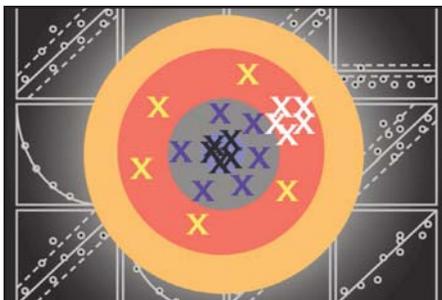


# SERIES



## Modern Analytical Facilities 2. A Review of Quality Assurance and Quality Control (QA/QC) Procedures for Lithochemical Data

Stephen J. Piercey  
*Department of Earth Sciences  
 Memorial University of Newfoundland  
 300 Prince Philip Drive  
 St. John's, NL, Canada, A1B 3X5  
 Email: spiercey@mun.ca*

### SUMMARY

Quality assurance and quality control (QA/QC) are critical components of modern analytical geochemistry. A properly constructed QA/QC program identifies both the source of analytical error and provides a means of establishing confidence in and assessing limitations of analytical data. A QA/QC program involves monitoring precision, accuracy, and potential contamination from sampling to analysis. Precision can be monitored via the systematic insertion of sample, pulp, and analytical duplicates, and reference materials; the resulting data are subsequently

evaluated using scatterplots, statistical tests (e.g. % relative standard deviation), Thompson-Howarth plots, and the average coefficient of variation ( $CV_{avg}$  (%)). Accuracy is determined through the submission of reference materials and monitored using statistical tests (e.g. % relative difference, t-test) and Shewart control charts. Blanks test contamination and results are monitored using Shewart control charts.

### SOMMAIRE

L'assurance de la qualité et le contrôle de la qualité (AQ-CQ) sont deux composantes essentielles à la géochimie analytique moderne. Un programme AQ-CQ bien conçu définit à la fois la source de l'erreur d'analyse et un moyen d'établir la confiance et d'évaluer les limites des données analytiques. Un programme AQ-CQ comprend le contrôle de la précision, de l'exactitude et de la contamination potentielle, de l'étape d'échantillonnage à l'analyse. La précision peut être contrôlée via l'insertion systématique d'échantillon, de pulpes, et de doublons d'analyse, et de matériaux de référence; les données obtenues sont ensuite évaluées en utilisant des diagrammes de dispersion, des tests statistiques (pourcentage d'écart type relatif, par ex.), des courbes de Thompson-Howarth, et des coefficients de variation moyens ( $CV_m$  (%)). La précision est déterminée par la soumission de documents de référence et de contrôle par des tests statistiques (différence relative en %, t-test, par ex.) et des graphiques de contrôle de Shewart. La contamination d'essais à blanc et les résultats sont contrôlés par des graphiques de contrôle Shewart.

### INTRODUCTION

The constant improvement in instrumental technology over the last 20 years has led to an increasing utilization of high precision lithochemical data in research and mineral exploration and development. Technology has also resulted in data acquisition in a fraction of the time and at much lower costs than even a decade ago. Despite the increase in utilization of lithochemical data, many individuals are uncertain about how to monitor and evaluate the quality of their data. Furthermore, regulations in some industrial sectors (e.g. NI-43-101 in the mining sector) require documentation of quality assurance and quality control (QA/QC) protocols before acquired data can be reported to the public. In addition, many in the research community suggest that there should be standardized information available about all analytical data reported, including well documented QA/QC protocols (e.g. Potts 2012).

In this manuscript various quality control methods are reviewed. The manuscript focuses primarily on lithochemistry; however, the methods shown are applicable to various disciplines of geochemistry (e.g. surficial, environmental). Furthermore, the paper is aimed at providing the general reader with a simplified, user-friendly approach to evaluating QA/QC data.

### SAMPLING, SAMPLE PREPARATION, AND ANALYTICAL METHODOLOGIES: OVERVIEW

While there are numerous reviews on analytical technologies used in lithochemistry (e.g. Jenner 1996; Sylvester 2001; Caughlin 2010), it is

important to briefly discuss analytical methods and sample collection/preparation for litho geochemistry. Sampling is the first step of any analytical program, and there should be a clear and documented 'chain of custody' for samples, with a limited number of individuals responsible for sampling and interaction with the laboratory (i.e. qualified persons or their designates). In some cases, the people who collect the samples will be responsible for undertaking the analyses, particularly in university-oriented research. In general, field samples for litho geochemistry should be roughly 1 kg in weight (i.e. ~two fist sizes); if the samples are coarse grained (e.g. megacrystic) it may require larger samples (e.g. Stanley 2007). Weathered edges of samples should be removed with a hammer in the field or with a rock saw when available. Drill core litho geochemical samples can be roughly 15–20 cm in length and should be unweathered, free of contamination from drill grease or other material. Both field and drill core samples should not cross-cut lithological boundaries. For primary petrology and chemostratigraphy (i.e. original geochemical signatures in the rock) the samples should be free of veins, alteration, and mineralization. Sampling for alteration or mineralization studies (i.e. superimposed effects) should contain the mineral phases and textures representative of the mineralization and alteration. To ensure proper 'chain of custody' all samples should be placed in bags, preferably high quality plastic/polyurethane bags, with each sample labeled with a unique tag, ideally with a UPC coding system, and the bags sealed with either zip ties or tape to prevent any potential tampering during transport from the field to the laboratory.

Samples should be pulverized with 90% passing through a 75  $\mu\text{m}$  mesh. Often this involves crushing in a steel jaw mill, riffing and sample splitting, with an aliquot of the riffled material being pulverized in a bowl and puck assembly or a ball mill pulverizer. Choosing the right pulverizer is important to prevent contamination of samples. An agate pulverizing mill results in the least contamination of samples (with the exception of  $\text{SiO}_2$ ), but it is often costly and some laboratories do

not provide this option. A valid compromise of cost and minimal contamination is a hardened carbon (low Cr and Fe) steel pulverizer. It is strongly suggested to avoid steel and tungsten-carbide mills as they often add contaminants to the rock powders, particularly Ta, W, Fe, Ni, Co, and Cr (Potts 1987).

During the sampling and sample preparation it is critical to ensure all equipment is cleaned in between samples. Cleaning of the cutting saw, crushing, and pulverizing equipment in between samples is imperative. Furthermore, pre-contamination of these instruments with your own sample is also a means of reducing contamination from external sources (e.g. previous samples). This involves taking part of your sample, putting it through the instrumentation, discarding this material, then processing another part of your sample to be used for analysis. These cleaning steps are particularly important because often the largest potential source of contamination in analytical programs occurs during the early stages of sample preparation (Abzalov 2008, 2011).

The powders created during the previous steps are subsequently analyzed using digestion methods, whereby the samples are dissolved either partially or fully and converted to a solution; or a solid-source method, where the sample powder is analyzed as-is or combined with a binding agent prior to analysis. The most common solid source methods include X-ray fluorescence (XRF) and instrumental neutron activation analysis (INAA). X-ray fluorescence provides major and select trace elements, albeit at a higher detection limit than many digestion methods (see below), whereas INAA provides an incomplete suite of trace elements (when compared to ICP-MS) including some rare earth elements (REE), high field strength elements (HFSE), and low field strength elements (LFSE). These methods, although still used, have largely been superseded by solution methods such as inductively coupled plasma emission spectroscopy (ICP-ES) and inductively coupled plasma mass spectrometry (ICP-MS). Both ICP-ES and ICP-MS require that the sample be dissolved or digested. For high quality REE, HFSE, and LFSE litho geochemical

data when using digestion methods, it is important to ensure that samples are completely digested; the utilization of total digestions, such as lithium metaborate fusions followed by acid dissolution, bomb dissolutions, or sinters (e.g.  $\text{Na}_2\text{O}_2$  sinter), ensures complete sample dissolution. In some cases, partial dissolutions (e.g. aqua regia, four-acid) may be appropriate for base metals and volatile metal species, particularly in exploration litho geochemistry when trying to elucidate metal contents associated with sulfide mineralization. The resultant solutions are analyzed using ICP-ES for major elements and high concentration trace elements (e.g. wt% to 10s of ppm) and by ICP-MS for trace elements (i.e. 100s of ppm to ppb). In some cases additional methods may be utilized, including atomic absorption spectroscopy (AAS), particularly for high concentration base metal species; gravimetric methods for Au-rich samples; infrared spectroscopy for sulfur, total carbon, and  $\text{CO}_2$ ; and cold vapour AAS (CV-AAS) or cold vapour flow injection mercury system (CV-FIMS) for Hg.

## QUALITY CONTROL AND QUALITY ASSURANCE PROTOCOLS

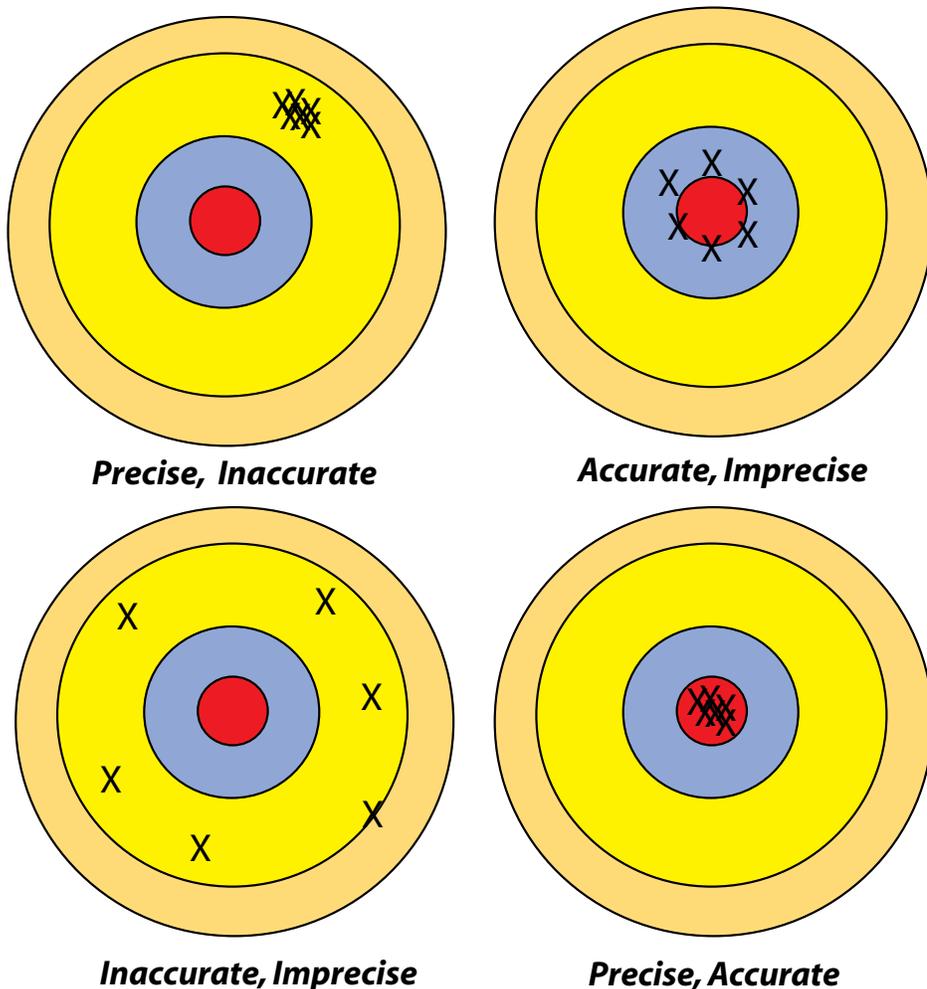
### Definitions

**Precision** – this is the measure of reproducibility of a measurement (Fig. 1).

**Accuracy** – how close a measured value is to a known or accepted value (Fig. 1).

**Lower Limit of Detection (LOD)** – this is the lowest concentration at which an instrument can detect a signal above background. The LOD is obtained by measuring the signal in a blank so as to determine the background signal. Assuming random and normally distributed errors of the background signal ( $X_B$ ) with a standard deviation  $S_B$  the  $\text{LOD} = X_B + S_B (3\sigma)$ ; Potts 1987; Jenner 1996).

**Limit of Quantification (LOQ)** – this is the concentration above the background where a result is considered quantifiable and trusted. Given a background with a concentration of



**Figure 1.** Graphical representation of precision and accuracy.

$X_B$  and a standard deviation of  $S_B$ , the LOQ is  $10S_B$  above the background (i.e.  $LOQ = X_B + 10S_B \approx 3.3 LOD$ ) (Potts 1987; Jenner 1996).

**Upper Limit of Detection (ULD)** – this is the highest concentration that an instrument can reliably quantify.

**Instrument Drift** – this is the variability in the measurement value of the same characteristic measured under the same conditions, same operator, at different points in time. Drift should be monitored and corrected for by using time-integrated calibration with reference materials.

**Bias** – this is the difference between the result that is expected and the result obtained. Bias can originate from a variety of sources, ranging from sampling, sample preparation, analysis, and statistical assumptions made. Bias is monitored using dupli-

cates, reference materials, and blanks.

**Reference Materials** – these are materials that have certified values for given elements of interest. They can be external reference materials from a commercial or government agency, or they can be created internally generally via a round robin analytical program. Furthermore, when testing accuracy using an internal reference material (or an external reference material) the material should be certified using the sample preparation, digestion(s) (if appropriate), and analytical method(s) utilized by the analytical program for which accuracy is being tested.

**Duplicate** – this is a repeat of a sample, sample pulp, coarse reject, or analysis taken at random to evaluate error introduced during sampling and preparation, and/or analytical bias and precision.

**Blank** – this is a sample (or sample solution) that does not contain detectable concentrations of the elements of interest using the selected analytical method, and is used to monitor contamination during the sample preparation and analysis. In some cases a blank is utilized to verify the lower limit of detection for an instrument.

**Matrix** – this is the general make-up of the sample to be analyzed. For example, a soil sample is a soil matrix, a mineralized sample is a mineralized matrix, a felsic rock has a felsic matrix. Matrix matching and choosing reference materials that are similar to the matrix being analyzed is a critical aspect of exploration geochemistry (see below).

**General Goal of a QA/QC Program**

The goal of a QA/QC program is to monitor the quality of data from sampling in the field through generation of the final results, so as to ensure that both the user and external parties are confident in the quality of the data obtained (Vallee 1998; Smee 2007; Potts 2012). Furthermore, a properly designed and implemented QA/QC program will also identify errors and the potential stage of the analytical program where the error occurred (e.g. sampling, crushing, sample preparation, instrumentation). Errors within an analytical program are additive, therefore, small errors throughout the sampling and analysis process can result in large errors by the time the actual analytical results are obtained (e.g. Stanley 2003a).

A QA/QC program is therefore aimed at understanding the following:

1. precision of results from sampling, sample preparation, and analysis;
2. accuracy of analytical results; and
3. contamination that results from sampling, sample preparation, and analysis.

Overall, the QA/QC program should identify bias during the analytical program through the utilization of duplicates, reference materials, and blanks.

**Measuring Precision**

Precision is tested using either refer-

ence materials or duplicate analyses. Reference materials have certified values, but limited ranges of values for given elements; therefore, they do not always measure the true precision as a function of concentration range. Random to semi-random analytical duplicates are alternatives to reference materials and provide a range of concentrations that allows precision to be monitored as a function of concentration. This is particularly important because precision decreases greatly as one approaches the LOD of a given analytical method. Duplicates should include both analytical duplicates (50%) and more specific duplicates (50%) that are chosen to reflect the range of concentrations of the element/commodity of interest (Abzalov 2008). Duplicate samples should be inserted at regular intervals when submitted to the laboratory.

### Reference Materials

Reference materials are used mostly to test accuracy, but can also be utilized to monitor precision. Reference materials should be inserted at regular intervals (e.g. one in every 20 samples); they should have a similar matrix, range of elements, and range of concentrations as the samples to be analyzed, and be certified with the same digestion and analytical technique as used during an analytical program. It is very important to note that compiled values for reference materials (e.g. Govindaraju 1994) will often have significant uncertainties because of the compilation of reported values from different sources with results obtained by numerous analytical methods.

Precision can be calculated from replicate analyses of the reference material using the relative standard deviation (%RSD; Jenner 1996):

$$\%RSD_i = 100 \times \frac{s_i}{\mu_i} \quad (\text{Eq. 1})$$

where  $\%RSD_i$  = percent relative standard deviation for element  $i$ ;  $s_i$  = standard deviation of the mean from the series of analytical runs for element  $i$ ; and  $\mu_i$  = mean value of element  $i$  over a series of analytical runs. Although %RSD varies as one approaches the LOD and ULD, and is dependent on the material in question (e.g. whole

rock versus nuggety gold), if an element is above the LOQ, then %RSD between 0 and 3% is excellent, between 3 and 7% is very good to good, 7–10% RSD is good, and >10% is not precise (Jenner 1996). Notably, %RSD will be greater, and precision poorer, as one approaches and goes below the LOQ towards the LOD.

An alternative method for testing precision is to use a chi-squared critical value (Abzalov 2008, 2011), whereby the standard deviation of the analytical results of a certified reference material ( $S_m$ ) is compared to the standard deviation reported for the certified reference material ( $\sigma_c$ ) using the following:

$$\left( \frac{S_m}{\sigma_c} \right)^2 \leq \frac{\chi_{(n-1);0.95}^2}{n-1} \quad (\text{Eq. 1.1})$$

where  $n$  = number of replicate analyses of the certified reference material, and is the critical value of the chi-square distribution at  $\alpha = 0.05$  (95% confidence level) for  $(n-1)$  degrees of freedom (ISO 1989; Abzalov 2008, 2011). If the values satisfy the above inequality then the data are precise at the 95% confidence level.

### Duplicate Samples: Scatterplots

Scatterplots provide a means of visualizing the duplicate data obtained during an analytical program and involve plotting the initial result on the X-axis ( $X_1$ ) and the duplicate data on the Y axis ( $X_2$ ) with control lines that represent a given level of precision (i.e. 5%, 10%, 20%, etc.; Figs. 2, 3). If all data lie within the control lines then the data are precise to the set precision (Figs. 2A, 3A). If the data lie outside the control lines they are not precise to said level (Figs. 2, 3B). Furthermore, one can view the relationships between precision and concentration and identify potential concentration-related problems affecting precision (Fig. 2).

### Duplicate Samples: Thompson-Howarth Plots

The Thompson-Howarth method of testing precision also uses duplicate samples. The Thompson-Howarth method has two forms: 1) the long method for when there are greater than 50 duplicates; and 2) the short

method for when there are less than 50 duplicates. In both the short and long methods the duplicate data ( $X_1$  = original analysis,  $X_2$  = duplicate analysis) are plotted in the form of the X-axis equal to  $(X_1+X_2)/2$  and the Y-axis equal to  $|X_1-X_2|$ ; these are the mean and absolute deviation, which is a proxy for the standard deviation, of the replicate analyses, respectively. The Thompson-Howarth method is discussed extensively in Thompson and Howarth (1978), Fletcher (1981), Stanley (2003b; 2006), Stanley and Lawie (2007), and Abzalov (2008, 2011).

### Thompson-Howarth Short Method

The short method of the Thompson-Howarth plots has the data plotted on a control graph with a control line and replicate data are compared to this control line. The major challenge in this plot is creating the control line, with the location of the control dependent on level of precision chosen by the user. For most major elements precision should be  $\leq 5\%$ ; for most trace elements it should be between 5–10%, and precision can be in excess of 20% for some metals (see *Average Coefficient of Variation* section below).

The Thompson-Howarth method assumes that all the errors in the duplicate data follow a normal distribution and are random. If the latter criteria are met, then precision at the 95% confidence level at a given concentration is given by ( $P_c$ ):

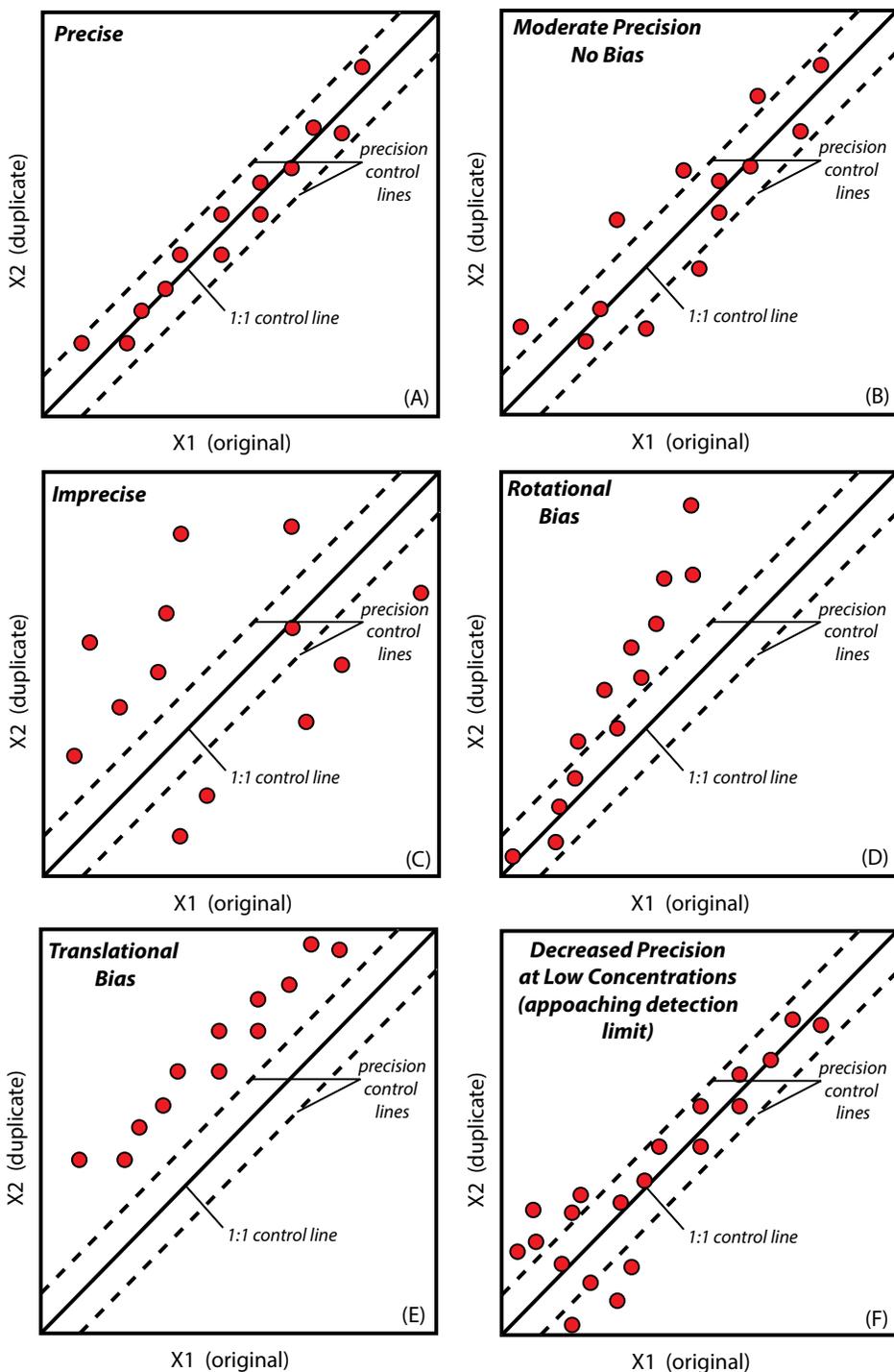
$$P_c = \frac{2 \times s \times 100}{\bar{X}} \quad (\text{Eq. 2})$$

where  $s$  is the standard deviation of the population and  $\bar{X}$  is the mean value of the population.

From the duplicate analyses the relationship between the median of the population (i.e.  $|X_1-X_2|$ ) and standard deviation is given by:

$$M_d = PF \times \sqrt{2} \times \sigma_c \quad (\text{Eq. 3})$$

where  $\sigma_c$  is the standard deviation of the population,  $M_d$  is the median difference (i.e.  $|X_1-X_2|$  is the estimate of the population's standard deviation), and  $PF$  is the percentile factor. The



**Figure 2.** Precision scatterplots that plot duplicate versus original analyses. Samples should lie within the defined precision envelopes. The geometric distribution of data on these plots provides insight into the nature of imprecision and also provides a means of monitoring the source of error. A) Precise data lying within the control lines. B) Moderately precise data with scatter suggesting no bias. C) Very imprecise data with a shot-gun distribution. D) Data with a rotational bias. E) Data with a translational bias (i.e. multiplier effect). F) Decreasing precision with concentration (heteroscedastic effect). Modified from Fletcher (2002).

percentile factor gives us an idea of how many of the sample population absolute differences (i.e.  $|X_1 - X_2|$ ) should lie below the control line on a

Thompson-Howarth plot. For example, if the percentile is the 50<sup>th</sup> we should expect that 50% of the population absolute differences lies below the

control line. If the 95<sup>th</sup> percentile is chosen then 95% of the population absolute differences should lie below the control line. Typically we use the 90<sup>th</sup> or 95<sup>th</sup> percentile as control lines (Thompson and Howarth 1978; Fletcher 1981, 1986; Stanley 2003b).

Using the relationship above, at the 95<sup>th</sup> percentile, the equation above becomes:

$$M_d = 1.960 \times \sqrt{2} \times \sigma_c \quad (\text{Eq. 4})$$

which is equivalent to:

$$M_d = 2.772 \times \sigma_c \quad (\text{Eq. 5})$$

At the 90<sup>th</sup> percentile equation (3) is rewritten to:

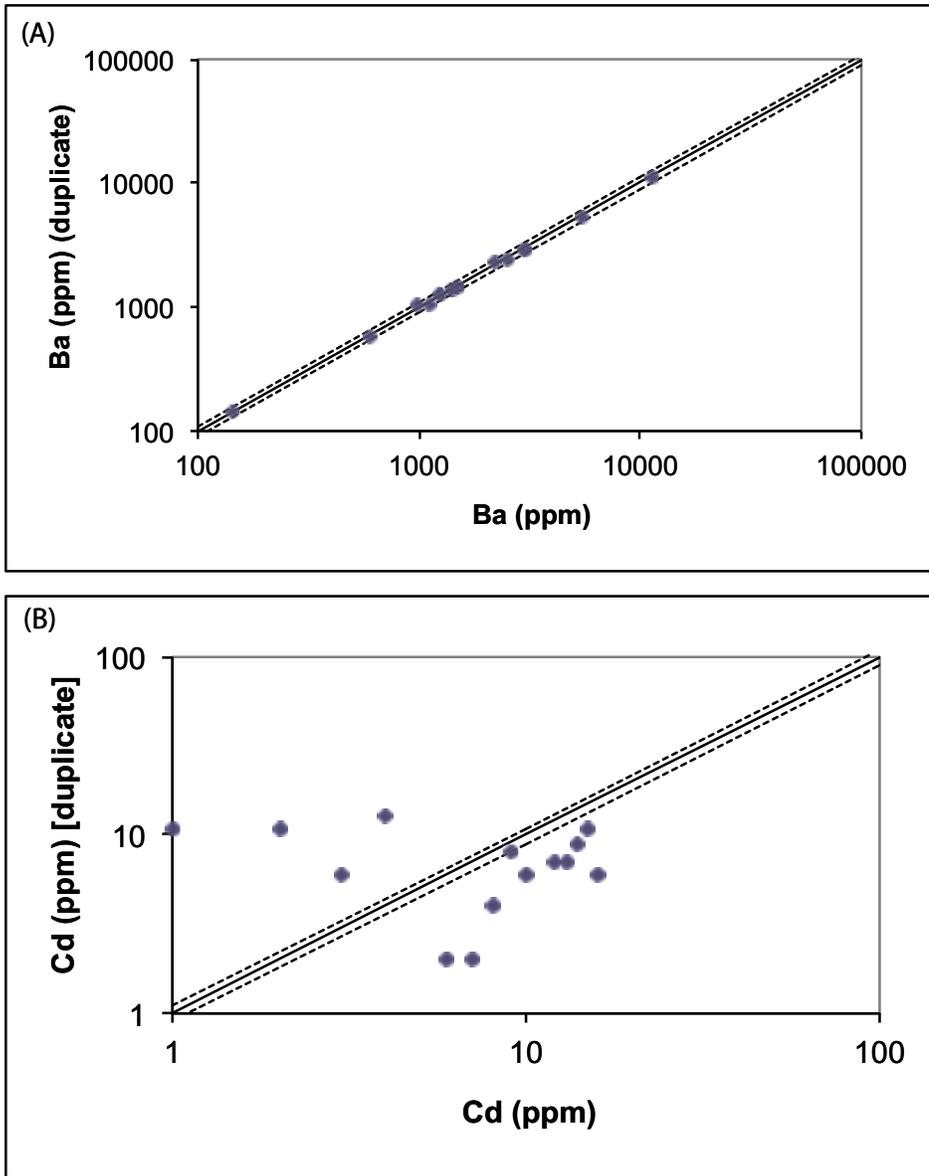
$$M_d = 2.326 \times \sigma_c \quad (\text{Eq. 6})$$

With equations (4) to (6),  $M_d = |X_1 - X_2|$  and  $\sigma_c$  is found by assuming that  $\sigma_c \approx s$  as in equation (2). Assuming we know the level of precision we desire (i.e.  $P_c$  is defined), equation (3) can be rearranged to find the value of  $s$  as a function of concentration:

$$s = \frac{P_c \bar{X}}{200} \quad (\text{Eq. 7})$$

Allowing  $\bar{x}$  to be the concentration of interest (i.e. input 0.001, 0.01, 0.1, 1, 10, 100, 1000, etc.) then  $s$  can be found at different concentrations and can be used in equation (5) so as to have  $|X_1 - X_2|$  as a linear relationship relative to  $s$  at a given percentile.

Upon creating a plot the next step is to interpret it. Figures 4 and 5 are typical Thompson-Howarth plots at the 90<sup>th</sup> percentile. If all pairs lie below the control line they are precise to that level. However, in cases where not all pairs are below the line it needs to be determined whether the points above the line ( $M$ ), relative to the number of analytical pairs ( $N$ ), are due to potentially random errors using a binomial probability chart (Table 1). For example, in Figure 4A, all points lie below the control line, therefore MgO is clearly precise to 15%. In Figure 4B, there is one point lying above the line



**Figure 3.** Scatterplots of real data. A) Barium illustrating precision to 10%. Note all data are within the control lines. B) Cadmium showing extremely poor precision.

( $M=1$ ) out of fifteen analytical pairs ( $N=15$ ), suggesting that there is a ~79% chance that this is a random error (Table 1) and  $\text{Fe}_2\text{O}_3$  is precise to 10%. In the case of Sb in Figure 4C, five pairs ( $M=5$ ) out of eight pairs ( $N=8$ ) lie above the line suggesting that the data are not precise to 15%; the probability that these points are above the line due to random errors is <1% (Table 1).

#### Thompson-Howarth Long Method

The long method is employed when there are greater than 50 pairs of duplicates. To create a long method

plot one must (Fletcher 1986):

1. calculate the mean values  $[(X_1+X_2)/2]$  and absolute differences  $[|X_1-X_2|]$  for the sample set;
2. arrange the list in increasing order of the concentration means;
3. from the first 11 results obtain the mean concentration  $[(X_1+X_2)/2]$  and median difference  $[|X_1-X_2|]$ ;
4. repeat this for each successive group of 11 ignoring any remainders <11;
5. calculate or obtain graphically the linear regression of the median difference of the means (Fig. 5A);
6. from the regression calculate the

y-intercept value of  $|X_1-X_2|$ , which can be taken as the standard deviation ( $s$ ) at concentration  $X$ . Since  $|X_1-X_2|$  is an estimate of the standard deviation ( $s$ ) and  $(X_1+X_2)/2$  is an estimate of the mean concentration ( $\bar{X}$ ), this linear regression yields an equation of the form:

$$s = m\bar{X} + c \quad (\text{Eq. 8})$$

- where  $m$  is the slope of the regression line and  $c$  the y-intercept of the regression (Fig. 5A); and
7. using the results for  $s$  in equation (8), we can calculate the precision ( $P_c$ ) as a function of concentration ( $\bar{X}$ ), where we input values of  $\bar{X}$  and obtain the precision as a function of concentration (e.g. Fig. 5B):

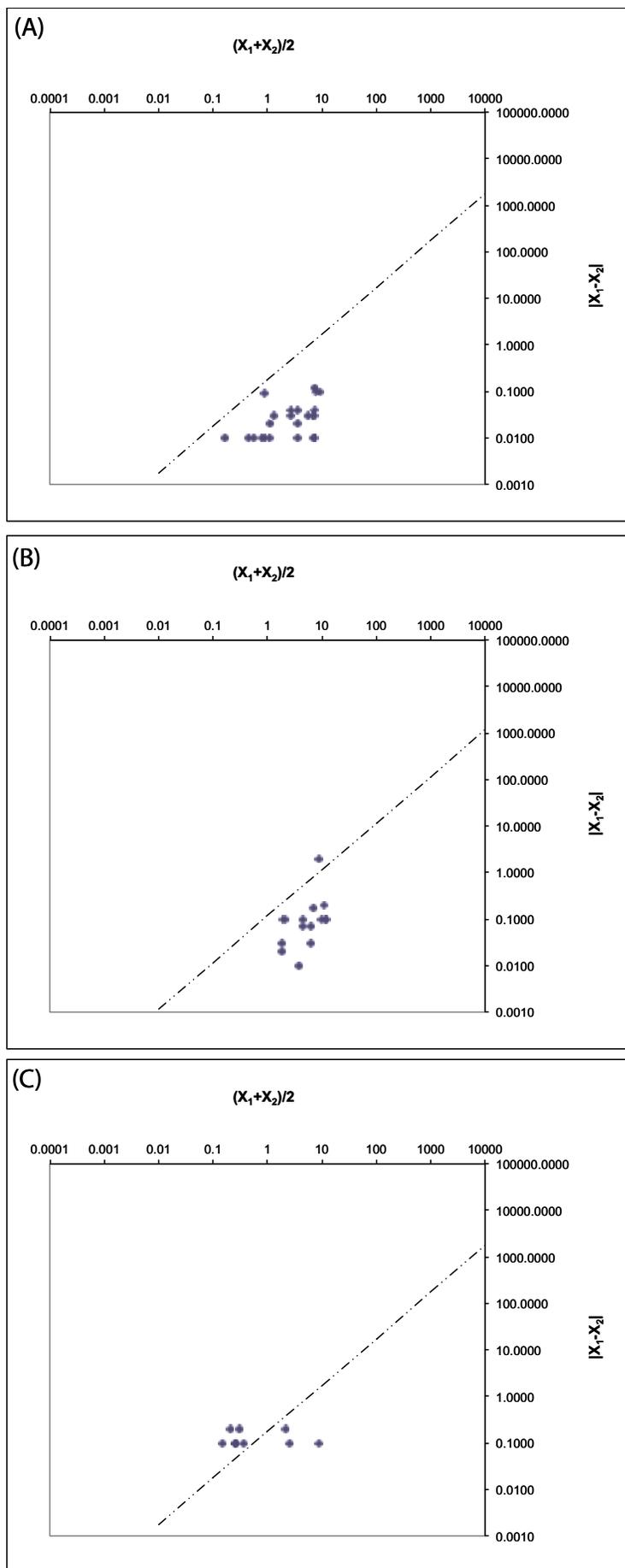
$$P_c = \frac{2 \times s \times 100}{\bar{X}} \quad (\text{Eq. 9})$$

#### Duplicate Samples: Average Coefficient of Variation

A major problem with the Thompson-Howarth method is that it is unsuitable for many datasets whose errors do not follow a normal distribution (Davis 2002), which is the case with most geochemical data, and is particularly so for Au and other elements that exhibit a nugget effect (Stanley 2006; Stanley and Lawie 2007). Because of the non-normality of lithochemical data, the Thompson-Howarth method consistently underestimates precision (Stanley and Lawie 2007; Abzalov 2008, 2011); to combat this problem, Stanley and Lawie (2007), and subsequently Abzalov (2008, 2011), recommended using the average coefficient of variation:

$$CV_{avg} (\%) = 100 \times \sqrt{\frac{2}{N} \sum_{i=1}^N \left( \frac{(a_i - b_i)^2}{(a_i + b_i)^2} \right)} \quad (\text{Eq. 10})$$

where  $CV_{avg} (\%)$  is the average coefficient of variation,  $N$  is the number of duplicate pairs, and  $a_i$  and  $b_i$  are the original and duplicate analyses, respectively, for the  $i$  different pairs.



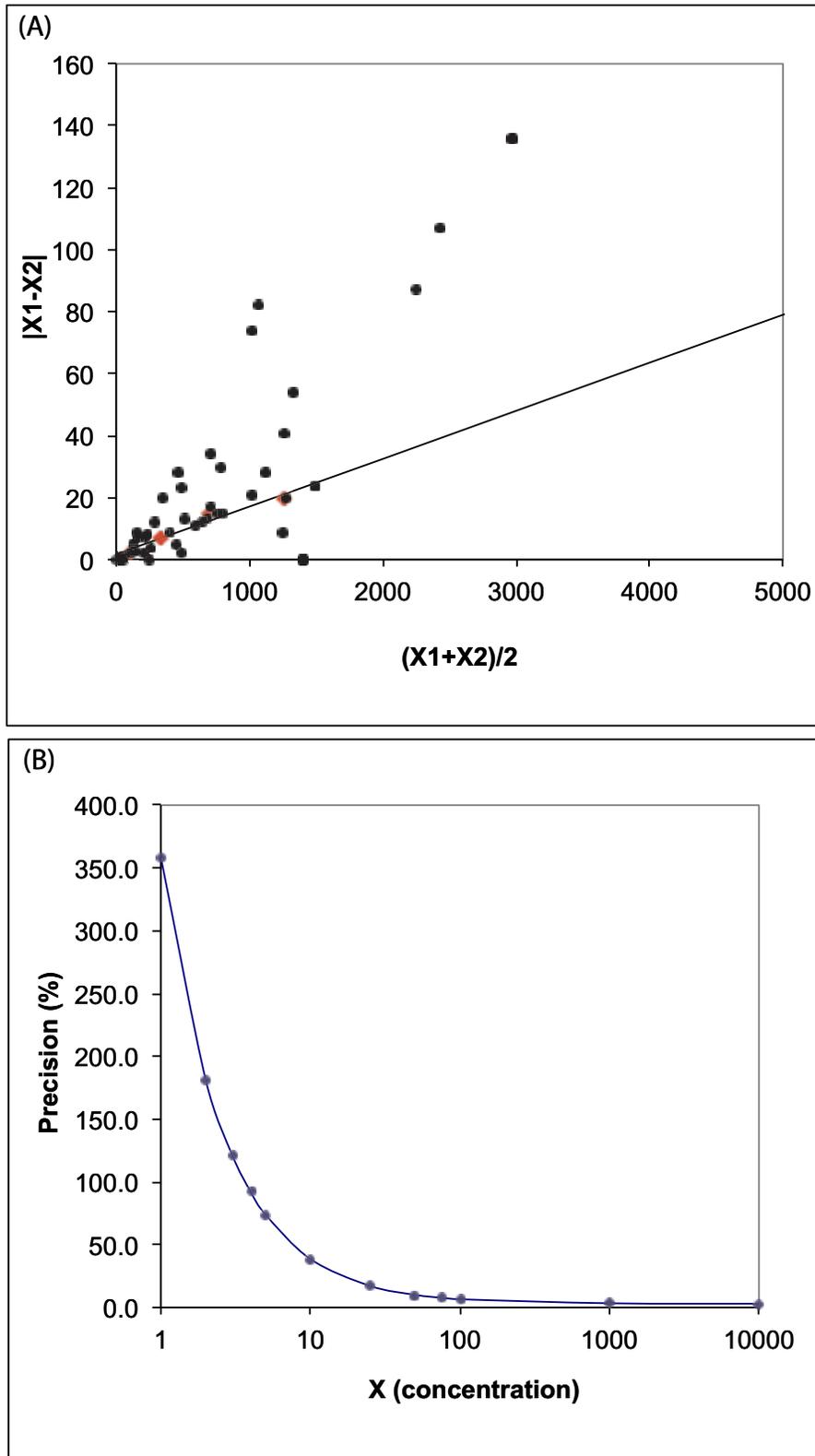
**Figure 4.** Thompson-Howarth short method plots. A) MgO, 15% precision. Notably, all samples lie below the control line, suggesting that precision for MgO is 15% or better. B) Fe<sub>2</sub>O<sub>3</sub>, 10% precision; only one point (M=1) out of fifteen pairs (N=15) lies above the line, therefore it is likely that the data are precise to 10% and the single point above the line is due to random error. C) Sb, 15% precision; five points (M=5) out of eight pairs (N=8) lie above the line, suggesting precision is worse than 15%. There is <1% chance that all these points are random errors.

Table 2 provides various estimates of the  $CV_{avg}$  (%) for various mineral deposit types and commodities for both field and analytical duplicates. For whole rock geochemistry, the  $CV_{avg}$  (%) values for elements that are >LOQ, the values for most standard elements (e.g. major, trace, REE, high field strength element, etc.), should be less than 10% for field duplicates, and less than 5–7.5% for pulp or analytical duplicates (Abzalov 2008).

The average coefficient of variation method is the best and most robust method of measuring precision using duplicate data, as it is independent of the nature of the distribution of errors (e.g. normal versus non-normal) and is equally applicable to elements that exhibit or do not exhibit the nugget effect.

**Measuring Accuracy**

Accuracy is measured using reference materials. Reference materials allow one to calculate accuracy using the percent relative difference and can be monitored using Shewart control charts. As with precision, accuracy decreases as the element value approaches the LOD of an analytical method; therefore, it is important to utilize reference materials with a range of concentrations to evaluate accuracy as a function of concentration for a given analyte.



**Figure 5.** A) Typical Thompson-Howarth long method plot for Ba by ICP-MS. The red points are the control points and the black circles are the duplicate data. The regression line through the control points provides an estimate of the standard deviation ( $s$ ) as a function of concentration ( $X$ ) (in this case  $s = 0.015X + 1.78$ ) and is used to construct the concentration-precision-type plots in (B). B) Thompson-Howarth long method concentration-precision plot. One can evaluate precision as a function of concentration using the above graph or by using equation (9).

### Relative Difference Method

Percent relative difference (%RD) is calculated from the replicate analyses of the reference materials using:

$$\%RD = 100 \times \frac{\mu_i - STD_i}{STD_i} \quad (\text{Eq. 11})$$

where  $\mu_i$  = mean value of element  $i$  in the standard over a number of analytical runs; and  $STD_i$  = 'known' or 'certified' value of  $i$  in the standard or reference material. Values for %RD can be negative or positive depending on whether values are less than the known value (i.e. %RD < 0) or greater than the known value (i.e. %RD > 0). In general %RD values of  $\pm 0-3\%$  are considered to have excellent accuracy, and values from 3–7% are considered to have very good accuracy; 7–10% have good accuracy; and values above 10% are not accurate (Jenner 1996). Accuracy, like precision, decreases as values approach the detection limit of an instrument, and is influenced by the element and type of material (e.g. Au has a nugget behavior and can be highly inaccurate).

### Statistical Tests

Another method of testing accuracy is by using the statistical t-test, given by (ISO 1989; CANMET 1998; Abzalov 2008, 2011):

$$|m - \mu| \leq \sqrt{\sigma_L^2 + \frac{S_w^2}{n}} \quad (\text{Eq. 12})$$

where  $\mu$  is the certified mean of the given standard sample,  $\sigma_L$  is the certified inter-laboratory standard deviation for the reference material (both  $\mu$  and  $\sigma_L$  are provided by the standard supplier),  $m$  is the arithmetic mean of the replicate analysis of the certified reference material,  $S_w$  is the estimated standard deviation of the replicate analysis of the certified reference material, and  $n$  is the number of replicate analyses of the certified reference material.

The above equation can be simplified using the relationship  $\sigma_L \sim 2S_w$  (CANMET 1998), and as long as  $n > 10$  then equation (10) can be re-written as (CANMET 1998; Abzalov 2008, 2011):

**Table 1.** Binomial probability chart for Thompson-Howarth plots (from Thompson and Howarth 1978; Fletcher 1981).

N	M=1	M=2	M=3	M=4	M=5	M=6	M=7
1	0.100000						
2	0.190000	0.010000					
3	0.271000	0.028000	0.001000				
4	0.343900	0.052300	0.003700	0.000100			
5	0.409510	0.081460	0.008560	0.000460	0.000010		
6	0.468559	0.114265	0.015850	0.001270	0.000055	0.000001	
7	0.521703	0.149694	0.025692	0.002728	0.000177	0.000006	0.000000
8	0.569533	0.186895	0.038092	0.005024	<b>0.000432</b>	0.000023	0.000001
9	0.612580	0.225159	0.052972	0.008331	0.000891	0.000064	0.000003
10	0.651322	0.263901	0.070191	0.012795	0.001635	0.000147	0.000009
11	0.686189	0.302643	0.089562	0.018535	0.002751	0.000296	0.000023
12	0.717570	0.340998	0.110870	0.025637	0.004329	0.000541	0.000050
13	0.745813	0.378655	0.133883	0.034161	0.006460	0.000920	0.000099
14	0.771232	0.415371	0.158360	0.044133	0.009230	0.001474	0.000181
15	<b>0.794109</b>	0.450957	0.184061	0.055556	0.012720	0.002250	0.000311

$$|m - \mu| \leq 2\sigma_L \quad (Eq. 13)$$

If  $\sigma_L$  is not known, equation (13) can be re-written as:

$$|m - \mu| \leq 4S_w \quad (Eq. 14)$$

Accurate data should satisfy the inequalities in (13) and (14).

**Monitoring Accuracy Using Shewart Control Charts**

Shewart control charts provide a very effective method to monitor the accuracy of a standard during a QA/QC program, as well as allowing one to address drift and bias (Croakin and Tobias 2006; Figs. 6, 7). The X-axis of a Shewart control chart contains the order of analysis of a reference material starting from the oldest on the left to the most recent on the right, and the Y-axis contains the values obtained for the standard (Figs. 6, 7). Also shown on the diagram is a horizontal line representing the mean value for the standard and the 2 standard deviations above and below the mean as control lines (Figs. 6, 7). These types of charts not only allow for continuous monitoring of data from each new analytical batch, but also allow monitoring of laboratory performance through time (Figs. 6, 7).

Interpretation of the graphs is relatively straightforward. If the

results for the standard are randomly distributed above and below the mean, but within the 2 standard deviations expected for a given reference material, it indicates that the lab is performing well and there is no analytical bias (Figs. 6A and 7A). If there are random outliers, this may indicate that the laboratory has not performed, or it may indicate that there are data transcription errors, such as a mislabeled standard or incorrect data input into the database (Figs. 6B and 7B). A shift at any point during the analytical program towards data that are consistently higher or lower than the certified values (Figs. 6C and 7B) indicates an introduced bias, likely due to poor instrument calibration, particularly if there are more than 2 points in a row that fail. A sudden decrease in the randomness of the errors and increase in the accuracy is a sign that data may have been tampered with or the reference materials are being handled differently (Fig. 6D). A continuous shift downward (or upward) of standard values indicates a drift of the results over time or the degradation of a standard over time (Fig. 6E).

Although the above are general interpretations, there are some rigorous rules to follow when monitoring data returned from the lab:

- 1) any single value that lies outside 2 standard deviations of the certified value, or four samples in a row that lie outside one standard deviation

with all either positively or negatively biased, should signal an alert to monitor the data and/or database more closely;

- 2) any time there are two values in a row that lie outside 2 standard deviations or one sample that lies outside 3 standard deviations of the mean of the certified value this should be cause for concern and should signal an audit of the sampling and laboratory methods.

**Choosing Reference Materials: Matrix and Concentration Matching**

The choice of reference materials is dependent on the medium, the elements of interest, the digestion and analytical methods, and the research or industrial question to be answered. There are a number of North American laboratories that provide outstanding reference materials, including:

- **Canadian Certified Reference Materials Program** (<https://www.nrcan.gc.ca/mining-materials/certified-reference-materials/7827>; wide range of reference materials for litho geochemistry, mineralogy, and ore metals).
- **United States Geological Survey** ([http://minerals.cr.usgs.gov/geo\\_chem\\_stand/](http://minerals.cr.usgs.gov/geo_chem_stand/); mostly for litho geochemistry).
- **CDN Resource Laboratories Ltd** (<http://www.cdnlabs.com/>; supplier of ore metal reference materials).

**Table 2.** Best practice and acceptable  $CV_{avg}$  (%) values for various commodities. This table provides an estimate of what is to be expected from certain types of deposits and can be extrapolated to other commodities of interest (e.g. Zn, Pb, ~ Cu).

Mineralization Type/Deposit	Metal	Best Practice	Acceptable Practice	Sample Type
Gold, very coarse grained and nuggetty	Au(g/t)	20(?)	40	Field duplicates
Gold, coarse to medium grained	Au(g/t)	20	30	Field duplicates
	Au(g/t)	10	20	Pulp duplicates
Cu–Mo–Au porphyry	Cu(%)	5	10	Field duplicates
	Mo(%)	10	15	Field duplicates
	Au(g/t)	10	15	Field duplicates
	Cu(%)	3	10	Pulp duplicates
	Mo(%)	5	10	Pulp duplicates
	Au(g/t)	5	10	Pulp duplicates
Iron Ore	Fe(%)	1	3	Field duplicates
	Al <sub>2</sub> O <sub>3</sub> (%)	10	15	Field duplicates
	SiO <sub>2</sub> (wt%)	5	10	Field duplicates
	LOI(wt%)	3	5	Field duplicates
Cu–As–Fe skarn and IOCG	Cu(%)	7.5	15	Field duplicates
	Au(g/t)	15	25	Field duplicates
	Cu(%)	5	10	Pulp duplicates
	Au(g/t)	7.5	15	Pulp duplicates
Ni–Cu–PGE-sulfides	Ni(%)	10	15	Field duplicates
	Cu(%)	10	15	Field duplicates
	PGE	15	30	Field duplicates
	Ni(%)	5	10	Pulp duplicates
	Cu(%)	5	10	Pulp duplicates
	PGE	10	20	Pulp duplicates
Detrital ilmenite sands	Total Heavy Minerals (%)	5	10	Field duplicates

• **Ontario Geoscience Laboratories**

([http://www.mndm.gov.on.ca/sites/default/files/laboratory\\_fees\\_schedule.pdf](http://www.mndm.gov.on.ca/sites/default/files/laboratory_fees_schedule.pdf) ; see page 20 for a number of lithochemical and PGE reference materials).

Any reference materials to be utilized during an analytical program should: 1) have the same matrix as the materials that are being analyzed (e.g.

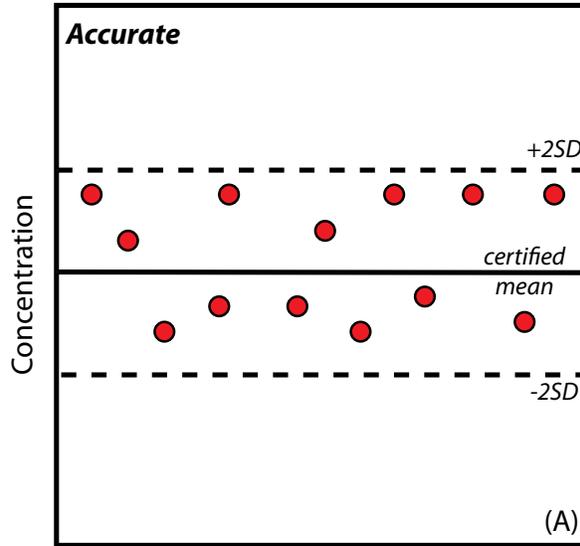
soil, sulfides, concentrates, rock samples); 2) contain the analytes of interest and concentration ranges similar to the materials to be analyzed; and 3) have been prepared in the same way and analyzed by the same methods. Although it is ideal to have them analyzed by the same method, this is not always possible, particularly for reference materials that have had a long shelf life and have been analyzed for

considerable periods of time by multiple methods (e.g. USGS reference materials).

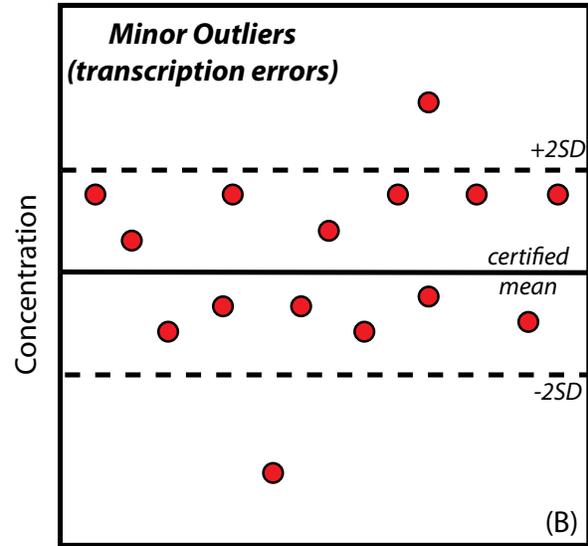
**Creating your Own Reference Materials**

An alternative to purchasing reference materials is to create an internal reference material. To create an internal reference material requires the choice of a material in sufficient quantities

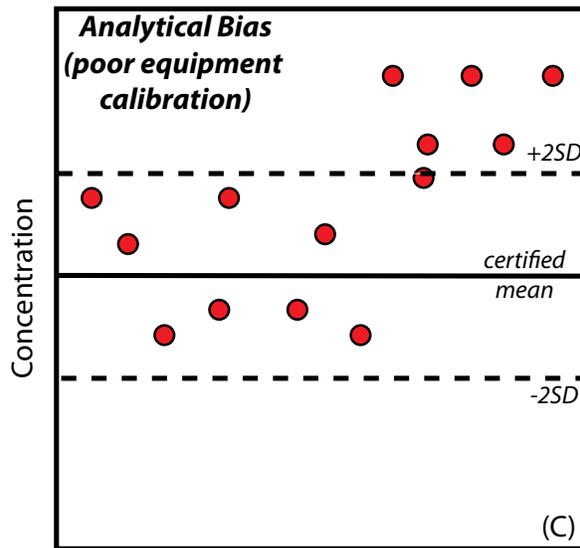
**Figure 6.** (opposite page) Schematic Shewart control charts illustrating the various patterns that can be obtained during the QA/QC monitoring program. A) Accurate data lying within two standard deviations of the mean with samples randomly distributed above and below the certified mean value. B) Random outliers indicate a potential data transcription error during sample preparation or database management. Alternatively, they may suggest random errors. C) Definite analytical bias indicating potentially poor calibration of instrumentation or a change in instrumentation, sampling, or preparation methodology. D) A rapid decrease in the magnitude of errors during the QA/QC program illustrates potential data tampering or different sample handling and data treatment. E) Systematic decrease (or increase) in values indicates potential drift and/or degradation of the reference material during the course of the QA/QC program. Diagram modified from Abzalov (2008).



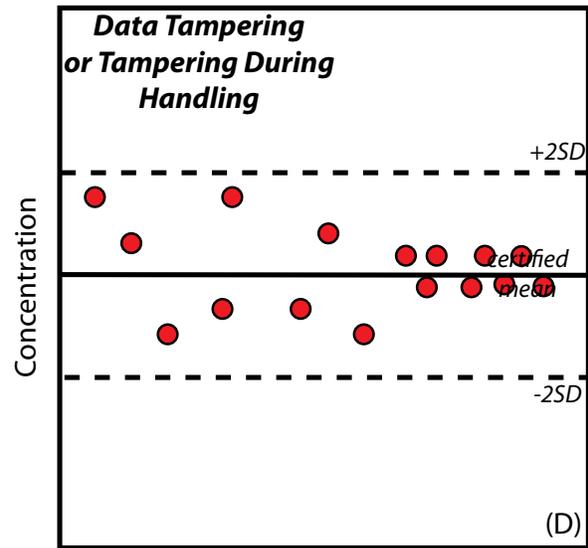
Order of Standard Analysis



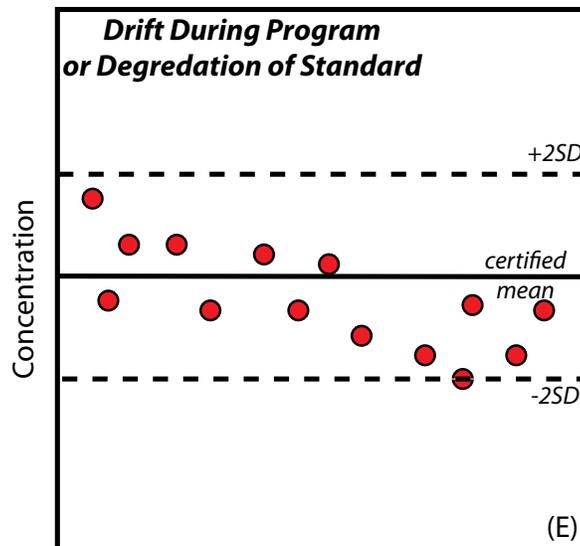
Order of Standard Analysis



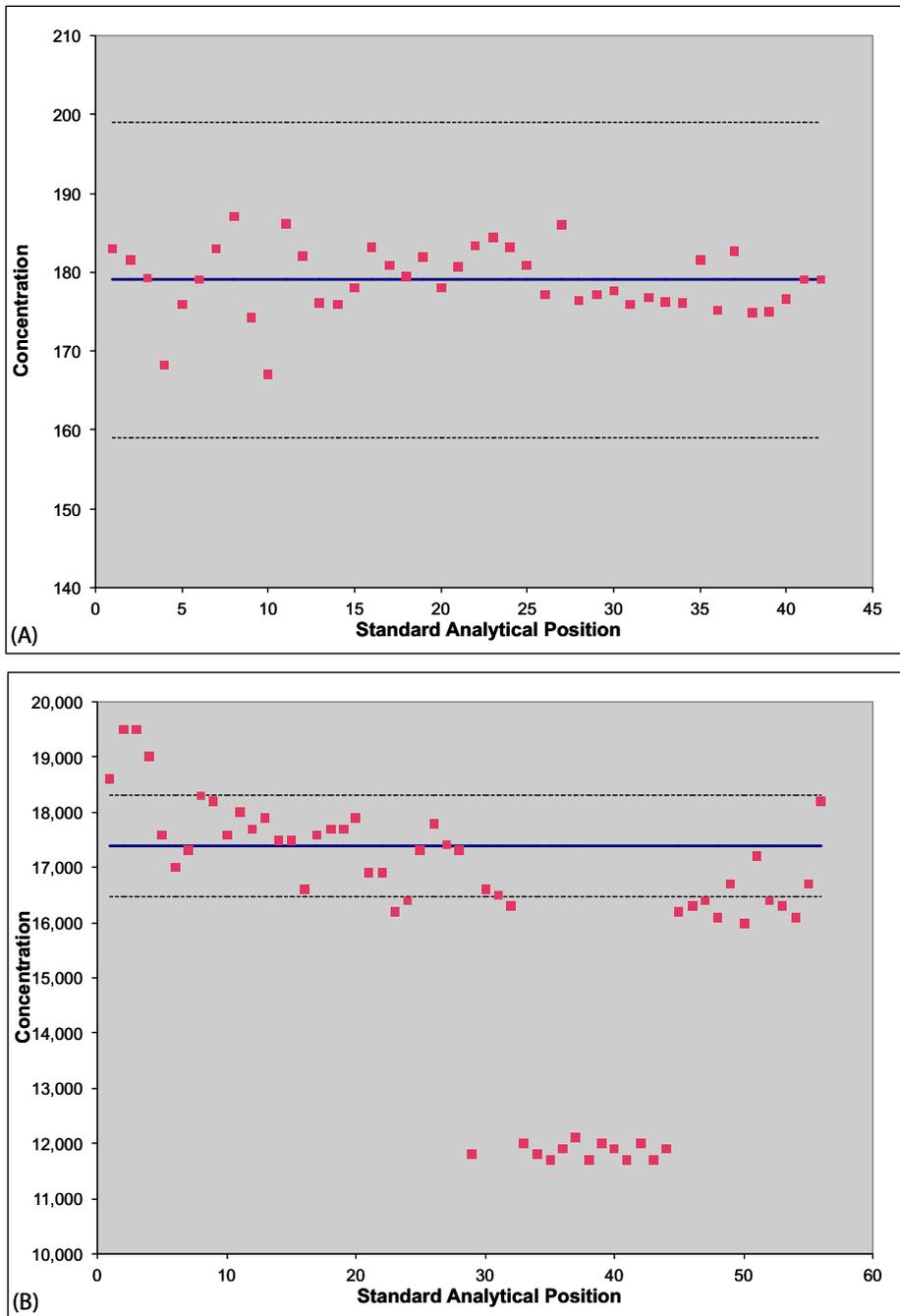
Order of Standard Analysis



Order of Standard Analysis



Order of Standard Analysis



**Figure 7.** Shewart control charts for testing accuracy of reference material utilizing real data. A) Zirconium from a commercial facility. Note the relatively random distribution of the data above and below the mean value and with all samples within 2 standard deviations of the certified value. This dataset is an example of good accuracy. B) Zinc from at a commercial facility. Note the highly erratic data with many samples lying outside of 2 standard deviations of the mean. Also note the periods of overestimation of results early in the program and severe underestimation later in the program. The latter underestimation signals a potential database transcription error. The fact that many points lie outside of the expected values for the reference materials suggests that the reliability of the data is poor, and should trigger sampling, laboratory, and database audits. This dataset is an example of poor accuracy.

(e.g. 50 kg or more), having this material homogenized, and then analyzed as part of a round robin testing program. A round robin testing program requires (Smee 2007): 1) a minimum of 60 analyses; and 2) the analyses must be obtained at a minimum of 6 different laboratories. It is highly recommended that more analyses be undertaken when creating internal reference materials, and that the values for internal reference material be continuously refined with increasing data. Some laboratories and facilities will produce reference materials for clients (e.g. Ontario Geoscience Laboratories; CDN Resources Laboratories).

### Blanks and Their Utilization

Blanks are used to test for contamination introduced during sample preparation and analysis. Contamination can occur at any stage during the sample preparation and analytical process, including contamination due to poor cleaning of crushing and pulverizing equipment, from unclean acids during sample preparation, or memory effects on instrumentation where the instruments are not sufficiently flushed with solution between analyses. A blank is a material that contains nil to extremely low concentrations of the element(s) of interest. In mineral exploration and development, blanks typically consist of relatively inert rocks such as quartzite, granite, or volcanic rocks, and some groups use coarse, clean sand (e.g. sandbox or gardening sand found at many building supply stores), as these materials do not typically contain base or precious metals. In litho-geochemistry, clean quartzite or other similar materials can be used. In some quartzite samples, however, there are often anomalous heavy minerals (e.g. monazite, zircon) and this can lead to erroneous values for HFSE and REE. Nevertheless, clean, pure quartzite is often a useful material for a trace-element field blank.

### Monitoring Contamination Using Shewart Control Charts

The Shewart control chart for blanks is similar to that for monitoring accuracy using reference materials. On the X-axis of the plot are the sequential analyses of the blank and on the Y-axis are the results obtained for the element

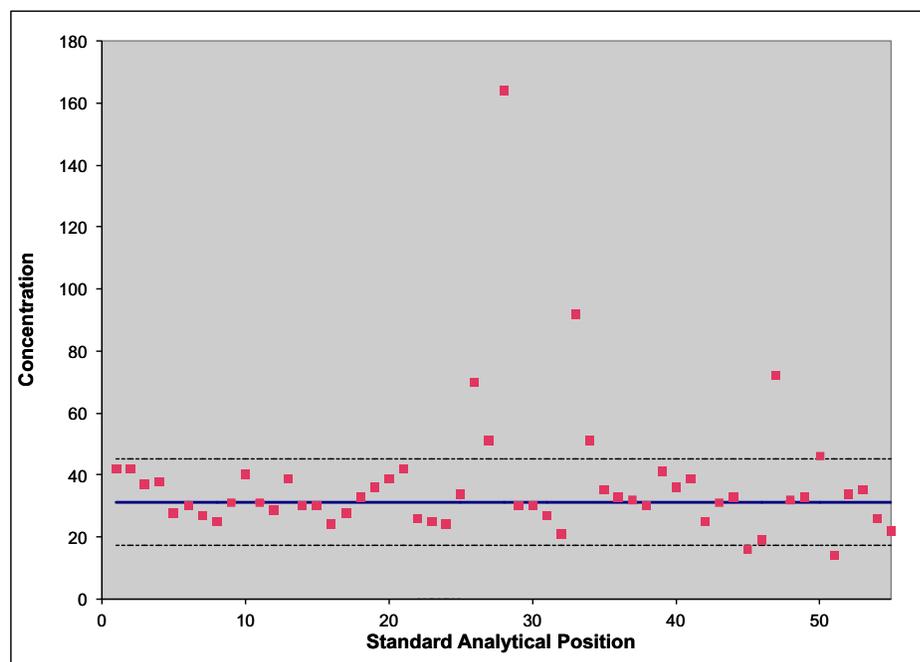
of interest (Fig. 8). Initially, the mean and standard deviation of the blank will be unknown, but after sufficient analyses an ongoing set of values will be constructed. When establishing the mean and standard deviation it is critical to not include the outliers in the dataset.

Control charts for the blanks are evaluated in the same way as data accuracy is monitored (Fig. 8). Either of the following cases should signal an alert to monitor the data and/or database more closely, and to audit the sampling and lab methods: 1) any single value lies outside 2 standard deviations of the certified value, or four in a row lie outside 1 standard deviation of the mean of the long-term accepted value; and 2) one value lies outside 3 standard deviations of the mean of the long-term accepted value.

### Frequency of Duplicate, Reference Materials, and Blank Insertion

The frequency of duplicate, reference material, and blank insertion depends on the nature of the research or industrial problem for which the data will be used. One should always remember that the data to be obtained will be used for future decision making. In both the industrial and research environments it is important to have regular insertion of blanks, duplicates, and reference materials to provide readers with an estimate of data quality and uncertainty in the data (Potts 2012).

Numerous estimates for the number of insertions of reference material, blank, and duplicate frequency have been proposed. Garrett (1969) suggested a frequency of 10%, whereas Sketchley (1998) suggested a frequency of 20%. More recently, Abzalov (2008) summarized various opinions and suggested a frequency of 5–10% for field duplicates and duplicates of the coarse rejects; 3–5% of the samples should be pulp duplicates, reference materials, and blanks; and ~5% of the duplicate samples should be processed and analyzed at an external laboratory. As a general rule for any analytical project the frequency for blank, duplicates, and reference materials should be 5–20%. Furthermore, it is important that smaller datasets have more quality control measures to ensure robust QA/QC statistics and



**Figure 8.** Shewart control chart for Zn content in a blank, showing results from a commercial facility. Notably, most data lie within the long-term accepted values for the blank. The outliers indicate potential contamination or database transcription errors at various points during the analytical program.

metrics of sampling, sample preparation, and laboratory performance.

### CONCLUSION

Any project, be it an industry project or a research program, hinges on the utilization of high quality data. In order to ensure confidence in reported data, a well-planned QA/QC program is recommended, utilizing duplicate samples, reference materials, and blanks inserted in a systematic fashion. Both field and coarse reject material should be used to test sampling precision and the early stages of sample preparation, and pulp duplicates should be utilized to test the analytical precision of the laboratory. Reference materials are utilized to monitor precision and accuracy, and blanks are inserted to monitor contamination. Various methods for monitoring precision are provided, but the most robust methods are utilization of duplicate scatterplots, statistical tests (e.g. %RSD), Thompson-Howarth plots, and the average coefficient of variation ( $CV_{avg}(\%)$ ). Accuracy is tested using statistical tests (t-test), relative differences compared to the reference materials, and Shewart control charts; blanks are also monitored using She-

wart control charts. Accuracy is determined chiefly by using recognized standards of the same composition and character as the unknowns, whereas blanks are evaluated by repeated analyses of material lacking the elements of interest. For both accuracy and blanks, having more than two values outside of two standard deviations of the mean, one sample three standard deviations from the mean, or four values in a row outside of one standard deviation of the mean of the certified value for an element (reference materials) or the long-term mean blank value, should instigate a QA/QC audit. Based on industry and research standards, it is recommended that the QA/QC program have between 5–20% duplicates of various types (e.g. field, coarse reject, pulp), blanks, and reference materials, and roughly 5% of the samples should be processed and analyzed in an external laboratory.

### ACKNOWLEDGEMENTS

I would like to thank numerous colleagues and students that have shaped my interest and knowledge of QA/QC and analytical geochemistry, including but not limited to: Kiel Arndt, K. Fletcher, Stu Hamilton, George Jenner,

Owen Lavin, Henry Longrich, and numerous industry colleagues who have shared their insight into the day-to-day application of QA/QC to the real world. Bob Linnen is thanked for encouraging me to submit this manuscript to *Geoscience Canada*. Two anonymous reviewers are thanked for comments and suggestions that have greatly improved the manuscript. This research was funded by an NSERC Discovery Grant; the NSERC–Altius Industrial Research Chair in Mineral Deposits at Memorial University funded by NSERC, Altius Resources Inc., the Research and Development Corporation of Newfoundland and Labrador; and the Canadian Mining Innovation Council (CMIC) Footprints Project, which is funded by CMIC and an NSERC Collaborative Research and Development Grant.

## REFERENCES

- Abzalov, M., 2008, Quality Control of Assay Data: A Review of Procedures for Measuring and Monitoring Precision and Accuracy: *Exploration and Mining Geology*, v. 17, p. 131–144, <http://dx.doi.org/10.2113/gsemg.17.3-4.131>.
- Abzalov, M., 2011, Sampling Errors and Control of Assay Data Quality in *Exploration and Mining Geology*, in Ivanov, O., *ed.*, *Applications and Experiences of Quality Control*: InTech Open Access Publishing, Rijeka, Croatia, p. 611–644, <http://dx.doi.org/10.5772/14965>.
- CANMET, 1998, Assessment of laboratory performance with certified reference materials: CANMET Canadian Certified Reference Materials Project Bulletin: Natural Resources Canada, Ottawa, ON, p. 5.
- Caughlin, B.L., 2010, Developments in analytical technology: *Geochemistry: Exploration, Environment, Analysis*, v. 10, p. 137–141, <http://dx.doi.org/10.1144/1467-7873/09-214>.
- Croakin, C., and Tobias, P., 2006, NIST/SEMATECH e-Handbook of Statistical Methods, National Institute of Standards and Technology/SEMATECH, (e-book, last updated April 2012): US Commerce Department's Technology Administration, <http://www.itl.nist.gov/div898/handbook/>.
- Davis, J.C., 2002, *Statistics and Data Analysis in Geology*, Third Edition: John Wiley and Sons, New York, 656 p.
- Fletcher, W.K., 1981, *Analytical methods in geochemical prospecting*: Elsevier, Amsterdam, 255 p.
- Fletcher, W.K., 1986, Analysis of soil samples, in Robertson, J. M., *ed.*, *Exploration geochemistry; design and interpretation of soil surveys: Reviews in Economic Geology Volume 3*, Society of Economic Geologists, p. 79–96.
- Fletcher, W.K., 2002, Quality control in exploration geochemistry, Mineral Exploration Research Centre, Unpublished Short Course Notes: Laurentian University, Sudbury, ON, p. 67.
- Garrett, R.G., 1969, The determination of sampling and analytical errors in exploration geochemistry: *Economic Geology*, v. 64, p. 568–569, <http://dx.doi.org/10.2113/gsecongeo.64.5.568>.
- Govindaraju, K., 1994, 1994 compilation of working values and sample description for 383 geostandards: *Geostandards Newsletter*, v. 18, p. 1–158, <http://dx.doi.org/10.1046/j.1365-2494.1998.53202081.x-i1>.
- ISO. International Organization for Standardization, 1989, *Guide 33: Uses of certified materials*, 1st Edition: Standards Council of Canada, Ottawa, ON, p. 12.
- Jenner, G.A., 1996, Trace element geochemistry of igneous rocks: *Geochemical nomenclature and analytical geochemistry*, in Wyman, D.A., *ed.*, *Trace Element Geochemistry of Volcanic Rocks: Applications for Massive Sulfide Exploration*, 12, Geological Association of Canada, Short Course Notes, p. 51–77.
- Potts, P.J., 1987, *A Handbook of Silicate Rock Analysis*: Blackie, Glasgow, UK, 622 p.
- Potts, P.J., 2012, A Proposal for the Publication of Geochemical Data in the Scientific Literature: *Geostandards and Geoanalytical Research*, v. 36, p. 225–230, <http://dx.doi.org/10.1111/j.1751-908X.2011.00121.x>.
- Sketchley, D.A., 1998, Gold deposits; establishing sampling protocols and monitoring quality control: *Exploration and Mining Geology*, v. 7, p. 129–138.
- Smee, B.W., 2007, Quality Control in Mineral Exploration: Controlling the quality of information from field to data base, in Coker, W.B., *ed.*, *Exploration 07, Short Course - Exploration Geochemistry: Basic Principles and Concepts*: Toronto, ON, Canada, [electronic short course notes].
- Stanley, C.R., 2003a, Estimating sampling errors for major and trace elements in geological materials using a propagation of variance approach: *Geochemistry: Exploration, Environment, Analysis*, v. 3, p. 169–178, <http://dx.doi.org/10.1144/1467-7873/03-008>.
- Stanley, C.R., 2003b, THPLOT.M: A MATLAB function to implement generalized Thompson–Howarth error analysis using replicate data: *Computers and Geosciences*, v. 29, p. 225–237, [http://dx.doi.org/10.1016/S0098-3004\(02\)00072-9](http://dx.doi.org/10.1016/S0098-3004(02)00072-9).
- Stanley, C.R., 2006, On the special application of Thompson–Howarth error analysis to geochemical variables exhibiting a nugget effect: *Geochemistry: Exploration, Environment, Analysis*, v. 6, p. 357–368, <http://dx.doi.org/10.1144/1467-7873/06-111>.
- Stanley, C.R., 2007, The fundamental relationship between sample mass and sampling variance in real geological samples and corresponding statistical models: *Exploration and Mining Geology*, v. 16, p. 109–123, <http://dx.doi.org/10.2113/gsemg.16.1-2.109>.
- Stanley, C.R., and Lawie, D., 2007, Average relative error in geochemical determinations: Clarification, calculation, and a plea for consistency: *Exploration and Mining Geology*, v. 16, p. 267–275, <http://dx.doi.org/10.2113/gsemg.16.3-4.267>.
- Sylvester, P., *ed.*, 2001, *Laser-ablation–ICPMS in the Earth Sciences: Principles and applications*: Mineralogical Association of Canada, Short Course Series, v. 29, 243 p.
- Thompson, M., and Howarth, R.J., 1978, A new approach to the estimation of analytical precision: *Journal of Geochemical Exploration*, v. 9, p. 23–30, [http://dx.doi.org/10.1016/0375-6742\(78\)90035-3](http://dx.doi.org/10.1016/0375-6742(78)90035-3).
- Vallee, M.A., 1998, Sampling quality control: *Exploration and Mining Geology*, v. 7, p. 107–116.

Received August 2013

Accepted August 2013

First published on the web

January 2014