N-way PLS applied to simultaneous spectrophotometric determination of acetylsalicylic acid, paracetamol and caffeine

Marcelo M. Sena, Ronei J. Poppi

Departamento de Química Analítica, Instituto de Química, Universidade Estadual de Campinas, P.O. Box 6154, CEP 13083-862, Campinas, SP, Brazil

Received 10 April 2003; received in revised form 11 August 2003; accepted 11 August 2003

Abstract

In this work, a simple and rapid analytical procedure was proposed for simultaneous determination of acetylsalicylic acid (ASA), paracetamol (PRC, also known as acetaminophen) and caffeine (CAF) in pharmaceutical formulations based on multivariate calibration and UV spectrophotometric measurements (210–300 nm). The calibration set was constructed with nine solutions in the concentration ranges from 10.0 to 15.0 μg ml⁻¹ for ASA and PRC and from 2.0 to 6.0 μg ml⁻¹ for CAF, according to an experimental design. The procedure was repeated at four different pH values: 2.0, 3.0, 4.0 and 5.0. Partial least squares (PLS) models were built at each pH and used to determinate a set of synthetic mixtures. The best model was obtained at pH 5.0. An N-way PLS model was applied to a three-way array constructed using all the pH data sets and enabled better results. This calibration model provided root mean squares errors of prediction (RMSEP) from 11.5 to 35% lower than those obtained with PLS at pH 5.0, depending on the analyte. The results achieved for the determination of these drugs in commercial tablets were in agreement to the values specified by the manufactures and the recovery was between 94.7 and 104.5%.

© 2003 Elsevier B.V. All rights reserved.

Keywords: N-way PLS; Acetylsalicylic acid; Paracetamol; Caffeine; Simultaneous spectrophotometric determination; Multi-way multivariate calibration

1. Introduction

In the last decade, three-way or, more generally, N-way analysis was introduced in the field of analytical chemistry [1]. A three-way array may be obtained by collecting data tables with a fixed set of objects and variables under different experimental conditions, such as sampling time, temperature, pH, etc. The tables collected under various conditions can be stacked providing a three-dimensional arrangement of data. In some situations, even higher dimensional arrays may be considered. These methods can be applied for exploratory analysis, curve resolution, analysis of variance and calibration purposes using spectrophotometric, spectrofluorimetric, chromatographic, flow injection, sensory analysis or experimental design data [2,3]. For N-way multivariate calibration, N-way partial least squares (N-PLS)
has recently received more attention than other tools. Its applications can be found on HPLC-DAD [4], molecular fluorescence spectrometry [5,6], 3D-QSAR [7], MS/MS [8] UV-Vis spectroscopic-kinetic data [9], image analysis [10], remote sensing [11] and industrial pharmaceutical batch process [12]. An example of a four-way-PLS application can be found on kinetic spectrofluorimetry data [13].

Acetylsalicylic acid (ASA), paracetamol (PRC) and caffeine (CAF) are frequently combined in a number of pharmaceutical formulations. ASA possesses antirheumatic, antipyretic and analgesic properties and is probably the major consumed drug in the world. PRC, also known as acetaminophen, is an antipyretic and analgesic drug, which, in contrast with ASA, has the advantage of not irritating the gastrointestinal mucosa. Their effect of pain relief can be enhanced by the stimulating power of CAF.

2. Theory

2.1. PLS

PLS [20] is a method for building regression models based on the latent variable decomposition relating two blocks, matrices \( X \) and \( Y \), which contain the independent, \( x \), and dependent, \( y \), variables, respectively. These two matrices are decomposed into a sum of \( f \) latent variables and two sets of models are obtained, of the form:

\[
X = TP^T + E = \sum_i p_i^T + E \tag{1}
\]
\[
Y = TQ^T + F = \sum_i q_i^T + F \tag{2}
\]

in which \( T \) is the score matrix; \( P \) and \( Q \) are the loading matrices for \( X \) and \( Y \), respectively, and \( E \) and \( F \) are the residual matrices. The superscript \( t \) indicates a transposed matrix. The product of \( T \) and \( P^T \) approximates to the independent variables (e.g. spectral data) and the product of \( T \) and \( Q^T \) to the dependent variables (e.g. concentrations). An important feature of PLS is that is possible to obtain scores matrix that is common to both the concentrations (\( Y \)) and measurements (\( X \)). The concentration of the new samples can be estimated from the new scores \( T^* \) and the model loadings \( Q \), which are substituted in Eq. (2), leading to Eq. (3):

\[
Y_{\text{new}} = T^* Q^T \tag{3}
\]

In this procedure, it is necessary to find the best number of latent variables, which normally is performed by using cross-validation, based on the determination
of the minimum prediction error. The difference between PLS1, in which the regression is carried out for each dependent variable individually \((Y, Q, F)\) and \(F\) column matrices, i.e. \(y, q, f\) vectors, and PLS2 (our case), in which all dependent variables are used simultaneously, should also be mentioned.

2.2. N-PLS

N-PLS is an extension of the PLS regression to multi-way data \([2,5,20]\). As for PLS, the term N-PLS2 is used for the case in which all dependent variables are predicted simultaneously \((an Y)\) for the case in which all dependent variables are used simultaneously, should also be mentioned.

\[N \text{-} \text{PLS}2 \text{ is firstly rearranged (matri}
\]
Table 1

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASA</td>
<td>+ + + − − − + − − −</td>
</tr>
<tr>
<td>PRC</td>
<td>+ + − + − + − − −</td>
</tr>
<tr>
<td>CAF</td>
<td>+ − + + − − − − −</td>
</tr>
</tbody>
</table>

Level (+): ASA, 15.00 μg ml⁻¹; PRC, 15.00 μg ml⁻¹; CAF, 6.00 μg ml⁻¹. Level (−): ASA, 10.00 μg ml⁻¹; PRC, 10.00 μg ml⁻¹; CAF, 2.00 μg ml⁻¹. Level (−−): ASA, 12.50 μg ml⁻¹; PRC, 12.50 μg ml⁻¹; CAF, 4.00 μg ml⁻¹.

was used for all measurements. An ultrasonic bath was also employed for sample extraction.

The data were handled using MATLAB software, 6.1 version (The MathWorks, Natick, USA). PLS routine came from “PLS Toolbox”, 2.0 version (Eigenvector Technologies, Manson, USA). N-PLS modelling was carried out by using “The N-way Toolbox for MATLAB”, 2.00 version (R. Bro, Foodtechnology, Copenhagen, Denmark) [28].

3.3. Procedure

3.3.1. Calibration set and synthetic mixtures

The calibration set was constructed according to a 2³ + 1 (three factors at two levels plus one central point) experimental design (Table 1). The ASA and PRC solutions were in the 10–15 μg ml⁻¹ range and the CAF solutions were in the 2–6 μg ml⁻¹ range. The synthetic mixtures used to validate the model were planned according to a 2³ experimental design similar to the calibration set (without a central point). For this validation set, the level (+) was 13.50 μg ml⁻¹ for ASA and PRC and 5.00 μg ml⁻¹ for CAF, and the level (−) was 11.50 μg ml⁻¹ for ASA and PRC and 3.00 μg ml⁻¹ for CAF. Seventeen standard solutions (calibration + validation sets) were prepared directly inside the cuvette, by the addition of 100 μl of each analyte intermediate solution in 2.70 ml of the respective buffer or ionic solution with the specified pH and 200 μl of deionised water. The dissolution was carried out with the aid of an ultrasonic bath (15 min). An aliquot of 100 μl of each sample was added into a cuvette containing 2.70 ml of the respective buffer or ionic solution at each pH value. Solutions prepared in the same way as the mixtures, but containing none of the analytes, were used as blank for each pH set. Spectra of pure ASA, PRC and CAF solutions were also recorded at each pH value.

3.3.2. Sample (tablet formulations) determination and recovery

The pharmaceutical preparations assayed had the following composition per tablet: Cibalena® (Novartis, Brazil), 200 mg of ASA, 150 mg of PRC and 50 mg of CAF; Excedrin Migraine® (Bristol-Myers Squibb Co., USA), 250 mg of ASA, 250 mg of PRC and 65 mg of CAF. Six tablets of each pharmaceutical formulation were weighed individually to obtain an average weight. The tablets were finely powdered and mixed, and a mass corresponding to one tablet for each formulation was weighed and dissolved in 500 ml of ethanol/water (20:80, v/v), in a volumetric flask. The dissolution was carried out with the aid of an ultrasonic bath (15 min). An aliquot of 100 μl of each sample was added into a cuvette containing 2.70 ml of the respective buffer or ionic solution at the specified pH and 200 μl of deionised water. The recovery method was performed by an addition of 100 μl of a 30 μg ml⁻¹ standard solution of each analyte. The spectra were obtained in the same conditions described previously. All these determinations were performed in triplicate.

4. Results and discussion

4.1. ASA, PRC and CAF ultraviolet spectra

Fig. 1 displays the UV absorption spectra for aqueous solutions of ASA, PRC and CAF at pH 2. As can be observed, there is a strong overlap among the spectra, which prevents the use of univariate calibration. The use of multivariate calibration for the resolution of these drugs has been performed in a pH range from 1 to 5.5 [22–24]. In this work, the determinations were performed at four different pH values, 2, 3, 4 and 5. The spectra of PRC and CAF did not change in this pH range, which was expected, since these species do not suffer ionisation in these conditions. In contrast, ASA spectrum presents λ_max at 229 nm at pH 2, which shifts to lower wavelengths at higher pH values.
At pH 5, this spectrum has a deformed shape, showing no clear maximum of absorbance. This can be attributed to the ionisation of ASA to acetylsalicylate ion, whose pK is approximately 3.4 [25].

4.2. Calibration and validation of PLS and N-PLS models

Multivariate calibration methods require a suitable experimental design of the standards belonging to the calibration set in order to provide good predictions. The calibration set was constructed using nine solutions according to an experimental design (Table 1). Another eight solutions were used as validation set according to a second experimental design, whose range was contained in the calibration design. One may argue that this two-level (plus one mid point) calibration design is inadequate and typically four or five concentration levels would be required for each compound [20]. However, this choice depends strongly on the nature of the system under calibration. For more complex matrices, such as soil samples analysed in IR, this design would certainly be insufficient, but for a simple matrix, such as in our case (synthetic aqueous solutions), it is appropriate, which is demonstrated by our good results and other successful applications found in the literature [21,23,24].

One PLS model was built for each pH data set. A three-way array (9 solutions × 91 wavelengths × 4 pH values) was constructed combining all the pH sets and it was used to build an N-PLS model. All the models were validated using cross-validation. The root mean squares errors of prediction (RMSEP) of the validation sets was the parameter employed for comparison among the models. RMSEP is given by:

$$\text{RMSEP} = \sqrt{\frac{\sum(y_r - y_p)^2}{n}}$$

where $y_r$ is the standard (real) value and $y_p$ is the value predicted by the model.

Table 2 shows the RMSEP values for all the calibration models. All the models were built using three latent variables. The best PLS2 model for the simultaneous determination of the three analytes was obtained at pH 5. This model presented the lowest RMSEP for ASA, but only the second one for CAF and the third one for PRC. It was also observed that...
as the pH increases, the RMSEP of ASA decreases. This observation can be explained by the decrease in the degree of overlap between ASA and PRC (Figs. 1 and 2). The incorporation of all the information obtained from the four different pH sets in a single three-way model enabled better predictions. This is one of the major advantages of N-way models. N-PLS2 presented the best predictions for ASA (RMSEP 17.5% lower than PLS2/pH 5) and CAF (RMSEP 35% lower than PLS2/pH 5 and 23% lower than PLS2/pH 4). Although it presented predictions for PRC worse than PLS/pH 2 model (RMSEP 14% higher), N-PLS2 was considered the best model for simultaneous determination of ASA, PRC and CAF.

In order to demonstrate the accuracy of the proposed method, the results of the predictions for synthetic mixtures taken as validation set are shown in Table 3. The most accurate results were obtained for ASA, whose errors never exceeded 2.3%. For PRC, the samples at the highest level (13.50 µg ml⁻¹) presented
somewhat higher errors, about 5-6%. The highest errors, about 10%, were observed for CAF samples at the lowest level (3.00 mg ml\(^{-1}\)). The proposed method can be considered appropriate for practical analysis taking into account the tolerance level of ±10% established in the US Pharmacopoeia [14] for this type of drugs.

### 4.3. Analysis of real samples

N-PLS2 and the best two-way calibration model, PLS2/pH 5, were applied to simultaneous determination of the three drugs in two different commercially available pharmaceutical formulations. The results are listed in Table 4 and are in agreement with those specified by the manufacturers. The highest difference between the claimed and the predicted amounts was 7% for CAF in Cibalena®. The standard deviations obtained for three replicated determinations with PLS2/pH 5 were higher, in some cases more than twice, than those obtained with N-PLS2. This indicated that N-PLS2 presented more precise results than PLS2/pH 5. The recovery method for standard additions of each analyte to two commercial pharmaceuticals was also performed, corroborating the efficiency of the proposed method. Table 5 presented the percentages of recovery for each addition.

<table>
<thead>
<tr>
<th>Pharmaceutical formulation</th>
<th>Label claim (mg/tablet)</th>
<th>Amount predicted with N-PLS2 (mg/tablet)a</th>
<th>Amount predicted with PLS2/pH 5 (mg/tablet)a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ASA</td>
<td>PRC</td>
<td>CAF</td>
</tr>
<tr>
<td>Cibalena®</td>
<td>200</td>
<td>150</td>
<td>50</td>
</tr>
<tr>
<td>Excedrin Migraine®</td>
<td>250</td>
<td>250</td>
<td>65</td>
</tr>
</tbody>
</table>

a Mean values and relative standard deviation of three determinations.

<table>
<thead>
<tr>
<th>Drug</th>
<th>Recovery (%)b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N-PLS2</td>
</tr>
<tr>
<td>Cibalena®</td>
<td>100.5 ± 0.5</td>
</tr>
<tr>
<td>Excedrin Migraine®</td>
<td>99.2 ± 0.6</td>
</tr>
</tbody>
</table>

b Mean values and relative standard deviation of three determinations.
5. Conclusions

The traditional and official methods for simultaneous determination of drugs in pharmaceutical preparations are based on chromatographic techniques, presenting relative high cost and time consumption. PLS multivariate calibration using UV spectrophotometric data can be considered a suitable method for an accurate, rapid and less expensive determination. This work applied this method to simultaneous determination of ASA, PRC and CAF in tablets at four different pH values and demonstrated that the use of a multi-way model, N-way PLS, improved the predictions, providing RMSEP values lower than those obtained with PLS2 at pH 5.0 (17.5% for ASA, 11.5% for PRC and 35% for CAF), which was considered the best PLS model. The improvement in the predictions can be attributed to the incorporation of one more dimension in the data, which implicates in more information used for modelling. Multi-way multivariate calibration using N-PLS is suitable to be combined with excitation-emission matrix (EEM) obtained in molecular fluorescence spectroscopy, HPLC-DAD and other hyphenated techniques. This combination can be considered useful and promising for developing routine quality control analysis of pharmaceuticals.

Acknowledgements

M.M.S. thanks CNPq for a fellowship. The authors thank Dr. Rasmus Bro for “The N-way Toolbox for MATLAB”.

References