ASSESSING THE ACCURACY OF ANALYTICAL METHODS USING LINEAR REGRESSION WITH ERRORS IN BOTH AXES

Jordi Riu* and F. Xavier Rius

Departament de Química. Universitat Rovira i Virgili.
Pl. Imperial Tàrraco, 1. 43005-Tarragona. CATALONIA, SPAIN.

17 Pages, 2 Tables, 3 Figures

Please address correspondence to:

Jordi Riu
Department of Chemistry,
Universitat Rovira i Virgili de Tarragona
Pl. Imperial Tàrraco, 1
E - 43005 TARRAGONA
CATALONIA, SPAIN

Tel.: 34 - 77 - 558187
Fax.: 34 - 77 - 559563
e-mail: rusell@quimica.urv.es

ABSTRACT

In this paper a new technique for assessing the accuracy of analytical methods using linear regression is reported. The results of newly developed analytical methods are regressed against the results obtained using reference methods. The new test is based on the joint confidence interval for the slope and the intercept of the regression line, which is calculated taking the uncertainties in both axes into account. The slope, intercept and variances which are associated to the regression coefficients are calculated with Bivariate Least Squares regression, BLS. The new technique was validated using three simulated and five real data sets. The Monte Carlo method was applied to obtain 100,000 data sets for each of the initial simulated data sets to show the correctness of the new technique. The application of the new technique to five real data sets enables differences to be detected between the results of the joint confidence interval based on the BLS method and the results of the commonly used tests based on Ordinary Least Squares or Weighted Least Squares regression.

KEY WORDS

Errors in both axes, joint confidence interval, uncertainties, linear regression.

INTRODUCTION

Assessing accuracy is a fundamental step in the method validation process. The analyte concentration value obtained with the new method is often compared with the reference method by replicating measurements and applying a t-test or an F-test. But if the validity of the new method is checked with a range of analyte concentrations, the linear regression also gives additional statistical information such as the presence of proportional errors, the need of including a blank correction and the calculation of confidence intervals for the regression coefficients.

The statistical test which compares the intercept and slope values obtained by linear calibration with the theoretical values of zero and unity, bearing in mind the correlation between the two regression coefficients, was applied by Mandel and Linnig[1] to analytical results. This procedure, extensively used up to now, is based on the linear regression hypotheses being fulfilled by Ordinary Least Squares (OLS) or, whenever heteroscedasticity is present in the dependent variable, by Weighted Least Squares, WLS.[2] But when applied to method comparison, this procedure has the drawback that it regards the reference method (usually represented on the abscissa axis) as not only being free of systematic errors but of random ones as well. This reference method often includes random errors of the same order of magnitude as the new method to be validated and, as a result, the bibliography is full of methods which are considered to be correct but which may contain systematic errors.

There are other approaches[3] which apply a type of weighted regression to calculate the regression coefficients taking into account errors in both axes. The individual confidence intervals of the regression coefficients are then applied to assess the accuracy. This approach, as well as leading to biased values of the regression coefficients, does not take into account the covariance between the slope and the intercept and, consequently, leads to erroneous final results.

This study proposes applying a joint confidence test for the intercept and the slope to assess the accuracy of new analytical methods. These regression coefficients are calculated here by applying calibration methods which consider errors in both axes, (Bivariate Least Squares, BLS), and which, therefore, take into account the uncertainty in the results which both methods may have.

To show the goodness of the new validation method, three simulated and five real data sets have been used. Random errors were added to the simulated data sets using the Monte Carlo method. The values obtained with the new approach are shown to agree with the theoretical results expected, while the results of using the joint confidence test based on OLS, WLS or BLS for the real data sets can lead to different conclusions about the correctness of the validation.

The main limitation of calibration methods which include errors in both axes is that the uncertainty associated to the values on both the x and the y-axis has to be estimated, which often means that data analysis will take longer.

BACKGROUND AND THEORY

Notation. The estimated regression coefficients of the calibration line will be denoted as a 1(intercept) and b2(slope), while the coefficients corresponding to the experimental calibration data pairs are denoted as a and b. The column matrix $\boldsymbol{b}$3is defined by the estimated regression coefficients, while the column matrix **b** is defined by the experimental values. The error, measured in terms of variance, for the set of n experimental data points (x_i,y_i) will be denoted as s^2, and its estimated value is defined by eq 1:[4-6]

$$\hat{s}^2 = \frac{1}{(n-2)}\sum_{i=1}^{n}\frac{1}{w_i}(y_i - \hat{y}_i)^2 = \frac{1}{(n-2)}\sum_{i=1}^{n}\frac{1}{w_i}(y_i - \hat{a} - \hat{b}x_i)^2 \qquad (1)$$

where w_i is the weighting factor for each data point and _i is the estimated value for the y predicted.

Bivariate Least Squares (BLS). Bivariate Least Squares is the generic name for a set of techniques used for regressing bivariate data, i.e. whenever a regression method is applied to data containing errors in both axes. Of all the different existing approaches for calculating regression coefficients and related statistical parameters considering errors in both axes,[7] Lisý's method[8] was selected because of its speed in estimating the correct results for the regression coefficients, because the variance-covariance matrix (which is useful in the subsequent development of statistical tests for assessing accuracy) is obtained with no extra effort, and because of the simplicity of programming its algorhythm. The method consists of minimizing the sum of the weighted residuals, S, expressed in eq 2:

$$S = \sum_{i=1}^{n}\frac{(y_i - \hat{y}_i)^2}{w_i} = (n-2)\hat{s}^2 \qquad (2)$$

This method uses the variance of the residuals ($s_{\varepsilon_i}^2$ 6), which can be expressed using the Taylor series, as a weighting factor even when the covariance between variables for each data pair is not zero, eq 3:

$$s_{\epsilon_i}^2 = w_i = s_{y_i}^2 + \hat{b}^2 s_{x_i}^2 - 2\hat{b}\,\mathrm{cov}(x_i,y_i) \tag{3}$$

where$s_{x_i}^2$ 8and$s_{y_i}^2$ 9respectively stand for the variances of each (x_i,y_i) individual data point.

By minimizing the sum of the weighted residuals in relation to the slope and the intercept, two non-linear equations are obtained and by including the partial derivatives of the squared residuals, eq 4 and the equivalent eq 5 can be written in matrix form:

$$\mathbf{R}\mathbf{b} = \mathbf{g} \tag{4}$$

$$\begin{bmatrix} \sum_{i=1}^{n} \frac{1}{s_{\epsilon_i}^2} & \sum_{i=1}^{n} \frac{x_i}{s_{\epsilon_i}^2} \\ \sum_{i=1}^{n} \frac{x_i}{s_{\epsilon_i}^2} & \sum_{i=1}^{n} \frac{x_i^2}{s_{\epsilon_i}^2} \end{bmatrix} \times \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \left(\frac{y_i}{s_{\epsilon_i}^2} + {}^1/_2 \left(\frac{R_i}{s_{\epsilon_i}^2} \right)^2 \frac{\partial s_{\epsilon_i}^2}{\partial \hat{a}} \right) \\ \sum_{i=1}^{n} \left(\frac{x_i y_i}{s_{\epsilon_i}^2} + {}^1/_2 \left(\frac{R_i}{s_{\epsilon_i}^2} \right)^2 \frac{\partial s_{\epsilon_i}^2}{\partial \hat{b}} \right) \end{bmatrix} \tag{5}$$

The slope and the intercept, which are components of vector$\boldsymbol{b}$12in eq 4 and eq 5, can be determined by carrying out an iterative process on the following matrix form, eq 6:

$$\mathbf{b} = \mathbf{R}^{-1}\mathbf{g} \tag{6}$$

With this method, and assuming that the straight line model is the correct one, the variance-covariance matrix of the calibration straight line coefficients, **B**, is obtained by multiplying the final matrix $\mathbf{R}^{-1}$ by the experimental error,[5-6] s^2. As the experimental error is unknown, the estimated value,$s^2$14, expressed in eq 1 should be used.[2]

It should be pointed out that if the situation were to be $s_{\varepsilon_i}^2$ 15= w_i = 1 (i.e, all errors are due to the experimental measurement in the ordinate axis) the expressions obtained would be the same as if the Ordinary Least Squares method were to be applied.

Joint confidence interval. When two analytical methodologies are compared using linear regression, the plot of the values obtained from samples which have different concentrations of the analyte and which are analysed by the two methods, should give a straight line of approximately unity slope and zero intercept if the results are not statistically different at a given level of significance. In general[9] the confidence region -joint confidence interval- for the straight line regression coefficients corresponds to their quadratic distribution, and is given by eq 7:

where $\chi^2_{(1-\gamma)}$ is the 1-γ level of the χ^2 distribution with two degrees of freedom and r_{ij} are the

$$\sum_{i=1}^{2}\sum_{j=1}^{2} r_{ij}(\hat{\mathbf{b}}_i - \mathbf{b}_i)(\hat{\mathbf{b}}_j - \mathbf{b}_j) = s^2\chi^2_{(1-\gamma)} \tag{7}$$

elements of the **R** matrix.

Eq 7 gives an ellipsoid which is defined by the regression coefficients and which has its centre at **Error!**17.

The experimental error, s^2, which is only known through its estimate,$s^2$18, appears together with the χ^2 distribution in eq 7. From the expression of the experimental error, eq 1, it can be shown[10] that:

$$\sum_{i=1}^{n} w_i^{-1}(y_i - \hat{y}_i)^2 = s^2\chi^2_{(1-\gamma)} \tag{8}$$

where eq 8 is independent of eq 7. w_i stands for the weighting factor used, and $\chi^2_{(1-\gamma)}$ is the 1-γ level of the χ^2 distribution with n-2 degrees of freedom.

The corrected ratio of the two χ^2-distributions has an F distribution with two and n-2 degrees of freedom:[9]

$$F = \frac{\sum_{i=1}^{2}\sum_{j=1}^{2} r_{ij}(\hat{\mathbf{b}}_i - \mathbf{b}_i)(\hat{\mathbf{b}}_j - \mathbf{b}_j)}{\sum_{i=1}^{n} w_i^{-1}(y_i - \hat{y}_i)^2} \frac{n-2}{2} = \frac{\sum_{i=1}^{2}\sum_{j=1}^{2} r_{ij}(\hat{\mathbf{b}}_i - \mathbf{b}_i)(\hat{\mathbf{b}}_j - \mathbf{b}_j)}{2\hat{s}^2} \tag{9}$$

The joint confidence test applied by Mandel and Linnig[1] consists of checking the presence of the theoretical point (0,1) within the limits of the joint confidence region spanned by eq 9 when the parameters of the regression line are calculated using the OLS method. However, as has been pointed out earlier, this expression does not take into account the fact that in the comparison of analytical methodologies the errors in both axes are comparable and so OLS cannot be applied without the possibility of committing considerable errors when assessing accuracy.

The estimated variance-covariance matrix obtained using the Lisý et al. method enables the r_{ij} coefficients to be determined taking into account the errors in both methodologies. The r_{ij} coefficients are the elements of the **R** matrix in eqs 4 and 5. By introducing these coefficients into eq 7 and by taking into account the corrected ratio of the two χ^2-distributions, eq 9, it is possible to develop the joint confidence distribution for the intercept and the slope taking into account the errors in both methodologies:

$$[\hat{a} - a \quad \hat{b} - b] \times \begin{bmatrix} \sum_{i=1}^{n} \frac{1}{s_{\epsilon_i}^{\ 2}} & \sum_{i=1}^{n} \frac{x_i}{s_{\epsilon_i}^{\ 2}} \\ \sum_{i=1}^{n} \frac{x_i}{s_{\epsilon_i}^{\ 2}} & \sum_{i=1}^{n} \frac{x_i^2}{s_{\epsilon_i}^{\ 2}} \end{bmatrix} \times \begin{bmatrix} \hat{a} - a \\ \hat{b} - b \end{bmatrix} = 2\hat{s}^2 F_{1-\alpha(2,n-2)} \tag{10}$$

$$\sum_{i=1}^{n}\frac{1}{s_{\epsilon_i}^2}(\hat{a}-a)^2 + 2\sum_{i=1}^{n}\frac{x_i}{s_{\epsilon_i}^2}(\hat{a}-a)(\hat{b}-b) + \sum_{i=1}^{n}\frac{x_i^2}{s_{\epsilon_i}^2}(\hat{b}-b)^2 = 2\hat{s}^2 F_{1-\alpha(2,n-2)} \tag{11}$$

where $F_{1-\alpha(2,n-2)}$ is the tabulated F-value at a significance level of α with two and n-2 degrees of freedom.

The limits of the ellipse depend on the experimental errors and on the significance level chosen, and its tilt, as in the case of OLS, is the result of the well known correlation between the slope and the intercept. The ellipsoid which defines the region of the space for the joint confidence test has its axes orientated in the direction of each of the two $\mathbf{V_i}$ eigenvectors of the R matrix. The length of the individual semiaxes is the same as $\sqrt{2\hat{s}^2 F_{1-\alpha,(2,n-2)}} \times \sqrt{\lambda_i}$23, where λ_i is each of the two eigenvalues of the **R** matrix.[11]

Validation process. The objective of the validation process is to assess whether the joint confidence interval test based on the BLS regression technique provides correct results, i.e. new methodologies which show no statistical differences with respect to the reference method at the level of significance chosen must be accepted and new methods which provide results that differ statistically from the results obtained using the reference method must be rejected. It will also be shown that, for several cases, the joint confidence interval test based on ordinary least squares (OLS) or weighted least squares (WLS) methods provides results which significantly differ from the ones obtained with the joint confidence interval test based on the BLS technique.

In order to assess the correctness of the test, a study will be made of three simulated and five real data sets for which uncertainties are considered in both axes and data pairs are produced with and without homoscedasticity which are differently distributed throughout the calibration range. The first two simulated data sets give rise to a calibration line characterized by a zero intercept and a unity slope. The third set gives a line in which there is a significant difference between the result of the method on the abscissa and the result on the ordinate axis, the slope being 1.1 and the intercept zero. The Monte Carlo method [12,13] was used to generate 100,000 different data sets for each original regression data set. So, a random error was added to each value of the data pairs giving rise every time to 100,000 regression lines. The validation involves checking whether the method for obtaining the joint confidence interval with errors in both axes, BLS, for an α significance level gives results such that the theoretical point zero intercept and unity slope falls into the joint confidence region in (1-α)% of cases. These results are compared with the ones obtained in a similar way using OLS and WLS methods.

The joint confidence interval tests based on the OLS, WLS and BLS methods were applied to the original data sets and it was confirmed that, in some cases, results are different.

EXPERIMENTAL SECTION

Data sets and software. The three simulated and five real data sets below were used to validate the methodology. In all cases, the three calibration lines obtained from applying the BLS, OLS and WLS methods are shown in Figures 1 and 2.

Data set 1: Homoscedastic data set containing 20 pairs of x,y values equally distributed within the range 2-40 on both axes. Standard deviations of 1 were considered for all x and y values (Figure 1a).

Data set 2: 20 data pairs randomly distributed throughout the range between 100 and 900 units. Heteroscedasticity with random standard deviation is present, in such a way that the standard deviation of any one point is higher than the standard deviation of the previous point. The standard deviation is at most 25% of the individual point (Figure 1b).

Data set 3: 20 data pairs with the x-values equally distributed within the range 1-20 and the y-values 10% higher than each corresponding x-value. The theoretical values of the regression coefficients are a=0, b=1.1. Heteroscedasticity with random standard deviation is present, in such a way that each standard deviation is at most 9% of each individual value (Figure 1c).

Data set 4: a comparative study of methods for analysing Ca^{2+} in waters[14] using atomic absorption spectroscopy (AAS, reference method) and the technique of sequential injection analysis (SIA, new method). The uncertainties in AAS are derived from the uncertainties associated to the calculated linear regression using a univariate linear calibration computer program[15.] As can be seen in Figure 2a there is considerable heteroscedasticity in this axis. The uncertainties in SIA, which are practically constant at all points, are calculated with a multivariate regression model and the PLS technique using the Unscrambler programme (Unscrambler-Ext. ver. 4.0, Camo A/S, Trondheim, Norway).

Data set 5: the resistance of the thicknesses (in μm) of various films of commercially available photoresists and silicon dioxide substrates is measured.[16] An ellipsometer and a Nanospec/AFT are used to get the results. The uncertainties for all measurements represent the variations in the film thicknesses across the wafer surface (Figure 2b).

Data set 6: concentrations of polycyclic aromatic hydrocarbons (PAHs) recovered from railroad bed soil after supercritical fluid extraction (SFE) using two different modifiers.[17] The standard deviations are based on triplicate supercritical fluid extraction at each point (Figure 2c).

Data set 7: percentage of recovery for several organochlorine pesticides after microwave-assisted extraction (MAE) using solvent (hexane/acetone 1:1) and solvent/soil suspensions spiked with the target compounds.[18] The standard deviations are the average of three determinations at each point (Figure 2d).

Data set 8: comparative study of atomic absorption spectroscopy (AAS) and emission spectrometry using emulsion formation (ES) in the determination of Ca in lubricant oils.[3] The standard deviations are the average of three determinations in each condition (Figure 2e).

Applying the OLS method to the different data sets does not take into account the uncertainties of each point in the x and y-axes, and only homoscedasticity with a unity standard deviation in the y-variable is considered. When applying the WLS method, the uncertainties in the x-axis are neglected and only the uncertainties in the y-axis are taken into account. In the first three data sets four levels of significance are used with α values of 0.001, 0.01, 0.05 and 0.1.

All computations were done with home-made Matlab subroutines (Matlab for Microsoft Windows ver. 4.0, The MathWorks, Inc., Massachusetts, USA). A computer program will be available from the authors shortly.[19]

RESULTS AND DISCUSSION

Data set 1. Table 1 summarizes the results of applying the joint confidence interval test derived from BLS and OLS regression techniques to the 100,000 data sets obtained by applying the Monte-Carlo simulation to data set 1. It can be seen that, very approximately, the percentage of data sets for which the theoretical point of zero intercept and unity slope falls within the region defined by the joint confidence intervals based on the BLS or OLS methods agrees with the theoretical values given by the level of significance. In this case it is evident that the same results are obtained with the OLS and WLS methods

Due to the fact that the uncertainties in both axes are homoscedastic and of the same size (s = 1 in this case) the data structure on the x-axis meets the theoretical OLS conditions. Therefore, in spite of there being uncertainties in the x-axis, the results obtained by applying the OLS method are very close to the ones obtained using BLS, with no significant differences in comparison to the theoretical values. It should be pointed out, however, that the results obtained using BLS are closer to the theoretical results than the ones obtained by OLS.

Data set 2. The results of applying BLS, OLS and WLS joint confidence interval tests for the intercept and the slope to the 100,000 data sets derived from data set 2 are summarized in Table 1. The percentage of data sets for which the theoretical point of zero intercept and unity slope falls within the region defined by the joint confidence intervals based on the BLS method agrees quite closely with the corresponding theoretical values given by the pre-established levels of significance. Similar results are obtained when using WLS conditions. However, as is to be expected from the heteroscedasticity in the y-axis, the percentage of data sets for which the theoretical point of zero intercept and unity slope fall within the region defined by the joint confidence interval based on the OLS method differ quite considerably (up to 10% with the present data set) from the theoretical values.

It should be pointed out that although there are heteroscedastic uncertainties on the abscissa axis, WLS gives good results (even though it does not take these uncertainties into account). This may be due to the fact that, in this data set, there is a constant relation between the uncertainties present in the two axes, so the heteroscedasticity considered only in the ordinate axis leads to results which are very similar to the ones obtained with the BLS method, the application of which is rigorously more correct.

Those users with data sets which have a similar structure to the ones analysed in this section are therefore compelled to use at least the joint confidence test based on WLS if they want to prevent systematic errors. The absence of readily available information and software to carry out such a test enhances the usefulness of the BLS test developed here.

Data set 3. The results of applying the BLS, OLS and WLS joint confidence interval tests for the intercept and the slope to the data sets derived from data set 3 are summarized in Table 1. Because 100,000 data sets are analysed all derived from data with a slope of 1.1 and because the points are perfectly fitted to a line, the percentage of data sets obtained using the Monte Carlo method in which the theoretical point of zero intercept and unity slope falls within the joint confidence interval is low in all three methods. This serves to confirm that if there is a significant difference between the results of the method represented on the abscissa and ordinate axis, the theoretical point of zero intercept and unity slope will also fall outside the region defined by the joint confidence interval when BLS is used.

Data set 4. The results of applying BLS, OLS and WLS methods to data set 4 are summarized in Table 2. According to the joint confidence test based on OLS conditions, AAS and SIA with multivariate detection techniques give results which are not statistically different at a significance level α=5% but this is not true for the test based on BLS and WLS regression techniques.

Figure 3a shows the joint confidence intervals obtained by applying the three regression techniques studied. It can be seen that the joint confidence interval obtained by BLS is lower than the one obtained by OLS. One factor that may have an influence is that the slopes obtained with both methods are very similar (OLS: b=0.73±0.11; BLS: b=0.74±0.10) but that the intercept obtained with OLS (a=43±18) is slightly larger than the one obtained with BLS (a=39±16). The negative correlation between the intercept and the slope means that if the data set has high or low intercepts and slopes, the theoretical point of zero intercept and unity slope is more likely to fall within the region defined by their joint confidence interval. So, OLS detects differences between the compared data sets while BLS does not.

What is more, in accordance with what could be expected of a descriptive analysis, the presence of considerable uncertainties in the x-axis makes the BLS regression line oscillate in a broader interval with the consequent broadening of the joint confidence interval. So, the ellipse obtained with BLS is larger than the one obtained with WLS which only takes into account the errors in the ordinate axis. Furthermore, in this data set there are no great differences in the values of the regression coefficients obtained by the three methods, and it is quite clear that the WLS and BLS methods give less importance to the points with most uncertainty, giving smaller ellipses in general. So, it is usually easier to find statistically significant differences between the two data sets tested.

Data set 5. The results of applying the joint confidence interval test obtained using BLS, OLS and WLS regression techniques to data set 5 are summarized in Table 2 and Figure 3b. The joint confidence interval test obtained with OLS indicates the lack of similarity between the two methods of measuring thickness but the theoretical point of zero intercept and unity slope falls into the joint confidence region for the BLS and WLS methods for a level of significance α=5%.

Although it is difficult to visually observe differences in the regression coefficients of the three regression lines in Figure 2b, the different conclusions reached using OLS and BLS/WLS techniques might be due to the different correlation between slopes and intercepts obtained using these three regression techniques. Although these differences are not considerable in absolute terms, the different values obtained for the regression coefficients are enough for ellipses to be constructed with significant differences in tilt, so giving conflicting conclusions about the similarity between measurement techniques, as can be observed in Figure 3b.

It is interesting to note the role of the correlation between the variables analysed (analytical methods compared) and their incidence in the statistical test which is clear from the last two sets of analysed data. There is a low correlation between variables in data set 4, just the opposite of data set 5. In contrast to OLS and the Mandel and Linnig test, a low correlation between variables, together with heteroscedasticity, makes it easier to find statistically significant differences between the methods studied.

Data set 6. The results of applying BLS, OLS and WLS methods to the data sets derived from data set 6 are summarized in Table 2 and Figure 3c. The effect of considerable heteroscedasticity in both axes is quite clear in this example. Three points of high uncertainty can be seen to have a strong effect on the calculations of the regression coefficients obtained with the OLS, WLS and BLS techniques. In this case, the result (Figure 3c) is that a statistically significant difference is found between the two methodologies tested with the joint confidence interval test derived from OLS and WLS but not with BLS for a level of significance α=5%. BLS clearly gives more importance to the four points with limited analytical results than to the three points which have a high degree of uncertainty. This tendency does not exist for OLS and is greatly exaggerated for WLS.

Data set 7. The conclusions reached with data set 6 are clearly confirmed in this data set. The presence of two points with high leverage but with a different degree of associated uncertainty (Figure 2d) is the reason for the clearly different behaviour of the tests based on BLS, OLS and WLS regression techniques i.e, while the joint confidence interval test based on OLS does not find statistically significant differences between the two methods compared, BLS and WLS do find these differences for a level of significance α=5%. So the presence of the point with high uncertainty means that the slope obtained by OLS is higher than unity and the intercept much smaller than zero and that the confidence intervals are much broader. These conditions make it more likely that the theoretical point of zero intercept and unity slope will fall within the region defined by the joint confidence interval. In contrast, the fact that BLS and WLS regression techniques take into account the uncertainties minimizes the effect of the aforementioned point, so slopes and intercepts are obtained which are smaller than unity and zero, respectively. These are not favourable conditions for the theoretical point of zero intercept and unity slope to fall within the region defined by the joint confidence interval (Figure 3d).

Despite the similarity in the conclusions obtained, the differences between this data set and data set 4 should be noted. Although the correlation between the variables on each axis is low in both cases, the origin of the discrepancy between the behaviour of the test based on OLS and the test based on BLS and WLS is quite different.

Data set 8. This real data set exemplifies the case in which the joint confidence interval test based on OLS or WLS regression analysis leads to results which are in conflict with the results given by the BLS regression technique (Figure 3e, Table 2). According to the latter, there is a significant statistical difference between the two methods compared whereas there is not according to the parameters obtained with OLS/WLS for a level of significance α=5%.

Although there is a slight difference in the regression coefficients obtained by the three methods which displaces the centre of the ellipse obtained by BLS and gives values which are different from the ones obtained by OLS and WLS (Figure 2e), the smaller joint confidence interval using BLS is much more important for obtaining these differing results between the three methods. It is smaller because there is a series of points with greater uncertainties than the others which means that for the WLS and BLS methods, these points are not so important when carrying out the linear calibration and so the uncertainty associated to the intercept and the slope is reduced.

CONCLUSIONS

The new joint confidence test for regression coefficients based on the BLS calibration method was applied to three simulated data sets and proved to give correct results. The new test is not always necessary for drawing accurate conclusions about method equivalency. However, as is shown by its application to five real data sets, if it is not used two overall conflictive situations may be caused:

i) the theoretical point defined by zero intercept and unity slope is located within the joint confidence region derived using the BLS method, but not using the OLS or WLS methods. This is important for those users who have developed a new method which is accurate but which it has not been recognised as such until now. The result of this is that good, new methods are being rejected if the joint confidence test based on OLS or WLS methods are used. The old method will probably still be used and the figures of merit of a new method will be unnecessarily wasted.

ii) the theoretical point defined by zero intercept and unity slope is located within the joint confidence region derived using the OLS/WLS methods, but not using BLS, i.e, there are new analytical methods that give results that do not statistically differ from the reference method using the OLS or WLS methods but they do differ if the more reliable BLS regression technique is used. In these cases, the new method is wrongly considered to be accurate when using the joint confidence test based on OLS or WLS methods, and may easily give biased results whenever it is used to analyse new samples.

The new method is for general use, with no restrictive statistical constraints. It has the additional advantage of being invariant upon switching axes for the regression coefficients. The presence of possible outliers containing considerable uncertainties has the effect of reducing the joint confidence intervals in BLS and WLS with respect to OLS. However, the user should be aware of two weaknesses; the first, general to most BLS techniques is the lack of robustness in the presence of outliers with low individual uncertainty; and the second is that the individual variances associated to each data point are needed, which may mean a longer analysis time. If these variances are omitted, the application of the BLS technique gives identical results to the ones obtained with classical OLS or WLS methods.

ACKNOWLEDGMENTS

The authors would like to thank the DGICyT (project num. BP93-0366) for financial support and the Spanish Ministry of Education and Science for providing Jordi Riu's doctoral fellowship.

LITERATURE CITED

(1)Mandel, J.; Linnig, F. J.; *Anal. Chem.* **1957**, *29*, 743-749
(2)Draper, N.; Smith, H. *Applied Regression Analysis, Second Edition*; Jonh Wiley & Sons: New York, 1981; pp 5-128
(3) De la Guardia, M.; Salvador, A.; Berenguer, V. *Ann. Quim.* **1980**, *77*, 129-132
(4)Myers, R. H. *Classical and Modern Regression with Applications (2nd Edition)*; Duxbury Press: Belmont, California, 1990; pp 18-21
(5)Wentworth, W. E. *J. Chem. Educ.* **1965**, *42*, 96-103
(6)Wentworth, W. E. *J. Chem. Educ.* **1965**, *42*, 162-167
(7)Riu, J.; Rius, F. X. *J. Chemom.* **1995**, *9*, 343-362
(8)Lisý, J. M.; Cholvadová, A.; Kutej, J. *Computers Chem.* **1990**, *14*, 189-192
(9)Mood, A. McF. *Introduction to the Theory of Statistics (2nd Edition)*; McGraw-Hill: New York, 1963; pp 112-301
(10)Scheffé, H. *The Analysis of Variance*, John Wiley & Sons, New York (1959); pp 3-41
(11)Meloun, M.; Militký, J.; Forina, M. *Chemometrics for Analytical Chemistry Vol II*; Ellis Horwood: London, 1994; pp 20-24
(12)Meier, P. C.; Zünd, R. E. *Statistical Methods in Analytical Chemistry*; John Wiley & Sons: New York, 1993; pp 145-150
(13)Güell, O.; Holcombe, J. A. *Anal. Chem.* **1990**, *60*, 529A-542A
(14)Ruisánchez, I.; Rius, A.; Larrechi, M. S.; Callao, M. P.; Rius, F. X. *Chemom. Int. Lab. Sys.* **1993**, *24*, 55-63
(15)Boqué, R.; Rius F.X.; Massart D.L. *J. Chem. Educ. (Computer Series)* **1994**, *71*, 230-232
(16)Gamsky, C. J.; Howes, G. R.; Taylor, J. W. *Anal. Chem.* **1994**, *66*, 1015-1020
(17)Langenfeld, J. J.; Hawthorne, S. B.; Miller, D. J.; Pawliszyn, J. *Anal. Chem.* **1994**, *66*, 909-916
(18)López-Avila, V.; Young, R.; Beckert, W. *Anal. Chem.* **1994**, *66*, 1097-1106
(19)Riu, J.; Rius, F. X. In preparation.

Table 1. Percentage of simulated data sets for which a statistical difference has not been found between the two methods compared at four different levels of significance. 100,000 data sets were obtained for each original data set.

Data set	**α-value (%)**	**BLS (%)**	**OLS (%)**	**WLS (%)**
1	**10**	90.00	89.35	89.35
	5	95.01	94.66	94.66
	1	98.94	98.86	98.86
	0.1	99.90	99.90	99.90
2	**10**	89.05	75.22	90.12
	5	94.21	84.16	94.98
	1	98.63	94.23	99.03
	0.1	99.81	98.73	99.90
3	**10**	0.13	0.14	0.00
	5	0.40	0.44	0.01
	1	3.45	3.73	0.14
	0.1	22.85	22.21	2.71

Table 2. Verification of whether the theoretical point of zero intercept and unity slope falls within the region defined by the joint confidence test using different calibration methods.

Data set	**BLS**	**OLS**	**WLS**
4	no	yes	no
5	yes	no	yes
6	yes	no	no
7	no	yes	no
8	no	yes	yes

FIGURE CAPTIONS

Figure 1. BLS, OLS and WLS coincident calibration lines for the three simulated data sets used in the text together with the individual points and their uncertainties.

Figure 2. BLS, OLS and WLS calibration lines for the five real data sets used in the text together with the individual points and their uncertainties. BLS, solid lines; OLS, dashed lines; WLS, dotted lines.

Figure 3. Joint confidence intervals based on BLS, OLS and WLS methods for the five real data sets studied in the text. BLS, solid lines; OLS, dashed lines; WLS, dotted lines.

Figure 1

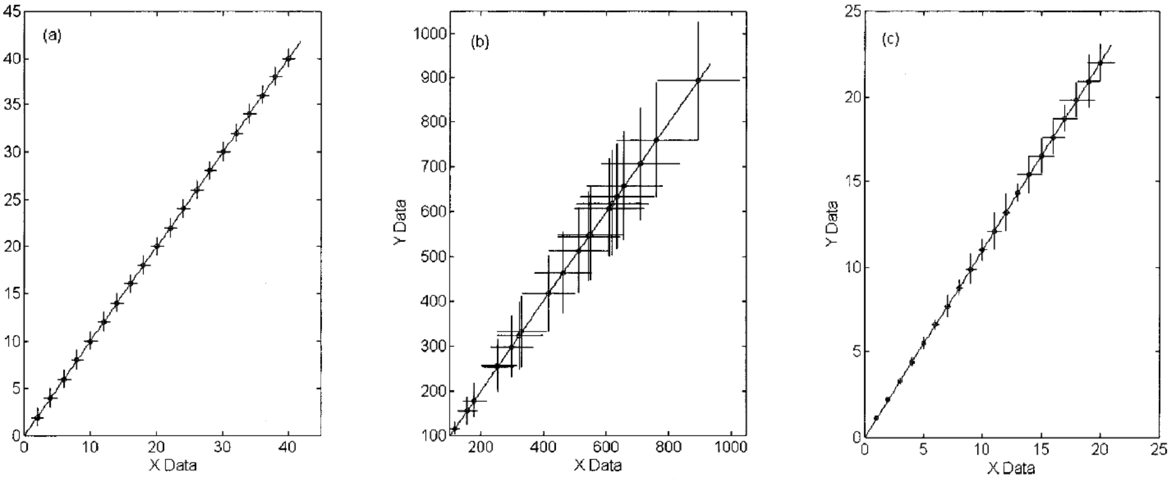
(a)
(b)
(c)
Y Data
X Data
0
5
10
15
20
25
30
35
40
45
100
200
300
400
500
600
700
800
900
1000

Figure 2

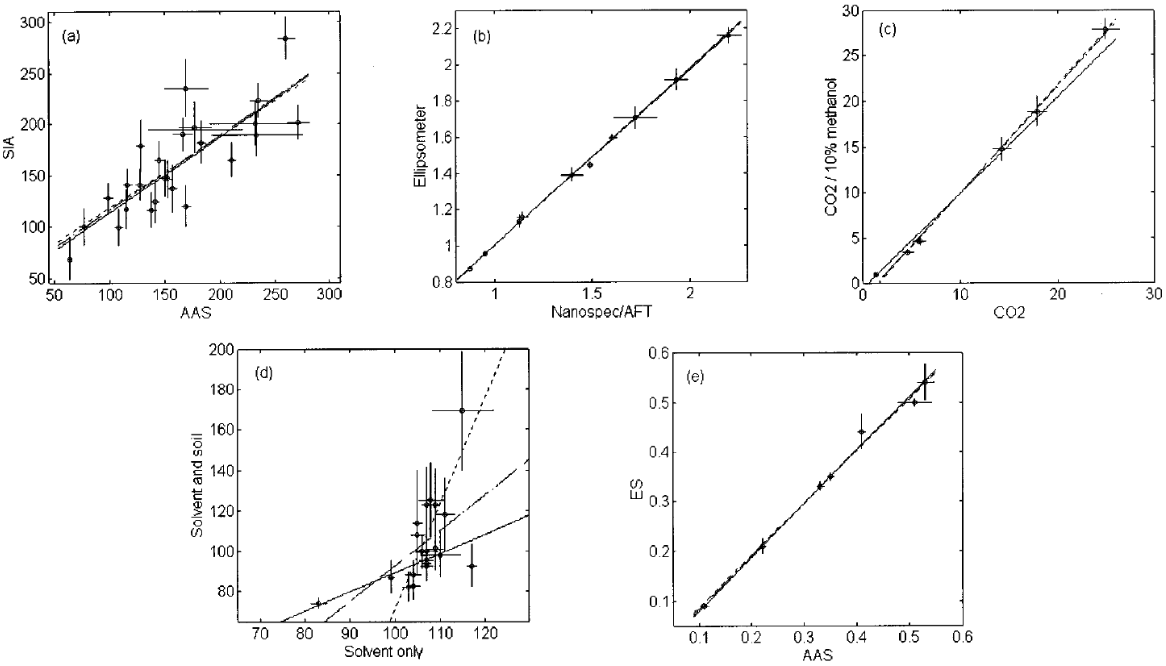
(a)
SIA
AAS
(b)
Ellipsometer
Nanospec/AFT
(c)
CO2 / 10% methenol
CO2
(d)
Solvent and soil
Solvent only
(e)
ES
AAS

Figure 3