

# A New Method for the Improvement of Data Production in Phytochemical Analysis

João Ponte-e-Sousa<sup>a\*</sup> and António Neto-Vaz<sup>a,b</sup>

## ABSTRACT:

**Introduction** – The use of the average analytical signal for the construction of curves by the least squares method (LSM) over the standard addition method (SAM) is widespread. It would be advantageous, however, to find a way to avoid intermediary averages, which are known to be the cause of significant increases in standard deviations (SD).

**Objective** – To develop a protocol that uses all gathered data to create curves by LSM over SAM. To use Excel<sup>®</sup> for the estimation of  $y = mx + b$  and  $R^2$  rather than using LSM equations for the SD of  $m$ ,  $x$  and  $b$ .

**Methodology** – The level of lead (II) in the bark (cork) of *Quercus suber* Linnaeus was determined using differential pulse anodic stripping voltammetry (DPASV). Three current samples were taken for each of the four standard additions. These signals were combined for adjustment by LSM. The results were compared with those obtained after averaging the current for each addition, and the expression of uncertainty in the measurements determined.

**Results** – The new method shows an expanded uncertainty of  $\pm 0.3321 \mu\text{g/g}$  (nearly 1.42%). The difference between the results obtained by the new and the old method is  $0.01 \mu\text{g/g}$  (23.41 and  $23.40 \mu\text{g/g}$ ). The limit of detection changed approximately from  $4.8$  to  $4 \mu\text{g/g}$  and the relative SD approximately from 9 to 6%.

**Conclusion** – The absence of intermediary averages in curves improved the determination of lead (II) in cork by DPASV. Estimation of SD only with LSM equations produced results that were significantly worse. The changes are large enough to transform an apparently internally non-validated procedure (repeatability for precision) into an internally validated procedure. Copyright © 2009 John Wiley & Sons, Ltd.

**Keywords:** Voltammetry; standard addition method; least squares method; cork

## Introduction

No doubt numbers play an important role in phytochemical analysis but 'the fact remains that many highly competent scientists are woefully ignorant of even the most elementary statistical methods. It is even more astonishing that analytical chemists, who practise one of the most quantitative of all sciences, are no more immune than others to this dangerous, but entirely curable, affliction' (Miller and Miller, 1988).

Using averages in the middle of data analysis reduces the information extractable from a set of measures. Ideally only final results should be shown as averages. The present paper shows how to combine measurements of an analytical signal in such a way that all possible final results inferable from a data population, such as the one presented in this study, will not be neglected. The use of this particular case of combinatorial analysis is compared with the earlier method for the estimation of the lead content in cork from *Quercus suber* L. by differential pulse anodic stripping voltammetry (DPASV). The effect that the new method has on the uncertainty of the estimation is emphasised. The new method extracts more information from the chemical analysis. In fact, instead of using just the curve of mean values of the analytical signals for each concentration, it uses  $r^s$  curves, where  $r$  is the number of repetitions of the analytical signal per concentration and  $s$  is the number of different concentration values. The curves correspond to all possible choices of one measurement for each concentration.

## Experimental

### Standards and chemicals

The supporting electrolyte solution, 0.1 M sodium chloride, was prepared by dissolving sodium chloride p.a. from Merck (Darmstadt, Hessen, Germany) in Milli-Q water; the 0.1 M nitric acid solution was obtained by dissolving nitric acid 65% p.a. (Merck) in Milli-Q water; the  $1.1 \times 10^{-2}$  M lead (II) stock solution was obtained by dissolving lead nitrate from Riedel (Seelze, Lower Saxony, Germany) in 0.01 M nitric acid; a standard solution of  $1.1 \times 10^{-5}$  M lead (II) was achieved by diluting the stock solution appropriately. The digester reagents used were: nitric acid 65% p.a. (Merck) and hydrogen peroxide 30% p/v p.a. from Panreac (Castellar del Vallés, Spain, Spain). The solutions were prepared daily.

\* Correspondence to: J. Ponte-e-Sousa, Department of Chemistry, University of Évora, Rua Romão Ramalho no. 59, 7000-671 Évora, Portugal. E-mail: jcps@uevora.pt

<sup>a</sup> Department of Chemistry, University of Évora, Rua Romão Ramalho no. 59, 7000-671 Évora, Portugal

<sup>b</sup> ICAM (Institute for Mediterranean Agricultural Sciences), Pólo da Mitra, University of Évora, 7000-Évora, Portugal

### Extraction and digestion of cork sample

The interior of the bark of cork-oak was granulated with a plastic granulator, and an aliquot (0.09 g) of the resulting bark powder was mixed with 2 mL nitric acid and 0.25 mL of hydrogen peroxide. Following incubation for 2 h, the mixture was digested in a closed recipient at 85–90°C for 8 h.

### DPASV analysis

The instrumentation used for DPASV comprised an Eco Chemie Autolab/pgstat 20 (Metrohm, Utrecht, Netherlands) potentiostat/galvanostat and a Metrohm VA stand 663 together with a hanging mercury drop electrode (HMDE), a glassy-carbon rod counter-electrode and an Ag–AgCl–KCl 3 M reference electrode. Eco Chemie GPES 4.9 software was employed to run the instruments and to obtain voltammograms. Voltammetric conditions were as shown in Table 1. For the voltammetric experiments, aliquots (0.1 mL) of the standard were added to 1 mL of the cork solution and 20 mL of the supporting electrolyte. Each measurement was made three times in each addition. Two-dimensional data treatment of current vs standard concentration was performed with the Excel 2003® line adjustment tool.

**Table 1.** Voltammetric conditions employed in the study

Voltammetric conditions	
Purge time	5 s
Deposition potential	−0.650 V
Deposition potential time	180 s
Stirrer during deposition	Yes
Equilibration time	30 s
Impulse time	0.05 s
Time interval	0.25 s
Initial potential	−0.650 V
Final potential	−0.1 V
Step potential	0.00195 V
Impulse amplitude	0.03 V
Scanning velocity	7.8 mV/s
Drop surface	0.25 mm <sup>2</sup> ± 10%
Stirrer rotation velocity	1500/min
Deaeration time	10–12 min

## Results and Discussion

Experimental measurements of the corrected electric intensity (*i*) produced an array of values that will be used as examples in the discussion in order to demonstrate their full exploitation. Table 2 shows *i* values measured three times for each of the four standard concentrations. Measurements reported refer to the sample material (a portion of the interior of the bark of cork-oak) and to

the blank assay, as usually employed to subtract any influence arising from all the solutions used for the correct treatment of the sample. The focus of this study is the random dependent variable *i* and four populations of three individual values resulting from the four standard concentrations tested are considered.

### Old algorithm: least squares equations

**Cork solutions.** In order to show any existing differences in the final results, previous methodologies have been compared with the new methodology proposed here. As an example, focus will be given to the experimental results obtained for cork as shown in Table 2. Three current measurements are registered for each standard concentration. The electrolytic cell only contained dissolved cork and supporting electrolyte when the first recording was performed. A new HMDE (drop 2) subsequently replaced the old one, which had fallen into the cell. The same occurred after *i* had been recorded for drop 2. Thus, three consecutive recordings were obtained for each of the four standard concentrations, as indicated in Table 2.

The first comment is related to the initial step of the treatment of experimental results because LSM equations are usually applied after the reduction of all analytical signal recordings to their average values so that the LSM adjustments can be performed on a set of pair values, namely, analytical signal vs standard concentration. Table 3 shows the average values for cork and successive calculations for the production of the necessary data for the least squares treatment.

The quantities  $S_{xx}$ ,  $S_{yy}$  and  $S_{xy}$  are needed to obtain all results allowed by LSM, where

$$S_{xx} = \sum x_i^2 - (\sum x_i)^2 / N, S_{yy} = \sum y_i^2 - (\sum y_i)^2 / N \text{ and} \\ S_{xy} = \sum x_i y_i - \sum x_i \sum y_i / N,$$

*N* being the number of data pairs used in the construction of the curves.

Substitution gives:

$$S_{xx} = 133.3 \times 10^{-16} \text{ M}^2, S_{yy} = 1.986 \times 10^{-16} \text{ A}^2, \text{ and} \\ S_{xy} = 16.25 \times 10^{-16} \text{ MA}.$$

Now the slope (*m*) of the adjusted line can be calculated as:

$$m = S_{xy} / S_{xx}, \text{ from which } m = 0.122 \text{ A/M}.$$

The intercept (*b*) with the analytical signal axis is:

$$b = \sum y_i / N - m \sum x_i / N, \text{ from which } b = 1.354 \times 10^{-8} \text{ A}.$$

**Table 2.** Current measured by DPASV for the null, first, second and third standard additions

Lead (II) concentration (10 <sup>−8</sup> M)	Electric intensity <i>i</i> (10 <sup>−8</sup> A)					
	Cork			Blank		
0.000	1.463	1.349	1.372	0.780	0.774	0.769
5.213	1.932	1.961	1.952	4.063	4.069	4.069
10.38	2.564	2.624	2.543	7.016	7.052	7.178
15.49	3.323	3.237	3.289	9.880	10.22	9.950

**Table 3.** Data obtained from cork solutions using the least squares method

$x_i^a$ ( $10^{-8}$ M)	$y_i^b$ ( $10^{-8}$ A)	$x_i^2$ ( $10^{-16}$ M <sup>2</sup> )	$y_i^2$ ( $10^{-16}$ A <sup>2</sup> )	$x_i y_i$ ( $10^{-16}$ MA)
0.000	1.395	0.000	1.945	0.000
5.213	1.948	27.18	3.796	10.16
10.38	2.577	107.7	6.640	26.75
15.49	3.283	239.9	10.78	50.85
$\Sigma 31.08$	$\Sigma 9.203$	$\Sigma 374.9$	$\Sigma 23.16$	$\Sigma 87.76$

<sup>a</sup>  $x_i$  = standard concentration of lead (II).

<sup>b</sup>  $y_i$  = mean corrected electric intensity.

The standard deviation  $s_y$  of the residuals is:

$$s_y = \sqrt{[(S_{yy} - m^2 S_{xx}) / (N - 2)]}, \text{ from which } s_y = 5.830 \times 10^{-10} \text{ A.}$$

The standard deviation ( $s_m$ ) of the slope ( $m$ ) is:

$$s_m = s_y / \sqrt{S_{xx}}, \text{ from which } s_m = 5.049 \times 10^{-3} \text{ A/M.}$$

The standard deviation ( $s_b$ ) of the intercept ( $b$ ) with the analytical signal axis is:

$$s_b = s_y \sqrt{[1 / (N - (\Sigma x_i)^2 / \Sigma x_i^2)]}, \text{ from which } s_b = 4.888 \times 10^{-10} \text{ A.}$$

The equation for the least squares line is:

$$y = mx + b, \text{ from which } y = (0.122 \text{ A/M})x + 1.354 \times 10^{-8} \text{ A.}$$

The coefficient of determination ( $R^2$ ) is:

$$R^2 = (\Sigma x_i y_i - \Sigma x_i \Sigma y_i / N)^2 / \{ [\Sigma x_i^2 - (\Sigma x_i)^2 / N] [\Sigma y_i^2 - (\Sigma y_i)^2 / N] \}$$

which, after appropriate substitution, gives  $R^2 = 0.997$ .

Final Pb (II) concentration in electrolytic cell result appears, according to the standard addition method, as

$$0 = (0.122 \text{ A/M})x + 1.354 \times 10^{-8} \text{ A} \leftrightarrow x = 1.111 \times 10^{-7} \text{ M.}$$

The standard deviation ( $s_c$ ) for the analytical results obtained with the standard addition method is given by (Miller and Miller, 1988):

$$s_c = (s_y / m) \sqrt{[(1/N) + (\Sigma y_i / N)^2 / (m^2 S_{xx})]}, \text{ from which } s_c = 8.181 \times 10^{-9} \text{ M.}$$

$s_c = (s_y / m) \sqrt{[(1/L) + (1/N) + (y_c - \Sigma y_i / N)^2 / (m^2 S_{xx})]}$  is the equation for the standard deviation of the analyte estimate for a calibration curve. In the case presented,  $y_c$  equals zero, since the result stems from the intercept of the adjusted line with the horizontal axis, and  $L$ , the number of replicate measures of the sample, appears as infinite, as if by extrapolating the adjusted line, thus making  $1/L$  equal to zero.

**Blank solutions.** Table 4 shows averages for the blank and successive calculations for the least squares treatment. The quantities  $S_{xx}$ ,  $S_{yy}$ ,  $S_{xy}$  that are required for LSM may be calculated in a manner analogous to that outlined above to give the following values:

$$S_{xx} = 133.3 \times 10^{-16} \text{ M}^2$$

$$S_{yy} = 47.29 \times 10^{-16} \text{ A}^2$$

$$S_{xy} = 79.38 \times 10^{-16} \text{ MA}$$

$$m = 0.595 \text{ A/M}$$

$$b = 8.584 \times 10^{-9} \text{ A}$$

$$s_y = 10.97 \times 10^{-10} \text{ A}$$

$$s_m = 9.497 \times 10^{-3} \text{ A/M}$$

$$s_b = 9.194 \times 10^{-10} \text{ A.}$$

The equation for the least squares line is thus  $y = (0.595 \text{ A/M})x + 8.584 \times 10^{-9} \text{ A}$ , and  $R^2 = 0.999$ . The final Pb (II) concentration in

**Table 4.** Data obtained from the blank using the least squares method

$x_i^a$ ( $10^{-8}$ M)	$y_i^b$ ( $10^{-8}$ A)	$x_i^2$ ( $10^{-16}$ M <sup>2</sup> )	$y_i^2$ ( $10^{-16}$ A <sup>2</sup> )	$x_i y_i$ ( $10^{-16}$ MA)
0.000	0.774	0.000	0.600	0.000
5.213	4.067	27.18	16.54	21.20
10.38	7.082	107.7	50.15	73.51
15.49	10.01	239.9	100.3	155.2
$\Sigma 31.08$	$\Sigma 21.94$	$\Sigma 374.9$	$\Sigma 167.6$	$\Sigma 249.9$

<sup>a</sup>  $x_i$  = standard concentration of lead (II).

<sup>b</sup>  $y_i$  = mean corrected electric intensity.

the electrolytic cell result appears, according to the standard addition method, as

$$0 = (0.595 \text{ A/M})x + 8.584 \times 10^{-9} \text{ A} \leftrightarrow x = 1.442 \times 10^{-8} \text{ M},$$

and the standard deviation ( $s_c$ ) for the analytical results obtained with the standard addition method is  $s_c = 1.734 \times 10^{-9} \text{ M}$ .

### New algorithm

Linear regression, as performed using Excel 2003<sup>®</sup>, is a fairly easy way to treat data from the standard addition method, and is used in laboratories throughout the world. Nevertheless, the Excel<sup>®</sup> package used to execute these calculations has no way of directly attaining SD for the slope,  $y$  intercept and  $x$  estimates, hence the need to develop an alternative way of obtaining such values. One can see that, in order to obtain such estimates, several slopes,  $y$  intercepts and  $x$  estimates are needed. However, the question remains how to attain these values with data from one sample. The problem was solved using all available data instead of just the  $i$  means for each concentration. The corrected electric current, compensated for dilution effects, was used.

The first step for the resolution of the problem was to verify whether every measurement was equally likely because none of them differ much from the others. That is why any combination of four values for  $y$ , each stemming from its group of three replicates, should produce an adjusted line as favourable as the one resulting from its averages.

The data set, composed of three replicates of the corrected electric intensity for each of the four standard concentrations, can be combined in 81 different ways ( $3^4$ ). This is shown in Fig. 1, where  $a$ ,  $b$  and  $c$  are current measurements for the null addition,  $d$ ,  $e$  and  $f$  are current measurements for the first addition,  $g$ ,  $h$  and  $i$  are current measurements for the second addition and  $j$ ,  $k$  and  $l$  are current measurements for the third addition. Tables 5 and 6 show these values, organised in the same order of the first matrix of Fig. 1, respectively, for the cork solution and for the blank. This algorithm shows all possible combinations of values able to adjust a line by LSM, without repetition or change of column in such a way that all lines will be different and will point to a different estimate of the lead (II) content of the cork sample after blank subtraction. The algorithm is absolutely general and expresses itself as  $r^s$  final equations,  $r$  being the number of repetitions of the measurement of the analytical signal and  $s$  the number of standard additions, including the null addition.

The substitution of each letter in the matrix at the top of Fig. 1 by the numbers inside Tables 5 or 6, written in the same relative position, followed by the routine described in Fig. 1, will produce 81 sets of four  $i$  values, which will be the dependent random variable in each of the 81 graphs. The independent variable will be the concentration, which takes the values 0.000, 5.213, 10.38 and  $15.49 \times 10^{-8} \text{ M}$ , as previously presented.

It is possible to write all 81 equations for the cork solution and the blank according to the combinatory scheme in Fig. 1, starting from the data inside Table 5 or 6, to which the scheme presented

$a$	$d$	$g$	$j$
$b$	$e$	$h$	$k$
$c$	$f$	$i$	$l$

  

$a$  $d$  $g$   
 $a$  $d$  $g$  $k$   
 $a$  $d$  $g$   
 $a$  $d$  $h$  $j$   
 $a$  $d$  $h$  $k$   
 $a$  $d$  $h$  $l$   
 $a$  $d$  $j$   
 $a$  $d$  $j$   
 $a$  $d$  $k$   
 $a$  $d$  $l$   
 $a$  $e$  $g$  $j$   
 $a$  $e$  $g$  $k$   
 $a$  $e$  $g$   
 $a$  $e$  $h$  $j$   
 $a$  $e$  $h$  $k$   
 $a$  $e$  $h$  $l$   
 $a$  $e$  $j$   
 $a$  $e$  $k$   
 $a$  $e$  $l$   
 $a$  $g$  $j$   
 $a$  $g$  $k$   
 $a$  $g$  $l$   
 $a$  $h$  $j$   
 $a$  $h$  $k$   
 $a$  $h$  $l$   
 $a$  $j$   
 $a$  $k$   
 $a$  $l$

$b$  $d$  $g$   
 $b$  $d$  $g$  $k$   
 $b$  $d$  $g$   
 $b$  $d$  $h$  $j$   
 $b$  $d$  $h$  $k$   
 $b$  $d$  $h$  $l$   
 $b$  $d$  $j$   
 $b$  $d$  $j$   
 $b$  $d$  $k$   
 $b$  $d$  $l$   
 $b$  $e$  $g$  $j$   
 $b$  $e$  $g$  $k$   
 $b$  $e$  $g$   
 $b$  $e$  $h$  $j$   
 $b$  $e$  $h$  $k$   
 $b$  $e$  $h$  $l$   
 $b$  $e$  $j$   
 $b$  $e$  $k$   
 $b$  $e$  $l$   
 $b$  $g$  $j$   
 $b$  $g$  $k$   
 $b$  $g$  $l$   
 $b$  $h$  $j$   
 $b$  $h$  $k$   
 $b$  $h$  $l$   
 $b$  $j$   
 $b$  $k$   
 $b$  $l$

$c$  $d$  $g$   
 $c$  $d$  $g$  $k$   
 $c$  $d$  $g$   
 $c$  $d$  $h$  $j$   
 $c$  $d$  $h$  $k$   
 $c$  $d$  $h$  $l$   
 $c$  $d$  $j$   
 $c$  $d$  $j$   
 $c$  $d$  $k$   
 $c$  $d$  $l$   
 $c$  $e$  $g$  $j$   
 $c$  $e$  $g$  $k$   
 $c$  $e$  $g$   
 $c$  $e$  $h$  $j$   
 $c$  $e$  $h$  $k$   
 $c$  $e$  $h$  $l$   
 $c$  $e$  $j$   
 $c$  $e$  $k$   
 $c$  $e$  $l$   
 $c$  $g$  $j$   
 $c$  $g$  $k$   
 $c$  $g$  $l$   
 $c$  $h$  $j$   
 $c$  $h$  $k$   
 $c$  $h$  $l$   
 $c$  $j$   
 $c$  $k$   
 $c$  $l$

**Figure 1.** The method used to obtain 81 regression lines ( $3^4$ ). The columns shown in the first matrix are the corrected electric intensities measured for each of the four concentrations measured. The first column is for the null addition, the second column is for the first addition, the third column is for the second addition and the fourth column is for the third addition. By choosing one measurement from each concentration (i.e. one value from each column), a particular four-point data set for adjusting the corresponding regression line is obtained. There are 81 ways (which are listed in the three boxes) of constructing these four-point data sets.

**Table 5.** Cork *i* data positioned as in the first matrix of Fig. 1

Cork, <i>i</i> ( $10^{-8}$ A)			
Addition 0	Addition 1	Addition 2	Addition 3
1.463	1.932	2.564	3.323
1.349	1.961	2.624	3.237
1.372	1.952	2.543	3.289

**Table 6.** Blank *i* data positioned as in first matrix of Fig. 1

Blank, <i>i</i> ( $10^{-8}$ A)			
Addition 0	Addition 1	Addition 2	Addition 3
0.780	4.063	7.016	9.880
0.774	4.069	7.052	10.22
0.769	4.069	7.178	9.950

in Fig. 1 is applied. Standard concentration sets are constant and are shown in the first column of Table 2.

Several analytical parameters were determined from the 24 current intensities measured, as already presented, and the results are shown in Table 7. The main result is the finding of an expanded uncertainty (95% confidence level) in the measurement according to ISO (1995) of  $U = 0.332 \mu\text{g/g}$ . Relative standard deviation falls from approximately 9 to 6%. Current data treatment allowed for the determination of the standard deviation of the coefficient of determination ( $R^2$ ), making the standard deviation of  $R^2$  a new analytic parameter. It also allowed for the estimation of a better limit of detection, falling from approximately  $4.7 \mu\text{g/g}$  to roughly  $3.9 \mu\text{g/g}$ .

In this manner, the uncertainty measurement in lead (II) concentration in cork can be measured using all available data instead of just the averages, which are unable to achieve the results presented here. Table 7 compares the values obtained using the new method presented in this paper with those from the older method.

In fact, one can see that, if all possible results from all data collected are treated, as has been done, their standard deviation is much smaller than the standard deviation predicted by least squares, and that the estimation of lead (II) concentration in the sample also differs (a difference of more than  $0.01 \mu\text{g/g}$  was found in our case study).

Phytochemical analysis is of great importance throughout the world, particularly in the agro-industry, but quality certification processes follow standard procedures that are not always fully understood by those who exercise them, thus making errors more likely to occur. The simplification of approach is the best way to improve performance. In fact, the present approach was developed based on a serious doubt as to how to calculate standard deviations of slope,  $xx$  and  $yy$  axis intercepts if least squares gave only one value for the slope and the  $xx$  and  $yy$  intercepts. It would require several measurements in order to calculate the averages and standard deviations. After all the equations for least squares are known (Miller and Miller, 1988), the knowledge of the foundations on which the estimation of those standard deviations is based would be complete. Nevertheless, if all least square

procedures commence by gauging the differences between  $y$  measured and  $y$  estimated, then the achievement of standard deviations should be possible through the study of all possible lines adjustable to experimentally measured  $yy$  data, specifically when it is possible to have a considerable number of repetitions of  $y$  values for each  $x$  value. The case studied above shows how to perform such an approach after three repetitions of each of the four  $y$  values, producing 81 lines. The variabilities of the slope and axis intercepts are revealed to be smaller than those estimated by least square equations, thus presenting a new way of estimating it which generates improved results.

The lead/cork case reported here can be used as a worthy model when alerting researchers to the possibility of significant errors inherent in the addition of all experimental parameters. Although some of these parameters are not yet definable by hydrodynamics or electroanalytical chemistry, it is assumed that they are quasi-linear in their empirical use, but this clearly involves some risk in the final result. The new method proposed warns of the need for a full understanding of the art of analytical science beyond the mere mechanical operation of equipment, with critical insight into the statistical procedure to be adopted for interpreting the results of the analysis of traces.

This method allows for the estimation of known analytical parameters after one single chemical analysis. The advantages of such an approach must be compared with its disadvantages; therefore this discussion will conclude with an enumeration of the main advantages and disadvantages.

The main advantage, and the underlying strength of the new approach, is to include all of the analytical signals. Methods are known for the elimination of extreme values, but there are no methods for the inclusion of coherent values in their absolute state, only their averages. In relation to the elimination of the extreme or incoherent values, one must refer to a criterion for the rejection of incoherent values: the  $Q$  test

assumes that the distribution of populational data is normal. [Nevertheless], this condition cannot be proven as true or false for distributions with less than 50 results. (...) This is why J. Mandel, talking about a small set of data, writes: "Those that believe that they can reject observations using an exclusion criterion integrated in a statistical rule for the rejection of incoherent values are just deluding themselves." (...) In the end, the only valid reason for the rejection of a result within a small group of data is the knowledge that some mistake was done during the measuring. Without this knowledge it is advised a careful approximation to the exclusion of values of incoherent appearance [for that a statistical rule is just an important auxiliary]. (Skoog *et al.*, 1988).

Using just the averages instead of all the information in the data set is also a process of exclusion but, contrary to the exclusion of outliers or incoherent data, it excludes quite coherent measurements. This clearly changes the final result, mostly its uncertainty, as is shown in the above example, and must be avoided at all times. The simplest way to do this is through the application of the new method, which considers all analytical signal data equally.

Also relevant is the possibility of gauging the uncertainty of measuring of the analytical signal for each set of such measurements. In the above practical application, 24 signals were needed and an exact estimation of the expanded uncertainty related to the measuring of the analytical signal variable was achieved.

The present approach also allows for the estimate of a minimum value for total uncertainty in an analysis from only one

**Table 7.** Analytical parameters

		New algorithm		Old algorithm	
		Cork	Blank	Cork	Blank
<i>m</i> (A/M)	<i>avg</i>	0.122	0.595	0.122	0.595
	<i>s</i>	0.004	0.009	0.005	0.010
<i>b</i>	<i>avg</i> (10 <sup>-8</sup> A)	1.354	0.858	1.354	0.859
	<i>s</i> (10 <sup>-10</sup> A)	3.632	3.045	4.914	9.232
<i>R</i> <sup>2</sup>	<i>avg</i>	0.995	0.999	0.997	0.999
	<i>s</i>	0.003	0.000	n.a. <sup>a</sup>	n.a.
[Pb (II)]	<i>avg</i> (10 <sup>-8</sup> M)	11.13	1.443	11.12	1.444
	<i>s</i> (10 <sup>-10</sup> M)	61.44	7.040	82.35	17.43
	<i>u<sub>n</sub></i> <sup>b</sup> (10 <sup>-10</sup> M)	6.826	0.782	n.a.	n.a.
LOD (10 <sup>-8</sup> M)		1.654 (3.998 µg/g)		1.967 (4.755 µg/g)	
LOL (10 <sup>-8</sup> M)		15.49		15.49	
CV (%)		5.521	4.879	7.403	12.07
<i>N</i>		81		1	
[Pb (II)] <sub>t</sub>	<i>avg</i> (10 <sup>-8</sup> M)	9.685		9.680	
	<i>s<sub>t</sub></i> <sup>c</sup> (10 <sup>-10</sup> M)	61.84		84.17	
	** <i>U</i> (10 <sup>-10</sup> M)	13.74		n.a.	
<i>D</i> <sup>d</sup>		241.7		241.7	
<i>C<sub>mass</sub></i> <sup>e</sup> (µg/g)	<i>avg</i>	23.41		23.40	
	<i>s<sub>t</sub></i> <sup>c</sup>	1.495		2.035	
	CV <sub>s</sub> (%)	6.385		8.696	
	<i>U</i> <sup>b</sup>	0.332		n.a.	
	CV <sub>U</sub> (%)	1.419		n.a.	
Δ (µg/g)				0.013	
Δ <sub>r</sub> (%) <sup>f</sup>				0.000(1)	

The line produced by the least squares method is applied to the standard addition method: *b* is the intercept with the *yy* axis, *m* is its slope (method sensitivity), *R*<sup>2</sup> is the coefficient of determination, *avg* is the average, *s* is the absolute standard deviation, *LOD* is the limit of detection, i.e. 'blank plus three times standard deviation of the blank' (Skoog and Leary, 1992), *LOL* is the limit of linearity and *CV* is the coefficient of variation, the ratio *s/avg* × 100 or *U/avg* × 100. Number of values (*N*) = 81 for each—cork and blank—for new algorithm and 1 for the old algorithm

<sup>a</sup> n.a. = not applicable.

<sup>b</sup> Partial uncertainty,  $u_n = s_n/\sqrt{n}$ ; combined standard uncertainty,  $u = \sqrt{(u_1^2 + u_2^2)}$ ; expanded uncertainty,  $U = u \times 2$  (\*\* confidence level of 95%) (ISO, 1995).

<sup>c</sup>  $s_t = \sqrt{(s_{\text{cork}}^2 + s_{\text{blank}}^2)}$ .

<sup>d</sup> 21 × 5 ml × 207.2 g/mol/0.09 g.

<sup>e</sup> Mass-based concentration [µg Pb (II)/g of cork (interior of the bark of *Quercus suber* L.)].

<sup>f</sup>  $\Delta_r (\%) = (\text{new algorithm} - \text{old algorithm})/\text{old algorithm} \times 100$ .

sample, as well as the characterisation of each analysis using several usual parameters estimated from a single sample. Since this approach does not need equations for the standard deviations of the slope and of the *yy* and *xx* axis intercepts, it can be implemented in Excel<sup>®</sup> much more easily using the line adjustment routines. Uncertainty can now be calculated.

Several disadvantages were, however, encountered: The first is that, for a constant number of repetitions of the measurement of the analytical signal, the number of values increased exponentially with the number of concentrations used in the standardisation process (a large number of values can therefore be achieved, even though this disadvantage is currently of little interest due

to the use of computers). This method using LSM is also inadequate when any of the assumptions of the method are not fulfilled, specifically when autocorrelation of the analytical signal appears. It is also very important to point out that the purpose of the method is to study a single sample. Using it to avoid the repetition of the measurement of as many samples as needed would be a real scientific disadvantage, even though the production of data that this new method permits can be used to simulate a great number of sample repetitions.

It is well understood that *t* statistics are being used as if averaged final concentration resulted from between 50 and 100 values (*t* = 2), which is true for the new but not for the older

approach. Nevertheless, the equation for the standard deviation of the estimated concentration by the standard addition method differs from that of the calibration curve because estimation pretends that an infinite number of analytical signal measures were performed ( $1/L$  equals zero). This would lead to the logical conclusion that uncertainty of the previous approach would be zero. The non-application of this conclusion is advised.

The new approach shows a different result by including all current intensity measurements. The new approach does not use the least squares equation for the standard deviation of the results. This equation produces a standard deviation of the results based on points of the line that are not treated as averages (with their standard deviation), but simply as points of the line. That is why the older approach, in which the standard deviation of the result is based solely on the differences between the least square estimation of the points of the line and their averages, is shown as being poor for the production of a final result. In the new method the standard deviation also uses the differences between original analytical signal measurements (resulting in their standard deviations). In the end, repeatability of analytical signals measurement and its relevance to the precision of the estimate of the lead (II) concentration are the present work's most important issues. All the results were presented with four significant digits, but the intermediary computations used as many digits as the Excel<sup>®</sup> allowed (30). Alternative ways to show how

standard deviations of medium points (used to construct an adjusted line) would affect the lead (II) concentration estimate were investigated.

### Acknowledgements

Thanks are due to (chronological order): grant number SFRH/BD/31420/2006 from the Foundation for Science and Technology of the Ministry of Science, Technology and Higher Education of Portugal; Russell Alpizar-Jara, from the Department of Mathematics, University of Évora; David Zellmer, from the Department of Chemistry, California State University, Fresno; Victor M. M. Lobo, from the Department of Chemistry, University of Coimbra; Carlos Braumann, from the Department of Mathematics, University of Évora.

### References

- ISO, 1995. *Guide to the Expression of Uncertainty in Measurements*. International Standard Organization: Geneva.
- Miller J, Miller J. 1988. *Statistics for Analytical Chemistry*. Ellis Horwood: Chichester.
- Skoog DA, Leary JJ. 1992. *Principles of Instrumental Analysis*. Saunders College: New York.
- Skoog D, West D, Holler F. 1988. *Fundamentals of Analytical Chemistry*. Saunders College: New York.