

Artificial Neural Network Combined with Principal Component Analysis for Resolution of Complex Pharmaceutical Formulations

Giuseppina IOELE,*^a Michele DE LUCA,^a Erdal DINC,^b Filomena OLIVERIO,^a and Gaetano RAGNO^a

^a Department of Pharmaceutical Sciences, University of Calabria; 87036 Rende, CS, Italy; and ^b Department of Analytical Chemistry, Faculty of Pharmacy, Ankara University; 06100 Tandoğan, Ankara, Turkey.

Received June 4, 2010; accepted October 26, 2010; published online October 28, 2010

A chemometric approach based on the combined use of the principal component analysis (PCA) and artificial neural network (ANN) was developed for the multicomponent determination of caffeine (CAF), mepyramine (MEP), phenylpropanolamine (PPA) and pheniramine (PNA) in their pharmaceutical preparations without any chemical separation. The predictive ability of the ANN method was compared with the classical linear regression method Partial Least Squares 2 (PLS2). The UV spectral data between 220 and 300 nm of a training set of sixteen quaternary mixtures were processed by PCA to reduce the dimensions of input data and eliminate the noise coming from instrumentation. Several spectral ranges and different numbers of principal components (PCs) were tested to find the PCA-ANN and PLS2 models reaching the best determination results. A two layer ANN, using the first four PCs, was used with log-sigmoid transfer function in first hidden layer and linear transfer function in output layer. Standard error of prediction (SEP) was adopted to assess the predictive accuracy of the models when subjected to external validation. PCA-ANN showed better prediction ability in the determination of PPA and PNA in synthetic samples with added excipients and pharmaceutical formulations. Since both components are characterized by low absorptivity, the better performance of PCA-ANN was ascribed to the ability in considering all non-linear information from noise or interfering excipients.

Key words principal component analysis; artificial neural network; partial least square

Nowadays, the spectrophotometric techniques remain largely used in the field of pharmaceutical analysis because of the easy interpretation and handling of the spectral data. Furthermore, UV analysis is inexpensive, simply to use and allows to obtain rapidly high accuracy and reproducibility from a small number of samples. The difficulty in the analysis of pharmaceutical preparations using spectroscopic methods depends on the complexity of the mixture and increases with the number of components, particularly when spectral features of these substances are similar or largely overlapped. In some cases, excipients in commercial formulations may give spectral bands overlapping those from drugs. Another problem in a spectral procedure can be the presence of non-linearity, or small deviations from linearity, due to interaction between drugs or excipients.

Multivariate calibration methods, *e.g.* CLS (Classical Least Squares), ILS (Iterative Least Squares), PCR (Principal Component Regression) and PLS (Partial Least Squares), have represented in the last years a valid answer to the multicomponent determination of several drugs in the absence of interferences from excipients. Nevertheless, these classical chemometric methods show some difficulties in the resolution of complex samples, especially when the number of components increases, owing to the above mentioned problems.

Several analytical applications indicate that the combination of mathematical techniques provides more accurate results than those obtained by the application of single classical methods.^{1,2)} In recent years, the hyphenated methods based on principal component analysis and artificial neural network (PCA-ANN) regression have been proposed as calibration tools for the analysis of mixtures when spectral data were complex or when noise interference, nonlinear and interaction interferences were present.^{3–6)} Although the ANN methods have found greater application in nonlinear calibra-

tion systems, they have the ability to model also linear relationships.^{7–9)} A neural network is composed of an input layer and an output layer with one or several hidden layers of neurons. In the ANN model, neurons are used to calculate coefficients or relationships between input and output layers by means of transfer functions. In a network, input and output correspond to the entry data and quantitative outcome, respectively.^{10–12)} The combined use of PCA and ANN usually improves the training speed, enhances the robustness of the model and reduces the calibration errors. Since the ratio between samples and variables in the ANN should be kept as high as possible, PCA is widely used in chemometrics to reduce the number of variables in a data matrix. Several ANN applications in quantitative analysis use PC scores obtained from PCA, as input variables.^{13,14)}

In this work, a new hyphenated chemometric approach based on PCA and ANN was developed for the simultaneous determination of caffeine (CAF), mepyramine (MEP), phenylpropanolamine (PPA) and pheniramine (PNA) in pharmaceutical formulations without any separation step. The drugs are combined all together only in one pharmaceutical specialty but they are also present in many analgesic and antipyretic formulations in different combination as ternary, binary or single-component mixtures.

CAF is a psychoactive stimulant drug having the effect of temporarily warding off drowsiness and restoring alertness. It has diuretic properties too, at least when administered in sufficient doses to subjects who do not have a tolerance for it.¹⁵⁾ MEP is a first generation antihistamine, targeting the H1 receptor. However, it rapidly permeates the brain often causing drowsiness. It is used in combination products to treat the common cold and menstrual symptoms.¹⁶⁾ PPA, also known as norephedrine or oxyamphetamine, is a psychoactive drug of the phenethylamine and amphetamine chemical classes which is used as a stimulant, decongestant and anorectic

* To whom correspondence should be addressed. e-mail: giuseppina.ioele@unical.it

agent.^{17,18)} It is commonly used in prescription and over-the-counter cough and cold preparations. PNA is an antihistamine used to treat allergic conditions such as hay fever or urticaria. It has relatively strong sedative effects, and may sometimes be used off-label as an over-the-counter sleeping pill. Usually, it's used in combination with other drugs.¹⁹⁾ The above drugs have been analyzed in previous studies by using different spectrophotometric^{20–22)} and chromatographic methods.^{23–27)} CAF and MEP have been also quantified in pharmaceuticals by multivariate procedures.^{28–30)}

UV analysis of the quaternary mixture showed an extensive overlap of the drug spectra. The spectral interference resulted very critical for PPA and PNA, because of their low absorptivity and, at the same time, the relatively higher absorptivity of the other two components. For this reason PPA and PNA could more suffer by interference from excipients or instrumental noise. ANN approach appeared so useful to minimize noisy and nonlinear interferences in such a way to obtain a more robust model.

Performance of ANN-PCA was compared with a simple ANN calibration, to investigate the ability of the PCA in extracting the most useful information from the raw data set and with a classical linear calibration technique such as PLS2. PLS2 was preferred to PLS1 because the data processing is made contemporarily for all the components and only one ANN-PCA model is used to predict all four drugs at once. It results extremely effective when applied to systems with component presenting not very different concentrations. On the contrary, PLS1 processing is made for one constituent at a time and is preferable when the concentrations of the components present a high difference.

Validation of all the calibration models was performed by analysis of synthetic mixtures of the target compounds added with tablets excipients, to verify their prediction ability in terms of accuracy and precision. The models were finally applied to the simultaneous quantitative prediction of the four studied compounds in commercial tablets.

Chemometric Techniques. Principal Component Analysis (PCA) Principal component analysis (PCA), also named eigenvector-based multivariate analysis, is widely used in statistics to reduce the number of the variables of a data matrix. Usually, one of the main problems in modern data analysis is precisely the reduction of dimensionality. In fact the multidimensional data sets are difficult to interpret, and their structure cannot be directly visualized.

PCA transforms a number of correlated variables into a smaller number of uncorrelated variables called principal components (PCs), or factors, which can explain sufficiently the data structure. The main idea of PCA is to project the data from a high dimensional space in a lower dimensional space. In addition, PCA helps to find out in what respect one sample is different from another and which variables contribute most to this difference. The data structure so obtained can be visualized directly in a graphical way by projection of objects onto the space defined by the selected PCs. The first PC contains most of the variability in the data and the succeeding PCs that are uncorrelated to former ones carry the remaining variability. The details about this argument is well described in literature.^{31,32)}

Artificial Neural Network (ANN) Artificial neural networks (ANN), also known as “neural networks” (NNs), is

among the most widely used mathematical algorithms for overcoming non-linearity.^{30,33)} In the last time, ANN has demonstrated high ability in acquiring useful information from complex systems, in presence of noise or instrumental fluctuations, providing robust models.^{34,35)}

Usually, ANN is a computer system able to establish relationships between independent and dependent variables directly from raw data and can be used to model complex relationships between inputs and outputs or to find patterns in data. A number of different ANN structures have been proposed.^{3,4)} The structure adopted in this study was the Multi-layer Perceptron (MLP) with back propagation. This technique has demonstrated to be a very powerful data modeling tool, able to capture and represent complex input/output relationships and has been already employed in pharmaceutical analysis.^{36,37)} MLP is a group of highly interconnected neurons arranged in layers with a feed forward structure. ANN architecture is based on input and output layers, named nodes, interconnected *via* one or more hidden layers. The nodes in the input and output layers represent the independent and dependent variables, respectively.

MLP operation is divided into two steps. In a first step of training, the analytical data are presented as input parameters. These data are transformed by a weighting factor and the output values are iteratively fed to the hidden layers. During this step, termed ‘learning,’ the MLP algorithm learns to associate the inputs with the expected outputs. In the last step, called ‘testing’, MLP generates an output signal as a response to previously unknown inputs. The network so develops a model capable to predict the properties of the analytical system.

Partial Least Squares 2 (PLS2) PLS2 is a factor method, that defines a linear relationship between a set of dependent (response) variables and a set of predictor variables. The processing of data is made contemporarily for all the components and the regression simultaneously uses spectral and concentration data. It is an evolution of PLS1 in which, on the contrary, the data processing is made for one constituent at a time. The original variables are transformed into a smaller number of PCs and a new matrix constituted by the new variables PCs and scores allows to build the model for the prediction of the component concentrations of new samples.

The number of PCs must be optimized because the prediction error decreases with the number of PCs used until to reach an optimal value. The most used validation method is the full cross-validation, in which one reference at a time is removed from the calibration set, after that the same sample is predicted by using the calibration built with the others references. The selection of number of PCs can be performed by adopting the minimum SEP, which represents an estimate of the error when other samples are predicted with that model. Detailed description of the PLS technique can be found in literature.^{38,39)}

Experimental

Instruments and Software Absorption spectra were recorded on a wavelength range of 220–300 nm in a 10 mm quartz cell, by a Perkin-Elmer Lambda 40P spectrophotometer at the following conditions: scan rate 1 nm/s; time response 1 s; spectral band 1 nm; data density 1 point/nm. The software UV Winlab 2.79.01 (Perkin-Elmer, Germany) was used for spectral acquisition and elaboration. The data were computed with programs written

in MATLAB 7.0 (Mathworks). Application of PLS2 algorithm was supported by the software package "The Unscrambler 9.7[®]" (Camo Process As., Oslo, Norway).

Chemicals Ethanol 95% was of spectrophotometric grade (J.T. Baker, Holland). PTFE 0.45 μm membrane filters were purchased from Supelco (Milan, Italy). CAF was a generous gift from Onga SpA, Italy; MEP, PPA and PNA were generous gifts from Novartis SpA, Italy.

The pharmaceutical specialty (Triaminic[®] Tablets, Novartis Consumer Health SpA, Italy) was obtained commercially. This formulation contained 25.00 mg CAF, 25.00 mg MEP, 25.00 mg PPA and 25.00 mg PNA per tablet.

Standard and Calibration Solutions Stock solutions of the studied compounds were separately prepared dissolving in ethanol nearly 20.00 mg of CAF, MEP, PPA and PNA in 100 ml calibrated flasks. A calibration set of 16 mixture solutions consisting of the drugs in the concentration range of 2.52–20.16 $\mu\text{g}/\text{ml}$ for CAF, 2.53–20.24 $\mu\text{g}/\text{ml}$ for MEP, 2.48–19.84 $\mu\text{g}/\text{ml}$ for PPA and 2.60–20.80 $\mu\text{g}/\text{ml}$ for PNA was randomly prepared from the stock solutions. In order to optimize the building of the ANN methods, an independent validation set consisting of 12 synthetic mixtures of the four compounds in the above concentration ranges was prepared. A similar second validation set consisting of 12 samples was prepared by adding excipients in order to reproduce as closely as possible the composition of the commercial specialities. The following excipients were used: sodium chloride, polyethylene glycol 4000, sucrose magnesium stearate microcrystalline cellulose and starch. This dataset was used for the external validation of the ANN and PLS2 models.

Pharmaceutical Sample Solutions Pharmaceutical formulations (Triaminic[®] Tablets) were assayed by weighing the content of five tablets and reducing them to a fine powder. An amount exactly corresponding to the average weight was suspended in ethanol and made up to a volume of 10 ml. The suspension was sonicated for 10 min and then filtered through a PTFE 0.45 μm membrane filter. Samples for analysis were obtained after serial dilution 1:10 of this filtrate with ethanol until the concentration ranges above reported and then analyzed. The developed PLS2, ANN and PCA-ANN calibrations were applied to the spectral data so obtained and the predicted amounts of CAF, MEP, PPA and PNA in tablets were carried out.

Results and Discussion

Building of ANN and PLS2 Models The UV spectra of pure CAF, MEP, PPA and PNA were individually recorded in the spectral region between 200–300 nm. A first examination of these spectra pointed out a significant instability of the spectral signals in the region under 220 nm. The study was so pursued by using just the wavelength region between 220 and 300 nm. Figure 1 shows the single spectra of the drugs, at the same concentration ratio of the pharmaceutical tablets, and the mixture spectrum resulting from their sum.

As can be seen in this figure, the spectra of all the compounds overlap along the entire spectral range. Moreover, the absorptivity values of PPA and PNA along the full wavelength range result much lower than the other two components, making so their determination even more difficult to carry out.

The complexity of the studied multicomponent system reduced the chances of success in the simultaneous determination of the components by classical spectroscopy methods, which use measurements at discrete wavelengths. On the contrary, the multivariate methods can provide a greater resolution power because they use simultaneously a large number of signals, allowing so to extract more analytical information.

In this paper, the possibility of predicting the concentration of the four components by UV spectroscopy was studied through the non-linear calibration technique ANN and compared with the linear method PLS2. The ANN approach was also coupled with a PCA procedure in such a way to reduce the number of variables in the data matrix.

A calibration set consisting of 16 quaternary mixture solu-

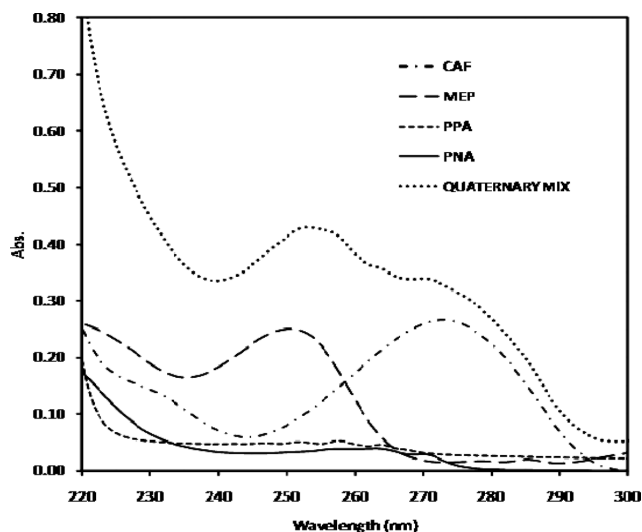


Fig. 1. Absorption Spectra in Ethanol of 4.92 $\mu\text{g}/\text{ml}$ CAF, 5.02 $\mu\text{g}/\text{ml}$ MEP, 5.32 $\mu\text{g}/\text{ml}$ PPA, 5.50 $\mu\text{g}/\text{ml}$ PNA and Their Quaternary Mixture

Table 1. Calibration Set Consisting of Quaternary Mixtures^{a)} in Ethanol with Drug Concentrations on Four Levels

| Sample | CAF | MEP | PPA | PNA |
|--------|-------|-------|-------|-------|
| 1 | 2.52 | 5.06 | 4.96 | 20.80 |
| 2 | 2.52 | 10.12 | 9.92 | 20.80 |
| 3 | 2.52 | 5.06 | 19.84 | 5.20 |
| 4 | 2.52 | 10.12 | 19.84 | 10.40 |
| 5 | 5.04 | 20.24 | 2.48 | 20.80 |
| 6 | 5.04 | 10.12 | 9.92 | 5.20 |
| 7 | 5.04 | 20.24 | 2.48 | 2.60 |
| 8 | 5.04 | 10.12 | 19.84 | 10.40 |
| 9 | 10.08 | 2.53 | 2.48 | 20.80 |
| 10 | 10.08 | 20.24 | 4.96 | 5.20 |
| 11 | 10.08 | 2.53 | 19.84 | 5.20 |
| 12 | 10.08 | 20.24 | 9.92 | 2.60 |
| 13 | 20.16 | 2.53 | 2.48 | 10.40 |
| 14 | 20.16 | 5.06 | 4.96 | 10.40 |
| 15 | 20.16 | 2.53 | 9.92 | 2.60 |
| 16 | 20.16 | 5.06 | 9.92 | 5.20 |

a) Concentration is expressed as $\mu\text{g}/\text{ml}$.

tions in ethanol with drugs at different concentrations was prepared. This training set is summarized in Table 1.

The absorbance values of these solutions in the wavelength set 220–300 nm were collected to build a data set consisting of a 16 \times 81 (samples \times wavelengths) dimension matrix. PLS2, ANN and PCA-ANN procedures were performed on these data. Special care was taken to ensure that, in the concentration ranges studied and for all the mixtures, either the total absorbance did not exceed the linear range of the spectrophotometer and the contribution of each component was additive.

ANN and PCA-ANN methods with different input sets were executed. Table 2 indicates the topology of these networks.

In a first stage, based on the use of the raw spectral data, ANN was directly applied to concentration and absorbance data. The dataset having dimension 16 \times 81 was presented as an input vector to the ANN algorithm. A validation set with dimension 12 \times 81 was also supplied during the training phase for early stopping of the training so to prevent over-fit-

Table 2. ANN Topologies in the Applied Networks

| | ANN | PCA-ANN |
|----------------------------------|----------------|----------------|
| Input | 81 | 4 |
| Hidden | 10 | 5 |
| Output | 4 | 4 |
| Transfer functions ^{a)} | tansig/purelin | tansig/purelin |

^{a)} Tansig and purelin are transfer functions used in Matlab.

ting.

There are no formal ways of determining the size of a neural network and the solutions are evaluated by means of empirical testing. The number of hidden neurons is a function of the complexity of a multicomponent mixture and other experimental factors, the number of input and output parameters and the number of training cases available. Hornik *et al.*⁴⁰⁾ recommend to start with only one hidden layer, and if the results are not good, the number of hidden layers will grow up. In the application of ANN and PC-ANN, a number of different configurations of hidden neurons was tested and the optimal number of neurons in the hidden layer providing the best results was selected.

A two layer ANN was built with the non-linear transfer function “tansig” in first hidden layer and the linear transfer function “purelin” in output layer. By using this structure, a rapid convergence of the inputs was reached and a linear mapping of the outputs provided to decrease the generalization error.

In applying the PCA-ANN approach, the calibration and validation datasets were reduced by PCA and three outputs were gathered. The first output was a loading matrix having size 81×4 (wavelengths \times loadings), each column consisting of coefficients for one PC. The columns were made in order of decreasing component variance. The second output was the matrix of the principal component scores having size 16×4 (samples \times scores). The scores were the data obtained from the transformation of the original data into the space of the PCs. The third output was a vector containing the 4 eigenvalues of the covariance matrix of input. The first four PCs were observed to represent 99.9% of the total variance. Therefore, the first four columns of the score matrix were used to compose new datasets for learning and testing. Both calibration and validation datasets were so presented to the PCA-ANN, which was a two-layered neural network with tansig and purelin transfer functions in hidden and output layers, respectively.

Performance of the two neural networks were compared by generalization accuracy and convergence speed. Figures 2 and 3 point out the calculated errors *versus* the number of iterations for the training step of the neural networks. Performance in terms of mean square error was obtained during the training process of the networks. The error normally decreased during the initial phase of training but, when the network begins to overfit the data, the error on the validation set will typically begin to increase. When the validation error increased for a specified number of iterations, training was stopped and saved.⁴¹⁾

Validation was performed by adopting the standard error of calibration (SEC) as a criterion to estimate the error when the samples are predicted.

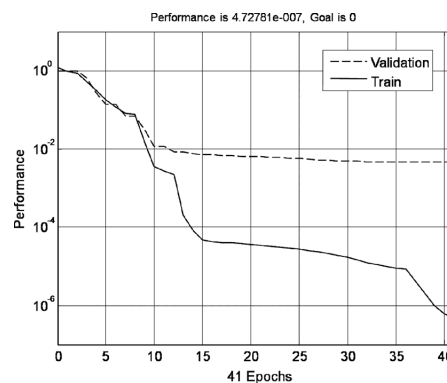


Fig. 2. Error Performance for Training of ANN with Raw Data Inputs

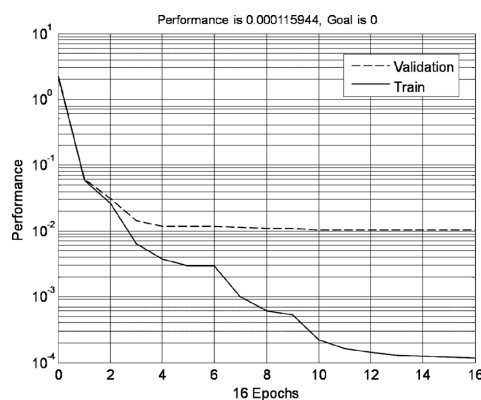


Fig. 3. Performance (Mean Square Error) for Training of ANN with PCA Applied Inputs

Table 3. Statistical Data Obtained from Internal Validation on the Reference Samples by ANN, PCA-ANN and PLS2

| Method | Parameter | CAF | MEP | PPA | PNA |
|---------|----------------|-------|--------|--------|--------|
| ANN | SEC | 0.355 | 0.478 | 0.348 | 0.379 |
| | Slope | 0.987 | 0.985 | 1.001 | 1.027 |
| | Intercept | 0.003 | -0.001 | 0.001 | 0.001 |
| | R ² | 0.984 | 0.980 | 0.987 | 0.975 |
| PCA-ANN | SEC | 0.254 | 0.386 | 0.392 | 0.213 |
| | Slope | 1.002 | 0.975 | 1.015 | 0.997 |
| | Intercept | 0.002 | -0.002 | -0.001 | -0.000 |
| | R ² | 0.998 | 0.999 | 0.992 | 0.986 |
| PLS2 | SEP | 0.219 | 0.243 | 0.360 | 0.262 |
| | Slope | 1.007 | 1.042 | 1.050 | 0.997 |
| | Intercept | 0.002 | -0.003 | -0.003 | -0.002 |
| | R ² | 0.993 | 0.999 | 0.998 | 0.990 |

A third calibration model was defined by PLS2 algorithm, by using concentration values and spectral data in the range 220–300 nm. The model was then validated by full cross-validation and the SEP values calculated each time after a new factor was added to the model. The optimum number of factors was found to be 4.

Validation furnished reliable assay results with SEC values for ANN and PCA-ANN methods not over 0.6 and SEP values for PLS2 not over 0.4. Table 3 shows SEC and SEP values, calculated from the difference between actual and predicted concentrations, and the statistical parameters from the regression analysis on the nominal and predicted concentrations. Correlation coefficients (R^2) confirmed an optimal cor-

Table 4. Statistical Data Obtained from External Validation on the Prediction Samples of ANN, PCA-ANN and PLS2 Methods

| Parameter | ANN | | | | PCA-ANN | | | | PLS2 | | | |
|-------------|--------|--------|--------|--------|---------|-------|--------|--------|--------|--------|--------|--------|
| | CAF | MEP | PPA | PNA | CAF | MEP | PPA | PNA | CAF | MEP | PPA | PNA |
| SEP | 0.148 | 0.308 | 0.295 | 0.323 | 0.094 | 0.130 | 0.216 | 0.194 | 0.153 | 0.076 | 0.279 | 0.253 |
| Slope | 1.011 | 1.016 | 1.032 | 1.027 | 0.998 | 0.992 | 1.028 | 1.021 | 1.049 | 0.972 | 0.955 | 0.968 |
| Intercept | 0.004 | -0.028 | -0.565 | -0.058 | 0.065 | 0.029 | -0.067 | -0.058 | -0.032 | -0.057 | 0.075 | -0.027 |
| Correlation | 1.000 | 0.999 | 1.000 | 0.998 | 1.000 | 0.999 | 0.999 | 1.000 | 0.999 | 0.999 | 0.998 | 0.999 |
| Recovery | 101.48 | 101.21 | 102.41 | 102.22 | 100.74 | 99.27 | 101.34 | 101.08 | 101.48 | 101.72 | 100.64 | 102.02 |
| S.D. | 1.93 | 2.07 | 1.93 | 2.43 | 1.35 | 1.46 | 2.08 | 1.21 | 1.63 | 1.05 | 2.12 | 1.53 |
| R.S.D. | 1.90 | 2.05 | 1.88 | 2.37 | 1.34 | 1.47 | 2.05 | 1.20 | 1.28 | 1.72 | 1.84 | 2.68 |

a) Concentration is expressed as $\mu\text{g/ml}$; S.D.: standard deviation; R.S.D.: relative standard deviation.

Table 5. Statistical Results by Applying ANN, PCA-ANN and PLS2 Methods to the Analysis of Pharmaceutical Formulations

| | ANN | | | | PCA-ANN | | | | PLS2 | | | |
|--------------------|--------|--------|-------|-------|---------|-------|--------|-------|-------|--------|--------|-------|
| | CAF | MEP | PPA | PNA | CAF | MEP | PPA | PNA | CAF | MEP | PPA | PNA |
| Mean ^{a)} | 25.74 | 25.82 | 24.38 | 24.12 | 24.92 | 25.45 | 25.82 | 24.03 | 24.53 | 26.24 | 27.53 | 23.87 |
| S.D. | 0.53 | 0.66 | 0.58 | 0.65 | 0.32 | 0.43 | 0.41 | 0.33 | 0.52 | 0.42 | 1.15 | 1.35 |
| R.S.D. | 2.06 | 2.56 | 2.90 | 3.24 | 1.28 | 1.69 | 1.49 | 1.57 | 2.12 | 1.60 | 3.89 | 4.06 |
| Recovery | 102.96 | 103.28 | 97.52 | 96.48 | 99.68 | 101.8 | 103.28 | 96.12 | 98.12 | 104.96 | 110.12 | 95.48 |

a) Results are expressed as mg/tablet and are referred to average of four replicates ($n=4$); S.D.: standard deviation; R.S.D.: relative standard deviation.

respondence between actual and predicted concentrations.

External Validation of the Models The defined chemometric models were validated through analysis of an external prediction set consisting of 12 quaternary mixtures with the same concentration ranges adopted in the calibration samples and spiked with excipients commonly used in tablet formulations. The values of SEP were calculated for all the methods and components and summarized in Table 4.

The parameters from the regression analysis between nominal and predicted concentrations, mean recoveries and relative standard deviations were also calculated and listed in the same table. Mean accuracy, expressed as % recovery (\pm R.S.D.), was found to be 101.58 (\pm 2.05), 100.61 (\pm 1.52) and 99.72 (\pm 2.42) for ANN, PCA-ANN, and PLS2, respectively.

A good coincidence was observed among the results obtained by the application of all the models in the estimation of CAF and MEP. PCA-ANN gave brilliant determination results also in the assay of PPA and PNA, slightly better than the results obtained by application of the simple ANN. Probably, the reduction of dimensionality of the input data allowed a selection of the most useful analytical information. In contrast, a significant difference between nominal amounts and predicted values were obtained in the prediction of PPA and PNA when the model PLS2 was applied. This inaccuracy was supposed due to the low absorptivity of these components, which made them more sensitive to instrumental noise or to interferences caused by some excipients. PCA-ANN gave the best results presumably because this method is particularly well suited to account for any non-linear information from a such complex analytical system.

Application of the Models to the Assay of Pharmaceuticals The above defined PLS2, ANN and PCA-ANN models were applied to the determination of CAF, MEP, PPA and PNA content in the commercial tablets, following the procedure above reported. The obtained results are shown in Table

5, also listing some statistical parameters, as standard deviation, percent relative standard deviation and standard error.

The measurements and the statistical parameters confirmed the results carried out in the external validation on the synthetic mixtures. A satisfactory coincidence between labeled and predicted amounts was observed in the estimation of all the drugs by applying the model PCA-ANN. Slightly less accurate results were obtained from application of the ANN model and significant errors were recorded in the prediction of PPA and PNA when the model PLS2 was used.

Conclusion

In this work, ANN and PCA-ANN multivariate methods demonstrated high resolution power in assaying a very complex pharmaceutical mixture, consisting in four drugs presenting a severe overlapping of their UV spectra. Great advantages of these methods are simplicity, rapidity and low cost, requiring neither sophisticated instrumentation nor any prior separation procedure. The ANN models were compared with a PLS2 model. The statistical parameters showed that both the ANN models were reliable, but PCA-ANN demonstrated a better performance in the prediction of the validation set samples, showing lower residual errors. Moreover, convergence speed for PCA-ANN was shorter than ANN, as verified in the training graphs. PLS2 also gave satisfactory results for the assay of CAF and MEP but proved to be less accurate in assaying PPA and PNA, probably for the non-linear interferences due to instrumental noise or the presence of excipients from tablet extraction. The successful application of the proposed methods to the quality control analysis of quaternary pharmaceutical preparations, demonstrated that the ANN procedures can be a valid alternative to the classical instrumental or multivariate procedures also for linear data systems.

Acknowledgements The authors thank the Ministry of University and

Research (MIUR) of Italy for financial support of this work.

References

- 1) Dinç E., Ragno G., Ioele G., Baleanu D., *J. AOAC Int.*, **89**, 1538—1546 (2006).
- 2) Dinç E., Kanbur M., Baleanu D., *Spectrochim. Acta A Mol. Biomol.*, **68**, 225—230 (2007).
- 3) Jin-Mei X., Xiao-Jian W., Ying-Jin Y., *Metabolomics*, **3**, 531—537 (2007).
- 4) Forshed J., Andersson F. O., Jacobsson S. P., *J. Pharm. Biomed. Anal.*, **29**, 495—505 (2002).
- 5) Yin C., Shen Y., Liu S., Yin Q., Guo W., Pan Z., *Comput. Chem.*, **25**, 239—243 (2001).
- 6) Absalan G., Soleimani M., *Anal. Sci.*, **20**, 879—882 (2004).
- 7) Khanchi A. R., Mahani M. K., Hajihosseini M., Maragheh M. G., Chalooosi M., Bani F., *Food Chem.*, **103**, 1062—1068 (2007).
- 8) Guibertau A., Galeano T., Mora N., Salinas F., Ortiz J. M., Virè J. C., *Comput. Chem.*, **25**, 459—473 (2001).
- 9) Ni Y., Zhang G., Kokot S., *Food Chem.*, **89**, 465—473 (2005).
- 10) Agatonovic-Kustrin S., Beresford R., *J. Pharm. Biomed. Anal.*, **22**, 717—727 (2000).
- 11) Tominaga Y., *Chemom. Intell. Lab. Syst.*, **49**, 105—115 (1999).
- 12) Plumb A. P., Rowe R. C., York P., Brown M., *Eur. J. Pharm. Sci.*, **25**, 395—405 (2005).
- 13) Gemperline P. J., Long R., Gregoriou V. G., *Anal. Chem.*, **63**, 2313—2323 (1991).
- 14) Zarei K., Atabati M., Kazemi L., *Il Farmaco*, **60**, 37—42 (2005).
- 15) Maughan R. J., Griffin J., Hum J., *Nutr. Diet.*, **16**, 411—420 (2003).
- 16) Parsons M. E., Ganellin C. R., *Br. J. Pharmacol.*, **147**, 127—135 (2006).
- 17) Flavahan N. A., *J. Pharmacol. Exp. Ther.*, **313**, 432—439 (2005).
- 18) Kernan W. N., Viscoli C. M., Brass L. M., Broderick J. P., Brott T., Feldmann E., Morgenstern L. B., Wilterdink J. L., Horwitz R. I., *N. Engl. J. Med.*, **343**, 1826—1832 (2000).
- 19) Korduba C. A., Veals J., Symchowicz S., *Life Sci.*, **13**, 1557—1564 (1973).
- 20) Abbaspour A., Mirzajani R., *J. Pharm. Biomed. Anal.*, **38**, 420—427 (2005).
- 21) Dinç E., Kökdil G., Onur F., *J. Pharm. Biomed. Anal.*, **26**, 769—778 (2001).
- 22) Shama S. A., Amin A. S., *Spectrochim. Acta A Mol. Biomol.*, **60**, 1769—1774 (2004).
- 23) Kartal M., *J. Pharm. Biomed. Anal.*, **26**, 857—864 (2001).
- 24) Abbasi K., Bhanger M. I., Khuhawar M. Y., *J. Pharm. Biomed. Anal.*, **41**, 998—1001 (2006).
- 25) Ferreyra C., Ortiz C., *J. Pharm. Biomed. Anal.*, **25**, 493—499 (2001).
- 26) Louhaichi M. R., Jebali S., Loueslati M. H., Adhoum N., Monser L., *Talanta*, **78**, 991—997 (2009).
- 27) Mikus P., Valásková I., Havránek E., *J. Pharm. Biomed. Anal.*, **38**, 442—448 (2005).
- 28) Ragno G., Ioele G., Risoli A., *Anal. Chim. Acta*, **512**, 173—180 (2004).
- 29) De Luca M., Oliverio F., Ioele G., Ragno G., *Chemom. Intell. Lab. Syst.*, **96**, 14—21 (2009).
- 30) Aksu Ö., Bozdogan A., Kunt G., *Anal. Lett.*, **39**, 751—761 (2006).
- 31) Meloun M., Militký J., Forina M., “Chemometrics for Analytical Chemistry,” Vol. 1, PC-Aided Statistical Data Analysis, Ellis Horwood, Chichester, 1992.
- 32) Beebe K. R., Pell R. J., Seasholtz M. B., “Chemometrics: a Practical Guide,” Wiley-Interscience Ed., New York, 1998.
- 33) Dinç E., Baleanu D., Ioele G., De Luca M., Ragno G., *J. Pharm. Biomed. Anal.*, **48**, 1471—1475 (2008).
- 34) Sun Y., Peng Y., Chen Y., Shukla A. J., *Adv. Drug Deliv. Rev.*, **55**, 1201—1215 (2003).
- 35) Takayama K., Fujikawa M., Obata Y., Morishita M., *Adv. Drug Deliv. Rev.*, **55**, 1217—1231 (2003).
- 36) Hussain A. S., Yu A., Johnson R. D., *Pharm. Res.*, **8**, 1248—1252 (1991).
- 37) Chen Y., Jiao T., McCall T. W., Baichwal A. R., Meyer M. C., *Pharm. Dev. Technol.*, **7**, 373—379 (2002).
- 38) Geladi P., Kowalski B. R., *Anal. Chim. Acta*, **185**, 1—17 (1986).
- 39) Wold S., Sjöström M., Eriksson L., *Chemom. Intell. Lab. Syst.*, **58**, 109—130 (2001).
- 40) Hornik K., Stinchcombe M., White H., *Neural Networks*, **2**, 359—366 (1990).
- 41) Haykin S., “Neural Networks: A Comprehensive Foundation,” Macmillan, New York, 1994.