ASSESSMENT OF DIFFERENT PARTIAL LEAST SQUARES VARIANTS FOR DETERMINATION OF BINARY-DRUG SYSTEM EXHIBITING INTENSE SPECTRAL OVERLAP

Ismail I. Fasfous*, Yahya S. Al-Degs* and Asmaa M. Mallah

(Received 25 June 2016) (Accepted 27 August 2016)

ABSTRACT

Amoxicillin (AMO) and clavulanic acid (CLA) are popular activate pharmaceutical ingredients that are widely used due to their efficient medical activity. However, this binary system suffers from intense spectral overlap (93.6%). Inspite of the intense spectral overlap and serious nonlinearity in the current system, both drugs were accurately quantified by multivariate calibration. The performance of different partial least squares PLS variants (NIPALS, SIMPLS, Kernel and Bidiagonalization) for accurate quantification of AMO-CLA in commercial formulation was outlined. Partial response and partial residual plots confirmed a serious nonlinearity in the binary system. Compared to other algorithms, PLS-Kernel exhibited a better performance for drugs quantification and seven latent variables were necessary for accurate quantification: 94.0(9.6%) and 95.6(5.2%) for AMO and CLA, respectively. The intense spectral overlap, nonlinearity, and non-modelled excipients are effectively handled by PLS-Kernel calibration.

Keywords: PLS-variants; NIPALS; SIMPLS; Kernel; Bidiagonalization; Amoxicillin–Clavulanic Acid; Spectral Overlap; Drug Formulation.

INTRODUCTION

Commercial drug formulations often contain more than one activate pharmaceutical ingredients APIs that present in variable levels to achieve the best pharmacological performance. The spectral overlap between the active ingredients is often moderate but intense overlap is also possible. In addition to spectral overlap and nonlinearity in the system, drug production stages like crystallization, drying, solid dosage form, added excipients, and tableting at different conditions can affect the spectral behavior of APIs. Accordingly, accurate analytical methods are always needed in this regard. Commercially, the combination of amoxicillin AMO and clavulanic acid CLA represents one of the potent broad spectrum antibiotics. The combination targets both Gram-positive and Gram-negative organisms especially those which showed resistance to beta-lactam antibiotics. AMO is often present in excess to CLA (or alone in some formulations) and acts on the bacterial cell walls by making them more porous. CLA, a mild antibacterial compound, helps AMO by competing and irreversibly binding to the bacterial cell wall. Consequently, the attacked bacteria cannot generate beta-lactamase and then become susceptible to AMO. Many analytical methods were proposed for quantification of AMO-CLA including liquid chromatography and PLS-NIPALS calibration. The main analytical problem in AMO-CLA is the intense spectral overlap and nonlinearity.

Partial least squares (PLS) modeling is an important tool in many fields including chemistry, medicine, pharmaceutical analysis, and process modeling. As an important multivariate calibration method, PLS can deal with collinearity of spectral data without losing the chemical information and offers an interactive diagnostic exploration of the data.

In general, the fast growing of PLS multivariate calibration is attributed to the availability of this algorithm, affordable running cost and short analysis time when compared to chromatographic-based methods. In pharmaceutical analysis, the drug of interest is often present with other ingredients and excipients which often absorb in the same spectral region. In such cases, spectral overlap is a serious analytical dilemma which may require initial sample clean-up and solute-separation before detection. For example, Martin and co-workers applied a liquid-liquid extraction step to extract spironolactone, canrenone, and hydrochlorothiazide from urine matrix prior to PLS calibration. PLS calibration was efficient for direct quantification of ternary drugs formulation in the presence of non-modeled excipients. Goicoechea and Olivieri applied PLS-regression with a suitable
wavelength-selection procedure to quantify tetracycline in human serum without applying any sample-cleaning or separation procedures\textsuperscript{11}. In case of intense spectral overlap, nonlinearity, and existence of interfering excipients, the resolution power of PLS would probably be affected.

Modern instruments can generate tremendous number of signals per sample, accordingly, the large \( X \) matrices (i.e., explanatory variables) become larger in size which needs more advance PLS-variants to end up with accurate results with short computation time and less-storage-capacity\textsuperscript{12,13}. In fact, NIPALS is the most adopted algorithm in pharmaceutical analysis\textsuperscript{5,7}. There are many PLS-variants that have no applications in pharmaceutical analyses including SIMPLS, Kernel and Bidiagonalization. It is known that the aforementioned algorithms are different in their mechanisms for running chemical analysis\textsuperscript{13}. See supplementary material for more details on the mechanisms of the variants. For instance, NIPALS is suitable for modeling many variables-\( X \) but it requires long computational time and more memory-storage\textsuperscript{13-15}. However, SIMPLS may be faster than NIPALS but it is not suitable for many variables-\( X \) matrices\textsuperscript{16}. On the other hand, PLS-Kernel is considered as adjustable algorithm which can fit systems of many variables or even many samples by creating condensed kernel-matrices\textsuperscript{14,15}. Bidiagonalization is an advanced variant, which decomposes \( X \) matrix into three smaller matrices of orthonormal vectors\textsuperscript{12,13}. The earlier algorithm deserves investigation as it has very limited application in pharmaceutical analysis. Compared to the other variants, NIPALS is extensively adopted in pharmaceutical analysis due to its availability in many commercial software packages.

The aims of this work are, a) assessment of resolving power of three PLS-variants (SIMPLS, Kernel, and Bidiagonalization) for handling binary drug systems of intense spectral overlap and high nonlinearity, and b) quick quantification of AMO-CLA in highly consumed formulations with minimum sample clean up by PLS-regression and this is necessary to assess the robustness of PLS-variants towards non-modeled interferences.

**MATERIALS AND METHODS**

**Chemicals**

The compounds amoxicillin and clavulanic acid were purchased from GlaxoSmithkline ( Worthing, UK). The commercial formulations (in tablet form) were obtained from local pharmacies. AMOCLAN\textsuperscript{®} (APIs: amoxicillin trihydrate 500 mg, clavulanic acid potassium form 125 mg, excipients: starch glycolate, magnesium stearate, silicon dioxide, cellulose, manufacturer HIKMA Pharmaceuticals, Amman, Jordan). All solvents used in chromatographic separation were HPLC grade and purchased from Sigma-Aldrich\textsuperscript{®}. Distilled water was used for preparation of solutions.

**Preparation of calibration and validation sets for MVC**

For each drug, a 200 mg/L standard solution was prepared using distilled water and used later to prepare working solutions. Stock solutions were stored at 4.0°C to ensure maximum stability of the drugs. Binary mixtures of AMO-CLA were prepared by mixing appropriate volumes of stock solutions and diluted with distilled water. Prior to final dilution, \( \text{pH} \) was adjusted to 3.5 for both systems using acetic acid/sodium acetate (0.50 M) buffer solution. All solutions were prepared in distilled water and sonicated for 2.0 min. Stability of drugs in solution was studied by scanning 5.0 ppm solutions of drugs after being exposed to ambient conditions over 24 hrs.

In multivariate calibration, calibration set and validation set are carefully prepared\textsuperscript{9,17,18}. Due to the intense spectral overlap between APIs, then a large number of mixtures is necessary to get accurate prediction of drugs by multivariate calibration. Univariate calibration was carried out to determine the working ranges for the drugs in order to set their appropriate levels in binary solutions. There are many strategies for designing calibration set. In pharmaceutical analysis, the full factorial designs are often adopted. Suitable designs for calibration mixtures can be obtained following Breton`s tables\textsuperscript{18}. The earlier tables are proposed for multilevel multifactor (multisolute) systems. According to Breton`s tables, nine mixtures are prepared based on an orthogonal design for both systems\textsuperscript{18}. The concentration levels of the drugs were 1.0, 6.0 and 12.0 mg/L. The validation set was randomly selected within the ranges of calibration concentrations. The selected levels of drugs in both sets are given in Table I.

For all solutions, the digitized spectra of solutions over the range 216-300 nm (85 spectral point/sample) were collected in \( X \) matrix. Calibration mixtures were prepared excluding excipients and this was necessary to assess the performance of multivariate calibration. As already known, drugs would show different extents of stability in solution. For the current binary systems, solutions of drugs were scanned after being exposed to laboratory conditions for 24 hours. In fact, AMO exhibited drastic spectral changes which necessitate scanning all solutions after preparation. The solutions of AMO were turned to pale yellow indicating the hydrolysis of the compound\textsuperscript{31}. It is worthy to mention that PCA-cluster...
analysis confirmed that old solutions of AMO are different from fresh solutions where two distinct clusters appeared in the plot. On the other hand, CLA manifested much better stability in solution as the variations in spectral shapes were insignificant over 24 hours.

**Multivariate calibration and drug quantification**

Before running PLS regression, the absorbance values were stored in $X$ matrix (9x85) while concentrations of the calibrated drug were stored in $y$ vector (9x1). Calibration vector $b$ for each drug was obtained by modeling $X$ and $y$ data using PLS. For Kernel algorithm, calibration matrix $B$ for both drugs was obtained in one step. The spectra of formulations were stored in $X_{AMOCLAN}$ (4x85) matrix in preparation for multivariate calibration. The preparation of formulations for spectral measurements is outlined in the following section.

**Assaying APIs by liquid chromatography**

Independently, APIs were quantified using high performance liquid chromatography HPLC. The drugs were separated using isocratic mode on $C_{18}$ column (Termo-Lot 8317 150mmx4.6 mm, particle size 5 μm) at 20°C. Separation was accomplished using methanol–tetraethylammonium acetate buffer (pH 4, 0.1 mM, 35:65, by vol.) as mobile phase. The flow rate and detection wavelength were set at 1.5 mL/min and 225 nm, respectively. In all runs, the injection volume was 20.0 μL. Calibration curves were generated and peak area was used for quantitative purposes. Triplicate determinations were made for each solution. Determination of drugs in commercial formulation was carried out as following. Ten tablets of AMOCLAN® were separately grounded, mixed and homogenized. A sample equivalent to the mass of one tablet was suspended in 50 mL distilled water, sonicated for 5.0 min, centrifuged for 5.0 min, filtered, and diluted to 1.0 L by distilled water. The optimum sonication time was also investigated by monitoring the UV spectra of extract over the period 1-10 min. Interestingly, after 5.0 min of sonication the absorbances did not show considerable changes indicating the complete dissolution within 5.0 min. The extract of the formulation was diluted to be within the analytical range (See Table I). A dilution factor of 50 was reasonable for drugs detection in the extract. pH of all solutions was adjusted to 4.0 for chromatographic analysis and 3.5 for multivariate calibration. Four identical measurements were carried out for commercial formulation by multivariate calibration and HPLC.

**RESULTS AND DISCUSSION**

In pharmaceutical analysis, the purpose of PLS calibration is to find the concentration of APIs in formulation without running other expensive and laborious measurements\(^\text{19}\). Many issues are related to the performance of PLS: a) non-selectivity; b) collinearity in spectral data; c) nonlinearity; d) designing calibration mixtures and selection of the informative spectral regions\(^\text{20}\); and e) presence of outliers\(^\text{8,19}\). To have accurate calibration results, the earlier issues should be handled in a proper way\(^\text{20}\).
Pearson product moment coefficient PPMC was employed to evaluate the effect of tablet's excipients on the spectral features of drugs\textsuperscript{22}. Moreover, the variations in APIs levels in formulation and the possible changes in spectral shapes during tablet preparation are assessed by PPMC test. This test measures the closeness between two chromatograms or spectra recorded for a chemical system\textsuperscript{22}. The similarity index $r$ between the spectrum of synthetic mixture/solution $x_i$ and the one recorded for the formulation $y_i$ recorded at $i^{th}$ wavelength is estimated as:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$ \hspace{1cm} \text{Eq 1}$$

Where $\bar{x}$ and $\bar{y}$ are the average of absorbance values. Based on the magnitude of $r$, the similarity between spectra is classified as strong ($|r| \geq 0.8$), moderate ($0.8 > |r| \geq 0.5$) and weak ($|r| < 0.5$)\textsuperscript{22}. Table II summarizes the results of PPMC test on spectra of synthetic solutions and extracts of formulations.

<table>
<thead>
<tr>
<th></th>
<th>AMO</th>
<th>CLA</th>
<th>Mixture</th>
<th>Tablet</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMO</td>
<td>1.0</td>
<td>0.86</td>
<td>0.73</td>
<td>0.62</td>
</tr>
<tr>
<td>CLA</td>
<td>1.0</td>
<td>0.59</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>Mixture</td>
<td></td>
<td>1.0</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>Tablet</td>
<td></td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a}Spectra were recorded at 5.0 and 1.2 mg/L for AMO and CLA, respectively. Spectra of the tablets were recorded for diluted extracts of the formulations (50). Each spectrum composed of 85 spectral points.

As indicated in Table II, a high correlation ($r=0.96$) was observed between the spectra of AMO–CLA mixture and the tablet one, which indicated the insignificant influence of excipients on the spectral feature of the drugs. The spectrum of formulation (Fig. 1B) was comparable to the one reported for mixture, however, with lower intensity due to intense dilution with water (dilution factor 50). Due to the intense spectral overlapping between drugs, a high $r$ value (0.86) was obtained by PPMC test. The results of PPMC test also indicated that the spectrum of AMO ($r=0.73$) is closer to the mixture than CLA ($r=0.59$) and this will reduce the selectivity of the proposed method for CLA. The higher selectivity for AMO was obtained for the distinct absorption at 230 nm. Moreover, the closeness between the spectra AMO and the tablet ($r=0.62$) was relatively higher compared to CLA ($r=0.56$). In summary,
the current system encountered intense spectral overlap (93.6%) and this will deteriorate the selectivity. The PPMC test indicated a poor correlation between CLA spectrum and the spectra of synthetic mixture and extract of formulation.

**Nonlinearity in binary system**

To detect nonlinearity in multivariate calibration of many highly collinear variables (like spectroscopic data), graphical and quantitative numerical tools are often adopted. For graphical tools, PC factors or PLS variables are used to generate four plots in order to detect nonlinearity in the analytical system. Due to the orthogonal nature of PC factors, then nonlinearity can be detected graphically. Partial response plot PRP and residual plot RP are generated to diagnose nonlinearity in the current analytical systems. PRP is generated by plotting \( y \) versus the first PC and RP is generated by plotting the predicted \( \hat{y} \) against \( e_{y,PC1-PCA} \), where \( e_{y,PC1-PCA} \) is the residual error in \( y \) estimated from measured and predicted values. Predicted values were estimated using the optimum PLS variables A. To detect nonlinearity in binary drug systems, the scores obtained by NIPALS were used in the diagnoses as presented in Fig. 2.

PRP would detect the intrinsic nonlinearity between explanatory variables \( X \) and \( y \) while RP can graphically depict the nonlinearity stems from the employed model. For PAR, the results are shown in Fig 2 and PRP indicated a linear behavior between PC1 and \( y \). The equal distribution of the residuals \( (e_{y,PC1-PCA}) \) around zero is strongly support the linear nature of multivariate system. Accordingly, PAR-CAF system would be handled by linear multivariate calibration algorithms.

Similar plots were obtained for CAF but not presented in the manuscript. On the other hand, AMO-CLA system suffers from high nonlinearity as inferred diagnosis analysis. PRP shows nonlinear behavior or curvature in PC1-\( y \) plot. Furthermore, nonlinearity is inferred from the unequal distribution of the residuals in the residual plot (see Fig. 2). When nonlinearity is negligible (like the case in PAR-CAF), then it is possible to rely on linear calibration models and handling unmodeled excipients by adding extra factors. In fact, Kernel-PLS model was able to handle nonlinearity in AMO-CLA stems from the high spectral overlap.

**Assessment of PLS-variants for APIs quantification**

Before assaying real formulations, the performance of different PLS-variants was assessed by analyzing excipients-free synthetic solutions. The performance of PLS calibration is further enhanced by proper selection of informative spectral regions and removal of outliers prior to model building. Generally, principal component analysis PCA is often applied to detect outliers in all solutions (calibration, validation and extracts). Using PCA, seven and six eigen values were retained upon decomposing PAR–CAF and AMO–CLA matrices, respectively. Statistical analysis revealed that the variability in the data (in both systems) was satisfactorily explained using the first two eigen values. The earlier fact is acceptable as both systems containing only two drugs. For both drug systems, projection of PC1 (first principal component vector) against PC2 (second principal component vector) was carried out to detect potential outliers. The diagnosis plot (not shown) indicated that all samples including synthetic and formulations constitute a uniform cluster and no outlier was detected. The optimum spectral ranges were determined using moving window partial least squares regression strategy MWPLSR. In this strategy, a spectral window begins at the \( p \)th wavelength and ends at the \( (p+h-1) \)th wavelength is constructed, where \( h \) is a pre-selected window size. The optimum spectral regions were 216-277 and 235-284 nm for PAR and CAF, respectively. For the other system, the regions...
215-278 and 219-294 nm were obtained for AMO and CLA, respectively. Using the informative spectral regions, PLS-Kernel always ended up with fewer prediction errors and lower latent variables.

Table III summarizes the calibration outputs of several PLS-variants along with the results of chromatographic techniques.

Table III: Drugs quantification in validation sets by several PLS-variants and independently by liquid chromatography

<table>
<thead>
<tr>
<th>PLS-Variant</th>
<th>AMO-CLA (spectral overlap 93.6%)</th>
<th>AMO</th>
<th>CLA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rec</td>
<td>REP%</td>
<td>RSD</td>
</tr>
<tr>
<td>NIPALS</td>
<td>92.6(10)</td>
<td>7.4</td>
<td>12.0</td>
</tr>
<tr>
<td>SIMPLS</td>
<td>89.4(10)</td>
<td>6.3</td>
<td>11.6</td>
</tr>
<tr>
<td>Kernel</td>
<td>97.5(8)</td>
<td>4.9</td>
<td>7.5</td>
</tr>
<tr>
<td>Bidiagonal</td>
<td>92.7(8)</td>
<td>5.4</td>
<td>10.6</td>
</tr>
<tr>
<td>HPLC*</td>
<td>101.1</td>
<td>1.4</td>
<td>3.1</td>
</tr>
</tbody>
</table>

a. Mean recovery of the drugs (validation sets, see Table I).
b. REP%: relative error of prediction.
c. Relative standard deviation
d. Number of PLS-variables were obtained by leave-one-out cross-validation technique (see section 2).
e. For chromatographic conditions, five-point calibration lines were obtained for each drug. The regression equations were obtained by plotting peak area versus drug concentration:

AMO: A = 367521C + 5239 (r² = 0.9923), dynamic range 0.5-10.9 mg/L, detection limit 0.02 mg/L. CLA: A = 395567 C + 41837 (r² = 0.9951), dynamic range 0.4 - 9.5 mg/L, detection limit 0.03 mg/L.

As shown in Table III, PLS-variants were workable for drug quantification even in the presence of intense overlap like AMO-CLA. In all variants, drugs were predicted using the optimum PLS-variables in the validations sets with reasonable relative error of prediction (<8% in all cases). PLS-Kernel achieved the best results and this is obvious from REP% and overall precisions. Both drugs in PAR-CAF system were predicted by PLS-Kernel with REP% of 1.3 and 1.6 for PAR and CAF, respectively. The other intensely-overlapped system, AMO and CLA were quantified with higher REP% values of 3.9 and 4.3, respectively. One more interesting point on PLS-Kernel is the lower PLS-variables (7 for both drugs) which reflected the less computational efforts compared to the rest of variants. NIPALS algorithm, the classical PLS-variant, is found applicable for both drugs but with extra regression steps as indicated from the higher variables (10 factors were needed to model AMO-CLA system). It is worth mentioning that all variants were more effective for handling PAR-CAF compared to AMO-CLA and this is attributed to the intense overlap in the former system. Moreover, PLS-Kernel was workable for modeling AMO-CLA system.

An interesting feature on the mechanism of PLS-variants was deduced from cross-validation plots. The optimum number of variables needed to build the model is a key step to yield a better external prediction. Cross-validation procedure was applied for this purpose, consisting of systematically removing one of the calibration samples in turn, and using the remaining samples for construction of the latent variables and regression equations. The optimum number of factors was carefully chosen according to the criterion proposed by Haaland and Thomas. The closeness between predicted and nominal concentrations was assessed by finding prediction error sum of squares PRESS which estimated as following:

\[ PRESS = \sum_{i=1}^{n} (C_{i,\text{pred}} - C_{i,\text{exp}})^2 \]

Where \( C_{i,\text{pred}} \) and \( C_{i,\text{exp}} \) are the predicted and experimental drug levels, respectively. Fig. 3 depicts PRESS versus PLS-variables plots as obtained by cross-validation technique.
For PAR, a stable performance was observed for all variants and the variability in the data was explained using 2 PLS variables. However, the behavior of PLS-Bidiagonal was inconsistent with other variants. A large fitting error was observed at higher variable, which is not often reported. Accordingly, the application of PLS-Bidiagonal is not recommended in this case. Similar results were reported for CAF (data not provided). For AMO-CLA system, more factors were needed (for all models) to explain the variability in spectral data. As shown in Fig 3B, the residual errors were stabilized at 6 variables for the variants except PLS-Bidiagonal which showed inconsistent behavior at higher variables. The optimum number of variables was provided in Table III.

In terms of computational efforts, Kernel model was the best one as it used less number of factors for modeling the intensely overlapped AMO-CLA system (8 factors compared to 10 for NIPALS). Except for PLS-Kernel, the drugs were modeled individually by estimating regression vectors in separate steps. This also indicates that Kernel-PLS required less computation efforts as it estimates all vectors in one step.

 Obviously, the proposed chromatographic procedure showed an excellent performance for quantifying both drugs binary systems with average recovery of 106.0–101.1 and REP% 1.3–1.4 for the highly overlapped AMO-CLA system. The excellence of chromatographic procedure is attributed to better selectivity and sensitivity for drugs after being separated by the stationary phase. However, the large consumption of organic solvents, using advanced and expensive instruments, high running/maintenance cost would support the applicability of PLS-regression in pharmaceutical formulations.

**Figures of merit for drugs quantification**

Net-analyte signal NAS analysis is a common multivariate calibration technique that often used to estimate the figures of merit for classical and inverse multivariate calibration methods. More details estimation of figures of merit by NAS are outlined in supplementary material. Figures of merit are summarized in Table IV.

The results showed that PAR and CAF have higher selectivities (0.89-0.92) and better sensitivities (0.41-0.53) and lower detection limits (0.22-0.31 mg/tablet) when compared to AMO and CLA values. The main conclusion is that multivariate calibration showed an outstanding resolution power for mixtures of moderate spectral overlap and this power seems to be affected by highly-overlapped systems or nonlinearities in response-concentration relationship, which often encountered in chemical analysis. Both inter-day and intra-day precision and accuracy of the Kernel-PLS was evaluated. Three binary mixtures of PAR-CAF (2.0, 6.0, and 8.0 mg/L) and AMO-CLA (2.0, 4.0, and 6.0 mg/L) were assayed within the same day and over three days. Six determinations were made for each solution. Drug recoveries and REP% were 98.3–101.6% and 2.7–3.1 for PAR–CAF and 97.6–98.5 and 4.6–5.4 for AMO-CLA system. Solutions were freshly prepared in the same day of analysis. The results are satisfactory and indicate the reproducibility of Kernel-PLS for pharmaceutical analysis.

**Quantification of AMO–CLA in commercial formulation**

Now, both PLS-Kernel and chromatographic methods are applied to quantify binary-drug systems in commercial formulations where many interfering substances are existed. Due to the poor stability of AMO, scanning of solutions was carried out directly after dissolution form formulation. Table V summarizes the results with some statistical parameters to validate the credibility of the proposed method.

Table IV: PLS-Kernel regression parameters, figures of merit, and selected spectral ranges by MWPLSR

<table>
<thead>
<tr>
<th>Parameter</th>
<th>AMO</th>
<th>CLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral region (nm)</td>
<td>AMO-CLA (spectral overlap 93.6%)</td>
<td></td>
</tr>
<tr>
<td>Dynamic range (mg/L)</td>
<td>215-278 nm</td>
<td>219-292 nm</td>
</tr>
<tr>
<td>Latent variables</td>
<td>1.8-12</td>
<td>2.0-12.0</td>
</tr>
<tr>
<td>Mean Recovery</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>REP%</td>
<td>98.4</td>
<td>101.5</td>
</tr>
<tr>
<td>RSD</td>
<td>3.9</td>
<td>4.3</td>
</tr>
<tr>
<td>LOD (mg/tablet)</td>
<td>6.8</td>
<td>7.4</td>
</tr>
<tr>
<td>LOQ(mg/tablet)</td>
<td>0.56</td>
<td>0.62</td>
</tr>
<tr>
<td>Selectivity</td>
<td>1.8</td>
<td>2.0</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.32</td>
<td>0.46</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>Spectral overlap% (NAS calculations)</td>
<td>68</td>
<td>54</td>
</tr>
</tbody>
</table>

Before discussing the results in Table V, it should be mentioned that preparation and dilution procedures of the formulations have a significant effect on the final results. Due to the poor stability of AMO, scanning of solutions was carried out directly after dissolution form formulation. Table V summarizes the results with some statistical parameters to validate the credibility of the proposed method.
earlier. The following conclusions may be drawn from Table V, a) both methods were satisfactory for quantification of PAR–CAF in formulation as indicated from the statistical t and F tests, however, chromatographic method furnished more accurate results (when compared with claimed levels), with relative errors of 1 and 1.5% for PAR and CAF compared to 4.4 and 4.6% for PLS1 for the same drugs. In general, the outputs of both methods would be convincing for pharmaceutical analysts where R.S.D. values (in all cases) less than 5%. For this system, the superiority of chromatography is mainly attributed to the perfect chromatographic separation of moderately-overlapped drugs (56.3%) and the influence of excipients on PLS-Kernel regression power, and b) liquid chromatography was more accurate than Kernel-PLS for determination of Amo–CLA in real formulation. For chromatographic methods, the estimated relative errors were 1.8 and 3.2% for Amo and CLA against 7.0 and 9.6% for PLS1 for the same drugs, respectively. On the other side, both methods were found of comparable precision as deduced from F-test. The superiority of chromatography for handling Amo–CLA was attributed to, a) the physical separation of drugs from excipients which improve overall detection, and 2) the nonlinearities stem from intense spectral overlap and non-modeled excipients has been reduced by the PLS calibration.

**CONCLUSIONS**

PLS-Kernel model has better calibration performance compared to the other PLS-variants for handling binary-drug systems of intense spectral overlapping and serious nonlinearity. Moreover, the model has shown more flexibility for modeling many variables-X matrices which is necessary for assaying formulations. Considering the high operational costs associated with advance liquid
chromatography, PLS-Kernel calibration is a handy substitute for separation-based methods chromatography. Due to the availability of most PLS-variants in commercial software packages, the analysts should select the proper PLS-algorithm to end up with the optimum results in a short time and with less computational efforts.

ACKNOWLEDGEMENT

The financial support from the Hashemite University/Deanship of Scientific Research is gratefully acknowledged.

Supplementary Material

Theoretical background of PLS variants

PLS is an effective algorithm for developing a quantitative relationship between several predictor variables X (spectral data) and a property of interest Y (drug content). The relationship between X and Y or y (for one single independent variable) is given as:

\[ y = Xb \]

where y, X, and b are drug standard concentration, absorbance data of standard solutions measured at different wavelengths, and calibration sensitivity which is necessary for external prediction. PLS is an algorithm that used to find b which is often accomplished using different variants. In general, the dimensions of the earlier quantities are X (samples×variables) and Y (samples×variables). In the following matrix equations, the symbols t, u, q, and r are stand for transpose of matrix, inverse of matrix, PLS-loading for y, PLS-loading for t, PLS-score for t, and PLS-score for y, respectively. Prediction of the target drug from unknown spectrum a_un is obtained as:

\[ a_{un} = c_{un}b \]

This algorithm is extensively applied in pharmaceuticals analysis.

SIMPLS

This algorithm is faster than NIPALS but it is not recommended for many variables X matrices. To find b, the following quantities are computed:

\[ s = Xy \]

\[ r = PLS-loading\ for\ y \]

\[ t = PLS-score\ for\ X \]

\[ p = PLS-loading\ for\ X \]

\[ q = PLS-loading\ for\ y \]

The quantities r, t, p, and q are stored in R, T, P and Q, respectively. Before estimating the next PLS-variable s is projected on a subspace of P. The above algorithm is stopped once all PLS-variables are computed and b is obtained as:

\[ b = Rw(PW)^{-1}q \]

Prediction of drug in a new sample is estimated as outlined in NIPALS. SIMPLS is faster than NIPALS as no deflation is needed. It also uses a fewer number of matrices to find b.
There are two versions of Kernel algorithm. The first one can handle matrices where \( i > J \) and the other one (which is suitable for the current drug system) is proposed for \( J > i \). In both algorithms, condensed matrices are created from \( X \) and \( Y \) (or \( y \)). In the current case, two condensed matrices are created \( XX^T \) and \( YY^T \) or \( yy^T \). Kernel matrix is then estimated as: \( XX^TYY^T \). The main steps of the algorithm are\(^{14-16} \):

The eigenvector of the Kernel matrix is taken as the first \( X \) score vector \( t_1 \). The \( Y \) score vector is then estimated as:

\[
u_1 = YY^T t_1.
\]

The next step is to update the association matrices by eliminating the explained variable as:

\[
G_i = I - t_i t_i^T \quad (I \text{ identity matrix})
\]

\[
X_i X_i^T = G_i X X^T G_i
\]

\[
Y_i Y_i^T = G_i Y Y^T G_i
\]

The matrices involved in the algorithm are smaller than the original ones and this reduces computation time and memory storage. The next \( t \) and \( u \) vectors are estimated as outlined above using the updated matrices. The calibration matrix is estimated from weight and loading matrices (\( W, P \) and \( Q \)) as following:

\[
W = XU
\]

\[
P = (T^T X)(T^T T)^{-1}
\]

\[
Q = (T^T Y)(T^T T)^{-1}
\]

The above step is repeated until the optimum number of variables is estimated. \( B \) is obtained as:

\[
B = W(PW)^{-1}Q
\]

The final prediction of drugs from unknown spectrum \( a_{un} \) is obtained as:

\[
C_{un} = a_{un} B
\]

**Bidiagonalization**

Basically, this advanced algorithm is started by decomposing \( X \) into three matrices\(^{13,32,33} \):

\[
X = U R V^T
\]

Where, \( U(i \times J) \) and \( V(i \times J) \) are matrices with orthonormal columns (i.e., \( U^TU = V^TV = 1 \)) and \( R(J \times J) \) is the bidiagonal matrix\(^{12} \). Once \( U, R \) and \( V \) are estimated with the optimum PLS-variables, the calibration vector is estimated as\(^{19} \):

\[
b = VR^T U y
\]

Prediction of the target drug in the new sample is carried out as mentioned earlier. It is important to mention that the authors showed that the common cross-validation technique would be adopted for SIMPLS, Kernel and Bidiagonal algorithms.

**Figures of merit of multivariate calibration**

The analytical performance of multivariate calibration is assessed by carrying out net-analyte signal NAS calculations. NAS is a suitable method to characterize the analytical figures of merit related to the multivariate calibration during drugs quantification\(^1 \). For classical multivariate calibration, the basic equation that is needed to estimated figures of merit is\(^{25} \):

\[
s_k^* = (I-S_k S_k^*) s_k
\]

where \( S \) is the matrix of sensitivities collected for all other solutes, \( s_k \) is the sensitivity vector of the analyte, and \( s_k^* \) is the estimated net part of the \( k \)th component that is orthogonal to the other constituents\(^ {25} \). NAS is necessary to find meaningful parameters to assess the analytical method like sensitivity \( S \text{EN} \), selectivity \( S \text{EL} \), limit of detection \( S \text{LoD} \), limit of quantification \( S \text{LoQ} \). SEN is estimated from the net signal of analyte \( k \) (\( s_k^* \)) as\(^ {30} \):

\[
S \text{EN} = \left\| s_k^* \right\|
\]

\( \text{SEL} \) which measures the extent of spectral overlapping is estimated as\(^ {1,30} \):

\[
S \text{EL} = \left\| s_k^* \right\| / \left\| s_k \right\|
\]

\( \text{LOD} \) which gives the minimum detectable amount of the solute \( k \) is given as\(^1 \):

\[
\text{LOD} = 3 \left\| \varepsilon \right\| / \left\| s_k^* \right\|
\]

The minimum quantifiable amount of the solute is estimated as\(^1 \):

\[
\text{LOQ} = 10 \left\| \varepsilon \right\| / \left\| s_k^* \right\|
\]

Where \( \left\| \varepsilon \right\| \) represents the instrumental noise which estimated by recording five spectra of the blank over the studied range. Then the norms of blank readings \( \left( \left\| \text{NAS}_{\text{blank}} \right\| \right) \) are estimated and \( \left\| \varepsilon \right\| \) is taken as the standard deviation of estimated norms\(^ {1,30} \).
REFERENCES


---

### IDMA PUBLICATIONS RATE CARD

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of Publications</th>
<th>Cost in ₹</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>IDMA BULLETIN (Annual Subscription – 48 Issues) (Published on 7th, 14th, 21st and 30th of every month)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Members</td>
<td>1000/- p.a.</td>
</tr>
<tr>
<td></td>
<td>• Government Research / Educational Institutions</td>
<td>2000/- p.a.</td>
</tr>
<tr>
<td></td>
<td>• Non-Members</td>
<td>4000/- p.a.</td>
</tr>
<tr>
<td>2.</td>
<td>INDIAN DRUGS (Annual Subscription – 12 Issues) (Published on 28th of every month)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Members</td>
<td>1000/- p.a.</td>
</tr>
<tr>
<td></td>
<td>• Students</td>
<td>1000/- p.a.</td>
</tr>
<tr>
<td></td>
<td>• Government Research / Educational Institutions</td>
<td>2000/- p.a.</td>
</tr>
<tr>
<td></td>
<td>• Non-Members</td>
<td>4000/- p.a.</td>
</tr>
<tr>
<td>3.</td>
<td>IDMA APA Forum</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Annual Membership</td>
<td>500/-</td>
</tr>
<tr>
<td></td>
<td>• Life Membership</td>
<td>5000/-</td>
</tr>
<tr>
<td>4.</td>
<td>TECHNICAL MONOGRAPHS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NO. 1: STABILITY TESTING OF EXISTING DRUG SUBSTANCES AND PRODUCTS</td>
<td>400/-</td>
</tr>
<tr>
<td></td>
<td>NO. 2: PRIMARY &amp; SECONDARY CHEMICAL REFERENCE SUBSTANCES</td>
<td>400/-</td>
</tr>
<tr>
<td></td>
<td>NO. 3: INVESTIGATION OF OUT OF SPECIFICATION (OOS) TEST RESULTS</td>
<td>400/-</td>
</tr>
<tr>
<td></td>
<td>NO. 4: PHARMACEUTICAL PREFORMULATION ANALYTICAL STUDIES</td>
<td>400/-</td>
</tr>
<tr>
<td></td>
<td>NO. 5: ENVIRONMENTAL MONITORING IN CLEANROOMS</td>
<td>400/-</td>
</tr>
<tr>
<td></td>
<td>NO. 6: CORRECTIVE/PREVENTIVE ACTIONS (CAPA) GUIDELINE</td>
<td>400/-</td>
</tr>
<tr>
<td>5.</td>
<td>IDMA MEMBERSHIP DIRECTORY</td>
<td>1,500/-</td>
</tr>
<tr>
<td>6.</td>
<td>IDMA ANNUAL PUBLICATION</td>
<td>1,500/-</td>
</tr>
</tbody>
</table>

**KINDLY NOTE:**
- Mailing of IDMA Bulletin and Indian Drugs by Post will commence prospectively only after receipt of payment.
- All payments may be made in advance by Cheque / DD / RTGS / NEFT only in favour of: “Indian Drug Manufacturers’ Association”.
- Courier charges for Publications under Serial Nos. 3 to 6 will be extra as applicable.
- Please intimate us details through email immediately after making the remittance through RTGS/NEFT, so as to enable us to do the needful promptly.

**INDIAN DRUG MANUFACTURERS’ ASSOCIATION**
102-B, ‘A’-Wing, Poonam Chambers, Dr A B Road, Worli, Mumbai 400 018. Tel: 2494 4624 / 2497 4308 Fax: 022-2495 0723
E-mail: admin@idmaindia.com/mail_idma@idmaindia.com, Website: www.idma-assn.org / www.indiandrugsonline.org