Available online at www.sciencedirect.com
SCIENCE DIRECT
Talanta 68 (2006) 1222–1229
Talanta
www.elsevier.com/locate/talanta

Net analyte signal-based simultaneous determination of antazoline and naphazoline using wavelength region selection by experimental design-neural networks

Bahram Hemmateenejad[a,b,*], Raoof Ghavami[a], Ramin Miri[b], Majtaba Shamsipur[c,*]

[a] *Department of Chemistry, Shiraz University, Shiraz, Iran*
[b] *Medicinal & Natural Products Chemistry Research Center, Shiraz University of Medical Science, Shiraz, Iran*
[c] *Department of Chemistry, Razi University, Kermanshah, Iran*

Received 27 January 2005; received in revised form 27 May 2005; accepted 15 July 2005
Available online 24 August 2005

Abstract

Net analyte signal (NAS)-based multivariate calibration methods were employed for simultaneous determination of anthazoline and naphazoline. The NAS vectors calculated from the absorbance data of the drugs mixture were used as input for classical least squares (CLS), principal component and partial least squares regression PCR and PLS methods. A wavelength selection strategy was used to find the best wavelength region for each drug separately. As a new procedure, we proposed an experimental design-neural network strategy for wavelength region optimization. By use of a full factorial design method, some different wavelength regions were selected by taking into account different spectral parameters including the starting wavelength, the ending wavelength and the wavelength interval. The performance of all the multivariate calibration methods, in all selected wavelength regions for both drugs, was evaluated by calculating a fitness function based on the root mean square error of calibration and validation. A three-layered feed-forward artificial neural network (ANN) model with back-propagation learning algorithm was employed to model the nonlinear relationship between the spectral parameters and fitness of each regression method. From the resulted ANN models, the spectral regions in which lowest fitness could be obtained were chosen. Comparison of the results revealed that the net NAS-PLS resulted in lower prediction error than the other models. The proposed NAS-based calibration method was successfully applied to the simultaneous analyses of anthazoline and naphazoline in a commercial eye drop sample.

Keywords: Naphazoline; Antazoline; Net analyte signal; Wavelength selection; Multivariate calibration; Artificial neural network

1. Introduction

Multivariate spectral calibrations, which are now become standard methods for performing quantitative spectral analysis, allow the simultaneous determination of several analytes in a given mixture [1]. Partial least squares (PLS) and principal component regression (PCR) are the most common multivariate calibration methods for quantitative spectral analysis [2–4]. These full spectrum multivariate calibration methods offer the advantages of speed in the determination of components of interests in mixtures and avoiding separation steps in the analytical procedures.

Recently, a new family of multivariate calibration methods based on the concept of net analyte signal (NAS) has been proposed [5–18]. NAS was defined by Lorber [5] as the part of a mixture spectrum that is useful for model building; this implies that NAS is part of the spectrum of mixture that is orthogonal to the spectra of interferences and background variations. The NAS calculations are used both for estimation of the figures of merit of an analytical method [6–8] and for the construction of multivariate calibration models [9–12]. The hybrid linear analysis (HLA) methods developed by Xu and Schechter (HAL/XS)

* Corresponding author. Tel.: +98 831 4274580; fax: +98 831 4274503.
E-mail address: mshamsipur@yahoo.com (M. Shamsipur).

0039-9140/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2005.07.045

[13] and Goicoechea and Olivieri (HLA/GO) [14] are constructed based on the NAS concept. Other NAS-based multivariate calibration methods have also been proposed in which the vectors of NAS of mixtures are used as input for other multivariate calibration methods such as classical least squares (NAS/CLS), principal component regression (NAS/PCR) and partial least squares regression (NAS/PLS) [15–18]. Moreover, in a simple case, the norm of the NAS vector can be used to construct a univariate calibration model, where this parameter is plotted against the analyte concentration and a linear relationship is observed, ideally.

Application of the mentioned methods to multicomponent spectroscopic analysis usually requires spectral variable selection for building well-fitted models and avoiding non-modeled interferences [19–21]. Training the multivariate calibration methods with selected spectral regions, rather than full-spectrum region, allows the informative part of the spectrum, which is related to the variation of concentration of analyte, to be modeled and, therefore, other parts of spectrum which are related to the variation of concentration of other analytes and/or background variations will be discarded. Hence, the performance of multivariate calibration models will be enhanced. Several approaches have been proposed for selection of optimal set of spectral regions for multivariate calibration such as generalized simulating annealing [22], genetic algorithms [23], artificial noise introduction in PLS modeling [24], wavelet transform [25], successive projections algorithm [26] and moving windows selection strategy [27]. Genetic algorithms (GA) are an interesting, flexible and widely used variable selection method among different proposed strategies. In recent years, the moving windows strategy (MWS) have been applied for wavelength region selection [15–18]. In this method, multivariate calibration methods are run on different selected wavelength windows and, consequently, those revealing better results are selected. Two spectral parameters (i.e. starting λ and ending λ) are varied in order to select the spectral windows. However, the interval between the sensors (i.e. wavelength interval, WI) in the selected windows is another MWS requirement that should be optimized to obtain most relevant results.

Artificial neural networks (ANN) are nonparametric nonlinear modeling techniques that have attracted increasing interest in recent years [28–30]. Nonlinear multivariate maps use a nonlinear transformation of the input variable space to project inputs onto the designated attribute values in output space. The strength of modeling with layered, feed-forward artificial neural networks lies in the flexibility of the distributed soft model defined by the weight of the network. Both linear and nonlinear mapping functions may be modeled by suitably configuring the network. Multilayer feed-forward neural network trained with back-propagation learning algorithm become increasingly popular techniques [31,32]. Recently, Capitán-Vallvey et al. applied Kohonen neural network for wavelength selection in PLS calibration [33].

In the present work, we applied an experimental design strategy for variation of the three spectral parameters [34]. Very different combinations of spectral parameters were selected and, in each combination, the performance of multivariate calibration models used was evaluated. An artificial neural network model with back-propagation learning algorithm was used to model the relationship existed between the prediction ability of the models and the corresponding selected spectral parameters. From the optimized ANN model, those spectral parameters that give most of the results were selected for each multivariate calibration model. Four different NAS-based multivariate calibration models including NAS/CLS, NAS/PCR, NAS/PLS and univariate extension of NAS calculation were used. The methods were used for simultaneous determination of antazoline and naphazoline in nasal solution.

Naphazoline hydrochloride (NAP) [2-(1-naphtylmethyl)-2-imidazoline monohydrochloride] is a sympathomymetic, belonging to the imidazole group. It is a vasoconstrictor of relatively long-lasting action that acts on the α receptors of the vascular smooth muscle [35]. Antazoline (ANT) is another imidazoline ligand which has alpha 2-adrenoceptor antagonistic properties. It is now widely recognized that imidazoline derivatives provided with α_2-adrenoceptor antagonist properties not only exhibit a high affinity for α_2-adrenoceptors but also bind to non-adrenergic imidazoline sites in various tissues [36]. Pharmaceuticals containing the two active ingredients (anthazoline and naphazoline) are currently commercialized in our country as nasal solution.

2. Theoretical background

CLS, PLS, ANN and experimental design are well documented in the literature and, therefore, these methods are not discussed. The readers can refer to the cited literature for more discussion about these methods [1–4]. Here, only those sections related to NAS calculations and calibrations are discussed.

2.1. Notations

In this section, the following notations will be used. A capital letter in boldface demonstrates a matrix and a lowercase letter in boldface denotes a vector. Lowercase italic letters denote the scalars. Meanwhile, the following matrices and vectors will be used through the present work: the $m \times n$ data matrix (**R**) composed of the absorbance data of m samples at n wavelengths (or sensors), a $1 \times n$ vector of the pure spectrum of analyte k ($\mathbf{s}_k$), a $m \times 1$ vector of calibration concentrations of analyte k ($\mathbf{c}_k$) and a $1 \times n$ vector of unknown sample absorbance data (**r**).

2.2. Net analyte signal calculations

Lorber has defined the NAS for an analyte k in a given mixture as the part of its spectrum, which is orthogonal to the space spanned by the spectra of all other analytes and interferences [5]. Some different algorithms have been proposed for NAS calculations [9–12]. In this work, the algorithm suggested by Lorber for inverse multivariate calibrations was employed [12]. First, the PLS regression was done on the calibration data ($\mathbf{R}$ and $\mathbf{c}_k$) and, at optimum number of factors, the absorbance data matrix was reconstructed ($\hat{\mathbf{R}}$). After that, the $\hat{\mathbf{R}}_{-k}$ matrix containing the absorbance data of all existing species in the calibration sample expect the analyte k was calculated by rank annihilation method:

$$\hat{\mathbf{R}}_{-k} = \hat{\mathbf{R}} - \alpha \hat{\mathbf{c}}_k \mathbf{d}_k \quad (1)$$

where $\mathbf{d}_k$ is the spectral information of the analyte of interest, $\hat{\mathbf{c}}_k$ is the estimated concentration of the analyte k by the following equation:

$$\hat{\mathbf{c}}_k = \hat{\mathbf{R}}\hat{\mathbf{R}}^{+}\mathbf{c}_k \quad (2)$$

the superscripts "+" denote the matrix pseudo-inverse and α is a scaling factor calculated by Eq. (3):

$$\alpha = \frac{1}{\mathbf{d}_k \hat{\mathbf{R}}^{+} \mathbf{c}_k} \quad (3)$$

A projection matrix, which is orthogonal to $\hat{\mathbf{R}}_{-k}$ is defined as:

$$\mathbf{H} = \mathbf{I} - (\hat{\mathbf{R}}_{-k})^{+}\hat{\mathbf{R}}_{-k} \quad (4)$$

Net analyte signal vector ($\mathbf{r}_k^*$) is calculated by post multiplying of projection matrix ($\mathbf{H}$) by the vector spectra of mixture ($\mathbf{r}$):

$$\mathbf{r}_k^* = \mathbf{H}\mathbf{r} \quad (5)$$

If $\mathbf{r}$ replaced by the matrix of absorbance data ($\mathbf{R}$), the left hand side of the Eq. (5) will be the matrix of NAS of the same size as $\mathbf{R}$.

2.3. NAS-based calibrations

The NAS vector ($\mathbf{r}_k^*$) is a characteristic of analyte of interest and its length (norm) is directly related to the analyte concentration:

$$\left\|\mathbf{r}_k^*\right\| = bc \quad (6)$$

This is a type of Bear–Lambert equation. || || denote the norm of a vector and is a scalar and b is a proportionality constant.

In Eq. (5), if $\mathbf{r}$ replaced by the pure spectrum of analyte k ($\mathbf{s}_k$), the left hand side of this equation will be the net pure spectrum of analyte k, i.e. a part of pure spectrum which is orthogonal to the spectra of other absorbing species ($\mathbf{s}_k^*$). The relationship between the matrix of NAS ($\mathbf{R}^*$), $\mathbf{s}_k^*$ and the concentration of analyte in the mixtures can be written in the classical least square (CLS) form:

$$\mathbf{R}^* = \mathbf{c}_k \mathbf{s}_k^* \quad (7)$$

In CLS regression, the major requirement is that the concentration of all absorbing components should be known and this limits the application of CLS in the simultaneous determinations. However, in the NAS/CLS model (Eq. (7)), $\mathbf{R}^*$ contains the spectral information of the analyte of interest (not the absorbance contribution of other coexisting components). Therefore, CLS can be successfully applied to the NAS data.

NAS/CLS is an ideal model in which it is assumed that calculated NAS only contains the spectral information of the analyte of interest. But, in some instances, the calculated NAS should be containing some spectral impurity of other sources rather than analyte because of interactions between components, some non-linearities or other sources. In these cases, the NAS/CLS is not the best choice and inverse calibration methods especially factor analysis based methods can overcome these problems. NAS/PCR and NAS/PLS are inverse regression methods in which the NAS is used as input of PCR and PLS regression methods.

3. Experimental

3.1. Apparatus and reagents

Antazoline and naphazoline (both from Merck Company) stock solutions were prepared by dissolving appropriate amounts of each drug in triply distilled water.

All spectra were recorded on an Ultrospec 3000 pro (Pharmacia Biotech) UV–vis spectrophotometer equipped with 10 mm quartz cells. The Swift(II) software was used to collect the absorbance data of the solution into a spreadsheet.

All necessary programs needed for NAS calculation, CLS, PCR and PLS regressions and ANN modeling were written in MATLAB 6.0 (MathWorks Inc.) and run on a Pentium IV personal computer with Windows XP operating system.

3.2. Methodology

Two sets of the standard solutions of mixtures of the two drugs were prepared (36 calibration solutions and 16 validation solutions). The calibration set solutions were prepared according to 6-level full factorial design. Using such a design, maximum information for each compound can be obtained by using only a few numbers of standard solutions. In Table 1, the concentrations of standard solutions are represented. In order to evaluate the performance of the employed models, 16 standard solutions were considered in the validation set whose concentrations were selected randomly (Table 2). For preparation of each standard solution, appropriate volumes of the stock solution of each drug were added to 10.0 ml

Table 1
Concentrations of drugs in the calibration set ($\mu g\,ml^{-1}$)

Solution no.	NAP	ANT	Solution no.	NAP	ANT
1	0.87	0.82	19	26.2	0.82
2	0.87	9.00	20	26.2	9.00
3	0.87	17.0	21	26.2	17.0
4	0.87	25.3	22	26.2	25.3
5	0.87	33.5	23	26.2	33.5
6	0.87	43.0	24	26.2	43.0
7	8.72	0.82	25	36.5	0.82
8	8.72	9.00	26	36.5	9.00
9	8.72	17.0	27	36.5	17.0
10	8.72	25.3	28	36.5	25.3
11	8.72	33.5	29	36.5	33.5
12	8.72	43.0	30	36.5	43.0
13	17.4	0.82	31	47.0	0.82
14	17.4	9.00	32	47.0	9.00
15	17.4	17.0	33	47.0	17.0
16	17.4	25.3	34	47.0	25.3
17	17.4	33.5	35	47.0	33.5
18	17.4	43.0	36	47.0	43.0

volumetric flasks and diluted with triply distilled water. The absorbance spectra of each solution was recorded in the wavelength region of 230–330 nm and digitized in 1.0 nm intervals. The order of recording of absorbance spectra of calibration and validation samples were chosen randomly in order to neglect the effect of recording time.

The NAS vectors of each standard solution were calculated according to the procedure discussed in the theoretical section. The NAS/CLS, NAS/PCR, NAS/PLS programs were written in MATLAB. Leave-one-out cross validation combining with the *F*-test criterion of Thomas and Haaland [37] were used to select optimum number of factors in the NAS/PCR and NAS/PLS models. The performance of each model was evaluated by root mean square error of cross validation (RMSECV) root mean square error of calibration (RMSEC), root mean square error of validation (RMSEV) and the multiple of determination coefficient for calibration and prediction (R_C^2 and R_V^2, respectively).

Table 2
Concentrations of drugs in the prediction set in $\mu g\,ml^{-1}$ (actual), their corresponding predicted values (pred) and percent of relative error (%RE) obtained by NAS/PLS model

Solution no.	NAP			ANT		
	Actual	Pred	%RE	Actual	Pred	%RE
1	1.60	1.65	3.13	17.6	18.0	2.27
2	6.40	5.95	−7.03	3.20	3.45	7.81
3	11.2	11.5	2.68	38.4	37.2	−3.12
4	16.0	15.2	−5.26	25.6	25.5	−0.39
5	20.8	21.6	3.85	22.4	22.3	0.45
6	20.8	20.1	−3.36	9.60	9.20	−4.15
7	22.4	23.3	4.01	33.6	31.5	−6.25
8	36.8	35.6	−3.26	14.4	13.5	−6.25
9	43.2	42.2	−0.23	27.2	27.2	0.00
10	14.4	14.6	1.39	16.0	15.6	−2.50
11	33.6	34.3	2.08	20.8	20.5	−1.44
12	8.00	7.85	−1.87	17.6	17.4	−1.14
13	27.2	27.7	1.69	24.0	24.1	0.42
14	22.4	21.0	−6.25	8.00	8.4	5.67
15	33.6	33.8	0.59	11.2	10.5	−6.25
16	4.80	4.56	−5.00	36.8	36.8	0.00

3.3. Experimental design-neural network for wavelength selection

Starting wavelength (SW), ending wavelength (EW) and wavelength interval (WI) are three common spectral parameters that define a window of wavelength region. These spectral parameters were used as factorial design factors. The SW was varied between 200 nm and 280 nm in 20 intervals and EW was selected between 230 nm and 320 nm in 10 nm intervals. The WI was varied between 1 nm and 7 nm. By using the full factorial design, different combinations of spectral factors were selected and, consequently, different wavelength regions were determined. Sub-matrices of the original absorbance data matrix were constructed considering the selected wavelength region properties. Different NAS-based multivariate calibration methods were run on the sub-matrices and in each case the RMSEC and RMSEV were calculated. The fitness function (η), was calculated using the RMSEC and RMSEV:

$$\eta = \left\{ \frac{[(m_c - n - 1)]\mathrm{RMSEC}^2 + m_v \mathrm{RMSEV}^2]}{(m_c + m_v - n - 1)} \right\}^{1/2} \quad (8)$$

where m_c and m_v are the number of standard solutions in the calibration and validation sets, respectively, and n represents the number of sensors (wavelengths) used in each spectral region.

Now, for each multivariate calibration model, there is a ($m_w \times n_w$) matrix of spectral factors (m_w being the number of selected spectral windows and n_w the number of spectral factors, i.e. $n_w = 3$) and a ($m_w \times 1$) vector of fitness. A three-layered feed-forward artificial neural network model with back-propagation of error algorithm was used to model the relationship between the matrix of spectral factors (**D**) as independent variables (i.e. input of network) and fitness function (**f**) as dependent variable. The network used here was one that has been discussed previously.

4. Results and discussion

Electronic absorption spectra of aqueous solutions of 20.0 $\mu g\,ml^{-1}$ of ANT and NAP and a mixture containing 10.0 $\mu g\ ml^{-1}$ of each drug are shown in Fig. 1. The high degree of overlapping of spectra of drugs is obvious from this figure.

The first step in any multivariate calibration method is finding the linear working range for each analyte, separately. For this purpose, a single-component calibration curve was plotted for each drug at its λ_{max} (i.e. 242 nm and 280 nm for ANT and NAP, respectively). It was found that the calibration curves were linear in the concentration ranges

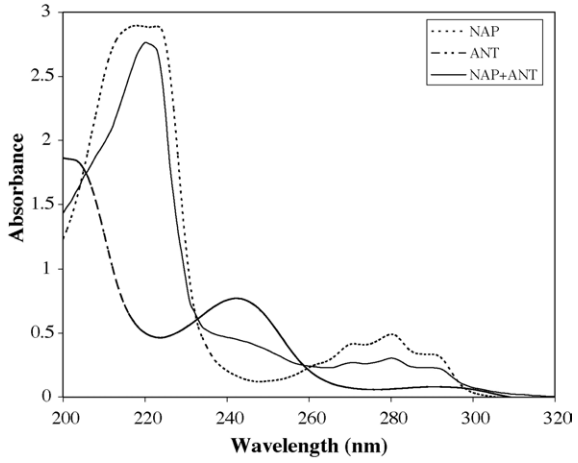

Fig. 1. Absorbance spectra of 20 μg ml^{-1} of ANT and NAP and a mixture of 10 μg ml^{-} of both drugs.

of 0.80–75.00 μg ml^{-1} for ANT and 1.00–65.0 μg ml^{-1} for NAP. However, in order to insure that the maximum absorbance of mixtures do not exceed 3.0, the maximum concentrations of the two drugs used for preparation of standard mixture solutions were chosen to be 50.0 μg ml^{-1}.

For multivariate calibration modeling, two sets of standard solutions were prepared. The calibration mixtures (36 samples) were used for model building and the validation mixtures (16 samples) were used to evaluate the performance of the constructed models.

4.1. Wavelength region selection

Application of the multivariate calibration methods to the multi-component spectroscopic analysis usually requires spectral variable selection for building well-fitted models and avoiding non-modeled interferences. Meanwhile, through the wavelength selection procedure, the non-informative parts of spectrum of collinear absorbance data are discarded from the original data. The absorbance spectra of the mixtures of ANT and NAP (Fig. 1) can be divided to three regions. These include below 235 nm, in which both components possess high absorbances so that the absorbance of some mixtures are exceeded that of the instrument reading, between 235 nm and 265 nm which contain more information about the absorbance data of ANT, and beyond 260 nm region, in which the spectral information due to NAP is predominant.

In this section, attempts were made to select the best wavelength range for each drug separately. For this purpose, an experimental design strategy was used to select different combinations of spectral parameters. The performances of the four different multivariate calibration methods employed at the selected regions were evaluated. Some different criteria are existed for evaluation of the multivariate calibration methods through wavelength selection regions including selectivity (SEL) proposed by Lorber, RMSCV, RMSECV and, recently, an error indicator function (EIF) proposed by Goicoechea and Olivieri [14]. Here, the authors used the concepts of selectivity, sensitivity and detection limit to measure the quality of the models. Therefore, they found spectral regions that had the optimum values for the selectivity and sensitivity as measured by the NAS calculations. In almost all approaches, the prediction ability of the optimized model was emphasized in order to avoid the over fitting problem. However, in some instances, under fitting problem could be obtained by using these criteria. Thus, in this work, we used the fitness function (η), which is calculated based on the use of both the RMSEC and RMSEV, to avoid both under-fitting and over-fitting problems for model evaluation (Eq. (8)).

The ANN model used here for modeling the nonlinear relationship between spectral parameters and fitness functions for each multivariate calibration method contains three layers: an input layer, a hidden layer and an output layer. The input layer has three nodes (spectral parameters) including SW, EW and WI and the output layer contained only one node (fitness function). A bias unit with a constant activation of unity was connected to each unit in the hidden and output layers. The number of nodes in the hidden layer was optimized through a learning procedure. Several network configurations were tested, each with a different number of hidden layer elements. The ANN models confined to a single hidden layer because network with more than one hidden layer would be harder to train. To insure that the over fitting and under fitting of the ANN model do not occur, the fitness function (η) was calculated from the training and validation data for each configuration. For each multivariate calibration model, the training was stopped when no further improvement was found in the fitness function. The network parameters including learning rate (η), momentum (α), number of nodes in the hidden layer (N_H) and transfer functions in the hidden and output layers was optimized to give a lower fitness function. It was found that, for all of the neurons, the sigmoid and linear transfer functions used for hidden and output layers, respectively, give better results. The network parameters

Table 3
Neuron architectures used for wavelength region optimization of different multivariate regression methods

Regression method	ANT				NAP			
	N_H	η	α	Epoches	N_H	η	α	Epoches
NAS/CLS	3	0.94	0.14	28000	2	0.98	0.22	23000
NAS/PCR	3	0.85	0.26	22000	3	0.0.91	0.19	16000
NAS/PLS	2	0.97	0.33	21000	2	0.87	0.24	18000

Table 4
Optimized spectral parameters obtained by ANN modeling

Regression method	ANT			NAP		
	SW	EW	WI	SW	EW	WI
NAS/CLS	240	280	6	280	320	5
NAS/PCR	250	270	1	270	320	4
NAS/PLS	240	260	1	280	300	3

of the optimized network for each multivariate calibration method are presented in Table 3.

After the weights and the structures of the optimized networks were determined, the spectral parameters were scanned in 1.0 nm intervals. In each case, the fitness of the multivariate calibration model was predicted by the ANN models and a spectral region with lowest fitness function was selected. The results are listed in Table 4.

4.2. NAS calculations

The NAS calculations were performed for each analyte separately. By using of the calibration data, the projection matrix (**H**) was calculated for each drug (Eq. (5)). The number of PLS latent variables required for the PLS estimation of the original data matrix ($\hat{\mathbf{R}}$) was found to be 3 for both drugs. The NAS vectors for each mixture, and for each drug, were obtained by post-multiplying of the absorbance vector of each mixture (both the calibration and prediction mixtures) by the projection matrix (Eq. (6)). In order to evaluate $\mathbf{s}_k^*$, the pure component spectrum of each drug was first calculated. In Fig. 2 are shown the calculated NAS of ANTH and NAPH for the prediction mixtures. The variations in the absorbance spectra of mixtures are not directly related to the concentration changes of any individual analyte. While, as it is obvious from the plots shown in Fig. 2, the variations of the calculated NAS vectors for each drug are related directly to the concentration changes of each analyte. The plots of $\mathbf{r}_k^*$ against $\mathbf{s}_k^*$ for

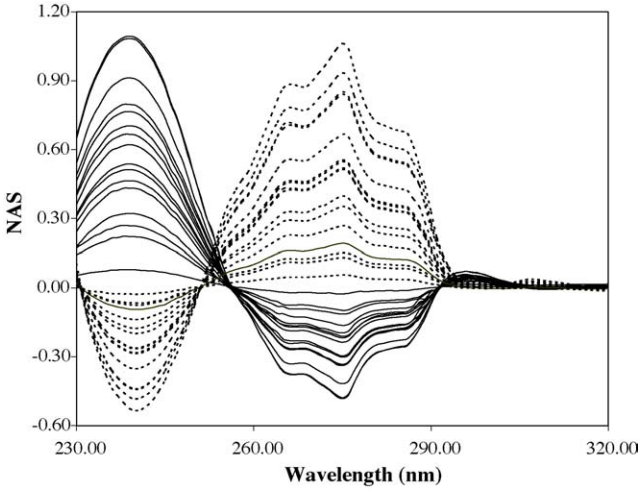

Fig. 2. The resolved NAS vectors of ANT (solid line) and NAP (dashed line) in the prediction set mixtures.

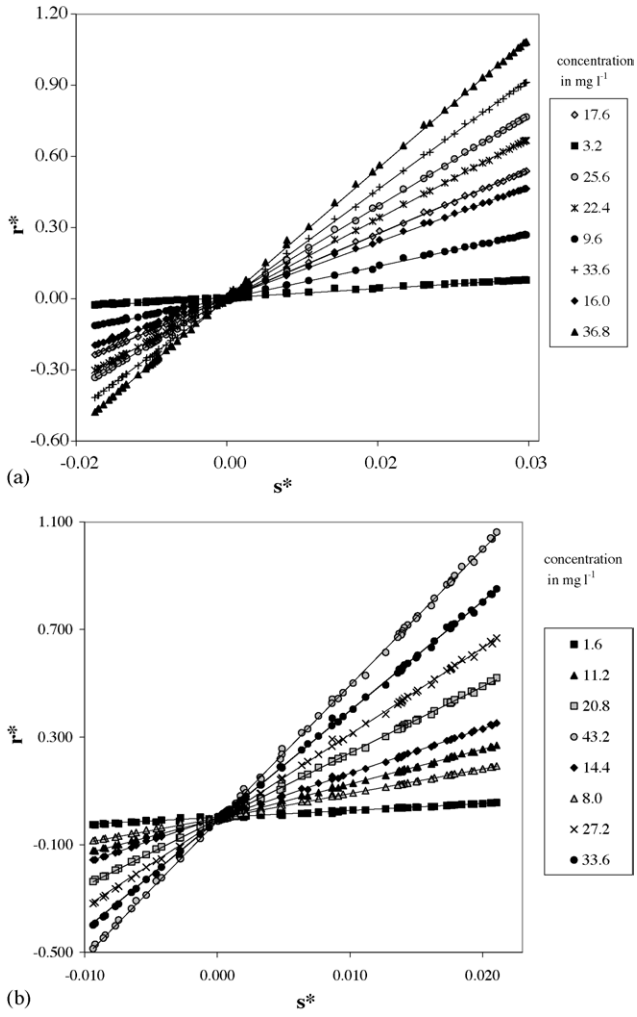

Fig. 3. Net analyte signal regression plots for (A) ANT and (B) NAP.

the calibration mixtures (net analyte signal regression plot, NASRP) are shown in Fig. 3. As it is obvious, the data of each plot shows a nice linear behavior. Referring to the concentration of drugs in Table 1, it is obvious that the slopes of these plots are directly related to the concentration of each drug in the mixtures. Plot of the norms of the NAS vectors against the respective concentrations of drugs in the prediction set produced a linear relationship with a high correlation coefficient.

4.3. NAS-based multivariate calibrations

In order to enhance the results obtained by the NAS calculations, the CLS, PCR and PLS multivariate calibration methods were used to model the relationship between the NAS vectors and the concentration of analytes. The results obtained by the three NAS-based models are represented in Table 5. In this table are included the number of factors (f), prediction residual errors sum of squares (PRESS), fitness function (η), percent relative errors for the calibration and prediction sets (REP_c and REP_p, respectively) and the square

Table 5
Statistical results of the NAS-based multivariate calibration models

Drug	Parameter	Regression method		
		NAS/CLS	NAS/PCR	NAS/PLS
ANT	F	–	3	2
	PRESS	0.652	0.580	0.573
	η	0.367	0.222	0.218
	REP_c	1.24	0.97	0.95
	REP_p	1.97	1.02	1.04
	r_c^2	0.9965	0.9991	0.9980
	r_p^2	0.9959	0.9964	0.9961
NAP	F	–	2	2
	PRESS	0.974	0.849	0.854
	η	0.284	0.153	0.168
	REP_c	2.41	1.44	1.42
	REP_p	2.83	1.84	1.85
	r_c^2	0.9971	0.9983	0.9989
	r_p^2	0.9957	0.9975	0.9982

Table 6
Statistical results of the conventional multivariate calibration models

Drug	Parameter	Regression method		
		CLS	PCR	PLS
ANT	F	–	3	3
	PRESS	8.27	0.658	0.628
	η	4.45	0.222	0.271
	REP_c	9.67	1.04	1.13
	REP_p	12.82	1.20	1.84
	r_c^2	0.9821	0.9951	0.9982
	r_p^2	0.9805	0.9952	0.9973
NAP	F	–	3	3
	PRESS	12.028	0.870	0.902
	η	5.13	0.169	0.175
	REP_c	12.36	1.58	1.67
	REP_p	14.82	2.03	2.01
	r_c^2	0.9901	0.9979	0.9984
	r_p^2	0.9894	0.9975	0.9980

of correlation coefficient for the calibration and prediction sets (r_c^2 and r_p^2, respectively). The optimum number of principal components of the factor analysis-based methods (i.e. PCR and PLS) was obtained by leave-one-out cross validation using PRESS of cross validation. The predicted values of the concentrations of drugs in the prediction set obtained by the NAS/PLS model together with the percent relative errors are shown in Table 2.

The results listed in Table 5 confirm the superiority of PLS and PCR methods over the CLS method. However, the difference between the results obtained by PLS and PCR models are not significant. If the information contents of the calculated NAS vectors are related to the concentration changes of the analyte of interest, CLS is expected to give results close to those of the PLS and PCR methods. However, as is obvious from Table 5, the results obtained by CLS are poor relative to those obtained by PCR and PLS. In addition, PCR used two and three factors for NAP and ANT, respectively, while PLS used two factors for both drugs. Thus, it can be concluded that, although NAS calculations give information that is unique for the analytes of interest and enhance the modeling power of the CLS, PLS and PCR methods, it contains some non-informative contents.

The wavelength regions used by different multivariate calibration methods are about the same. The respective spectral regions between 240–90 nm and 270–20 nm were used by different modeling methods for ANT and NAP. The major difference is that the spectral regions used by CLS are wider than those of PCR and PLS methods. Moreover, the optimum wavelength interval for CLS is larger than that for the two other methods. This may be due to the importance of co-linearity problem in the CLS modeling method.

4.4. Conventional multivariate calibration methods

In order to check the importance of NAS calculations as a preprocessing method in different multivariate calibration methods, the CLS, PCR and PLS procedures were also used to process the absorbance spectra of mixtures of antazoline and naphazoline at the optimum wavelength regions. The results are included in Table 6. By comparing the results given in Tables 5 and 6, it can be concluded that the NAS calculations are highly improved the performance of the CLS method. The performances of the PLS and PCR methods are also enhanced in the case of NAS preprocessing, although, the difference is not very significant.

4.5. Analyses of ANT and NAP in a commercial eye drop

To examine the applicability of the proposed NAS-based calibration methods in the analysis of the real samples, ANT and NAP in a commercial eye drop (Sina Darou Company, Tehran, Iran) were analyzed by the NAS/PLS regression method at the optimum wavelength regions. First, 2.0 ml of the sample was taken and subjected to the successive dilution with doubly distilled water to obtain concentrations in the region of calibration curves. The absorbance spectrum of the solution was then recorded and transferred to the NAS/PLS calibration models to predict the concentrations of the ANT and NAP. In the other trials, different amounts of ANT or NAP or a mixture of the drugs were spiked into the original eye drop and the procedure for the analysis of ANT and NAP were repeated. Each analysis was repeated seven times. The results are listed in Table 7. In this table, the resulting values of ANT and NAP as the mean of seven replicates, the relative standard deviation of seven replicate analyses and the mean recoveries are represented. The R.S.D.s for ANT and NAP are varied between 0.90–4.32% and 2.95%–6.60%, respectively, indicating good reproducibility of the proposed analysis method. Indeed, the recoveries (relative to the declared amounts of drugs in the eye drop and the spiked amounts) varied between 98.5–103.8% and 97.4–105.8% for ANT and NAP, respectively. This confirms the high accuracy of the proposed method for simultaneous analyses of ANT and NAP in a commercial formulation.

Table 7
Results for analyses of antazoline and naphazoline in a commercial sterile eye drop (anthazoline+naphazoline drop, Sina Darou Company Tehran, Iran) by the proposed method

No.	Drug	Added to the eye drop (mg ml^{-1})	Found (mg ml^{-1})[a]	R.S.D. (%)	Recovery (%)
1	ANT	0.00	4.93	4.15	98.5[b]
	NAP	0.00	0.51	2.95	101.0[b]
2	ANT	4.00	9.15	1.61	103.8
	NAP	0.00	0.48	6.60	96.3
3	ANT	0.00	5.15	4.32	102.9
	NAP	1.33	1.90	3.14	105.8
4	ANT	2.70	7.76	0.90	102.3
	NAP	2.70	3.12	3.15	97.4

[a] Average concentration over seven replicates.
[b] These recovery values were calculated relative to the declared amount of the eye drop sample: ANT, 5.00mg ml^{-1} and NAP, 0.50mg ml^{-1}.

5. Conclusion

Net analyte signal-based multivariate calibration methods (i.e. NAS/CLS, NAS/PCR and NAS/PLS) were employed for simultaneous determination of antazoline and naphazoline. An experimental design-artificial neural network approach was used for the selection of spectral regions that showed to get the best results. The wavelength regions selected for NAS/CLS model were wider than those obtained by the other methods. All models provided excellent results with relative prediction errors lower than 2%. The results confirmed the superiority of the NAS/PLS and NAS/PCR over the NAS/CLS. Comparison of the NAS pre-proceeded multivariate calibration models with the conventional ones revealed that the NAS calculations greatly highly affected the results of CLS model and improved the results obtained by this model. The NAS calculations also enhanced the factor analysis based methods PLS and PCR, although the differences were not very significant. Application of the proposed NAS/PLS method to an anthazoline + naphazoline commercial eye drop revealed that the method could analyze the drugs in real samples with quantitative recoveries.

References

[1] H. Martens, T. Naes, Multivariate Calibration, John Wiley, New York, 1992.
[2] P.D. Wentzell, L.V. Montoto, Chemometr. Intell. Lab. Syst. 65 (2003) 257.
[3] M. Barker, W. Rayens, J. Chemometrics 17 (2003) 166.
[4] S. Ren, L. Gao, Talanta 50 (2000) 1163.
[5] A. Lorber, Anal. Chem. 58 (1986) 1167.
[6] N.M. Faber, Anal. Chem. 71 (1999) 557.
[7] R. Boqué, J. Ferré, N.M. Faber, F.X. Rius, Anal. Chim. Acta 451 (2002) 313.
[8] N.M. Faber, J. Ferre, R. Boque, J.H. Kalivas, Trends Anal. Chem. (2003) 22.
[9] N.M. Faber, Anal. Chem. 70 (1998) 5108.
[10] A. Lorber, K. Faber, B.R. Kowalski, Anal. Chem. 69 (1997) 1620.
[11] K. Faber, A. Lorber, B.R. Kowalski, J. Chemometr. 11 (1997) 419.
[12] J.T. Olesberg, M.A. Arnold, B. Shih-Yao, B. Shih-Yao Hu, J.M. Wiencek, Anal. Chem. 72 (2000) 4985.
[13] L. Xu, I. Schechter, Anal. Chem. 69 (1997) 3722.
[14] H.C. Goicoechea, A.C. Olivieri, Chemometr. Intell. Lab. Syst. 56 (2001) 73.
[15] N.R. Marsili, M.S. Sobrero, H.C. Goicoechea, Anal. Bioanal. Chem. 376 (2003) 126.
[16] A.E. Mansilla, I.D. Meras, M.J.R. Gomez, A. Munoz de la Pena, F. Salinas, Talanta 58 (2002) 255.
[17] G.V. Truyols, J.R. Torres-Lapasio, M.C. Garcýa-Alvarez-Coque, J. Chromatogr. A 991 (2003) 47.
[18] A. Munz de la Pena, A.E. Mansilla, M.I.A.A. Valenzuela, H.C. Goicoechea, A.C. Olivieri, Anal. Chim. Acta 463 (2002) 75.
[19] J.H. Jiang, R.J. Berry, H.W. Siesler, Y. Ozaki, Anal. Chem. 74 (2002) 3555.
[20] C.H. Spiegelman, M.J. McShane, M.J. Goetz, M. Motamedi, Q.L. Yue, G.L. Cote, Anal. Chem. 70 (1998) 35.
[21] H.C. Goicoechea, A.C. Olivieri, J. Chem. Inf. Comput. Sci. 42 (2002) 1146.
[22] J.H. Kalivas, N. Roberts, J.M. Sutter, Anal. Chem. 61 (1989) 2024.
[23] R. Leardi, J. Chemometr. 8 (1994) 65.
[24] V. Centner, D.L. Massart, O.E. deNoord, S. Jong, B.M. Vandeginste, C. Sterna, Anal. Chem. 68 (1996) 3851.
[25] B.K. Alsberg, A.M. Woodward, M.K. Winson, J.J. Rowl, D.B. Kell, Anal. Chim. Acta 368 (1998) 29.
[26] M.C.U. Araujo, T.C.B. Saldanha, R.K.H. Galvao, T. Yoneyama, H.C. Chame, V. Visani, Chemometr. Intell. Lab. Syst. 57 (2001) 65.
[27] H.C. Goicoechea, A.C. Olivieri, Analyst 124 (1999) 725.
[28] F. Despagne, D.L. Massart, Analyst 123 (1998) 157R.
[29] S. Seculic, M.B. Seasholtz, Z. Wang, B.R. Kowalsky, S.E. Lee, B.R. Holt, Anal. Chem. 65 (1993) 835A.
[30] M. Shamsipur, B. Hemmateenejad, M. Akhond, Anal. Chim. Acta 461 (2002) 147.
[31] B. Hemmateenejad, M.A. Safarpour, F. Taghavi, J. Mol. Struct. (Theochem) 635 (2003) 183.
[32] B. Hemmateenejad, M. Akhond, R. Miri, M. Shamsipur, J. Chem. Inf. Comput. Sci. 43 (2003) 1328.
[33] L.F. Capitán-Vallvey, N. Navas, M. del Olmo, V. Consonni, R. Todeschini, Talanta 52 (2000) 1069.
[34] S.N. Demin, S.L. Morgan, Experimental Design: A Chemometrics Approach, Elsevier Health Sciences, 1993.
[35] A. Goodman-Hillman, T. Rall, A. Nier, P. Taylor, The Pharmacologycal Basis of Therapeutics, McGraw-Hill, New York, 1996.
[36] D. Berdeu, R. Puech, G. Ribes, M. Loubatieres-Mariani, G. Bertrand, Eur. J. Pharmacol. 324 (1997) 233.
[37] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193.