply all six Dixon tests to their data. Because of the limited availability of the report by Villeneuve *et al.* (2002), selected data for the illustration of these tests are summarized in Table 9. With the availability of new critical values, the Dixon tests could now be applied to the six hydrocarbon compounds in IAEA-417 compiled here. The tests (at the strict 99% confidence level) detected outlier values for 4 of the 6 hydrocarbon compounds, and the final statistics for these 4 cases showed a considerable improvement (Table 9) as compared to the original statistics on raw data. These results should be compared with the Box-and-Whisker plot method only *after* the application of all other discordance tests (Barnett and Lewis, 1994; Verma *et al.*, 1998; Verma, 2005) to the data under evaluation, which will be done in future after extending the critical value tables for the remaining dozens of test variants.

Paleontology

Our first example is the I.A. (índice de anchura – width index) data on *Cuvieronius* – one of the most common genera of the Gomphotgheriidae family recorded in Mexico during Pliocene and Pleistocene – compiled by Alberdi and Corona-M. (2005; see table 4 of this paper). Application of Dixon tests to these paleontology data did not show the presence of any outlier in these data at least with respect to the six Dixon tests and, therefore, these data can be interpreted using standard statistical techniques (outlier-based methods), although as suggested in the previous subsection, before doing so we must apply the other discordance tests to these data (work in progress).

The second example is for two different associations of ammonoids from Lower Jurassic sediments from Mexico to test the above mentioned hypotheses H_0 and H_1 for the diameter data of these two sets of ammonoids (see table 2 of Esquivel-Macías *et al.*, 2005 paper). The initial statistical data were: for association 1, mean ± standard deviation 36 ± 38 (n = 35); for association 2, mean ± standard deviation 17 ± 20 (n = 59). The initial data, thus, showed a rather large variability for both associations. The Dixon tests demonstrated that each association had 2 outliers (two largest values) at the 99% confidence level, rendering the final statistics as: for association 1, mean ± standard deviation 28 ± 22 (n = 33); for association 2, mean ± standard deviation, 14 ± 9 (n = 59). The application of Dixon tests, thus, provides additional information for the interpretation of these data (see Verma, 2005 for more details).

The third example is for Maastrichtian shallow-water ammonites of northeastern Mexico to test if the WB/WH (whorl breadth to height ratio) of 16 samples (not considering "uncertain" values within brackets) reported by Ifrim *et al.* (2005) throughout their paper (*i.e.*, not in a single table). The Dixon tests showed that in terms of the WB/WH variable there were no outliers in these samples, assuming that they were drawn from a normal population.

Finally, we present the fourth example from paleontology for Upper Jurassic ammonites from Sonora, Mexico (W/H data for 16 samples reported by Villaseñor *et al.*, 2005 throughout their paper). Once again, for these ammonites the Dixon tests also showed no outliers, on the assumption that the data were drawn from a normal population.

Structural Geology

As an example of geology, we applied the Dixon tests to the inclination data of fault planes reported by Dávalos-Álvarez *et al.* (2005) in their Appendix C. Six inclination data sets showed a normal distribution; only for one set (FYB), one of the six Dixon tests detected outlier values.

Isotope Geology

As a further example of geology, we applied the Dixon tests to Sr isotope data on Tertiary volcanic sequences from Taxco-Quetzalapa region of southern Mexico (Morán-Zenteno *et al.*, 1998; see nine $(^{87}Sr\text{-}^{86}Sr)_i$ data in their table 4). According to the six Dixon tests, these data showed no outliers.

Geochemistry

As the final example for geosciences, we applied the Dixon tests to SiO_2 concentration data (100% adjusted data on an anhydrous and volatile-free basis using the SINCLAS computer program of Verma *et al.*, 2002) of mantle-xenolith-bearing basic and ultrabasic rocks from the Eastern Alkaline Province of Mexico, recently presented by Treviño-Cázares *et al.* (2005). This application of the Dixon tests demonstrated that the SiO_2 concentration data assumed to come from a normal distribution in these 19 samples, showed no discordant outliers.

Linear regressions

This is an important area of research in almost all science and engineering fields such as for instrumental calibrations (*e.g.*, Santoyo and Verma, 2003; Guevara *et al.*, 2005) and for exploring relationships between two or more variables, *e.g.*, the "inverse modeling" of trace element data (Verma, in press). Outliers in linear models can be detected and eliminated, using studentized or weighted residuals with respect to the regression equations (see pp. 315-325 in Barnett and Lewis, 1994, or pp. 40-41, 67, and 718-719 in Shoemaker *et al.*, 1996), and the above mentioned applications of linear regressions can thus be much improved. Although Shoemaker *et al.* (1996) commented on the application of only one Dixon test (N7), all six tests (N7 and N9-N13) will be of much use in detecting discordant outliers in such linear models. New critical values extended to sizes of up to 100 data augment the usefulness of this approach in many more scientific and engineering problems than the ones mentioned as examples in this subsection. The new critical values for n up to 100 have paved the way for a wider application of the Dixon tests.

Table 9. Results of petroleum hydrocarbons in IAEA-417 sediment sample (Villeneuve *et al.*, 2002) and application of Dixon tests.

Lab Code #	Phenanthrene	Chrysene	Fluoranthene	Pyrene	Benz[a] Anthracene	Benz[a] Pyrene
3	1538	---	5464	4635	---	1583
5	3580	3380	5030	5750	2970	2880
6	2830	180	5360	3240	60	6.3
10	3900	4600	8600	6200	3600	3300
12	4690	20950	13510	10570	---	---
13	3857	3501	7366	5640	3013	2743
16	4315.07	2453.1	7885.97	5230.23	3966.37	1496
18	4961	5503	9143	7110	4261	3979
20	4300	2900	11680	9800	2500	2200
22	3710	4220	11200	8130	3320	3480
25	2383	1964	4675	4615	2003	1765
26	4297	4100	9113	6441	4371	3588
27	---	3350	7470	4440	---	3200
30	4389	2474	8583	5769	1738	199
31	4477	370	883	462	2000	3400
32	3960	2230	4990	4970	1930	2620
33	4426	4405	8339	4933	---	2674
34	1559.92	2355.21	7131.98	10027.8	9204.76	1288.83
39	1090	---	3050	2980	2730	1920
42	5040	6220	14540	8830	4510	910
43	2050	2465	3940	3195	2345	1510
44	5559	3528	9372	6389	4746	4848
45	3440	2580	5410	4070	2230	2020
46	6500	5410	12000	8500	4600	4860
50	5170	5473.3	7373	6016.7	4756.7	---
51	3752	2101	4532	6252	3577	1988
52	3727	5069.5	8465	7638	3417	3474
53	5470	5310	12290	7880	3610	4160
57	1196	1170	2477	1475	1018	---
58	---	4900	9000	6800	3600	3900
60	2731	4217	5801	4380	3266	3302
61	4786.666	5116.667	9196.667	7460	3970	3176.667
62	---	---	5710	---	---	2610
65	4967	5000	9700	7200	3767	3400
66	3750	---	8790	7630	1820	2610
67	6310	140	6940	15000	3220	1900
68	15120	15800	30530	20100	---	---
70	2959	1659	5891	17884	2033	2018
72	16400	22500	36250	28950	15000	
77	---	6870	10020	6760	5370	5160
78	7572	4234	12029	18823	3236	3699
79	852	2305	5664	4527	---	1980
80	3843	3753	8415	6293	3102	2871
84	3930	4290	8520	6130	4010	3850
86	4307.775	4532.45	9061.45	6691.8	4368.85	4273.25
92	1935.3	1633.3	3151.7	2649.3	1606	721.1
94	4300	4700	9000	9700	3300	3800
95	5803	6270	9820	8167	4963	4427
96	2833	3457	4887	4109	2355	3110
Statistical information before and after the application of Dixon tests						
$n_{initial}$	45	45	49	48	42	44
$\overline{x}_{initial}$	4400	4700	8700	7500	3600	2800
$s_{initial}$	2900	4500	5900	5100	2300	1200
$n_{outlier}$	2	0	2	2	2	0
n_{final}	43	45	47	46	40	44
$\overline{x}_{final}$	3900	4700	7700	6800	3200	2800
s_{final}	1500	4500	3000	3600	1200	1200

Note the data were reproduced as such from the original report, without questioning their (statistically speaking) correct or incorrect presentation. n:number of data; $\overline{x}$: mean; s: standard deviation; the subscripts initial and final refer to the initial (before applying Dixon tests) and final (after applying Dixon tests) data; $n_{outlier}$: number of discordant outlier detected by the Dixon tests. --- in the data column means that no data were provided by a given lab.

Other applications

Because of the applicability of the Dixon tests to a larger number of chemical elements (up to $n = 100$), it will be possible in future to use the method of Velasco *et al.* (2000) to empirically assess the relative efficiency of these tests by comparing their performance with that of other discordance tests. We will also be able to use the simulation procedure to assess their relative efficiency and, thus, compare the two assessments (empirical and numerical) to arrive at more definite conclusions concerning these tests.

In fact, these discordance tests (N7 and N9-N13) should be applicable to experimental data in many other scientific and engineering fields (besides the ones mentioned above), such as ecology (Yurewicz 2004), geodesy (Kern *et al.*, 2005), medical science and technology (Tigges *et al.*, 1999; Hofer and Murphy 2000; Reed *et al.*, 2002; Stancak *et al.*, 2002), and water resources (Buckley and Georgianna, 2001).

Finally, users of a number of internet sites (*e.g.*, San Francisco State University http://squall.sfsu.edu/courses/geo475/stats.htm; Statistics for chemists – nonparametric hypothesis tests http://www.webchem.science.ru.nl/cgi-bin/Stat/HypT/nphypt.pl; database http://www.wormbase.org; and Environmental sampling and monitoring primer http://ewr.cee.vt.edu/environmental/teach/smprimer/outlier/outlier.html) will also benefit from the incorporation of these new tables of critical values into these systems.

CONCLUSIONS

In synthesis, the new, precise, and accurate critical values computed for all six Dixon discordance tests offer a great advantage for diverse applications in univariate data sets, because (i) the higher precision (four significant digits instead of only three in the earlier literature values) should reduce the errors in the application of these tests because they are applied at certain significance levels of 0.05 or 0.01 in most science and engineering applications; (ii) Dixon tests (N7 and N9-N13) now have precise critical values similar to the Grubbs tests; and (iii) the increment of n up to 100 extends the application to data sets of larger sizes than was possible earlier (n was only up to 30). Finally, we must emphasize that these new critical values will open more extensive applications of these six Dixon discordance tests for normal univariate data in a variety of scientific and engineering fields, including earth sciences.

ACKNOWLEDGEMENTS

This research was partly supported by the "Sistema Nacional de Investigadores" (Mexico), through a scholarship to A. Quiroz-Ruiz as the first author's (SPV's) "Ayudante de Investigador Nacional Nivel 3". The first author (SPV) is also indebted to the Editor-in-Chief Susana Alaniz-Álvarez for her kind invitation to contribute one of his papers to the journal "Revista Mexicana de Ciencias Geológicas". We are also grateful to three reviewers –who opted to remain anonymous– for providing us valuable suggestions for improvement of our earlier manuscript.

REFERENCES

Alberdi, M.T., Corona-M., E., 2005, Revisión de los gonofoterios en el Cenozoico tardío de México: Revista Mexicana de Ciencias Geológicas, 22(2), 246–260.

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

Bartlett, J.M., Dougherty-Page, J.S., Harris, N.B.W., Hawkesworth, C.J., Santosh, M., 1998, The application of single zircon evaporation and model Nd ages to the interpretation of polymetamorphic terrains: an example from the Proterozoic mobile belt of south India: Contributions to Mineralogy and Petrology, 131(2-3), 181–195.

Batjes, N.H., 2005, Organic carbon stocks in the soils of Brazil: Soil Use and Management, 21(1), 22–24.

Box, G.E.P., Muller, M.E., 1958, A note on the generation of random normal deviates: Annals of Mathematical Statistics, 29(2), 610–611.

Buckley, J.A., Georgianna, T.D., 2001, Analysis of statistical outliers with application to whole effluent toxicity testing: Water Environment Research, 73(5), 575–583.

Bugner, E., Rutledge, D.N., 1990, Modelling of statistical tables for outlier tests: Chemometrics and Intelligent Laboratory Systems, 9(3), 257-259.

Dávalos-Álvarez, O.G., Nieto-Samaniego, A.F., Alaniz-Álvarez, S.A., Gómez-González, J.M., 2005, Las fases de deformación cenozoica en la región de Huimilpan, Querétaro, y su relación con la sismicidad local: Revista Mexicana de Ciencias Geológicas, 22(2), 129–147.

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.

Dybczyński, R., 1980, Comparison of the effectiveness of various procedures for the rejection of outlying results and assigning consensus values in interlaboratory programs involving determination of trace elements or radionuclides: Analytica Chimica Acta, 117(1), 53–70.

Dybczyński, R., Tugsavul, A., Suschny, O., 1979, Soil-5, a new IAEA certified reference material for trace element determinations: Geostandards Newsletter, 3(1), 61–87.

Dybczyński, R., Polkowska-Motrenko, H., Samczynski, Z., Szopa, Z., 1998, Virginia tobacco leaves (CTA-VTL-2) - new Polish CRM for inorganic trace analysis including microanalysis: Fresenius Journal of Analytical Chemistry, 360(3-4), 384–387.

Ebdon, D., 1988, Statistics in Geography: Oxford, Basic Blackwell, 232 p.

Esquivel-Macías C., León-Olvera, R.G., Flores-Castro, K., 2005, Caracterización de una nueva localidad fosilífera del Jurásico Inferior con crinoides y amonites en el centro-oriente de México: Revista Mexicana de Ciencias Geológicas, 22(1), 97–114.

Freeman, B.D., Quezado, Z., Zeni, F., Natanson, C., Danner, R.L., Banks, S., Quezado, M., Fitz, Y., Bacher, J., Eichacker, P.Q., 1997, rG-CSF reduces endotoxemia and improves survival during E-coli pneumonia: Journal of Applied Physiology, 83(5), 1467–1475.

Gawlowski, J., Bartulewicz, J., Gierczak, T., Niedzielski, J., 1998, Tests for outliers; a Monte Carlo evaluation of the error of first type: Chemia Analityczna (Warshaw) Chemical Analysis, 43(4), 743–753.

Gibbons, R.D., 1994, Statistical Methods for Groundwater Monitoring: New York, John Wiley, 279 p.

Gill, U., Covaci, A., Ryan, J.J., Emond, A., 2004, Determination of persistent organohelogenated pollutants in human hair reference material (BCR 397); an interlaboratory study: Analytical and Bioanalytical Chemistry, 380(7-8), 924–929.

Gladney, E.S., Roelandts, I., 1988a, 1987 compilation of elemental concentration data for USGS BIR-1, DNC-1 and W-2: Geostandards Newsletter, 12(1), 63–118.

Gladney, E.S., Roelandts, I., 1988b, 1987 compilation of elemental concentration data for USGS BHVO-1, MAG-1, QLO-1, RGM-1, SCo-1, SDC-1, SGR-1, and STM-1: Geostandards Newsletter, 12(2), 253–262.

Gladney, E.S., Jones, E.A., Nickell, E.J., Roelandts, I., 1991, 1988 compilation of elemental concentration data for USGS DTS-1, G-1, PCC-1, and W-1: Geostandards Newsletter, 15(2), 199–396.

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

Govindaraju, K., Potts, P.J., Webb, P.C., Watson, J.S., 1994, 1994 Report on Whin sill dolerite WS-E from England and Pitscurrie microgabbro PM-S from Scotland; assessment by one hundred and four international laboratories: Geostandards Newsletter, 18(2), 211–300.

Govindaraju, K., Potts, P.J., Webb, P.C., Watson, J.S., 1995, Correction to "1994 Report on Whin sill dolerite WS-E from England and Pitscurrie microgabbro PM-S from Scotland; assessment by one hundred and four international laboratories": Geostandards Newsletter, 19(1), 97.

Graybeal, D.Y., DeGaetano, A.T., Eggleston, K.L., 2004, Improved quality assurance for historical hourly temperature and humidity; development and application to environmental analysis: Journal of Applied Meteorology, 43(11), 1722–1735.

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

Guevara, M., Verma, S.P., Velasco-Tapia, F., 2001, Evaluation of GSJ intrusive rocks JG1, JG2, JG3, JG1a, and JGb1: 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; example of X-ray fluorescence spectrometry: Geostandards and Geoanalytical Research, 29(3), 271-284.

Hanson, D., Kotuby-Amacher, J., Miller, R.O., 1998, Soil analysis; Western States proficiency testing program for 1996: Fresenius Journal of Analytical Chemistry, 360(3-4), 348–350.

Harcourt, A.H., Coppeto, S.A., Parks, S.A., 2005, The distribution-abundance (density) relationship; its form and causes in a tropical mammal order, Primates: Journal of Biogeography, 32(4), 565–579.

Hofer, J.D., Murphy, J.R., 2000, Structured use of the median in the analytical measurement process: Journal of Pharmaceutical and Biomedical Analysis, 23(4), 671–686.

Holcombe, G., Lawn, R., Sargent, M., 2004, Improvements in efficiency of production and traceability for certification of reference materials: Accreditation and Quality Assurance, 9(4-5), 198–204.

Ifrim, C., Stinnesbeck, W., Schafhauser, A., 2005, Maastrichtian shallow-water ammonites of northwestern Mexico: Revista Mexicana de Ciencias Geológicas, 22(1), 48–64.

Ihnat, M., 2000, Performance of NAA methods in an international inter-laboratory reference material characterization campaign: Journal of Radioanalytical and Nuclear Chemistry, 245(1), 73–80.

Imai, N., Terashima, S., Itoh, S., Ando, A., 1996, Database on internet for geological survey of Japan geochemical reference samples: Geostandards Newsletter, 20(2), 161–164.

In't Veld, P.H., 1998, The use of reference materials in quality assurance programmes in food microbiology laboratories: International Journal of Food Microbiology, 45(1), 35–41.

Kern, M., Preimesberger, T., Allesch, M., Pail, R., Bouman, J., Koop, R., 2005, Outlier detection algorithms and their performance in GOCE gravity field processing: Journal of Geodesy, 78(9), 509–519.

Kinderman, A.J., Ramage, J.G., 1976, Computer generation of normal random variables: Journal of American Statistical Association, 71(356), 893–896.

King, E.P., 1953, On some procedures for the rejection of suspected data: Journal of American Statistical Association, 48(263), 531–533.

Langton, S.D., Chevennement, R., Nagelkerke, N., Lombard, B., 2002, Analysing collaborative trials for qualitative microbiological methods; accordance and concordance: International Journal of Food Microbiology, 79(3), 175–181.

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

Lin, Z., Inn, K.G.W., Filliben, J.J., 2001, An alternative statistical approach for interlaboratory comparison data evaluation: Journal of Radioanalytical and Nuclear Chemistry, 248(1), 163–173.

Linkosalo, T., Hakkinen, R., Hari, P., 1996, Improving the reliability of a combined phenological time series by analyzing observation quality: Tree Physiology, 16(7), 661–664.

Lozano, R., Bernal, J.P., 2005, Assessment of eight new geochemical reference materials for XRF major and trace element analysis: Revista Mexicana de Ciencias Geológicas, 22(3), 329-344.

Luedeling, E., Nagieb, M., Wichern, F., Brandt, M., Deurer, M., Buerkert, A., 2005, Drainage, salt leaching and physico-chemical properties of irrigated man-made terrace soils in a mountain oasis of northern Oman: Geoderma, 125(3-4), 273–285.

Lugo-Ospina, A., Dao, T.H., Van Kessel, J.A., Reeves III, J.B., 2005, Evaluation of quick tests for phosphorus determination in dairy manures: Environmental Pollution, 135(1), 155–162.

Marsaglia, G., 1968, Random numbers fall mainly in the planes: National Academy of Science Proceedings, 61(1), 25–28.

Marsaglia, G., Bray, T.A., 1964, A convenient method for generating normal variables: Society for Industrial and Applied Mathematics, SIAM Review, 6(3), 260–264.

Matsumoto, M., Nishimura, T., 1998, Mersenne Twister; A 623-dimensionally equidistributed uniform pseudorandom number generator: Association for Computing Machinery, ACM Transactions of Modelling and Computer Simulations, 8(1), 3–30.

Morabito, R., Massanisso, P., Cámara, C., Larsson, T., Frech, W., Kramer, K.J.M., Bianchi, M., Muntau, H., Donard, O.F.X., Lobinski, R., McSheehy, S., Pannier, F., Potin-Gautier, M., Gawlik, B.M., Bowadt, S., Quevauviller, P., 2004, Towards a new certified reference material for butyltins, methylmercury and arsenobetaine in oyster tissue: Trends in Analytical Chemistry, 23(9), 664–676.

Morán-Zenteno, D.J., Alba-Aldave, L.A., Martínez-Serrano, R.G., Reyes-Salas, M.A., Corona-Esquivel, R., Angeles-García, S., 1998, Stratigraphy, geochemistry and tectonic significance of the Tertiary volcanic sequences of the Taxco-Quetzalapa region, southern Mexico: Revista Mexicana de Ciencias Geológicas, 15(2), 167–180.

Muranaka, K., 1999, Teaching statistical methods: Journal of Chemical Education, 76(4), 469–469.

Patriarca, M., Chiodo, F., Castelli, M., Corsetti, F., Menditto, A., 2005, Twenty years of the Me.Tos. Project; an Italian national external quality assessment scheme for trace elements in biological fluids: Microchemical Journal, 79(1-2), 337–340.

Reed, D.S., Smoll, J., Gibbs, P., Little, S.F., 2002, Mapping of antibody responses to the protective antigen of Bacillus anthracis by flow cytometric analysis: Cytometry, 49(1), 1–7.

Roelandts, I., 1994, Nickel and nickel alloy reference materials: Spectrochimica Acta, 49B(10), 1039–1048.

Rorabacher, D.B., 1991, Statistical treatment for rejection of deviant values; critical values of Dixon's "Q" parameter and related subrange ratios at the 95% confidence level: Analytical Chemistry, 63 (2), 139-146.

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 sensitivities and detection limits: Journal

of Chromatography A, 997(1-2), 171–182.
Schaber, J., Badeck, F.W., 2002, Evaluation of methods for the combination of phenological time series and outlier detection: Tree Physiology, 22(14), 973–982.
Sevransky, J., Vandivier, R.W., Gerstenberger, E., Correa, R., Ferantz, V., Banks, S.M., Danner, R.L., Eichacker, P.Q., Natanson, C., 2005, Prophylactic high-dose Nω-monomethyl-L-arginine prevents the late cardiac dysfunction associated with lethal tumor necrosis factor-α challenge in dogs: Shock, 23(3), 281–288.
Shoemaker, D.P., Garland, C.W., Nibler, J.W., 1996, Experiments in Physical Chemistry. 6th edition: New York, McGraw Hill, 778 p.
Sieber, J., Broton, D., Fales, C., Leigh, S., MacDonald, B., Marlow, A., Nettles, S., Yen, J., 2002, Standards reference materials for cements: Cement and Concrete Research, 32(12), 1899–1906.
Stancak, A., Hoechstetter, K., Tintera, J., Vrana, J., Rachmanova, R., Kralik, J., Scherg, M., 2002, Source activity in the human secondary somatosensory cortex depends on the size of corpus callosum: Brain Research, 936(1-2), 47–57.
Stevens, R.J., O'Bric, C.J., Carton, O.T., 1995, Estimating nutrient content of animal slurries using electrical conductivity: Journal of Agricultural Science, 125(2), 233–238.
Taylor, B.J., 2000, A statistical analysis of the metallicities of nine old superclusters and moving groups: Astronomy and Astrophysics, 362, 563–579.
Thomulka, K.W., Lange, J.H., 1996, A mixture toxicity study employing combinations of tributyltin chloride, dibytyltin dichloride, and tin chloride using the marine bacterium vibrio harveyi as the test organism: Ecotoxicology and Environmental Safety, 34(1), 76–84.
Tigges, M., Iuvone, P.M., Fernández, A., Sugrue, M.F., Mallorga, P.J., Laties, A.M., Stone, R.A., 1999, Effects of muscarinic cholinergic receptor antagonists on postnatal eye growth of rhesus monkeys: Optometry and Vision Science, 76(6), 397–407.
Treviño-Cázares, A., Ramírez-Fernández, J.A., Velasco-Tapia, F., Rodríguez-Saavedra, P., 2005, Mantle xenoliths and their host magmas in the Eastern Alkaline Province (NE Mexico): International Geology Review, 47(12), 1260-1286.
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., Guevara, M., Verma, S.P., 2001, Evaluation of concentration data in geochemical reference materials: Chemie der Erde, 61(1), 69–91.
Verma, M.P., 2004, A revised analytical method for HCO_3^- and CO_3^{2-} determinations in geothermal waters; an assessment of IAGC and IAEA interlaboratory comparisons: Geostandards and Geoanalytical Research, 28(3), 391–409.
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., 1998, 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., 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., in press, Extension-related origin of magmas from a garnet-bearing source in the Los Tuxtlas volcanic field, Mexico: International Journal of Earth Sciences (Geologische Rundschau).
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., Torres-Alvarado, I.S., Sotelo-Rodríguez, Z.T., 2002, SINCLAS; standard igneous norm and volcanic rock classification system: Computers & Geosciences, 28(5), 711–715.
Villaseñor, A.B., González-León, C.M., Lawton, T.F., Aberhan, M., 2005, Upper Jurassic ammonites and bivalves from Cucurpe Formation, Sonora (Mexico): Revista Mexicana de Ciencias Geológicas, 22(1), 65–87.
Villeneuve, J.-P., de Mora, S.J., Cattini, C., 2002, World-wide and regional intercomparison for the determination of organochlorine compounds and petroleum hydrocarbons in sediment sample IAEA-417: Vienna, Austria, Analytical Quality Control Services, International Atomic Energy Agency, 136 p.
Villeneuve, J.-P., de Mora, S., Cattini, C., 2004, Determination of organochlorinated compounds and petroleum in fish-homogenate sample IAEA-406: results from a worldwide interlaboratory study: Trends in Analytical Chemistry, 23(7), 501–510.
Wang, X.-D., Söderlund, U., Lindh, A., Johansson, L., 1998, U-Pb and Sm-Nd dating of high-pressure granulite- and upper amphibolite facies rocks from SW Sweden: Precambrian Research, 92(4), 319–339.
Woitge, H.W., Scheidt-Nave, C., Kissling, C. Leidig-Bruckner, G., Meyer, K., Grauer, A., Scharla, S.H., Ziegler, R., Seibel, M.J., 1998, Seasonal variation of biochemical indexes of bone turnover: Results of a population-based study: Journal of Clinical and Endocrinological Metabolism, 83(1), 68–75.
Yurewicz, K.L., 2004, A growth/mortality trade-off in larval salamanders and the coexistence of intraguild predators and prey: Oecologia, 138(1), 102–111.
Zaric S., Niketic S.R., 1997, The anisotropic π-effect of the nitro group in ammine-nitro cobalt(III) complexes: Polyhedron, 16(20), 3565–3569.

Manuscript received: September 21, 2005
Corrected manuscript received: March 10, 2006
Manuscript accepted: March 16, 2006