The Analyst

Calibration of Gas Chromatography–Mass Spectrometry of Two-component Mixtures Using Univariate Regression and Two- and Three-Way Partial Least Squares

Cevdet Demir and Richard G. Brereton*
School of Chemistry, University of Bristol, Cantock's Close, Bristol, UK BS8 1TS

Univariate calibration and two-way and three-way partial least squares (PLS) were applied to a series of GC–MS results for 21 mixtures of two closely eluting compounds, salbutamol and clenbuterol. Steps in the analysis, including baseline correction, alignment of chromatograms, mass selection, unfolding (for three-way data), standardizing and centring, are described, appropriately modified for the problem in hand. Both mass spectral and, for three-way data, time dependent loadings can be calculated. The quality of quantitative predictions was determined using a leave one out cross validation method. For PLS slightly better predictions were obtained compared with the best predictions for univariate single ion monitoring. Three-way PLS provides a wealth of extra information.

Keywords: *Partial least squares; calibration; gas chromatography–mass spectrometry; chemometrics*

GC–MS is frequently employed as a quantitative technique.[1,2] The traditional approach is to record a chromatogram of a mixture in the presence of a known concentration of one or more reference compounds,[3] and then determine the ratio of the peak areas of each component in the mixture at a selected mass number to the peak area of the reference compound.

This approach of using selective ion monitoring is then used to estimate concentrations, for example in environmental monitoring.[4] However, GC–MS contains information over the entire mass range and, hence, is a multivariate technique. Single masses provide only univariate information, whereas using a range of masses allows multivariate approaches such as partial least squares (PLS).[5–8] There are two ways in which PLS could be employed. The first, or two-way, method simply involves performing PLS on a set of masses simultaneously, for example, selecting 20 masses, and then doing calibration. However, GC–MS can also be considered as a candidate for three-way calibration. Rather than sum the masses up over all points in time, the intensity at each point in time and each mass number can be recorded to give a two-way matrix for each experimental mixture, and a three-way matrix (tensor)[9,10] when a series of mixtures are analysed.

Although PLS supposedly should provide better results than univariate methods, key problems have to be overcome first. For example, the majority of masses in GC–MS are redundant, containing no useful information. In this respect GC–MS differs crucially from, for example, HPLC with diode-array detection, where each wavelength contains some useful information. Hence mass selection[11,12] is a key step. For three-way data, chromatograms have to be aligned, since the true offset of each chromatogram differs slightly according to run. Scaling or standardization[13] has an important role to play in multivariate calibration, as each mass may have different significance, especially if the spectrum of one component in a mixture is dominant by a prominent ion.

In this paper, we propose approaches for multivariate calibration of GC–MS data and look at the effectiveness in relation to univariate methods. In addition, multivariate calibration can reveal extra trends and information about the data.

Experimental

Previous studies established a mixture of salbutamol and clenbuterol as good reference standards because both peaks are partially overlapping and each compound has several ions in common.[11] A mixture design resulting in 21 chromatograms at six concentration levels was used for the study in this paper. Salbutamol and clenbuterol were purchased from Sigma (Poole, Dorset, UK) and quinine from Fluka (Gillingham, Dorset, UK). In order to improve volatility, trimethylsilyl (TMS) derivatives were prepared by adding *N,O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA) (Sigma) to the standard samples. The samples were heated at 80 °C for 1 h. Reagents were removed under nitrogen (40 °C). The derivatized samples were redissolved in toluene–MSTFA (99 + 1 v/v). Stock standard solutions of salbutamol, clenbuterol and quinine were prepared at concentrations of 4.22, 4.04 and 4.10 mg ml^{-1}, respectively. From these solutions a 153.75 mg ml^{-1} quinine solution was prepared and for salbutamol and clenbuterol six working standard solutions were prepared by diluting 20, 10, 6.667, 5, 4 and 3.333 times. The theoretical mixture design is given in Table 1, and the concentrations in Table 2. A typical chromatogram is illustrated in Fig. 1.

Mass spectra recorded on a Fisons (Loughborough, UK) MD800 mass spectrometer were obtained at a scan rate of 200 min^{-1} for the datasets. The GC–MS conditions were the same as in previous studies.[12] The region where clenbuterol and salbutamol elute was reduced to 18 datapoints in time. The spectra of pure salbutamol and clenbuterol are given in Figs. 2 and 3 for reference.

Table 1 Mixture design for salbutamol and clenbuterol. The numbers 0.5, 1, *etc.*, refer to relative concentrations

Experiment No.*	Salbutamol	Clenbuterol
1	0.5	0.5
2(a) and (b)	1.0	1.0
3(a) and (b)	1.0	2.0
4(a) and (b)	1.0	3.0
5	1.5	1.5
6(a) and (b)	2.0	1.0
7(a) and (b)	2.0	2.0
8(a) and (b)	2.0	3.0
9	2.5	2.5
10(a) and (b)	3.0	1.0
11(a) and (b)	3.0	2.0
12(a) and (b)	3.0	3.0

* (a) and (b) represent replicates; 21 experiments were performed in total.

Preprocessing, Selection of Significant Masses and Ratioing to Internal Standard

Before performing any data analysis, it is important to preprocess the data to obtain useful information. The first step is to perform baseline correction. A baseline region which represents only noise can be selected visually or by plotting the logarithm of the sum of squares of the data. After the selection of noise regions, the mean intensity at each mass number is calculated, and these means are subtracted from the data to give a baseline corrected dataset.

Owing to the variability of GC conditions, the compounds may not elute at exactly the same time for each run. Therefore, it is necessary to adjust the elution time for all the samples so that the salbutamol and clenbuterol peaks are centred at the same datapoint. After shifting the elution time the clusters of peaks for each chromatogram are aligned.

Table 2 Injected amounts of salbutamol and clenbuterol

Experiment No.*	Salbutamol/ ng μl^{-1}	Clenbuterol/ ng μl^{-1}
1	52.75	50.5
2(a) and (b)	105.5	101
3(a) and (b)	105.5	202
4(a) and (b)	105.5	303
5	158.25	151.5
6(a) and (b)	211	101
7(a) and (b)	211	202
8(a) and (b)	211	303
9	263.75	252.5
10(a) and (b)	316.5	101
11(a) and (b)	316.5	202
12(a) and (b)	316.5	303

* (a) and (b) represent replicates; 21 experiments were performed in total.

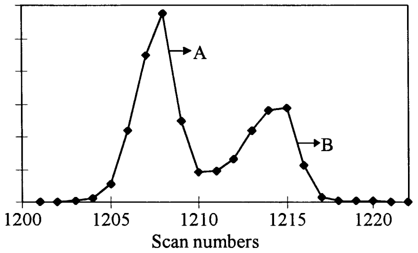

Fig. 1 A typical mixture chromatogram of (A) salbutamol and (B) clenbuterol.

The next step is to select a number of significant masses for calibration. Of the mass numbers between 100 and 400 only certain diagnostic masses are useful. The majority of masses are either not detected or represent noise. Previously we described a method for the selection of masses in an individual chromatogram, using the variance/mean ratio for each mass. In this study, however, there are 21 chromatograms, each of which results in different diagnostic masses.

For each of the 21 chromatograms, the 50 most significant masses ranked by the variance/mean ratio over the cluster containing salbutamol and clenbuterol are computed, using the procedure described previously. Masses that are common to all 21 sets of significant masses are then selected, resulting in 20 significant masses, 10 of which primarily arise from salbutamol and 10 from clenbuterol. Using significantly fewer than 20 masses results in a distribution biased towards salbutamol, presumably because the spectrum of salbutamol dominates in intensity. The order of these masses is determined as follows. If k_{ij} is the rank of mass j in chromatogram i (with rank 1 being most significant), the total rank over all chromatograms is given by

$$K_j = \sum_{i=1}^{I} k_{ij}$$

The smaller this number, the more significant is the mass.

For the internal standard (quinine), the total peak area is an indication of the amount of quinine present; since an equal amount of quinine standard was introduced into each mixture, this number acts as an internal calibrant. The total integral over the quinine peak and all masses (not just the 20 significant masses) was determined and is given, for chromatogram i, by q_i. Because this varies, the ratio $d_{ijn} = z_{ijn}/q_i$, where z_{ijn} is the intensity at mass j over the salbutamol and clenbuterol cluster at time n in chromatogram i was calculated as an indicator of intensity as ratioed to an internal standard at each mass. Calibration was performed relative to z rather than the raw data in all cases.

Univariate Calibration

The simplest method of regression is univariate calibration,[14] in which peak areas at selected masses are calibrated to concentrations. In order to carry out a univariate calibration, a single mass would be chosen which measures the response that corresponds to the concentration of a compound in a sample. The choice of

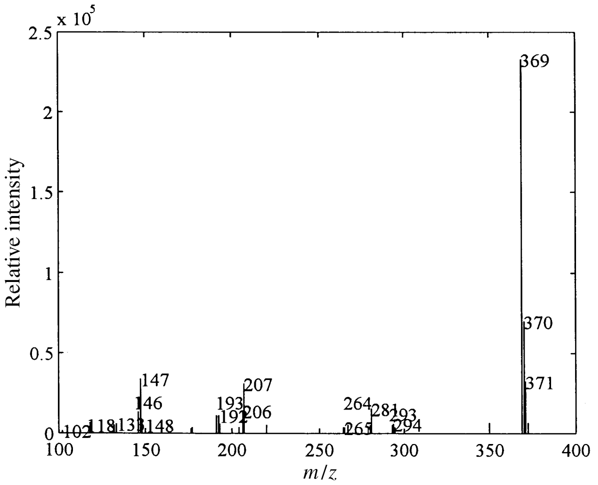

Fig. 2 Mass spectrum of pure salbutamol.

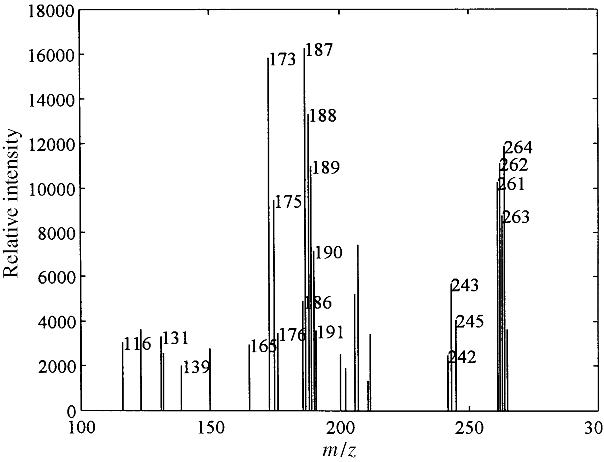

Fig. 3 Mass spectrum of pure clenbuterol.

the mass is important and represents the compound with no interference from other masses that may be present. For mass j, the total peak area (ratioed to quinine) is given by x_{ij}:

$$x_{ij} = \sum_{n=N_1}^{N_2} d_{ijn}$$

Integrating over the region of salbutamol/clenbuterol gives a total area at mass j. A model of the form

$$\hat{y}_{il} = b_{0jl} + b_{1jl} x_{ij}$$

where b_{0jl} is an intercept term, b_{1jl} is the slope, x_{ij} is the total integrated ratioed peak area at mass j, $\hat{y}_{il}$ is the predicted concentration of a standard l and i is the sample number, is then computed, using inverse regression,[15] where the two terms are given by

$$b_{1jl} = \frac{\sum_{i=1}^{I} (x_{ij} - \bar{x}_j)(y_{ij} - \bar{y}_l)}{\sum_{i=1}^{I} (x_{ij} - \bar{x})^2}$$

$$b_{0jl} = \bar{y}_l - b_{1jl}\bar{x}_j$$

This approach assumes all errors are in the measurement of intensity by GC–MS, which is probably the main source of error in this study.

The quality of calibration for compound l can then be determined using the root mean square error of prediction:

$$R_l = \sqrt{\frac{1}{I} \sum_{i=1}^{I} (\hat{y}_{il} - y_{il})^2}$$

where y_{il} is the known concentration of compound l for sample i. This error term is used for comparisons among significant masses and calibration models for compound l.

PLS Calibration

Notation

In this paper three-way matrices are denoted by uppercase bold italic underlined characters, two-way matrices by uppercase bold italics, vectors by lowercase bold italics and scalars by normal italics.

For three-way PLS, the raw data matrix $\underline{\boldsymbol{D}}$ has dimensions $I \times J \times N_0$ with I samples, J masses and N_0 ($N_2 - N_1 + 1$) points in time. The matrix $\boldsymbol{D}$ arises from this when unfolded, having dimensions $I \times JN_0$ as described below.

For two-way PLS, the raw data matrix is $\boldsymbol{X}$ having dimensions $I \times J$. Each column corresponds to the total integrated intensity between times N_1 and N_2 for each of the J masses. N_1 and N_2 were estimated as the elution times at the beginning and end of the peak cluster. No compounds are eluted before and after these times. Note that the two-way matrix $\boldsymbol{X}$ is different from the matrix $\boldsymbol{D}$.

PLS can be three-way or two-way, and to distinguish this a superscript $\boldsymbol{w}$ (2 or 3) is placed above the appropriate scores, loadings and error matrices, vectors and scalars. The left-hand side superscript l (1 or 2) refers to compound number.

Concentration vectors are denoted by y with an appropriate left-hand side superscript.

Further details are given in the Appendix.

Unfolding

There are many possible three-way calibration methods in the literature. An easy way of performing three-way calibration is to unfold the three-way data matrix into a two-way matrix.[16–20] This is performed by concatenating the rows of a matrix to give a row vector. If the objects for one sample form a $I \times N$ matrix, it becomes a length IN row vector. The two-way matrix is established by performing the same operation for all the samples. Fig. 4 shows the unfolding of a three-way data matrix $\boldsymbol{D}$ with I (21) rows and $N_0 \times J$ (360 = 18 × 20) columns. PLS is then used to calculate the matrices $^{l}\boldsymbol{T}$ and $^{l}\boldsymbol{P}$ for component l. The concentration vector $^{l}\boldsymbol{y}$ is the same for both two- and three way PLS.

Standardization

In this paper we consider the influence of standardization on PLS predictions. Standardizing is a form of scaling. For two-way PLS this process is simple and involves the transformation

$$^{s}x_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{\sum_{i=1}^{I} \frac{(x_{ij} - \bar{x}_j)^2}{(I-1)}}}$$

For three-way PLS, the procedure is somewhat more complex than in the two-way case. The aim of standardization in this study is to ensure that all masses have equal influence over the calibration problem. The mass at *m/z* 369 is very intense and would dominate the analysis, so salbutamol would be predicted much better than clenbuterol in the absence of standardization. However, it is still important to retain the relative influence of readings between samples as per two-way data; if the data are standardized after unfolding, the variation in intensity in different times is lost; each time and mass is independently standardized. The aim of standardization should *not* be to weight each time equally but only to weight each mass equally. Therefore, the procedure involves standardizing the entire mass readings at all times over all samples prior to unfolding. Standardizing for three-way PLS analysis can be described as follows:

$$\bar{d}_j = \frac{\sum_{n=N_1}^{N_2} \sum_{i=1}^{I} d_{ijn}}{IN_0}$$

$$^{s}d_{ijn} = \sqrt{\frac{\sum_{n=N_1}^{N_2} \sum_{i=1}^{I} (d_{ijn} - \bar{d}_j)^2}{IN_0}}$$

Note that this procedure is nevertheless different from that for two-way data, since there will be 378 (= 18 × 21) datapoints arising from each mass, rather than only 21 datapoints.

The scaled (standardized) samples are then unfolded. Note that the unfolded data will no longer be mean centred, which operation has to take place again after unfolding.

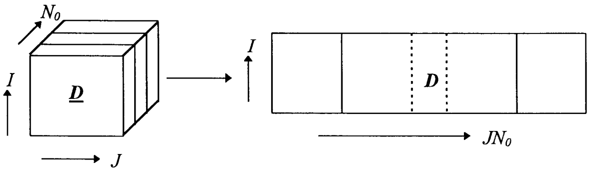

Fig. 4 Unfolding a three-way data matrix into a two-way data matrix to prepare for the PLS analysis.

Centring

Mean centring the data is the second important operation for PLS calibration. For two-way PLS, it involves subtracting the column mean from each element:

$$ {}^{c}x_{ij} = x_{ij} - \bar{x}_j $$

or

$$ {}^{c}x_{ij} = \sum_{n=N_1}^{N_2} d_{ijn} - \frac{\sum_{n=N_1}^{N_2}\sum_{i=1}^{I} d_{ijn}}{IN_0} $$

Note that for three-way data, the process of standardization as described above does not result in mean centring at each point in time.

Mean centring can be carried out by unfolding the three-way matrix to a two-way, $I \times JN_0$, matrix, and then centre this matrix as in ordinary PLS. For three-way PLS, the data are mean centred as follows:

$$ {}^{c}d_{ijn} = d_{ijn} - \bar{d}_{jn} $$

PLS

An application of two- and three-way PLS is presented using GC–MS data. The methods are well described in the literature. For two-way data, PLS is performed on the $\boldsymbol{X}$ matrix with dimension $I \times J$ so that

$$ \boldsymbol{X} = {}^{l}\overline{\boldsymbol{T}}\,{}^{l}\overline{\boldsymbol{P}} + {}^{l}\overline{\boldsymbol{E}} $$

$$ {}^{l}\boldsymbol{y} = {}^{l}\overline{\boldsymbol{u}}\,{}^{l}\overline{\boldsymbol{q}} + {}^{l}\overline{\boldsymbol{f}} $$

where $\overline{\quad}$ denotes two-way scores and loadings. If there are K components ${}^{l}\boldsymbol{T}$, for example, will have I rows and K columns. The corresponding unfolded PLS estimates the following model:

$$ \boldsymbol{D} = {}^{l}\boldsymbol{T}'''\,{}^{l}\boldsymbol{P}''' + {}^{l}\boldsymbol{E}''' $$

$$ {}^{l}\boldsymbol{y} = {}^{l}\boldsymbol{u}'''\,{}^{l}\boldsymbol{q}''' + {}^{l}\boldsymbol{f}''' $$

For two-way PLS, each mass has a corresponding loading for each component and each compound. The magnitude of these loadings can be presented graphically as a function of mass number. Although some loadings are negative, in order to be consistent with the result for three-way PLS, they are presented as absolute values.

The interpretation of loadings for three-way data is more complex. PLS is performed on the unfolded data matrix. The resulting loadings are folded back to form a loadings matrix for each component. These loadings can be plotted as a function of time for each mass, *i.e.*, graphs of ${}^{l}\mathbf{P}'''_{jn,m}$ versus n for each j and m. Another way is to plot a graph of the root mean square loadings:

$$ {}^{l}\boldsymbol{P}'''_{j,m} = \sqrt{\frac{1}{N}\sum_{n=N_1}^{N_2} {}^{l}\boldsymbol{p}'''^{2}_{jn,m}} $$

which shows how the loadings are influenced by mass number. These graphs can be compared with graphs for the corresponding two-way data.

Cross Validation

An important aspect is to evaluate how well the calibration model predicts unknown samples. Finding a minimum root mean square error can be used for this purpose.

Cross validation is frequently used to determine the optimum number of principal components for multivariate calibration.[21,22] Increasing the number of PLS components will always result in a closer fit to the data. However, this does not imply that the later components are significant.

There are several methods for testing the significance of each component, but most are based on cross validation. In this paper, we use the method of taking one sample out at a time.[23] Each of the 21 samples is removed in turn and the two errors are calculated by

$$ {}^{v}f_s = (y_s - \hat{y}_s)^2 $$

$$ {}^{m}f_s = \frac{\sum_{i \ni g_s} (y_i - \hat{y})^2}{(I-1)} $$

where s is the sample removed, g_s is the group of 20 samples excluding sample s, ${}^{v}f$ is the validation error and ${}^{m}f$ is the modelling error.

Over all 21 samples, there will be 21 values of ${}^{v}f$ and ${}^{m}f$, so the overall root mean square validation and modelling errors are given by

$$ {}^{v}F = \sqrt{\frac{1}{I}\sum_{s=1}^{I} {}^{v}f_s} $$

$$ {}^{m}F = \sqrt{\frac{1}{I}\sum_{s=1}^{I} {}^{m}f_s} $$

The prediction error, without cross validation, is given by $f^2 = \boldsymbol{f}'\,\boldsymbol{f}$, where $\boldsymbol{f}$ is the vector of prediction errors for each sample.

The cross validation error is likely to be considerably larger than the modelling error, which in turn is of the same approximate size as the prediction error. The prediction error will reduce according to the number of PLS components used but, if the data behave well, the validation error should decrease until the true number of components is found, and then increase, since the latter components are simply noise and so do not correctly model the validation samples.

To perform cross validation correctly, it is necessary to centre and, where appropriate, scale the 20 samples separately each time, *i.e.*, if standardization is required, this is performed 21 times. The transformation on the 20 samples is then repeated on the 21st. It is not correct to standardize and/or centre the entire dataset of 21 samples and then model the remaining 20 samples, as the components will be neither centred nor standardized.

Results and Discussion

Univariate Calibration

The univariate calibration errors for the 20 selected masses are given in Table 3. In addition, a crude form of validation is performed by constructing a model using 20 samples and determining the prediction error on the 21st. This is repeated leaving one sample out each time.

As can be seen from Table 3, the prediction, modelling and validation errors are comparable in size for this example. Interestingly, some masses result in substantial errors. For clenbuterol, all masses are relatively low in intensity compared with *m/z* 369 for salbutamol and so result in high errors. It would be hard to predict from first principles that *m/z* 166 results in a low error.

Only four masses result in errors around 15 ng μl^{-1}. This suggests that using selective ion monitoring as a means of univariate calibration is risky and highly dependent on the chosen mass, and it would be hard to predict which mass is best from first principles. For each component, the 10 masses are ranked in descending order of significance as suggested by the ranking criterion in the third section. For clenbuterol, it is clear that the more significant masses do not necessarily lead to better calibration predictions.

PLS Predictions

The two-way PLS errors for prediction, modelling and validation are given in Tables 4 and 5. A large number of conclusions may be drawn. The first relates to the number of significant PLS components in the model. Although the prediction error reduces dramatically in all cases, this does not imply that the later components are real. One criterion for significance is that the true number of components is reached when the validation error is a minimum. Extending the calculations to 10 components for clenbuterol suggests that this is reached at around four components (Table 6). However, a flat graph of validation errors often indicates that the safest result is at the beginning of the plateau, suggesting two components for the standardized data and four for the unstandardized data. The difference between the results for standardized and unstandardized data is striking, especially for clenbuterol. A reason is that the salbutamol peak at *m/z* 369 dominates the analysis so that the clenbuterol peaks are relatively minor in size, meaning a poor fit if only a few components are used. This is also reflected in the univariate calibration models since the less intense peaks are dominated by noise. Standardizing results in comparable predictions for both salbutamol and clenbuterol provided two components are used for salbutamol and four for clenbuterol as suggested by the validation errors. It is interesting that standardizing results in a worse model for salbutamol if only one PLS component is used. This is because *m/z* 369 is very intense and diagnostic, but if there is standardization this advantage is removed.

Three-way PLS predictions of the unstandardized data (Tables 7 and 8) result in errors comparable in size for both salbutamol and clenbuterol. In some cases the models are actually slightly worse. This probably is because of the problems of exactly aligning data. Because there is an unpredictable offset when a chromatogram is run, the precise position of a peak will change relative to the beginning of acquisition for an intense peak. This change may be significant for two-way data; the intensities are simply summed for each mass, so this offset is irrelevant.

When the data are standardized this problem apparently disappears, presumably because the intense tops of peaks have less influence on the model. The predictions for clenbuterol are now fairly good, even for one component.

Loadings

A great deal of information can come from the loadings plot. The easiest to interpret are the loadings plotted as a function of time using three way analysis. The loadings for the first two PLS components are given in Figs. 5 and 6. For the

Table 3 Root mean square errors of univariate calibration (ng μl^{-1}). Prediction error refers to the error when the entire 21 samples are used; validation and modelling errors are the average errors when the data set is split into 1 test and 20 training samples

Salbutamol				Clenbuterol			
Mass	Prediction	Modelling	Validation	Mass	Prediction	Modelling	Validation
369	14.51201	14.5034	15.2465	262	52.9003	52.3151	54.7844
370	15.9678	15.9253	16.7844	243	45.523	45.5224	46.6973
147	23.5113	20.5732	24.1903	264	42.2964	42.042	43.8859
207	13.8473	13.8339	14.5452	186	31.8329	28.9176	33.644
265	60.5821	60.5802	61.4897	173	26.4489	26.103	27.2311
133	28.4676	27.4523	29.2518	212	36.0138	35.4542	38.0682
294	24.1052	24.0685	24.6289	166	14.0452	13.1963	14.8421
281	24.8886	24.813	26.2671	188	34.3321	33.4795	35.831
220	20.2818	20.2816	21.2602	245	23.7456	21.4552	24.4513
177	51.3573	50.8283	52.6191	277	40.8727	40.2014	42.5105

Table 4 Root mean square error of two-way PLS calibration of salbutamol (ng μl^{-1})

	Unstandardized			Standardized		
Component	Prediction	Modelling	Validation	Prediction	Modelling	Validation
1	14.5	14.7907	15.5489	30.0977	29.218	34.3195
2	13.8364	14.1351	15.459	10.8668	10.705	12.454
3	12.6539	12.8975	15.0545	8.8071	8.4721	12.3484
4	10.3475	10.2772	14.7892	6.8891	6.5467	11.2131

Table 5 Root mean square error of two-way PLS calibration of clenbuterol (ng μl^{-1})

	Unstandardized			Standardized		
Component	Prediction	Modelling	Validation	Prediction	Modelling	Validation
1	65.8421	65.6303	69.8464	30.7642	30.7846	34.4136
2	29.2507	29.0469	31.1823	20.6712	20.9968	22.7378
3	22.6302	23.023	28.2061	13.2479	12.4302	22.1795
4	13.1725	13.2623	19.4711	11.8193	10.9652	19.1694

unstandardized data there is little difference for the first PLS component between salbutamol and clenbuterol, except a slightly larger maximum at fast elution time, reflecting the position of clenbuterol. This is presumably because the ion at *m/z* 369 dominates no matter which component is used for calibration. Even though *m/z* 369 is not well correlated with clenbuterol, it will, nevertheless, exhibit a small correlation because of the design as chromatograms 1, 5 and 9 are correlated for both components so there will be a slight correlation between salbutamol and clenbuterol concentrations which is picked up. The second component is necessary for a good model, particularly for clenbuterol, as can be seen from Tables 7 and 8. Here there is a dramatic difference, the earlier eluting salbutamol having a negative influence on clenbuterol calibration. The reverse is the case, but less strongly, for salbutamol.

The standardized time-dependent loadings show a much more obvious trend, the maxima for both salbutamol and clenbuterol being in the correct place. The second loadings provide an appropriate negative balancing effect. These are in good agreement with the expectations from the prediction error. The influence of standardization is especially obvious from the first PLS component of Fig. 6. Note that the small correlation with the second compound is almost certainly a result of the design used, demonstrating the special importance of design when doing calibration experiments. This is likely to disappear if an orthogonal design is employed.

The mass spectral loadings are harder to interpret primarily because they are displayed in the absolute value mode. Only the three-way spectral loadings are illustrated for sake of brevity (Figs. 7 and 8) but similar conclusions also come from two-way analysis. The root mean square loadings over all points in time are illustrated for this purpose. For example, the loadings for unstandardized salbutamol show maxima at peaks for salbuta-

Table 6 Root mean square error of two-way PLS cross validation of clenbuterol for 10 PLS components (ng μl^{-1})

Component	Unstandardized	Standardized
1	69.8464	34.4136
2	31.1823	22.7378
3	28.2061	22.1795
4	19.4711	19.1694
5	22.1532	21.9661
6	23.258	22.7049
7	24.7529	24.1972
8	26.2817	26.6376
9	27.2957	27.1828
10	30.8872	28.1535

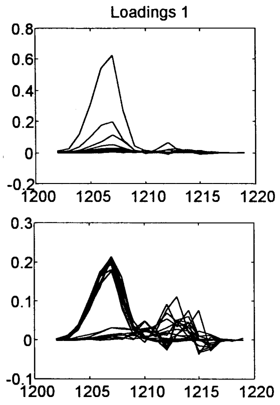

Fig. 5 Time-dependent loadings for the first two components of three-way PLS data for salbutamol. The top figures represent the unstandardized data and the bottom figures the standardized data.

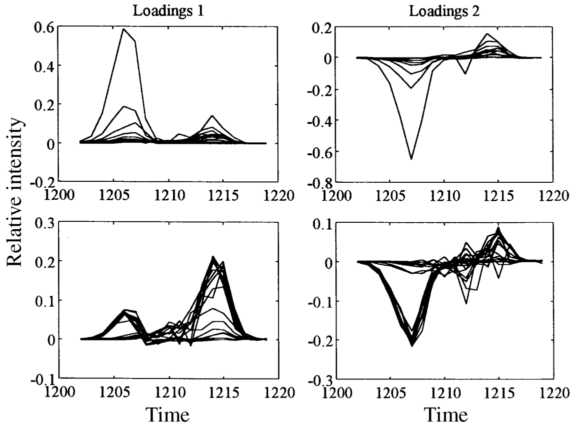

Fig. 6 Time-dependent loadings for the first two components of three-way PLS data for clenbuterol. The top figures represent the unstandardized data and the bottom figures the standardized data.

Table 7 Root mean square error of three-way PLS calibration of salbutamol (ng μl^{-1})

	Unstandardized			Standardized		
Component	Prediction	Modelling	Validation	Prediction	Modelling	Validation
1	15.3467	15.5379	21.7641	16.7973	15.7835	22.202
2	14.3254	14.5242	19.9991	14.0457	14.0227	20.3149
3	12.3116	12.1632	18.6847	10.9006	10.1287	18.8127
4	9.4362	9.5854	16.4184	7.414	7.0539	17.0949

Table 8 Root mean square error of three-way PLS calibration of clenbuterol (ng μl^{-1})

	Unstandardized			Standardized		
Component	Prediction	Modelling	Validation	Prediction	Modelling	Validation
1	57.4241	55.2667	65.7378	17.8093	17.7842	23.1488
2	25.6571	26.7913	31.1776	15.8925	15.9249	22.1179
3	17.7387	17.832	23.8823	9.7123	10.0147	21.119
4	11.9961	11.6328	22.3824	5.7637	5.6139	19.7164

mol in the first PLS component, being a good reconstruction of the salbutamol spectrum since the time-dependent loadings also suggest that the first component primarily picks up salbutamol. The second component is influenced more by clenbuterol, although in negative manner, so the loadings primarily relate to the clenbuterol spectrum with a more prominent maximum at *m/z* 262, although some salbutamol (*m/z* 369) still remains.

Standardizing the data distorts the relative mass spectral intensities but the trends are still there. For example, the time-dependent loadings for the first standardized component for clenbuterol and second for salbutamol are dominated by clenbuterol and the equivalent mass spectral loadings show low intensity at *m/z* 369 and 370 but relatively high intensities at *m/z* 262. Visually the loadings for the three-way first standardized PLS components of salbutamol and clenbuterol can be divided into two groups, those of value around 0.07 and 0.02. These are given in Table 9. Remarkably, the high value loadings all arise primarily from the calibrant and the low value loadings from the second component. There is a high degree of correspondence between the order of these loadings and the order of univariate calibration errors as given in Table 3. For example, the *m/z* 265 and 177 ions have the highest calibration errors for salbutamol and, therefore, appear intermediate between the salbutamol and clenbuterol group in clenbuterol (0.0542 and 0.0413). The *m/z* 166 and 245 ions have lowest univariate calibration errors for clenbuterol (and are the only two ions that could be employed to provide an estimate of clenbuterol concentration with a high degree of confidence in the univariate models) and both have the highest loadings. The results for salbutamol are slightly less clear, but the ion with the highest univariate calibration error (*m/z* 265) has, nevertheless, the lowest value of the loadings for the salbutamol first component. A graph of the time-dependent loadings for salbutamol explains this. The second component is not as well defined as it is for clenbuterol.

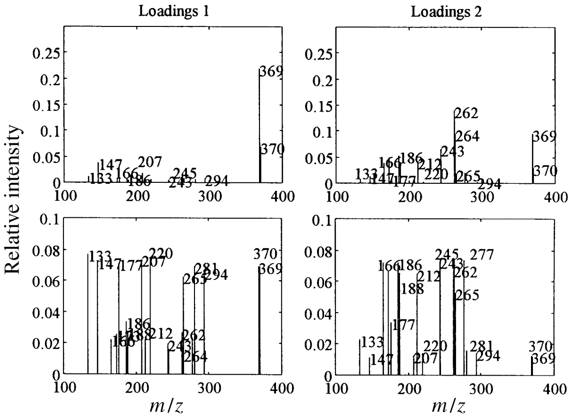

Fig. 7 Spectral loadings of three-way PLS data for salbutamol. The top figures represent the unstandardized data and the bottom figures the standardized data.

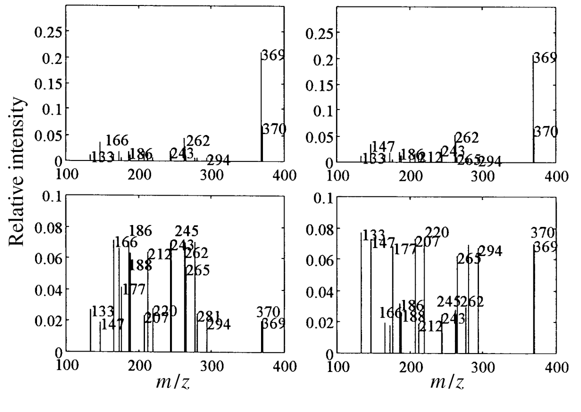

Fig. 8 Spectral loadings of three-way PLS data for clenbuterol. The top figures represent the unstandardized data and the bottom figures the standardized data.

Conclusions

This paper has described two methods for PLS calibration of GC–MS data. The improvement in prediction error is small but significant over univariate calibration if PLS is properly performed. However, the major advantage of PLS is that 20 masses can be used, and the prediction error is better than the best single mass. The difficulty with single ion approaches is that a correct choice of mass must be made. For at least one of the components, it is not clear, from first principles, which mass is most suitable, so selective ion monitoring risks poor results unless great care is taken to select a range of ions, some of which may not be obvious to the mass spectrometrists at first glance.

Three-way PLS does not have major advantages over two-way methods as far as prediction errors are concerned. One problem is correctly aligning chromatograms. If digital resolution is not very high (typically a peak is defined by around 10 datapoints in GC–MS), small shifts in detector offsets can result in difficulties exactly aligning chromatograms, so introducing extra errors into the calibration[24] which counterbalance the advantages of using the extra dimension. However, three-way analysis does result in more diagnostic information such as time-dependent loadings and, certainly for more complex problems of three or four components eluting at slightly different times, will reveal much more than two-way procedures.

The absolute importance of data scaling such as standardizing and mean centring at the correct step in the procedure has been discussed. In this paper, we have only reported meaningful results. Incorrectly applying data preprocessing methods results in meaningless output, so the user of such approaches must take great care.

We thank Dr. F. Burden for helpful discussions relating to cross validation and P. Hindmarch for writing the program for

Table 9 Root mean square loadings of the first standardized three-way PLS component of salbutamol and clenbuterol and their assignments (C and S)

Salbutamol	*m/z*	Corresponding compound	Clenbuterol	*m/z*	Corresponding compound
0.0769	133	S	0.0713	245	C
0.0740	207	S	0.0713	166	C
0.0739	220	S	0.0705	277	C
0.0719	147	S	0.0705	186	C
0.0713	177	S	0.0702	243	C
0.0695	281	S	0.0700	264	C
0.0694	369	S	0.0661	173	C
0.0692	370	S	0.0649	262	C
0.0658	294	S	0.0640	212	C
0.0626	265	S	0.0635	188	C
0.0341	186	C	0.0542	265	S
0.0270	188	C	0.0413	177	S
0.0266	262	C	0.0278	220	S
0.0257	173	C	0.0274	133	S
0.0232	277	C	0.0247	281	S
0.0224	212	C	0.0235	207	S
0.0224	166	C	0.0206	370	S
0.0196	243	C	0.0201	369	S
0.0176	264	C	0.0198	294	S
0.0165	245	C	0.0190	147	S

decoding GC–MS data. We are grateful for support from Uludag University (Bursa, Turkey).

Appendix

List of Notations Used

$\underline{\boldsymbol{D}}$	Three-way GC–MS data matrix
$\boldsymbol{D}$	Unfolded GC–MS data matrix
$\boldsymbol{X}$	Two-way GC–MS data matrix
i	Sample number
I	Total number of samples
j	Mass number
l	Compound number
K_j	Total rank over all chromatograms
k_{ij}	Rank of mass j
q_i	Quinine peak area for sample i summed over time and all masses
z_{ijn}	Intensity for mass j, sample i and time n
d_{ijn}	Peak area ratio for mass j, sample i and time n
n	Time
N_1	Point in time at the beginning of a cluster of peaks
N_2	Point in time at the end of a cluster of peaks
x_{ij}	Peak area ratio for mass j and sample i summed over all time
${}^{s}x_{ij}$	Standardized data matrix for two-way PLS
${}^{c}x_{ij}$	Mean centred data matrix for two-way PLS
${}^{s}d_{ijn}$	Standardised data matrix for three-way PLS
${}^{c}d_{ijn}$	Mean centred data matrix for three-way PLS
$\hat{y}_{il}$	Predicted concentration
$\underline{y_{il}}$	Concentration of compound l
$\underline{y_{\mathrm{l}}}$	Mean concentration of compound l
x_j	Mean peak area ratio of mass j summed over time
b_{ojl}	Intercept of univariate calibration line
b_{1jl}	Slope of univariate calibration line
R_l	Root mean square error of compound l
${}^{\mathbf{l}}\mathbf{y}$	Concentration vector for compound l
${}^{l}\boldsymbol{T}^{w}$	PLS scores matrix for compound l
${}^{\mathrm{l}}\boldsymbol{t}_{m}^{w}$	PLS scores vector for compound l
${}^{\mathrm{l}}\boldsymbol{t}_{im}^{w}$	PLS scores for sample i, component m and compound l
${}^{l}\boldsymbol{P}^{w}$	PLS loadings matrix for compound l
${}^{l}\boldsymbol{p}_{m}^{w}$	PLS loadings vector for component m and compound l
$\boldsymbol{w}$	Order of PLS (2 or 3)
${}^{l}\boldsymbol{p}_{mj}^{w}$	Loadings for mass j for two-way PLS
${}^{l}\boldsymbol{p}_{mjn}^{w}$	Loadings for mass j and time n for three-way PLS
${}^{l}\boldsymbol{P}_{mj}^{w}$	Sum of loadings for mass j over all time for three-way PLS
m	Component number
M	Total number of components
N_0	Number of data points in time over significant peaks ($N_2 - N_1 + 1$)
${}^{l}\boldsymbol{E}''$	Error matrix of $\boldsymbol{X}$ for two-way PLS and compound l
${}^{l}\boldsymbol{E}'''$	Error matrix of $\boldsymbol{D}$ for three-way PLS and compound l
${}^{l}\boldsymbol{f}''$	Error vector of $\boldsymbol{y}$ for two-way PLS and compound l
s	Removed sample
${}^{v}f$	Validation error
${}^{m}f$	Modelling error
g_s	Group of samples excluding sample s
${}^{v}F$	Root mean square validation error
${}^{m}F$	Root mean square modelling error

References

1 Boqué, R., and Rius, F. X., *Chemom. Intell. Lab. Syst.*, 1996, **32**, 11.
2 *Practical Mass Spectrometry*, ed. Middleditch, B. S., Plenum Press, New York, 1979.
3 Millard, B. J., *Quantitative Mass Spectrometry*, Heyden, London, 1978.
4 Cirovic, D. A., Brereton, R. G., Walsh, P. T., Ellwood, J. A., and Scobbie, E., *Analyst*, 1996, **121**, 575.
5 Höskuldsson, A., *J. Chemom.*, 1988, **2**, 211.
6 Wold, S., Geladi, P., Esbensen, K., and Ohman, J., *J. Chemom.*, 1987, **1**, 41.
7 Kowalski, B. R., and Seasholtz, M. B., *J. Chemom.*, 1991, **5**, 129.
8 Martens, H., and Naes, T., *Multivariate Calibration*, Wiley, New York, 1989.
9 Sanchez, E., and Kowalski, B. R., *J. Chemom.*, 1988, **2**, 247.
10 Sanchez, E., and Kowalski, B. R., *J. Chemom.*, 1988, **2**, 265.
11 Demir, C., Hindmarch, P., and Brereton, R. G., *Analyst*, 1996, **121**, 1443.
12 Hindmarch, P., Demir, C., and Brereton, R. G., *Analyst*, 1996, **121**, 993.
13 Smilde, A. K., *Chemom. Intell. Lab. Syst.*, 1992, **15**, 143.
14 Booksh, K. S., and Kowalski, B. R., *Anal. Chem.*, 1994, **66**, 782A.
15 Araujo, P. W., Cirovic, D. A., and Brereton, R. G., *Analyst*, 1996, **121**, 581.
16 Henrion, R., *Chemom. Intell. Lab. Syst.*, 1994, **25**, 1.
17 Ståhle, L., *Chemom. Intell. Lab. Syst.*, 1989, **7**, 95.
18 Smilde, A. K., and Doornbos, D. A., *J. Chemom.*, 1991, **5**, 345.
19 Bro, R., *J. Chemom.*, 1996, **10**, 47.
20 Geladi, P., *Chemom. Intell. Lab. Syst.*, 1989, **7**, 11.
21 Zhang, P., and Littlejohn, D., *Chemom. Intell. Lab. Syst.*, 1996, **34**, 203.
22 Sharaf, M., Illman, D. L., and Kowalski, B. R., *Chemometrics*, Wiley, New York, 1986.
23 Gemperline, P. J., *J. Chemom.*, 1989, **3**, 549.
24 Cirovic, D. A., Jacobsen, R. M., and Brereton, R. G., *Anal. Commun.*, 1996, **33**, 231.

Paper 6/08245I
Received December 6, 1996
Accepted March 10, 1997