

Quantitative estimates of precision for molecular isotopic measurements

John P. Jasper*

Molecular Isotope Technologies, LLC, 8 Old Oak Lane, Niantic, CT 06357-1815, USA

Received 5 July 2001; Accepted 5 July 2001

SPONSOR REFEREE: John M. Hayes, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA

At least three methods of calculating the random errors or variance of molecular isotopic data are presently in use. The major components of variance are differentiated and quantified here into least three to four individual components. The measurement of error of the analyte relative to a working (whether an internal or an external) standard is quantified via the statistical pooled estimate of error. A statistical method for calculating the total variance associated with the difference of two individual isotopic compositions from two isotope laboratories is given, including the variances of the laboratory (secondary) and working standards, as well as those of the analytes. An abbreviated method for estimation of error typical for chromatographic/isotope mass spectrometric methods is also presented. Copyright © 2001 John Wiley & Sons, Ltd.

Molecular isotopic analysis is now employed or is being evaluated in a variety of fields—from reconstructing ancient environmental conditions in organic geochemistry^{1,2} through wide-ranging applications³ to isotopic integrity of pharmaceuticals.⁴ The analytical measurements that underpin these applications are of little use without some quantitative estimate of random errors and understanding of potential systematic errors. Merritt and Hayes⁵ address in detail many matters of accuracy and precision in molecular isotope mass spectrometry. Particular problems of accuracy discussed include the background correction necessary for data production and establishment of the non-coelution of standard peaks. Beyond the problem of peak definition,⁵ there are at least three to four differentiable sources of variance that contribute to the total variance in any given molecular isotopic measurement. In addition, pooled estimates of error derived from large data sets^{6,7} permit more statistically relevant estimates of error than would calculations made on small numbers of replicates (e.g., 2, 3, 4 data points). In many practical cases, multiple sets of replicates (e.g., triplicates) are available. Information from these data sets can be pooled to provide improved error estimates. Here, we delineate the major sources of variance that contribute to the total variance of a given sample and focus on the pooled estimates of error relevant to the isotopic measurement of an analyte relative to its proximal (working) standard. We also point out the existence of systematic, typically incremental, errors caused by changes of standards, though this matter is not fully developed here.

*Correspondence to: J. P. Jasper, Molecular Isotope Technologies LLC, 8 Old Oak Lane, Niantic, CT 06357-1815, USA.
E-mail: JPJasper@MolecularIsotopes.com
Contract/grant sponsor: N.S.F. Ocean Sciences.
Contract/grant sponsor: N.A.S.A.

THEORY AND PRACTICE

Estimates of precision: a conceptual decomposition of variance

The total error (σ_{tot}) in the isotopic composition of a given molecular isotopic peak is typically a function of at least three or four error terms ($\sigma_1 \dots \sigma_n$) representing the differences of four or five measured isotopic compositions (Fig. 1). Isotopic compositions are typically reported relative to that of a primary standard such as the International Atomic Energy Authority's (IAEA) Vienna PeeDee Belemnite (VPDB) carbonate standard. A typical secondary standard used for carbon isotope measurements in many laboratories is the graphitic National Bureau of Standards-20 standard or NBS-20. In most molecular isotopic measurements, a $\text{CO}_2(\text{g})$ rectangular wave is used as an external standard. The CO_2 is injected into the gas stream flowing into the isotope ratio mass spectrometer (irmMS) and thus appears as another peak in the chromatogram. Finally, there is the analyte peak itself whose isotopic composition is defined relative to some or all of the preceding standards, depending on the use of external or internal standards.

The error terms ($\sigma_1 \dots \sigma_n$) are defined as follows with typical range compositions given by:

$$\sigma_1 = \delta_{\text{VPDB}} - \delta_{\text{NBS}} (\sim 0.02\text{‰})$$

$$\sigma_2 = \delta_{\text{NBS}} - \delta_{\text{XS}} (= 0.01\text{--}0.1\text{‰})$$

$$\sigma_3 = \delta_{\text{XS}} - \delta_{\text{IS}} (= 0.1\text{--}0.3\text{‰})$$

$$\sigma_4 = \delta_{\text{IS}} - \delta_{\text{A}} (= 0.10\text{--}0.3\text{‰}), \text{ and}$$

$$\sigma_5 = \delta_{\text{XS}} - \delta_{\text{A}} (= 0.1\text{--}0.3\text{‰})$$

where δ_i = the $\delta^{13}\text{C}$ of i , and i represents either the VPDB (primary) standard, the NBS (secondary) standard, an external standard (XS), an internal standard (IS), or the analyte (A). For reference, carbon isotopic results are typically expressed as δ -values (parts per thousand differ-

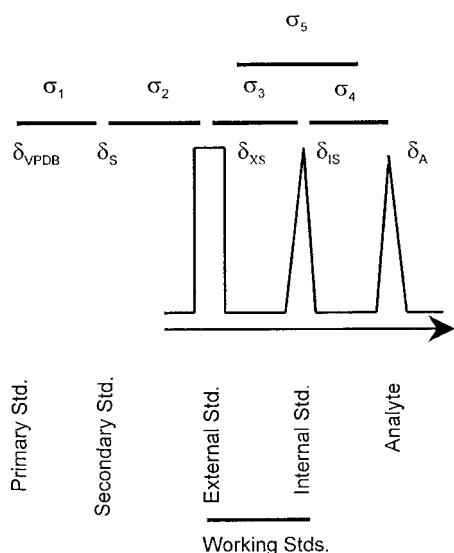


Figure 1. Diagram of the individual variances (σ_{1-5}) that constitute the total variance (σ_{tot}) of a molecular isotopic measurement. The symbols δ_i represent the stable carbon isotopic composition ($\delta^{13}\text{C}$) of the primary (e.g., VPDB) standard, secondary (e.g., NBS) standard, the working [external (XS) or internal (IS)] standard(s), or the analyte (A). The alternative variances (σ_{3+4} versus σ_5) represent alternative pathways for using different proximal (i.e., internal versus external) standards. Typically, the variance results are given relative to the proximal standards.

ences from a standard) defined as:

$$\delta^{13}\text{C}(\text{‰}) \equiv 1000[(R_{\text{smp}}/R_{\text{std}}) - 1] \quad (1)$$

where $R =$ the $^{13}\text{C}/^{12}\text{C}$ ratio and the subscripts denote the sample and the isotopic standard. Most commonly, and practically in all cases where comparisons between laboratories are involved, the standard is VPDB, which defines the zero point on the δ scale of carbon isotopic abundances.

Thus, the total error of molecular isotopic measurements can be expressed as errors added in quadrature in one of the following ways depending whether or not the internal standard is employed. In the first case, the internal standard is employed:

$$\sigma_{\text{tot}}(\text{IS}) = (\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2)^{0.5} \quad (2)$$

Using the error ranges given above, typical internal standard compositions ($\sigma_{\text{tot}}(\text{IS})$) span from 0.14 to 0.44‰. Individual compositions may vary depending on analytical conditions and chromatogram complexity.

In the second case, an internal standard is not used, but an external standard is:

$$\sigma_{\text{tot}}(\text{XS}) = (\sigma_1^2 + \sigma_2^2 + \sigma_5^2)^{0.5} \quad (3)$$

Similar to the preceding case, typical external standard compositions ($\sigma_{\text{tot}}(\text{XS})$) span from 0.10 to 0.32‰, with the same caveats as given for internal standards.

While the total error ranges given above are characteristic for the preceding formulation of total variance, in many

cases investigators report only the proximal error estimation (i.e., the estimation of error relative to either the internal standard (σ_4) or the external standard (σ_5)). When comparisons between samples are being made, it is only necessary to track errors back to the nearest common standard. Within a given laboratory or project, therefore, estimation of either σ_4 or σ_5 can be fully adequate. For comparisons between laboratories or over long intervals of time, σ_{tot} is usually pertinent.

Pooled estimates of error: the statistical constitution of error

The common expression for variance in a data set is

$$\sigma^2 = \Sigma d^2 / (n - 1) \quad (4)$$

where d is the difference between any given observation and the mean of all observations and n is the total number of observations. More generally, when a single procedure has been used to obtain sets of replicate analyses from multiple samples, and when σ can be assumed to be a characteristic of the procedure rather than of each sample,⁶ we can write:

$$\sigma_p^2 = \Sigma d_i^2 / \Sigma (n_i - 1) \quad (5)$$

where σ_p refers to a 'pooled standard deviation' based on i sets of replicates. In pooling, values of d_i , the difference between each observation and its corresponding mean, are summed across all data sets. The denominator is a summation of the number of 'degrees of freedom', which is equal to the total number of observations minus the number of means.

In the present application, we can imagine that four similar samples have been analyzed using the same method, sample **a** in triplicate, **b** in duplicate, **c** in quadruplicate, and sample **d** only once. Using standard methods, we could calculate σ_a , σ_b , and σ_c . All would be estimates of the precision of the same analytical method. Based only on duplicates, σ_b would be a poor estimate. Based on four observations, σ_c would be a better estimate. We would have no *direct* estimate of the error associated with sample **d**, but common sense indicates that the precision of its analysis is related to those of samples **a**, **b**, and **c**. Our best estimate of the precision of the method would come from pooling. Following Eqn. (5), we would write:

$$\sigma_p^2 = (d_{a,1}^2 + d_{a,2}^2 + d_{a,3}^2 + d_{b,1}^2 + d_{b,2}^2 + d_{c,1}^2 + d_{c,2}^2 + d_{c,3}^2 + d_{c,4}^2) / [(n_a - 1) + (n_b - 1) + (n_c - 1)] \quad (6)$$

where $d_{a,1}$ is the difference between the first observation of sample **a** and the mean value for sample **a**, etc. The numbers of replicates in each data set are represented by n_a , n_b ,... Equivalently, we could have written:

$$\sigma_p^2 = [(n_a - 1)\sigma_a^2 + (n_b - 1)\sigma_b^2 + (n_c - 1)\sigma_c^2] / [(n_a - 1) + (n_b - 1) + (n_c - 1)] \quad (7)$$

This form emphasizes that σ_p is, in effect, a weighted-average standard deviation.

The pooled standard deviation is the best estimate of the standard deviation of a single analysis. Standard errors to be

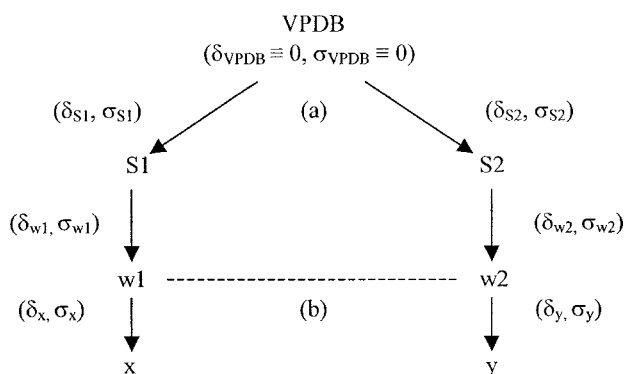


Figure 2. Diagram illustrating the total variance in the isotopic difference between two samples *x* and *y*. The $\delta^{13}\text{C}$ compositions (δ_i) and their 1σ standard deviations (σ_i) are denoted as follows: δ_{VPDB} and σ_{VPDB} represent the primary standard; δ_{S1} and σ_{S1} represent one secondary standard (viz., left; e.g., NBS-23); δ_{S2} and σ_{S2} represent another secondary standard (viz., right); δ_w and σ_w represent working standards; and δ_x, σ_x and δ_y, σ_y represent a first and second sample (see text). All of the intermediate variance terms (σ_i) contribute to the total variance, given by $\Delta'\delta$. Note that the working standards (w1, w2) may each represent either one external standard or an external-plus-an-internal standard as illustrated in Figure 1 and described in the text. Equations describing the total (line a) and abbreviated (line b) variances (given in the text) allow practical estimations of variance from each construction.

associated with the various mean values can be computed from it. In this example, we would report standard errors of $\sigma_p/\sqrt{3}$, $\sigma_p/\sqrt{2}$, $\sigma_p/\sqrt{4}$, and σ_p for the mean analytical results for samples **a**, **b**, **c**, and **d**.

In a typical isotopic calculation, results are reported in terms of δ vs the VPDB standard. Nevertheless, the pooled standard deviation calculated as outlined above is based only on variations relative to the standard used to calibrate the δ scale in the actual analysis. It is, therefore, an estimate only of either σ_4 or σ_5 .

Total variance between two isotopic compositions

When isotopic compositions of samples are to be compared between laboratories or, even within a single laboratory, over long periods of time, the nearest common standard will probably be VPDB. In that case, the relationships between the analytical results are summarized graphically by line *a* in Fig. 2. The variance in Δ (where $\Delta = \delta_x - \delta_y$) will be given by:

$$\sigma_{\Delta}^2 = \sigma_x^2 + \sigma_{w1}^2 + \sigma_{S1}^2 + \sigma_{S2}^2 + \sigma_{w2}^2 + \sigma_y^2 \quad (8)$$

where the σ_w terms refer to the calibrations of the working (internal or external) standards in each laboratory or procedure and the σ_s terms refer to the calibrations of the secondary standards. Secondary standards are often well calibrated (e.g., 0.02–0.03‰⁸). Values of σ_w depend on procedures used in individual laboratories. Since standards are analyzed more carefully than routine samples, we can expect $\sigma_{w1} < \sigma_x$ and $\sigma_{w2} < \sigma_y$. If procedures used for the

analyses of *x* and *y* are similar, so that $\sigma_x \approx \sigma_y$, we can then replace σ_x and σ_y with a single error term (σ_m) and summarize the result of these considerations as:

$$\sigma_{\Delta} \leq 2\sigma_m \quad (9)$$

An abbreviated solution for chromatography/isotope mass spectrometry

For the cases of gas or liquid chromatographic/isotope mass spectrometric methods, the total variance can be estimated in an abbreviated, yet very precise, method by accumulating the individual variances of specific compounds (σ_x, σ_y) and their respective standards (σ_{w1}, σ_{w2}), as indicated by line *b* in Fig. 2. The variances of the primary and secondary standards are ignored because this estimation of error focuses on the differential isotopic measurements between working standards and analytes. Thus, spanning from the isotopic composition of sample *x* to that of sample *y* ($\Delta'\delta = \delta_x - \delta_y$) directly via the internal standards w1 and w2, the total variance ($\sigma_{\Delta'\delta}^2$) is given by:

$$\sigma_{\Delta'\delta}^2 = \sigma_x^2 + \sigma_{w1}^2 + \sigma_{w2}^2 + \sigma_y^2 \quad (10)$$

In the preceding examples, since little variance (0.0004) is associated with the secondary (or laboratory) standard, its exclusion makes relatively little difference in estimation of the total error ($\sigma_{\Delta'\delta}$): 0.15‰ – 0.14‰ = 0.01‰.

A note on systematic errors

The two most typical kinds of systematic errors associated with isotope ratio-monitoring/GCMS are incomplete chromatographic resolution of compound peaks (coelution) and incomplete combustion. Coelution can result in varying degrees of mixing between the isotopic compositions of the relevant peaks. Four means by which to resolve coelution observed under a given set of analytical conditions are to: (i) chemically remove the problematic peak by preparative techniques (e.g., hydrolyze an ester⁴), (ii) chemically alter the peak of interest, thus changing its retention into another non-problematic region of the chromatogram, (iii) replace the chromatographic column with another with different retention properties, or (iv) when physical or chemical techniques do not effect peak separation, apply statistical techniques that may deconvolute the mass and isotopic coelution of peaks.⁹ Incomplete (or partial) combustion of organic material can isotopically fractionate it, giving anomalous isotopic results. Comparison of the observed isotopic compositions with known standard compositions will obviate this problem.

Systematic errors caused by the use of either internal standard compounds or of external standard gases are largely a function of the long-term availability of the standard and of the stability of the irmMS over the duration of data collection. When long-term (e.g., many-year) precision is required for large data sets, internal standard compounds with structural characteristics similar to those of the analytes are used to insure that standard and analytes undergo as nearly similar conditions of separation, combustion, and mass spectral analysis as possible.² For extended studies where high data quality (low variance) is required, systematic errors caused by the long-term intrinsic quality of

standards (viz., continuity) may typically become an issue. For example, an investigator should consider how many tanks of external standard CO₂ gas or milligrams of internal standard compound will be consumed during the course of the study and how much error will be introduced into the data with changes of standard materials. Such questions should be considered in the planning stage of long-term molecular isotopic studies, such as those of paleoceanographic time-series² or of isotopic product integrity.⁴

To minimize systematic errors in studies of long duration, a choice has to be made as to whether the analytes' working standard should be an internal standard compound or an external standard gas. This choice is largely dependent on the long-term compositional stability of the standard relative to the duration of the study. If an external standard CO₂ gas is suitably invariant during the course of the study, it should be used. If there are concerns about the stability of the external standard gas or the chromatographic or combustion steps of the analysis over the course of the study, then one should rely on internal standard compounds to minimize variance.

SUMMARY

For molecular isotopic data to be quantitatively useful, they must have meaningful estimates of random errors or variance. While in many cases only the proximal sources of variance are reported (i.e., the isotopic compositions versus the working standard), there are as many as four to five individual significant contributions to the variance in any given isotopic measurement. Those sources of variance are

explicitly described here. There are two main frames of reference for molecular-isotopic data: within a given laboratory or between laboratories. In the first case, the uncertainty reflects only variations relative to a local standard. In the second, the uncertainty must reflect variations all the way back to whatever standard has been shared between the laboratories. In *both* cases, assessment of the uncertainties requires that the standard deviation of replicate analyses be reported. The best estimate of that standard deviation can often be obtained by pooling observations from the analyses of multiple samples.

Acknowledgements

The author thanks J. M. Hayes for providing essential information and discussion of this manuscript. He also thanks N.S.F. Ocean Sciences (Chemical Oceanography) and N.A.S.A. for their support of this research.

REFERENCES

1. Hayes JM, Freeman KH, Popp BN, Hoham CH. *Org. Geochem.* 1990; **16**: 115.
2. Jasper JP, Mix AC, Prah FG, Hayes JM. *Paleoceanography* 1994; **9**: 781.
3. Lichtfouse E. *Org. Geochem.* 2000; **14**: 1337.
4. Jasper JP. *Pharm. Tech.* 1999; **23**(10): 106.
5. Merritt DA, Hayes JM. *Anal. Chem.* 1994; **66**: 2336.
6. Davies OL (ed). *Statistical Methods in Research and Production* (3rd edn). Hafner: New York, 1969.
7. Peters DG, Hayes JM, Hieftje GM. *Chemical Separations and Measurements: Theory and Practice of Analytical Chemistry*. Saunders Golden Series, 1974.
8. Coplen TB, Kendall C. *Anal. Chem.* 1982; **54**: 2611.
9. Goodman KJ, Brenna JT. *Anal. Chem.* 1994; **66**: 1294.