

Statistical methods in geochemical analysis: an important tool for evaluation and quality control of analytical methods and results

ANDREAS M. GRIMSTVEDT

Andreas M. Grimstvedt, Norges geologiske undersøkelse, P.O.Box 3006-Lade, 7002 Trondheim, Norway

Introduction

Statistical methods are important tools for evaluation and quality control of analytical methods and results. Typical applications are determination of central terms like precision, accuracy and detection limits, which define the quality of a method. Other typical applications include the control of the analytical process by means of control charts, and the improvement of existing methods. For a chemist working in the laboratory such knowledge is of vital importance. However, it is also important for users of the results to understand the meaning of the different terms and quantities. This will give them a better background for the choice of analytical methods and also for the treatment of the results. Laboratories have an important task in the presentation of analytical methods as well as results.

Central terms and quantities

Like any chemical analysis, geochemical analyses are also subject to errors. Errors can occur in most of the steps in an analytical process, from the sample weighing to reporting of the final result. Errors are usually classified as random, systematic or gross errors (Miller & Miller 1988). The systematic errors affect the accuracy of the results while the random errors influence the precision. The random errors are usually assumed to be normal distributed. This is a central factor in the statistical treatment of data. Central terms in describing an analytical method are the determination of detection limits, precision and accuracy.

The purpose of using detection limits is to avoid reporting results caused by background, e.g. instrument noise, signal caused by sample holder, solvent, etc. A detection limit is the lower limit where a signal can be stated with a certain probability to be larger than the background signal. The most common definition of LLD (Lower Limit of Detection) is signals that are equal to the 99.87% (3σ) confidence-level of the background signal (Kateman & Pijpers 1981, Miller & Miller 1988). The most common method to determine

LLD is to analyse replicate blank samples, and from these results calculate the mean and standard deviations. The detection limit is given by:

$$X_{\text{LLD}} = \bar{X}_0 + 3\sigma \quad [1]$$

In eq.[1] X_{LLD} , is the detection limit, \bar{X}_0 is the mean value of the blank/background, and σ is the standard deviation. The concentration level of the LLD is also dependent on the sensitivity of the actual method. A related term used in the literature (Kateman & Pijpers 1981, Miller & Miller 1988) is the lower limit of determination, which is the lower limit for quantitative determination, usually given by:

$$X_{\text{LLD}\cdot} = \bar{X}_0 + 10\sigma \quad [2]$$

Precision is an important criterion for the quality of an analytical method. Precision describes the spread of results. Two terms that are directly related to precision are repeatability (within-run precision) and reproducibility (between-run precision). The precision of a method is usually determined from repeated measurements of a sample, and presented as the standard deviation of the results.

Accuracy is usually defined as the difference between obtained value and true value. The concept of accuracy, however, is one of the most difficult topics in analytical chemistry. There are no standard methods for calculation, and the interpretation is highly dependent on the analytical methods used (Kateman & Pijpers 1981). Most analytical methods include some kind of calibration technique, i.e. relating signals to known concentrations (standards). This often involves fitting data to a straight line by means of the principle of least squares. In its most common use, this is based on the assumption that all responses are distributed with the same uncertainty (variance), and is denoted as an unweighted regression model. In many cases, however, the uncertainties will be dependent on the concentration level, and on uncertainties in the standard

used (Rollinson 1993). In these cases, the use of a weighted regression model will be a better choice. For more information about weighted regression, see Montgomery & Peck (1982), Hwang & Winefordner (1988) and Miller & Miller (1988). The accuracy is directly related to the calibration.

An important use of statistics is in the evaluation and improvement of analytical methods. Typical applications are tests for systematic errors, the effects of different treatments, calibration, magnitude of different sources of error, etc. Central methods are hypothesis testing, variance analysis and regression analysis (Kateman & Pijpers 1981, Montgomery & Peck 1982, Montgomery 1984, Miller & Miller 1988). To enable proper use of these techniques it is necessary to have a good understanding of basic statistical principles. Another useful application is in controlling the analytical process by means of control charts (Kateman & Pijpers 1981, Montgomery 1985, Miller & Miller 1988). The most commonly used in analytical chemistry are the so-called X-charts (Shewart control charts), which give a direct control of the precision. This is based on preparing and measuring a control sample frequently. The results are plotted in a chart which contains lines for mean value, warning-limits (mean ± 2 standard deviations) and action-limits (mean ± 3 standard deviations). If results fall outside these limits, certain procedures are followed to maintain the precision of the process (Grimstvedt 1994).

Examples of analytical results, precision and accuracy

At the Geological Survey of Norway, analyses of precious metals in geochemical materials have for a long time been sent to external laboratories. With these series, some international standards

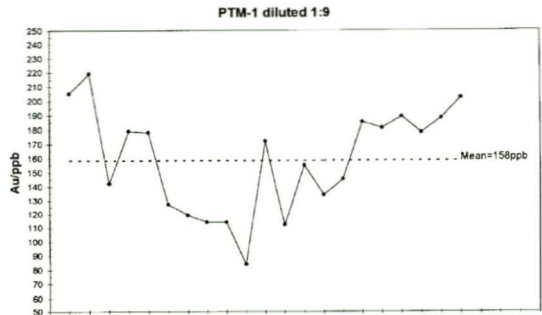


Fig. 2. Au analysis of standard PTM-1 (diluted 1:9, recom.value = 180 ppb).

have been included. The results of these analyses contain useful information, illustrating long-term precision and accuracy. Results for two standards, PTM-1 and PTC-1, are presented here. A detailed description of these standards has been given by Mcadam et al. (1973a,b). An evaluation of the specific analytical method or the actual laboratory is beyond the scope of this work. However, it can be mentioned that the reported detection limits for Au, Pt and Pd are below 10 ppb. Both standards have been diluted 1:9 with olivine, to reduce the concentrations to a level nearer that of the unknown samples. The dilution error (weighing error) has been estimated and found not to be significant compared with the total error. The results for Au are plotted in Figs.1 and 2, and all results are summarised in Tables 1 and 2.

The accuracy of the results can be tested by a hypothesis test, i.e. testing the hypothesis:

$$H_0 : X = \mu \text{ versus } H_1 : X \neq \mu \quad [3]$$

In eq.[3] H_0 is the null hypothesis, i.e. the results are equal to the recommended value, μ is the recommended value, and x is the analysed value. Neglecting the uncertainties in the recommended value, and assuming that all X_i values are independent and normal distributed with the same unknown standard deviation, the test criteria can be based on the student-t distribution (Høyland 1986). This results in the rejection of H_0 if:

$$\bar{X} - \mu \geq t_{\alpha/2, v} \cdot \frac{\hat{\sigma}}{\sqrt{n}} \quad [4]$$

Where $t_{\alpha/2, v}$ is the $\alpha/2$ percentage points of the student-t distribution with $v=(n-1)$ degrees of freedom, σ is the estimated standard deviation, and \bar{X} is the mean value.

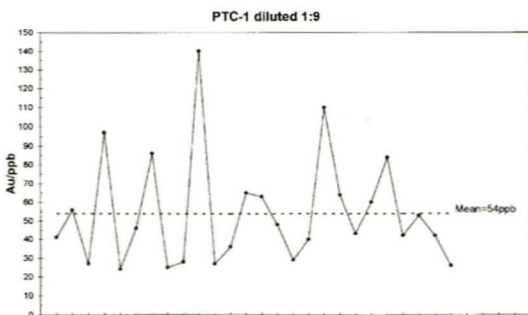


Fig. 1. Au analysis of standard PTC-1 (diluted 1:9, recom.value = 65 ppb).

At a 95% confidence-level, the null hypothesis is rejected for Pd in PTC-1, and Au and Pt in PTM-1 (at 90%-level H_0 is rejected for all elements except Pd in PTM-1). The relative standard deviation ($100 \cdot \frac{\hat{\sigma}}{\bar{X}}$) is in the range 20-50% for PTC-1, and approximately 20% for all elements in PTM-1. In statistical terms these results are imprecise and mainly inaccurate. These results are obtained over a long period, and the precision will normally be much better for shorter time intervals. The results for Au show a typical normal distributed behaviour (see Figs.1 & 2), while the calculated results for Pt and Pd could be highly influenced by outliers. Similar tendencies can be expected for many other analytical methods. The uncertainties should be taken into consideration, especially when comparing results obtained over a long time period.

Discussion

Documentation of accuracy is in many cases limited by the number of good, available, certified reference materials (standards). Standards used for geochemical materials are usually so-called secondary standards. Standards are particularly a problem in techniques that involve extraction in acids, because the contents are usually reported as a total. Accuracy of a method is frequently illustrated by reporting obtained and recommended values for a secondary standard. This can often be misleading, because the accuracy is dependent on the concentration level and the sample type (matrix). An extreme case will be if this standard is used alone for calibration. Therefore, in

techniques based on single standard calibration, the accuracy should always be tested with an independent standard. In geological applications, however, precision is usually more important than accuracy (Rollinson 1993). Accuracy can be very important in some cases, e.g. in environmental analysis where results should be compared with certain limits; and whether the results are above or below these limits could have major economical consequences. The laboratories can improve and control the precision by use of control charts. The users can also improve their results by including control samples in their series (Reimann 1989). This can be particularly effective when different laboratories or even different methods are used. Results for the control sample can, for example, be used to normalise results relative to the control sample. In cases where within-run precision is important, such as trend analysis within a series, the use of duplicates can increase the reliability of the results.

Detection limits actually mean that above these limits an element or compound is found significant. However, it does not say anything about quantification or uncertainties. It is important to note that detection limits are usually highly dependent on the sample type. Reported detection limits from commercial laboratories as well as instrument suppliers are commonly based on optimal conditions, and practical detection limits in routine analysis can often be higher. This is in agreement with observations made by Reimann (1989). Generally, methods with detection limits close to expected concentrations should not be used. The choice of method should always be based on the purpose of the investigations.

Table 1. Analysis of standard PTC-1 over a time period of four years.

The standard has been diluted 1:9. The mean value is \bar{X} , $\hat{\sigma}$ is the calculated standard deviation, n is the number of observations, and μ is the recommended value.

Element	\bar{X} /ppb	Max/ppb	Min/ppb	Median/ppb	$\hat{\sigma}$ /ppb	n	μ
Au	54	140	24	44.5	29	26	65
Pt	261	483	181	236	88	18	300
Pd	1140	1610	750	1151.5	220	18	1270

Table 2. Analysis of standard PTM-1 over a time period of four years.

The standard has been diluted 1:9. The mean value is \bar{X} , $\hat{\sigma}$ is the calculated standard deviation, n is the number of observations, and μ is the recommended value.

Element	\bar{X} /ppb	Max/ppb	Min/ppb	Median/ppb	$\hat{\sigma}$ /ppb	n	μ
Au	158	219	84	172	37	21	180
Pt	452	578	295	463	104	13	580
Pd	749	917	489	804	152	13	810

Conclusion

The application of statistics to geochemical analysis provides an important tool for decision making both for the chemist working in the laboratory and for the geologist or geochemist in the further use of analytical results. The users of analytical results cannot take full advantage of these statistical methods if there is a lack of information from the laboratory, or if the users do not have a good understanding of the uncertainties in the analytical methods and results. Thus, a good collaboration between the laboratory and the customers is important.

Acknowledgements

Special thanks to David Segar for help with the English presentation of the manuscript. Also, thanks to Kristian Bjerkli and Magne Ødegård for valuable discussions during the preparation of the manuscript.

References

- Grimstvedt, A.M. 1994: *Veiledning og rutiner ved statistisk behandling av data. NGU-Labs Kvalitetssystem, Felles Dokumenter, Geological Survey of Norway, Division for Test Laboratories, Trondheim, Norway* (unpublished).
- Hwang, J.D. & Winefordner J.D. 1988: Regression methods in analytical chemistry. *Progr. Analyt. Spectrosc.* 11, 209-249.
- Høyland, A. 1986: *Sannsynlighetsregning og statistisk metode-lære*, 4. utgave, TAPIR, Trondheim, 253pp.
- Kateman, G. & Pijpers, F.w. 1981: *Quality control in analytical chemistry, Chemical Analysis*, 60. John Wiley, New York, 276pp.
- Mcadam, R.C., Sutarno & Moloughney, P.E. 1973a: Noble-metals-bearing sulphide concentrate PTC: Its characterization and preparation for use as a standard reference material, *Technical Bulletin 176, Department of Energy; Mines and resources Mines Branch Ottawa, Canada*, 1-26.
- Mcadam, R.C., Sutarno & Moloughney, P.E. 1973b: Noble-metals-bearing nickel-copper matte PTM: Its characterization and preparation for use as a standard reference material. *Technical Bulletin 182, Department of Energy; Mines and resources Mines Branch Ottawa, Canada*, 1-33.
- Miller, J.C. & Miller, J.N. 1988: *Statistics for analytical chemistry*, 2nd. ed, John Wiley, New York, 227pp.
- Montgomery, D.C. 1984: *Design and analysis of experiments*. 2nd. ed., John Wiley, New York, 538pp.
- Montgomery, D.C. 1985: *Introduction to Statistical Quality Control*. John Wiley, New York, 520pp.
- Montgomery, D.C. & Peck, E.A. 1982: *Introduction to linear regression analysis*. John Wiley, New York, 504pp.
- Reimann, C. 1989: Reliability of geochemical analysis: recent experiences. *Trans. Instn Min. Metall (Sect. B: Appl. earth sci.)*, 98, 123-129.
- Rollinson, H.R. 1993: *Using Geochemical Data: Evaluation, Presentation, Interpretation*. 1st. ed., Longman Scientific & Technical, 352pp.