



NRC Publications Archive Archives des publications du CNRC

Uncertainty of relative sensitivity factors in Glow Discharge Mass Spectrometry

Meija, Juris; Methven, Brad; Sturgeon, Ralph E.

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below. / Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<http://doi.org/10.1088/1681-7575/aa83d6>

Metrologia, 2017-08-31

NRC Publications Record / Notice d'Archives des publications de CNRC:

<http://nparc.cisti-icist.nrc-cnrc.gc.ca/eng/view/object/?id=2cbae90a-128b-457a-b054-20bf7422c8f>

<http://nparc.cisti-icist.nrc-cnrc.gc.ca/fra/voir/objet/?id=2cbae90a-128b-457a-b054-20bf7422c8f6>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at <http://nparc.cisti-icist.nrc-cnrc.gc.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<http://nparc.cisti-icist.nrc-cnrc.gc.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.



Uncertainty of relative sensitivity factors in Glow Discharge Mass Spectrometry

Juris Meija, Brad Methven and Ralph E. Sturgeon

Measurement Science and Standards, National Research Council Canada
1200 Montreal Road, Ottawa ON K1A 0R6, Canada

E-mail: juris.meija@nrc-cnrc.gc.ca

Abstract. The concept of relative sensitivity factors required for the correction of the measured ion beam ratios in pin-cell Glow Discharge Mass Spectrometry (GD-MS) is examined in detail. We propose a data-driven model for predicting relative response factors which relies on nonlinear least squares adjustment and analyte/matrix interchangeability phenomena. The model provides a self-consistent set of response factors for any analyte/matrix combination of any element that appears as either analyte or matrix in at least one known response factor.

Keywords: Glow Discharge Mass Spectrometry; Calibration model; analyte/matrix interchangeability; Relative Sensitivity Factors; Least squares adjustment

Copyright

© Crown copyright in right of Canada, 2016

1. Introduction

Glow discharge mass spectrometry (GD-MS) is one of the most comprehensive and sensitive techniques currently available for determination of the composition of solid materials [1, 2]. As a combined tandem source [3], separation of the sputtering (sampling) and ionization processes in the GD provides response that is, to a first approximation, independent of sample matrix [4–6]. It is thus well-recognized that pin-cell GD-MS is relatively free of matrix effects, i.e., the slope of a calibration plot of amount content of an isotope of an impurity element present in different host matrices versus its detected ion beam intensity typically varies by no more than 30 to 100 % [5, 7]. The implications of this have recently been discussed with regard to utilization of GD-MS for establishment of the purity and traceability of high purity metals [8].

Mass fractions for each impurity element are directly derived from ion beam ratio data (i.e., the ion beam currents generated at the detectors for impurity isotope and the matrix isotope) corrected for small reproducible variations in sensitivities among different elements through application of relative sensitivity factors [2, 5].

The relative sensitivity factor for determination of element E (impurity) in host matrix of metal M and herein denoted with symbol $s(E, M)$, is the reciprocal slope of a virtual calibration function, based on the assumption of linearity with zero intercept and is defined as:

$$s(E, M) = \frac{w(E)}{w(M)} \cdot \frac{I(M)}{I(E)} \quad (1)$$

where $I(E)$ and $I(M)$ are the (corrected) ion beam intensities generated for the element E (impurity) and the reference element M (standard), $w(E)$ is the mass fraction of element E in the analysed standard material and $w(M)$ is the mass fraction of the matrix metal M in the analysed standard material. For example, the relative sensitivity factor for copper in a gold matrix is denoted $s(\text{Cu}, \text{Au})$. Corrections to ion beam intensities include the relative abundances of the measured isotopes as well as detector cross-calibration. In most publications, authors work with high-purity materials and assume $w(M) \approx 1$. This, in turn, means that the ratio $w(E)/w(M)$ is approximated with $w(E)$.

When quantitative analysis is desired, calibration of instrument response is typically undertaken using available Certified Reference Materials (CRMs) or,

potentially, gravimetrically spiked matrices prepared from high-purity substrates similar in composition to the material under study [9, 10], with which determined s values can then be rigorously calculated and applied. Such matrix matched approaches are readily capable of generating bias of typically no greater than 30–100 % relative and only require maintenance of reproducible sample geometry and discharge conditions for both sample and standard.

A more frequent situation entails semi-quantitative analysis in the absence of matched standards and in such case advantage is taken of the concept of the interchangeability of the s values amongst elements/matrices. Here, $s(E, \text{Fe})$ is frequently used as a normalization factor to derive other $s(E, M)$ values for a specific measurand in a different matrix in accordance with the expression:

$$s(E, M) = \frac{s(E, \text{Fe})}{s(M, \text{Fe})} \quad (2)$$

Normalizing the s values for GD-MS response to iron is a historical practice that evolved due to the relatively larger number of CRMs available for iron and steel matrices. Thus, s values provide a set of universal coefficients from which one s value can be converted to the s for another elemental impurity in a different matrix. This manuscript evaluates the limits of the interchangeability principle underlying Eq. (2).

Theoretical modeling of s values for the VG 9000 pin-cell GD-MS instrument has been undertaken by accounting for diffusional transport of the sputtered atoms in the glow discharge as well as their Penning, electron impact and asymmetric charge transfer ionization [5, 11]. Despite the comprehensive treatment of the subject, the best semi-empirical models can predict the s values only to about a factor of three (see Figure 1). Importantly, no corresponding analysis of estimated associated uncertainties is available.

GDMS measurements are often conducted to assess the purity of metals by measuring the levels of all *other* elements present at trace levels. The purity of the metal is then calculated as one minus the sum of all impurity levels. For a 99.99 % pure metal sample having typical levels of other trace elements (impurities), as an example, each of these impurities can be determined to a 50 % relative precision while still providing a fit-for-purpose purity estimate of the main element.

A recent assessment of uncertainty derived from the GD-MS calibration function highlighted the overwhelming contribution arising from s values to the estimated standard uncertainty in returned mass fraction for every element [8]. In light of this, a closer examination of s values is warranted and, in this contribution, a statistical analysis of s values compiled at NRC over the course of three decades of GD-MS characterization of high purity metals is presented. Herein, we propose a data-driven model for evaluating and predicting GD-MS relative sensitivity factors which relies on nonlinear least squares adjustment and the redanalyte-matrix interchangeability phenomenon. Particular attention is paid to the analysis of uncertainty, presented with the aim of assessing the typical global uncertainty associated with metal purity analysis using this technique.

2. Interchangeability of response factors

It is clear from Eq. (2) that if a single “comprehensive” CRM of a given matrix was available certified for mass fractions of all elements of interest, such a unique CRM could be used to derive s values for every element in every matrix via a multiplicative chain relationship. Such calculated values will be denoted herein as $s_c(E, M)$. In the absence of such a universal CRM, advantage is taken of the interchangeability relationships provided by Eq. 2. The simplest of such expressions is the reciprocal identity:

$$s_c(E, M) \doteq s^{-1}(M, E) \quad (3)$$

The dotted equal sign means that the left and right sides of the equation are not necessarily equal. The multiplicative property of s values can be extended into general combinatorial relationships of length $l = 2$:

$$s_c(E, M) \doteq s(E, X) s(X, M) \quad (4)$$

$$s_c(E, M) \doteq s(E, X) s^{-1}(M, X) \quad (5)$$

$$s_c(E, M) \doteq s^{-1}(X, E) s(X, M) \quad (6)$$

$$s_c(E, M) \doteq s^{-1}(X, E) s^{-1}(M, X) \quad (7)$$

We also consider eight combinatorial sequences of length $l = 3$:

$$s_c(E, M) \doteq s^{\pm 1}(E, X) s^{\pm 1}(X, Y) s^{\pm 1}(Y, M) \quad (8)$$

Together, these expressions form a set of equations that provide estimates for s values that have not been determined directly. In this article we consider equations (3) – (8) along with the identity expression

$$s_c(E, M) \doteq s(E, M) \quad (9)$$

For the NRC dataset containing 245 determined s values spanning 59 elements and 16 matrices, one can

establish 3000+ combinatorial expressions (with length $l = 1$ to 3). These expressions were generated using the generalized outer product of element and matrix lists. In programming language R, for example, one can use `outer(E, M, paste)` to obtain all element/matrix combinations from vectors E and M .

3. Experimental

3.1. Instrumentation

Measurements were performed with a VG 9000 (VG Microtrace, Windford, UK; currently supported by Thermo Fisher Scientific) reverse Nier-Johnson magnetic sector high-resolution mass spectrometer fitted with a pin-source tantalum cell that is cooled to near liquid nitrogen temperature to minimize out-gassing as the discharge heats. Mass resolution up to $m/\Delta m = 10\,000$ is available although the instrument is routinely operated at $m/\Delta m = 4000$ resolution (here m is the atomic mass of the relevant ion and Δm is the smallest mass difference of any two ions that can be separated, typically expressed as the width of the peak at a 5% height). A combination of Faraday and Daly detector systems that are conveniently cross-calibrated through use of argon isotopes from the discharge gas (^{38}Ar and ^{40}Ar) permits impurity elements to be quantitated at the levels of several parts in 10^9 . In our laboratory, the Faraday ion currents lie in the interval from 10^{-14} A to 10^{-9} A, whereas the Daly detector is operated between 5×10^{-19} A and 5×10^{-13} A, yielding some ten orders of magnitude of linear range.

Following shaping of the sample to accommodate the cryo-cooled pin cell and external chemical cleaning of the surface, a pre-burn sputtering of the sample is then undertaken using the GD plasma to etch the outer surface layers of the material and achieve a final cleaning of the surface prior to data acquisition. This typically requires about 20 min in a 5 mA, 1.1 kV discharge in a 1–2 mL min^{-1} flow of high-purity argon. During this period, the entire mass spectrum is interrogated to ensure that steady-state response has been achieved for each element (isotope) of interest, indicating that all surface contamination from the sample and the surrounding chuck holding the sample has been removed and bulk composition of the sample is being examined. After this step, the analytical data are acquired. These processes typically require 60–90 min at discharge conditions of 3.5 mA and 1.1 kV.

3.2. Computations

Computations were performed using Wolfram *Mathematica* [12] and R, a programming language and software environment for statistical computing and graphics [13]. The calculations discussed in this work have

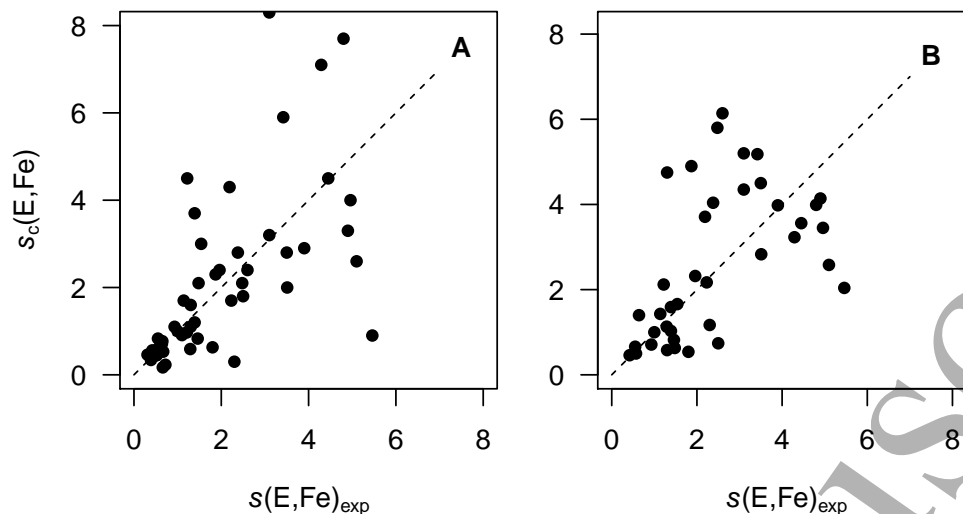


Figure 1. Predictive power of two notable theoretical models for relative sensitivity factors: study by (A) Vieth and Huneke [5] and (B) Bogaerts *et al* [11]. In both cases, the discrepancy between the experimental and calculated s values is significant with neither study assessing relative uncertainty.

Table 1. Main features of $s(E,M)$ datasets analyzed in this work.

Dataset	$s(E,M)$ values	Elements	Matrices	Time period
NRC	245	59	16	30+ years
Ref. [5]	112	38	6	unknown

been compiled in a self-contained R code [14].

Noteworthy is that treatment of s values for carbon, nitrogen, and oxygen have been specifically overlooked as these elemental impurities appear to not follow the usual pattern of matrix independence response and, coupled with the critical shortage of Certified Reference Materials with which to calibrate these elements, a less than desirable initial data base is available for these impurities [2,7]. We also have omitted available data for other nonmetallic elements such as iodine, bromine, chlorine and sulfur.

The NRC dataset relies on Certified Reference Materials for which the average relative standard uncertainty is 10 %. In the absence of an explicit uncertainty statement for certain relative sensitivity factors, we have imputed this average value of 10 % relative standard uncertainty. In addition to our own dataset we have analysed that of Vieth and Huneke [5]. The main features of these two datasets are summarized in Table 1.

4. Discussion

4.1. Dependence of the relative sensitivity factors on matrix composition

As noted above, s values derived from one (calibration) matrix can be used for the calculation of s values applied to the determination of elements in a different matrix. This forms the basis for the universality of GD-MS since the effect of sample matrix can be addressed using a model equation and no need arises for empirical correction factors. When the analyte and matrix elements are interchanged, e.g., when determining copper in silver as opposed to silver in copper, there is a strong inverse relationship between the corresponding s values as generalized above in Eq. (3). Figure 2A confirms this relationship amongst the 22 pairs of experimentally determined s values involving 13 elements.

4.2. Generalized equations for interchanging relative sensitivity factors

Equations (3)–(9) can be used to generate other combinatorial relationships between the determined s values. In this work, equations involving one, two, and three s values are considered, although there is no impediment to include combinatorial sequences of higher length. The general validity of such expressions can be shown with our data. For example, Eq. (4) is portrayed in Figure 2B with experimentally determined s values involving 41 elements.

With the use of these expressions, combinatorial reconstruction of yet-unknown s values from a given

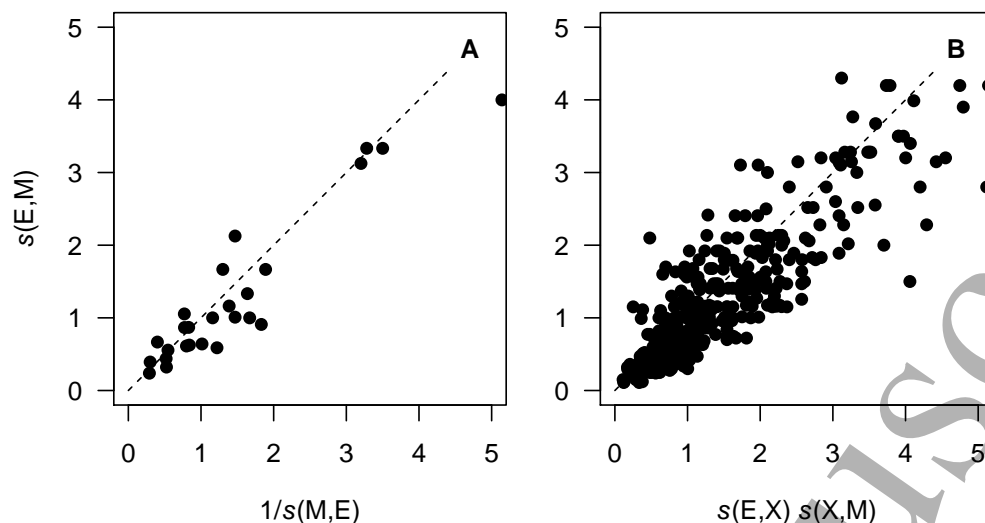


Figure 2. Relationships between experimental values of $s(E,M)$. Panel A shows the validity of the measurand-matrix interchangeability Eq. (3) whereas Panel B pertains to Eq. (4). Dotted lines represent function $y = x$.

set of known s values can be undertaken. For example, given two s values with one common element, $s(\text{Cu}, \text{Ag})$ and $s(\text{Cu}, \text{Au})$, four additional s values can be obtained: $s(\text{Ag}, \text{Cu})$, $s(\text{Au}, \text{Cu})$, $s(\text{Ag}, \text{Au})$, and $s(\text{Au}, \text{Ag})$.

The interchangeability of analyte and matrix enables quantitation using GD-MS in the absence of direct standards. As noted earlier, a fictitious calibration standard certified for mass fractions of all elements in a single matrix could be used to perform reliable quantitation of all elements in all matrices. Efficient use of the multiplicative interchangeability law, however, requires more knowledge of the statistical properties of this model, which are discussed below.

Consider a multiplicative interchangeability law for self-canceling sequences of s . The simplest of such sequences can be written as follows:

$$s(E_1, E_2) \times s(E_2, E_1) = s_0(E_1) \doteq 1 \quad (10)$$

Likewise, one example of a self-canceling sequence of length $l = 3$ is:

$$s(E_1, E_2) \times s(E_2, E_3) \times s(E_3, E_1) = s_0(E_1) \doteq 1 \quad (11)$$

One has to exclude the trivial sequences and therefore $E_1 \neq E_2 \neq E_3$ in Eqs. (10)–(11). In the absence of measurement or model errors, such sequences should yield the identity $s_0(E_i) = 1$ for all elements E_i , and this property can be used to evaluate the performance of the multiplicative interchangeability law. For each element, E_i , a set of calculated values, $s_0(E_i)$, can be established and presented as a histogram. Considering relationships involving two and three s values, a total of 3000+ values are generated for $s_0(E_i)$ where E_i

comprises 25 elements. A histogram of these values is shown in Figure 3 along with the results derived from the dataset of Vieth and Huneke [5].

It is evident from Figure 3 that deviations from the multiplicative interchangeability law are within a 95 % probability coverage interval of approximately [0.5, 2.0]. Broadly speaking, Figure 3 is in agreement with the often-quoted “factor of two” expanded uncertainty for s values [2, 8]. Figure 3A appears to support the traditional use of multiplicative uncertainty statements in GD-MS which tacitly imply lognormal distribution of errors. In addition, the 200 % magnitude of expanded uncertainty is in general agreement with these results which show a factor of $k_{95\%} = 2.5$ and 1.8 for data in Figure 3A. In addition to the NRC dataset, statistical analysis of the analyte/matrix interchangeability model can be performed using published data, such as the 1991 dataset from Vieth and Huneke [5]. Although this dataset is considerably smaller than the current NRC one ($N = 304$ with 6 matrices for Figure 3B), results similar to those shown in Figure 3A are obtained, as illustrated with Figure 3B, yielding a corresponding 95 % confidence interval of [0.6, 1.8].

The appearance of skewed distributions in Figure 3 can be explained without recourse to asymmetric probability density functions of the experimentally determined s values. Rather, skewed error profiles are a direct consequence of *large* symmetric uncertainties in s values and the fact that the variable of interest here is a nonlinear combination of these quantities, as illustrated in Figure 4. As noted recently by Possolo, the

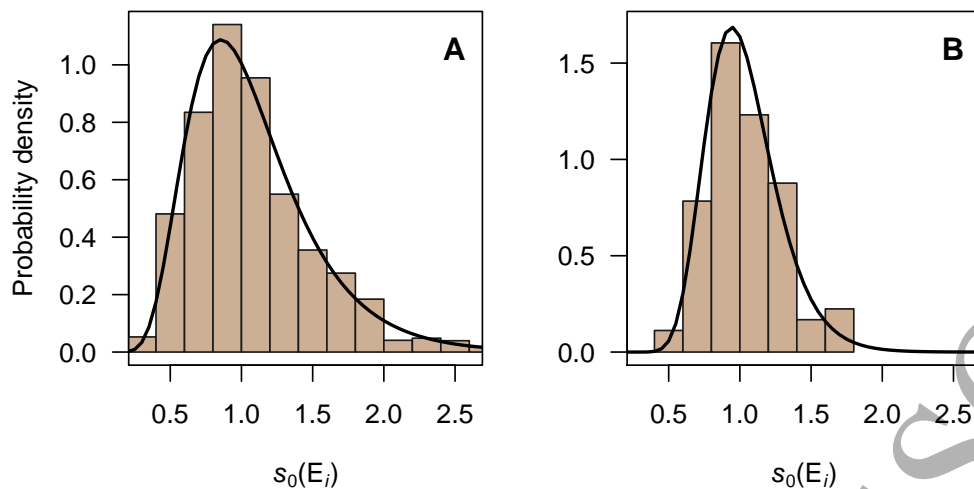


Figure 3. Histogram of $s_0(E_i)$ values from self-cancelling sequences ($l = 1, 2, 3$) of s from (A) NRC dataset and (B) from dataset of Vieth and Huneke [5]. Deviations from the multiplicative interchangeability law can be visualized using self-cancelling products or ratios of the experimentally determined values as shown in Eq. 10–11. These nonlinear combinations of the experimentally determined sensitivity factors can be modeled using the lognormal distribution (black curves).

use of a lognormal distribution remains a convenient alternative to approximate the ratio (or product) of normal random variables [15].

In summary, Figure 3 illustrates that the determined s values for a wide variety of elements and matrices are self-consistent to within a factor of two.

This consistency, however, can be improved for many relative sensitivity factors by performing least squares adjustment of the experimental data.

4.3. Least squares adjustment

As evident from Figure 3, there is ample inconsistencies between the experimentally determined s values and those inferred from the various combinatorial expressions such as the Eq. 10. The procedure for reducing such inconsistencies was pioneered in the late 1920s by Birge [16]. Much like individual measurement results in calibrations are commonly interpolated with a line by least squares fitting, interpolation over all experimentally determined s values can be undertaken in order to obtain a self-consistent set of s_c values. In this work, we adopt nonlinear least-squares adjustment of s values in line with the procedure that is employed to obtain the consensus values for the fundamental physical constants by CODATA [17].

In short, all experimentally determined values of $s(E, M)$ are recast in terms of $s_c(E, Fe)$ and $s_c(M, Fe)$ via equation

$$s(E, M) \doteq \frac{s_c(E, Fe)}{s_c(M, Fe)} \quad (12)$$

The set of all $s_c(E, Fe)$ is the output vector for the multivariate least squares fitting. The system of nonlinear equations is linearized using Taylor series expansion of functions in Eq. (12). Both fixed effects and random effects models for the least squares fitting were used.

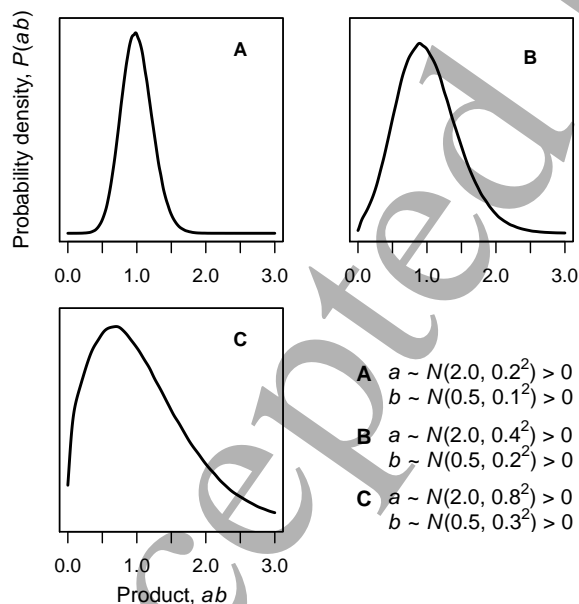


Figure 4. Product of two normally distributed random variables gives skewed probability density distribution in the presence of large relative uncertainties. The notation $a \sim N(2.0, 0.2^2)$ denotes random variable a that is drawn from a normal distribution with the mean 2.0 and standard deviation 0.2.

Relative sensitivity factors in GD-MS

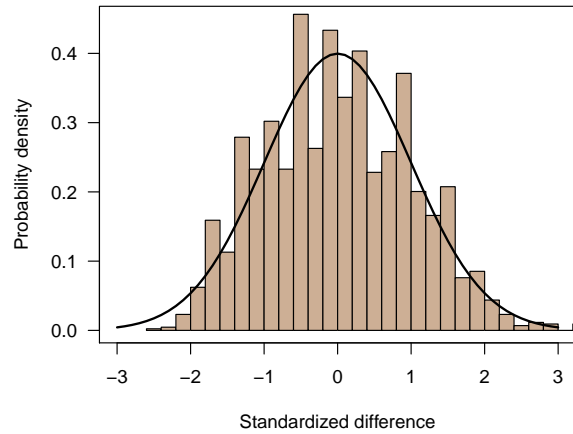


Figure 5. Histogram of standardized differences in the results of a copper reference material analyses (BCR 075B; IRMM, Belgium). The histogram combines determinations of nine elements (Al, Bi, Cd, Co, Mn, Ni, Pb, Sb, and Zn) all present at levels 1000–4000 ng/g. A total of 240 measurements were available for each element from data collected over the last 15 years at NRC. Results of each element are normalized such that their mean is 0 and standard deviation is 1 (standardized difference).

4.4. Least squares adjustment: fixed effects model

Suppose there are M determined values q_i of various s ($i = 1 \dots M$). For example, q_1 could be an experimentally determined value of $s(\text{Ag}, \text{Cu})$. Each input datum q_i can be expressed as a function f_i of one or more of the adjusted constants z_j ($j = 1 \dots N$) through the set of N observational equations ($N < M$). For example, z_1 could be $s_c(\text{Ag}, \text{Fe})$ and z_2 could be $s_c(\text{Cu}, \text{Fe})$. Then, $q_1 \doteq z_1/z_2$. The choice of adjusted constants is arbitrary and we have selected the set $s_c(\text{E}, \text{Fe})$ where E includes all analytes of the input data. The corresponding set of functions f_i is obtained using equations (3), (5), (6), and (9). We establish the following statistical model:

$$q_i = f_i(s) + e_i \quad (13)$$

where $e_i \sim N(0, \sigma_i^2)$. The adequacy of normally-distributed measurement errors is supported by the results obtained over a period of 15 years for the analysis of a copper quality control standard in our laboratory (see Fig. 5).

Most of the observational equations are nonlinear and we linearize expressions f_i using first-order Taylor series around starting values s_j that are nearly equal to the expected values of the adjusted constants z_j :

$$q_i \doteq f_i(s) + \sum_{j=1}^M \frac{\partial f_i(s)}{\partial s_j} (z_j - s_j) \quad (14)$$

In matrix notation, the least squares adjustment of the input data, \mathbf{Q} , provides the adjusted values, \mathbf{Z} , via a

variance-weighted linear least squares expression

$$(\mathbf{Q} - \mathbf{F}) \doteq \mathbf{A}(\mathbf{Z} - \mathbf{S}) \quad (15)$$

where \mathbf{A} is a $N \times M$ matrix with derivatives of functions f_i around s_j , $a_{ij} = \partial f_i(s)/\partial s_j$, and \mathbf{F} is a vector of values $f_i(s)$. Thus, \mathbf{Q} and \mathbf{F} are vectors of length N whereas \mathbf{Z} and \mathbf{S} are vectors of length M . The value of \mathbf{Z} is obtained using the Gauss-Newton method of least squares minimization:

$$\mathbf{Z} = \mathbf{S} + (\mathbf{A}^T \mathbf{W}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W}^{-1} (\mathbf{Q} - \mathbf{F}) \quad (16)$$

Here, \mathbf{W} is a diagonal matrix containing the variances of q_i , $w_{i,i} = u^2(q_i)$. The least squares minimization is an iterative process because nearly all combinatorial expressions, f_i , are nonlinear. Thus, the minimization starts with an approximation that elements s_j are equal to the corresponding input values q_i and continues until the convergence criterion,

$$\sum_{j=1}^N \frac{(z_j - s_j)^2}{u^2(z_j)} \quad (17)$$

is less than some arbitrary value, say 10^{-3} (in the absence of initial estimate we take $s_j = 1$ for all j). This process takes typically less than five iterations. After each iteration, vector \mathbf{S} is updated using expression $\mathbf{S} = \mathbf{Z}$ and then \mathbf{A} and \mathbf{F} are recalculated using the new values of \mathbf{S} . The uncertainty of the adjusted variables, $u(z_j)$, is obtained from the diagonal elements of the covariance matrix of \mathbf{Z} ,

$$\text{cov}(\mathbf{Z}) = s_e^2 (\mathbf{A}^T \mathbf{W}^{-1} \mathbf{A})^{-1} \quad (18)$$

where s_e is the standard error of the regression and is also known as the Birge ratio R_B :

$$s_e^2 = R_B^2 = \frac{1}{M - N} (\mathbf{Q} - \mathbf{F})^T \mathbf{W}^{-1} (\mathbf{Q} - \mathbf{F}) \quad (19)$$

4.5. Least squares adjustment: random effects model

Least squares adjustment presented above is a location-scale model which recognizes that each determined s value is perturbed only by a measurement uncertainty. With a random effects model, the possible inconsistencies in the data are modeled by an additional source of uncertainty which depends on the sample matrix (the random effect):

$$q_i = f_i(s) + e_i + t(M) \quad (20)$$

where $e_i \sim N(0, \sigma_i^2)$ and $t(M) \sim N(0, \tau_M^2)$ is the random effect which depends on the matrix element M in the measured quantity q_i . The random effects estimate of \mathbf{Z} is obtained as follows:

$$\mathbf{Z}_r = \mathbf{S} + (\mathbf{A}^T \mathbf{\Omega}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{\Omega}^{-1} (\mathbf{Q} - \mathbf{F}) \quad (21)$$

Relative sensitivity factors in GD-MS

8

Values for \mathbf{A} and \mathbf{F} are obtained from the fixed effects model. The covariance matrix of \mathbf{Z}_r is given as

$$\text{cov}(\mathbf{Z}_r) = (\mathbf{A}^T \boldsymbol{\Omega}^{-1} \mathbf{A})^{-1} \quad (22)$$

The matrix $\boldsymbol{\Omega}$ is identical to \mathbf{W} but with heterogeneity variances τ_M^2 appropriately added to all its diagonal elements. We estimate the values of τ_M using the method of restricted maximum likelihood by maximizing $L(\tau_M)$ as described by Chen *et al* [18]:

$$2L(\tau_M) = -\ln |\boldsymbol{\Omega}| - \ln |\mathbf{A}^T \boldsymbol{\Omega}^{-1} \mathbf{A}| - \mathbf{r}^T \boldsymbol{\Omega}^{-1} \mathbf{r} \quad (23)$$

where \mathbf{r} is the vector of regression residuals,

$$\mathbf{r} = (\mathbf{Q} - \mathbf{F}) - \mathbf{A}(\mathbf{Z}_r - \mathbf{S}) \quad (24)$$

The overdispersion of the results of the random effects model, R_B , can be calculated as in Eq. 19 but with \mathbf{W} replaced by $\boldsymbol{\Omega}$. Owing to the large number of degrees of freedom in our data set, however, we do not further expand the adjusted values for any remaining overdispersion (hence, Eq. 22 does not include the term s_e^2). In this manner, we believe that the remaining source of inconsistency may originate largely from the inconsistencies amongst the various CRMs and not the GD-MS analysis.

4.6. Performance test using Monte Carlo simulation

To test the performance of the least squares adjustment, we performed Monte Carlo simulation on a small scale test model consisting of five relative sensitivity factors in a system of three elements: Ag, Al, and Cu, whose values derive from hypothetical “true” values $s(\text{Al}, \text{Ag}) = 2$, and $s(\text{Cu}, \text{Ag}) = 5$. All of these values are perturbed with a 50 % relative uncertainty due to the measurement and an additional uncertainty component is added for each of the four matrices. This results in relative uncertainty of 60–100 % for each of the five measured sensitivity factors $s(\text{Al}, \text{Ag})$, $s(\text{Al}, \text{Cu})$, $s(\text{Ag}, \text{Cu})$, $s(\text{Cu}, \text{Al})$, and $s(\text{Cu}, \text{Ag})$. The resulting values for all relative sensitivity factors are modeled as random draws from zero-truncated normal distributions.

We believe that such a Monte Carlo test involves high relative uncertainties and therefore allows assessment of the performance of the least squares adjustment. The values of the five sensitivity factors were sampled 10 000 times from the statistical model described above and the adjusted values were obtained in each cycle using the least squares adjustment. The Monte Carlo simulation results are summarized in Fig. 6 and they show acceptable performance of the random effects model for our purposes.

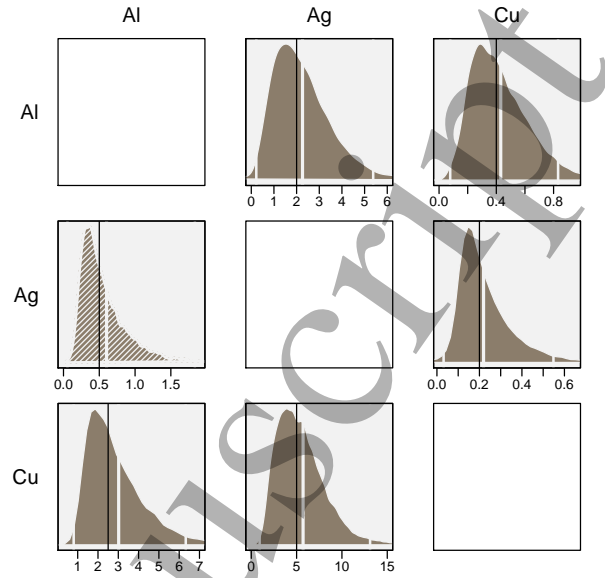


Figure 6. Monte Carlo simulation of the random effects least squares adjustment on a three element system. The shaded curves represent the probability densities of the adjusted sensitivity factors. The vertical white lines represent the mean values and 95 % shortest coverage intervals whereas the black lines correspond to the true values derived from $s(\text{Al}, \text{Ag}) = 2$ and $s(\text{Cu}, \text{Ag}) = 5$. The histogram of $s_c(\text{Ag}, \text{Al})$ is shaded because it is not part of the input variables.

4.7. Results of least squares adjustment

Application of the fixed effects model to the NRC dataset produces a self-consistent set of s_c values. However, the resulting Birge ratio is $R_B = 5.6$ for the NRC dataset and $R_B = 1.8$ for the dataset of Vieth and Huneke. Both values of R_B largely exceed the critical value of $R_{B, \text{crit}} \approx 1$, thus indicating the presence of an additional source of uncertainty. Consequently, the random effects model was chosen to fit the data. With the random effects model, however, we obtain $R_B = 1.6$ for the NRC dataset which we find as an acceptable level of unexplained heterogeneity.

Fig. 7 compares the input s values with the corresponding random effects least squares adjusted values of the NRC dataset. The overall agreement between the input $s(\text{E}, \text{M})$ and output $s_c(\text{E}, \text{M})$ values is promising, especially when compared to Fig. 1. In addition, one can see that the differences between the adjusted and experimental values do not exhibit asymmetric tendencies and the residuals follow a t -distribution with a zero mean and degrees of freedom $v \approx 2$. The expanded uncertainties ($k = 2$) of the residuals include zero 70 % of the time (such underestimation of the uncertainty is to be expected from the value of Birge ratio $R_B > 1$). Despite

Relative sensitivity factors in GD-MS

the somewhat low coverage probability, we find these results encouraging. Indeed, it is impossible to reconcile every discrepancy in a dataset that relies on multiple independent CRMs which have been analyzed over several decades and for which some inconsistencies might well be due to inherent biases in the certified values themselves.

The predictive power of the least squares adjustment has already served us well. For example, inspection of the regression residuals, $\mathbf{Q} - \hat{\mathbf{F}}$, showed that most of the largest residuals were from s values in a matrix of nickel. Subsequent inspection of the input data for nickel revealed that they were based on corresponding CRMs which contained only 50–60 % nickel thus raising the need for a consideration of modeling alloy matrices in the future.

The random effects least squares analysis of the NRC dataset provides an overall (mean) relative standard uncertainty of 25 % for the values of $s_c(\text{E}, \text{Fe})$. This summary is in agreement with the work of Matschat *et al* who report average deviations from the known values of less than 30 % for most of the investigated analytes in spiked pressed powders [19].

4.8. Sensitivity analysis

Nonlinear least squares adjustment of the input data (s) provides an estimate of uncertainty for each of the adjusted values (s_c). Regression diagnostics allows identification of influential data for better understanding of the dataset. Consider the projection (hat) matrix which links the observed data to the estimate of the regression parameters:

$$\mathbf{P} = \mathbf{A}(\mathbf{A}^T \boldsymbol{\Omega}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\Omega}^{-1} \quad (25)$$

Given that $(\hat{\mathbf{Q}} - \hat{\mathbf{F}}) = \mathbf{P}(\mathbf{Q} - \hat{\mathbf{F}})$, the matrix elements $p_{i,j}$ express the degree of leverage that an observation q_i has on the corresponding fitted value $f_j(z)$. We use Cook's distance to measure the overall effect on the regression estimates from the i^{th} observation, q_i [20]:

$$D_i = \frac{h_i}{1 - h_i} \cdot \frac{r_i^2}{N + 1} \quad (26)$$

where r_i is the standardized regression residual, $r_i = (q_i - \hat{q}_i)/(s_e \sqrt{1 - h_i})$, and h_i summarizes the leverage of q_i on all fitted values: $h_i = \sum_j^N p_{i,j}^2$. Experimental values of q_i for which the Cook's distance $D > 4/(M - N - 1)$ are deemed influential data and they merit a detailed reevaluation by the analyst [21].

4.9. Extrapolation

The advantage of the matrix interchangeability relationship lies in its ability to reconstruct values of $s_c(\text{E}, \text{M})$ which are otherwise unavailable in the

experimental dataset. In addition, a concerted least squares adjustment of all determined $s(\text{E}, \text{M})$ values provides a comprehensive assessment of the reliability of the calculated values without relying on any individual experimental data point.

The least squares adjustment provides a set of $s_c(\text{E}, \text{Fe})$ along with their uncertainties. The values of $s_c(\text{E}, \text{M})$ are ratios of $s_c(\text{E}, \text{Fe})$ and $s_c(\text{M}, \text{Fe})$ where E and M represent any element contained in the input dataset. Hence, $s_c(\text{E}, \text{M})$ can be viewed as a ratio of two normal random variables from the least squares adjustment. Given the possibly large relative uncertainties for some of the sensitivity factors, their ratio may take an asymmetric probability density which is a nontrivial topic of mathematical statistics in its own right [22, 23].

Due to the aforementioned complexities of analytical solutions for approximating normal ratio variables, we have employed the Monte Carlo method to obtain the best estimates and uncertainties for $s_c(\text{E}, \text{M})$. To achieve this, we draw random samples of $s_c(\text{E}, \text{Fe})$ from the multivariate normal distribution with N -dimensional mean vector, \mathbf{Z} , and $N \times N$ covariance matrix, $\text{cov}(\mathbf{Z}_r)$. Then, the full set of values for $s_c(\text{E}, \text{M})$ is obtained and such process is repeated many times. We use robust measures to extract the summary statistics of the obtained datasets: Tukey's trimean for location and the median absolute deviation (multiplied by 1.4826) for standard uncertainty.

Moreover, the least squares adjustment procedure provides a set of adjusted $s_c(\text{E}, \text{M})$ values for all possible element/matrix combinations along with their covariances which are particularly useful when evaluating uncertainty of a combination of these values, for example, in assigning overall purity for a given material.

The resulting uncertainty for a particular combination $s_c(\text{E}, \text{M})$ depends on the number of available s values that contain elements E and M, N_E and N_M . In particular, if either E or M is present in only one measured reference material, then one cannot take advantage of the analyte-matrix interchangeability property and the resulting uncertainty for $s_c(\text{E}, \text{M})$ is unaffected by the least squares adjustment. This is evident from Fig. 8 where the case at $\min(N_E, N_M) = 1$ compares well with the 95 % confidence interval seen in Fig. 3, i.e., the expanded relative uncertainties for these s_c range from 50 % to 200 % which is in line with the NRC's long-held view that these values are reliable to within a factor of two. It is evident that the relative uncertainty of s_c values decreases significantly with the availability of multiple s values per element, which attests to the utility of analyte-matrix interchangeability property. Fig. 8 also shows that the expanded uncertainty of NRC s_c values cannot be improved below

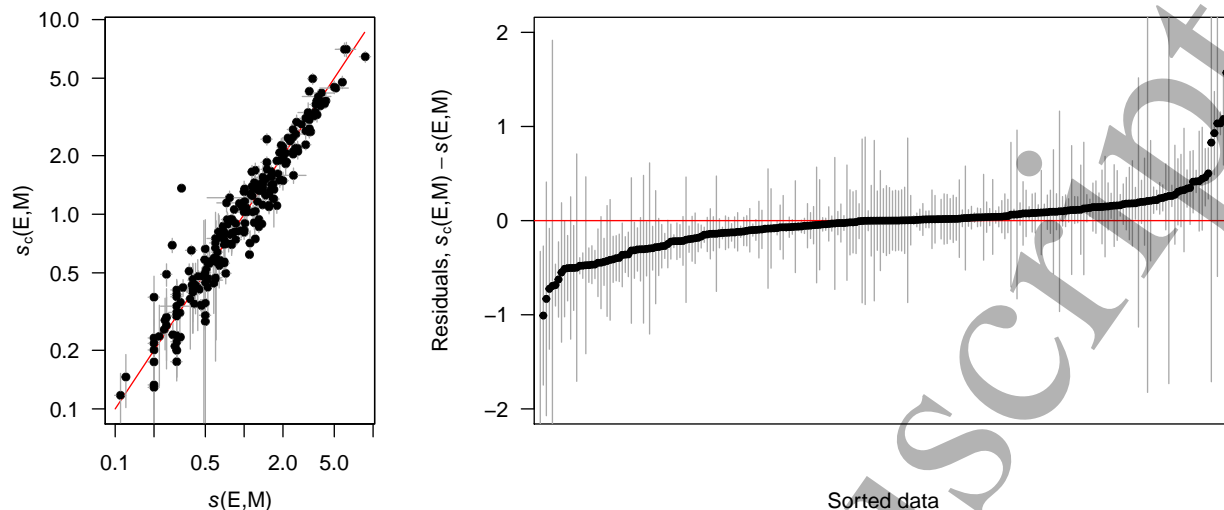


Figure 7. Comparison of the experimentally determined values of $s(E, M)$ with the corresponding least squares adjusted values of the NRC dataset, $s_c(E, M)$, using the random effects model (left panel). Residuals, $s_c(E, M) - s(E, M)$, are shown in the right panel where the x-axis corresponds to each input data point. The length of each vertical bar represents the expanded uncertainty (coverage factor $k = 2$) for each residual. This uncertainty is obtained by combining the uncertainties of the adjusted output values and the corresponding input values for the relative sensitivity factors.

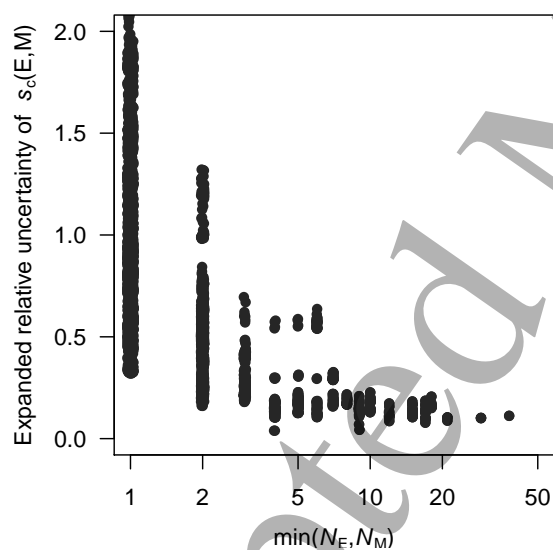


Figure 8. Expanded relative uncertainty of $s_c(E, M)$ as a function of the minimum number of experimental $s(E, M)$ values available for elements E and M (NRC dataset).

10 %, which seems to be limited by the analyte-matrix model uncertainty.

4.10. Periodic trends

When the adjusted s_c values are obtained using theoretical means, as several authors have done in the past, it is important to understand the origins of such trends. Although the use of the combinatorial

interchangeability law does not require any such knowledge, we nevertheless would like to note a few general observations that emerge from the NRC dataset, the implications of which are not immediately relevant at this time.

While the relative uncertainty of individual $s_c(E, Fe)$ values is large for some elements, one can nevertheless observe periodic trends in the changes of these values as a function of the atomic number of the element E. For example, the highest $s_c(E, Fe)$ values occur for zinc, cadmium, and mercury – all of which are the last members of the f-block elements. Broadly speaking, as one traverses through the periodic table of elements, the s_c values increase along the f-block elements and then decrease along the p-block elements.

5. Conclusion

In 2011 the CCQM noted the lack of current research programmes to develop and extend metrological applications of established techniques for high purity metals for measurements by glow discharge mass spectrometry. In response to this remark, in this work we analyse the relationships and patterns between the measured GD-MS relative response factors. From the existing set of 245 values covering 59 analytes in 16 matrices, we are able to obtain 3422 relative response factors for 59 analytes in 59 matrices. The predictive power of the analyte-matrix interchangeability model in conjunction with the least squares adjustment has proven useful at NRC and this approach provides a firm footing of the uncertainty evaluation for the s values

as it has heretofore been unavailable. Furthermore, the proposed method enhances confidence in the measured s values and in the calculation of s_c values for analyte-matrix combinations that are not accessible via Certified Reference Materials used for calibration.

Acknowledgments

This work benefited from the technical expertise of Kate Swider (NRC), Indumathi Pihillagawa Gedara (NRC), Dr. Bussaraporn Patarachao (NRC), and Dr. Mike Hinds (Royal Mint, Ottawa). We also wish to express our appreciation to an anonymous referee whose comments helped to improve this manuscript.

References

- [1] Harrison W W, Hess K R, Marcus R K and King F L 1986 *Analytical Chemistry* **58** 341A–356A URL <http://dx.doi.org/10.1021/ac00293a002>
- [2] Hoffmann V, Kasik M, Robinson P K and Venzago C 2004 *Anal Bioanal Chem* **381** 173–188 URL <http://dx.doi.org/10.1007/s00216-004-2933-2>
- [3] Kantor T and Hieftje G M 1995 *Spectrochimica Acta Part B: Atomic Spectroscopy* **50** 961–962 URL [http://dx.doi.org/10.1016/0584-8547\(95\)01391-Q](http://dx.doi.org/10.1016/0584-8547(95)01391-Q)
- [4] King F L, Teng J and Steiner R E 1995 *Journal of Mass Spectrometry* **30** 1061–1075 URL <http://dx.doi.org/10.1002/jms.1190300802>
- [5] Vieth W and Huneke J C 1991 *Spectrochimica Acta Part B: Atomic Spectroscopy* **46** 137–153 URL [http://dx.doi.org/10.1016/0584-8547\(91\)80017-W](http://dx.doi.org/10.1016/0584-8547(91)80017-W)
- [6] Nelis T and Pallosi J 2006 *Applied Spectroscopy Reviews* **41** 227–258 URL <http://dx.doi.org/10.1080/05704920600620345>
- [7] Robinson K and Hall E F H 1987 *JOM* **39** 14–16 URL <http://dx.doi.org/10.1007/BF03258853>
- [8] Sturgeon R E, Methven B, Willie S N and Grinberg P 2014 *Metrologia* **51** 410–422 URL <http://dx.doi.org/10.1088/0026-1394/51/5/410>
- [9] Gusarova T, Hofmann T, Kipphardt H, Venzago C, Matschat R and Panne U 2010 *Journal of Analytical Atomic Spectrometry* **25** 314–321 URL <http://dx.doi.org/10.1039/B921649A>
- [10] Gusarova T, Methven B, Kipphardt H, Sturgeon R, Matschat R and Panne U 2011 *Spectrochimica Acta Part B: Atomic Spectroscopy* **66** 847–854 URL <http://dx.doi.org/10.1016/j.sab.2011.12.001>
- [11] Bogaerts A, Temelkov K A, Vuchkov N K and Gijbels R 2007 *Spectrochimica Acta Part B: Atomic Spectroscopy* **62** 325–336 URL <http://dx.doi.org/10.1016/j.sab.2007.03.010>
- [12] Wolfram Research Inc 2014 *Mathematica 10.0* Champaign, Illinois URL <https://www.wolfram.com>
- [13] R Core Team 2015 *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing Vienna, Austria URL <https://www.R-project.org/>
- [14] Meija J 2016 *GDMS Relative Sensitivity Factor Calculator* National Research Council Canada Ottawa, Canada
- [15] Possolo A 2015 *Metrologia* **53** S17–S24 URL <http://dx.doi.org/10.1088/0026-1394/53/1/S17>
- [16] Birge R T 1929 *Reviews of Modern Physics* **1** 1–73 URL <http://dx.doi.org/10.1103/RevModPhys.1.1>
- [17] Mohr P J and Taylor B N 2000 *Reviews of Modern Physics* **72** 351–495 URL <http://dx.doi.org/10.1103/RevModPhys.72.351>
- [18] Chen H, Manning A K and Dupuis J 2012 *Biometrics* **68** 1278–1284 URL <http://dx.doi.org/10.1111/j.1541-0420.2012.01761.x>
- [19] Matschat R, Hinrichs J and Kipphardt H 2006 *Anal Bioanal Chem* **386** 125–141 URL <http://dx.doi.org/10.1007/s00216-006-0645-5>
- [20] Cook R D 1979 *Journal of the American Statistical Association* **74** 169–174 URL <http://dx.doi.org/10.1080/01621459.1979.10481634>
- [21] Altman N and Krywinski M 2016 *Nature Methods* **13** 281–282 URL <http://dx.doi.org/10.1038/nmeth.3812>
- [22] Marsaglia G 2006 *Journal of Statistical Software* **16**
- [23] Hayya J, Armstrong D and Gressis N 1975 *Management Science* **21** 1338–1341