Comparison of partial least squares and artificial neural network chemometric techniques in determination of sulfamethoxazole and trimethoprim in pharmaceutical suspension by ATR–FTIR spectrometry

M. Khanmohammadi [a,*], N. Dallali [b], A. Bagheri Garmarudi [a,c], M. Zarnegar [b] and K. Ghasemi [a]

[a] *Department of Chemistry, Faculty of Science, Imam Khomeini International University, Qazvin, Iran*
[b] *Department of Chemistry, Faculty of Science, Zanjan University, Zanjan, Iran*
[c] *Department of Chemistry and Polymer Laboratory, Engineering Research Institute, Tehran, Iran*

Abstract. Partial Least Square (PLS) and Artificial Neural Network (ANN) techniques were compared during development of an analytical method for quantitative determination of sulfamethoxazole (SMX) and trimethoprim (TMP) in Co-Trimoxazole® suspension. The procedure was based on Attenuated Total Reflectance Fourier Transform Infrared (ATR–FTIR) spectrometry. The 800–2500 cm^{-1} spectral region was selected for quantitative analysis. R^2 and relative error of prediction (REP) in PLS technique were (0.989, 2.128) and (0.986, 1.381) for SMX and TMP, respectively. These statistical parameters were improved using the ANN models considering the complexity of the sample and the speediness and simplicity of the method. R^2 and RMSEC in modified method were (0.997, 1.064) and (0.997, 0.634) for SMX and TMP, respectively.

Keywords: Artificial neural network, quantitative analysis, sulfamethoxazole, trimethoprim, ATR–FTIR spectrometry, partial least squares

1. Introduction

Sulfamethoxazole, 4-amino-N-(5-methyl-3-isoxazolyl) benzene sulfonamide, (SMX) is a sulfonamide antibiotic drug, which acts as a competitive antagonist of p-amino benzoicacid (PABA) and thus as an inhibitor of folic acid synthesis, since PABA is an integral component of the structure of folic acid [3]. Trimethoprim, 2, 4-diamino-5-(3, 4, 5-trimethoxybenzyl) pyrimidine, (TMP) is also a well-known antibacterial agent, used as a potentiator in combination with several sulfonamides, e.g., SMX for treatment of a number of bacterial infections [6]. Binary mixtures of TMP and SMX are common

---

*Corresponding author: M. Khanmohammadi, P.O. Box 34149-1-6818, Qazvin, Iran. Tel./Fax: +98 281 3780040; E-mail: mrkhanmohammadi@gmail.com.

in commercial pharmaceutical preparations, also known as TMP-SMX and Co-trimoxazole® [11]. Their double directed action is effective against a range of Gram-positive and Gram-negative organisms, and less resistance is encountered than when either agent is used alone [12].

There are several routes for determination of these antibiotics. Researchers have considered the analysis of slightly soluble compounds in micellar media [7] fluorometric determination of sulfamethoxazole [5] and analysis of trimethoprim [1]. Also, analytical methods for the determination of TMP range from liquid chromatography [9] spectrophotometry [8], potentiometry [1] capillary zone electrophoresis and capillary electrophoresis with amperometric detection at carbon electrodes [15] to NMR [13] and electro analysis [2]. Methods used to determine TMP in pharmaceutical preparations include also HPLC [14]. Focusing on combination of SMX and TMP, some analytical methods have been introduced for determination of these two compounds such as applying a high-performance liquid chromatographic method which determines TMP, SMX, its metabolite and a series of structurally related sulfonamides. The first derivative and classical least squares methods have been applied for comparative purposes to analyze UV-spectra of the methanolic solutions of SMX and TMP components in synthetic binary mixtures. The method has been based on a pure standard full-spectra treatment ranged from 350 to 200 nm at 2 nm intervals to compute the concentration of unknowns while the former techniques were based on measuring the absorbance at zero-crossing wavelengths, 288 and 240 nm for the two drugs, respectively [10]. Most of the introduced methods for this aim are time consuming, not reliable enough or requiring high amounts of chemicals. The main aim of this study was to develop a FTIR spectroscopic procedure for fast, accurate and direct determination of SMX and TMP in their binary mixture suspension, which could be used in the quality control of their pharmaceutical products. Also, both the quality and the quantity of the analytes are possible to be analyzed by using ATR–FTIR spectroscopy. Additional advantages are that the sample treatment is significantly reduced or avoided, and the time of analysis is also significantly reduced, while the sample analysis frequency is improved. Nowadays, utilizing multivariate chemometrics techniques, e.g., PLS for process of spectral data has found extended applications. Even though PLS assumes a linear relationship between the measured sample parameters and the intensity of its absorption bands, several authors have postulated that small deviations from linearity are acceptable as they can readily be suppressed by including additional modeling factors. However, in the presence of substantial non-linearity, PLS tends to give large prediction errors and calls for more suitable models. Analogous considerations can be made when modeling complex and overlapped signals. Intrinsically nonlinear calibration techniques such as non-linear partial least squares (NPLS), locally weighted regression (LWR), alternating conditional expectations (ACE) and artificial neural networks (ANN) [4,16] are applicable in the latter cases. However, it is important to state that these methods are computationally more complex than linear methods, and they heavily depend on the amount and quality of available data. The object of this research was to develop a simple alternative to existing approaches for quantitative determination of SMX and TMP in pharmaceutical suspension samples.

2. Experimental section

2.1. Materials and apparatus

A Bomem, MB-100 FTIR spectrometer (Canada) equipped with DTGS-(D_3IB) mid-IR detector, SiC source, Sb_2S_3 coated KBr beam splitter and Spectratech (Warrington, UK) in-compartment contact with a sampler and a horizontal attenuated total reflector (with a 45° ZnSe trough plate) was used for spectral

analysis. GRAMS/32 software (Galactic Ind. Co.) was utilized to process the absorbance data. A GBC UV-Visible (Cintra 6) spectrophotometer, attached to a Pentium (IV) computer, with 10 mm glass cell was used for reference method.

All of the reference reagents were of analytical grade from British Pharmacopia®. Standard solutions of SMX (0.5–2.5 %w/v) and TMP (0.3–1.4 %w/v) were prepared in sodium hydroxide (0.1 N) and chloroform respectively. All chemicals and solvents were from MERCK®. Some ingredients of Co-Trimoxazole® oral suspension such as polysorbate 80, saccharin sodium, sucrose, flavoring essence, sorbitol, RC 591 and methyl paraben are all soluble in both sodium hydroxide (0.1 N) and chloroform and may interfere analytes. SMX, TMP and all other ingredients were supplied by Alborz Daru Co., Iran.

2.2. Preparation of real samples (*extraction procedure*)

A 25 ml volume of suspension and 30 ml sodium hydroxide (0.1 N) were transferred to a separating funnel and mixed well. The mixture was then extracted by 4 equal quantities (20 ml) of chloroform. The organic phase (80 ml) was separated and 4 times washed by 5 ml of NaOH (0.1 N) in order to perform the reverse extraction. Both aqueous and organic phases were then diluted up to 100 ml by NaOH (0.1 N) for determination of SMX and by chloroform for determination of TMP respectively.

2.3. UV-VIS spectrophotometry

2.3.1. Determination of SMX

The extraction step is the same as which discussed above, while the aqueous phase was diluted to 200 ml by distilled water (A) and also 0.5 ml of solution A was again diluted to 20 ml by distilled water (B). Then 2 ml of solution B, 0.5 ml of HCl (4 M) and 1 ml of $NaNO_2$ (0.1 %w/v) were all mixed in a volumetric flask. After 2 min, 1 ml of $(NH_4)_2SO_4$ (0.1 %w/v) was added and after 3 min, 1 ml of N-naphtyl ethylene diamine dihydrochloride (0.1 %w/v) was added. The solution was diluted up to 25 ml by distilled water after 10 min and its absorbance was measured at 538 nm (SMX is 2.5×10^{-5} g ml^{-1} concentrated in this solution). The standard solution was made of dissolving 0.25 g of reference SMX in 50 ml of NaOH (0.1 N) and dilution by distilled water up to 250 ml (A′). Then 0.5 ml of (A′) solution was diluted by distilled water up to 20 ml (B′) and again prepared by the foresaid procedure. The absorbance was again measured at 538 nm.

2.3.2. Determination of TMP

The chloroform solution obtained in SMX determination step was 4 times reversely extracted by 50 ml of acetic acid (1 M). Extracted solutions were 4 times eluted by chloroform, diluted by acetic acid (1 M) up to 1 liter and 1 ml of this solution was mixed with 10 ml of acetic acid (1 M) being diluted to 100 ml by distilled water. The absorbance of this solution was measured at 271 nm (TMP is 2.5×10^{-5} g ml^{-1} concentrated in this solution).

The standard solution was made of solving 50 mg of reference TMP in acetic acid (1 M) up to 250 ml. Then 10 ml of this solution was diluted to 100 ml by distilled water. The absorbance was again measured at 271 nm.

2.3.3. Determination of SMX and TMP in real samples

For analysis of real samples, TMP standard solutions were provided in concentration range of 0.30×10^{-4}–0.41×10^{-4} g ml^{-1}. The absorbance of these samples were determined at 271 nm. TMP real sample (0.20×10^{-4} g ml^{-1}) was prepared as detailed in Section 3.2. Using this solution, standard

addition method was employed to make 0.30×10^{-4}, 0.31×10^{-4} and 0.35×10^{-4} g ml^{-1} solutions which their absorbance was also determined at 271 nm being compared to the calibration model. The analysis of SMX real samples was performed in the same procedure while its standard solutions were prepared by acetic acid in concentration range of 0.50×10^{-4}–1.30×10^{-4} g ml^{-1}. The absorbance was determined at 538 nm. Real samples (0.80×10^{-4} and 1.00×10^{-4} g ml^{-1} concentrated) were prepared by the procedure detailed in Section 3.1.

3. Results and discussion

3.1. Solvent selection

TMP is well dissolved in chloroform, freely soluble in benzyl alcohol, a little soluble in sodium hydroxide solution and insoluble in diethyl ether and carbon tetrachloride. SMX is soluble in sodium hydroxide solution and acetone while it is a little soluble in chloroform and practically insoluble in water. It was necessary to use solvents which are able to solve at least 1.0×10^{-2} g ml^{-1} of analytes in order to achieve acceptable detection limit. According to our survey, there was no appropriate solvent which dissolves both SMX and TMP. Thus, chloroform and sodium hydroxide (0.1 N) were selected for SMX and TMP, respectively.

3.2. ATR–FTIR spectra of SMX and TMP

One of the advantages of Fourier transformation is the possibility of accumulating a large number of scans, which could provide better limit of detection for IR measurements. An increase in the number of accumulated scans does not affect the absorbance of the analytes but reduces the noise level of the spectra obtained and, therefore, the S/N ratio is improved. Using higher resolution means to have more data points, but in a longer time. Nominal resolution of 8 cm^{-1} and number of 12 scans seems to be eligible for the present study. The FTIR spectra were recorded in 800–2500 cm^{-1} spectral region.

One of the important parameters in ATR–FTIR spectrometry is the transparency of solvent in the region in which the sample is being studied. Comparing the ATR–FTIR spectra of chloroform and TMP in chloroform while air is the background showed an interference which causes some problems in determination of TMP. This problem also occurs while comparing the spectra of sodium hydroxide (0.1 N) and SMX in sodium hydroxide (0.1 N) with air background, and in both situations, spectra show that main bands are completely covered by solvents. Using solvent as the background, some intensive bands would appear which seem to be useful in determination of SMX and TMP (Figs 1 and 2).

3.3. Perturbing effect of other ingredients

As mentioned before, there are several ingredients in Co-Trimoxazole® which may interfere with aimed SMX or TMP. Therefore all of these components were solved in both chloroform and NaOH (0.1 N) and their spectra were recorded. On the other hand, SMX is a little soluble in chloroform as TMP is a little soluble in NaOH (0.1 N) thus it was needed to record their spectra (by the maximum possible concentration) to know if there is any interference or not. All of the ingredients were in a low level of response while being recorded by ATR–FTIR showing insignificant signals except flavoring essence and polysorbate 80 which are highly soluble in NaOH (0.1 N), showing intensive bands and cover the signals due to SMX (Fig. 3). Thus it was necessary to add these 2 components to the standard solutions of SMX, and then try to determine the SMX in accordance with the calibration model.
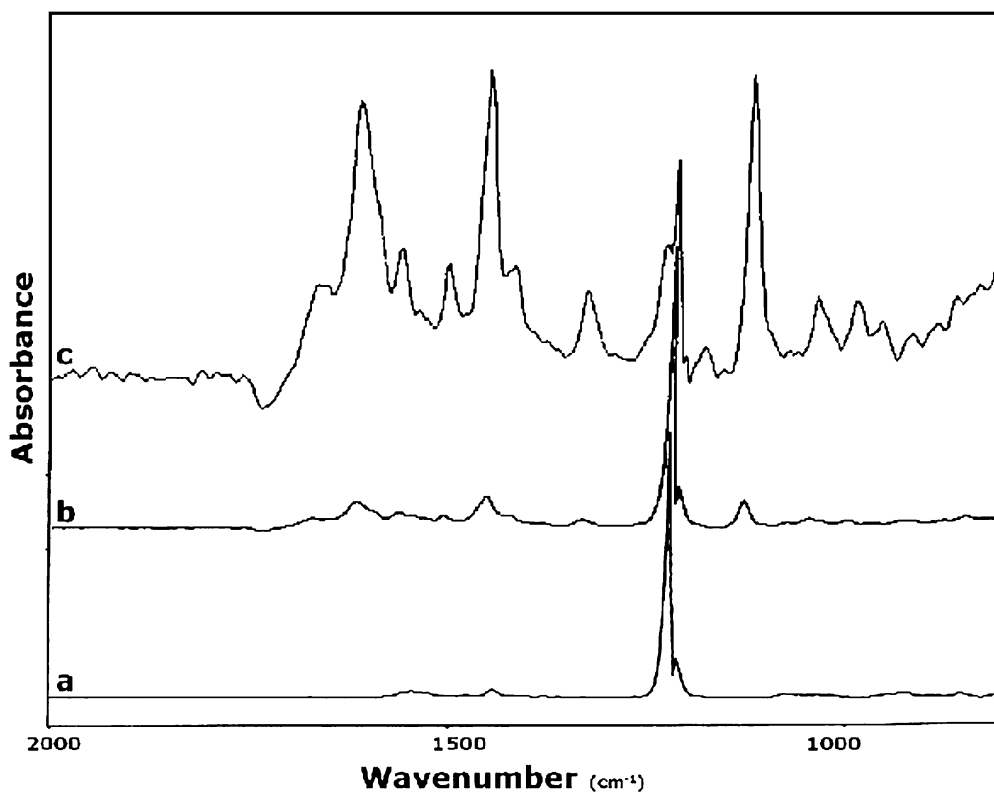
Fig. 1. ATR–FTIR spectra of (a) chloroform with air background, (b) TMP (0.009 g ml^{-1} in chloroform) with air background and (c) TMP (0.009 g ml^{-1} in chloroform) with solvent background.

3.4. Standard solution preparation

It was recognized that the 1100–1160, 1440–1480 and 1580–1642 cm^{-1} spectral regions are reasonable for TMP determination while 1073–1111, 1390–1430 and 1443–1480 cm^{-1} regions are noticeable for determination of SMX because they show the most relation between concentration and absorbance. Standard solutions of TMP in chloroform where prepared in concentration range of 0.003–0.014 g ml^{-1} and standard solutions of SMX in NaOH (0.1 N) were prepared in concentration range of 0.005–0.025 g ml^{-1}. ATR–FTIR spectra of all standard solutions were recorded in 800–2500 cm^{-1} spectral region.

3.5. Applying PLS in determination of SMX and TMP

PLS is often presented as the major regression technique for multivariate data. PLS has seen an unparalleled application success, both in chemometrics and other fields. Amongst other features, the PLS approach gives superior interpretation possibilities, which can best be explained and illustrated by examples. In spectroscopy we usually expect linear additivity, and this is especially important for chemical instrumental data. PLS is always an important tool when there is partial knowledge of the data. PLS can be very robust provided that future samples contain similar features to the original data, but the predictions are essentially statistical. An important feature of PLS is that it takes into account errors in both
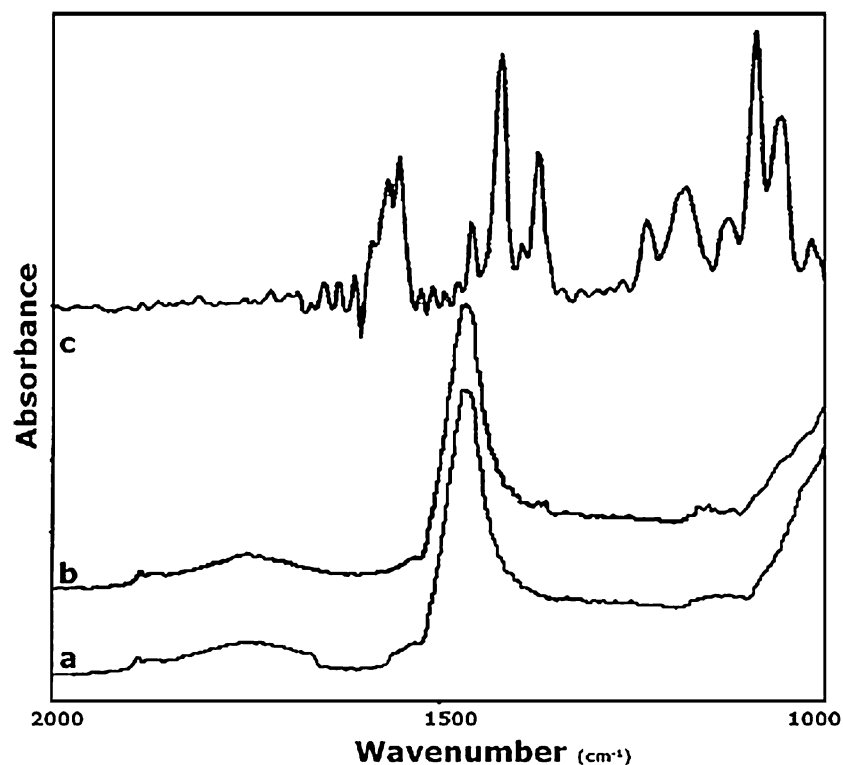
Fig. 2. ATR–FTIR spectra of (a) NaOH (0.1 N) with air background, (b) SMX (0.020 g ml^{-1} in NaOH) with air background and (c) SMX (0.020 g ml^{-1} in NaOH) with solvent background.

the concentration estimates and the spectra. In the other words, PLS is capable of considering the errors in the different steps of the analysis, e.g., solution preparation, spectrum recording and concentration prediction. The most widespread approach is often called PLS1. In the application of the PLS algorithm, it is generally known that the spectral range and the number of PLS factors are critical parameters. The spectral range determines the location and quality of spectral information, and the number of PLS factors should be optimally selected to avoid an overfitting. Each successive PLS component approximates both the concentration and spectral data better. A key issue in calibration is to determine how well the data have been modeled. Cross-validation is used if and when there are not enough objects for a separate test set or when it is not possible to pick out a representative test set. Cross validation makes the calibration set as large and representative as possible. Comparing the predicted and measured values gives us an expression of the modeling error which is calculated for each object.

In order to perform the calibration method, all standard solutions of both SMX and TMP were provided by the appropriate solvent, in presence of all interfering components and ATR–FTIR spectra were recorded with solvent background. The suitable region was optimized among the selected regions which had showed the best relation between concentration and absorbance. PLS1 was applied in different selected spectral regions, statistical parameters such as REP (relative error of prediction), R^2, RMSEC (root mean square error of calibration) and "optimum number of factors" were investigated to find the best spectral region for calibration. It is mentionable that, leave one out cross validation was applied to select the optimum number of factors. Each time, one of the standard solutions was left out, the model
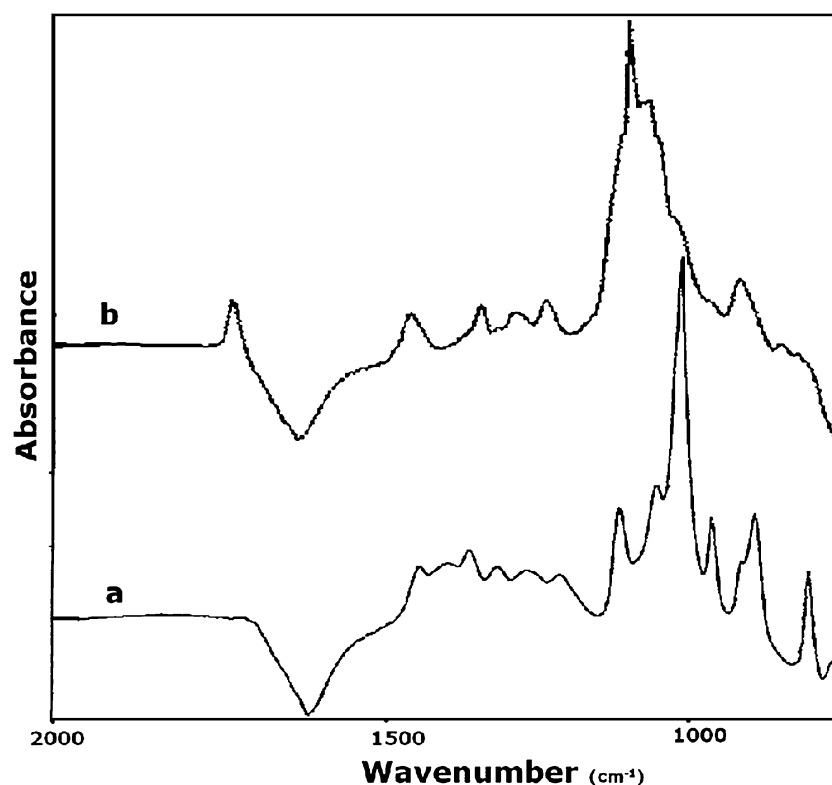
Fig. 3. FTIR–ATR spectra of interfering components: (a) polysorbate 80 and (b) flavoring essence.

Table 1

Comparing the results obtained in analyzing standard solution of SMX and TMP

	PLS		ANN	
	SMX	TMP	SMX	TMP
Spectral region (cm^{-1})	1426.56–1396.21	1641.55–1580.49		
Number of factors	5	3		
RMSEC	0.018	0.023	0.009	0.011
REP	2.128	1.381	1.064	0.634
R^2	0.989	0.986	0.997	0.997

was formed for the remaining samples and then the left sample was predicted by model. This process was repeated for all of standard samples and RMSEC was calculated in all models (Table 1).

The number of factors was varied from 1 to 6 and the same procedure was accomplished for each "number of factor", calculating RMSEC according to:

$$\text{RMSEC} = \frac{\sqrt{\sum_{i=1}^{n}(y_{ki} - y_{pi})^2}}{n - f - 1},\tag{1}$$

where y_{ki} is actual concentration, y_{pi} is predicted concentration, f is optimum number of factors and

n is the number of standard solutions. Also the prediction error sum of squares (PRESS) was calculated each time a new factor was added. One reasonable choice for the optimum number of factor would be that number which yields the minimum PRESS.

Thereby, considering RMSEC, REP and R^2 values best spectral region and optimum number of factors for TMP were 1580–1641 cm^{-1} and 3, respectively, while they were 1396–1426 cm^{-1} and 5 for SMX. Finally, all of samples were quantitatively determined by the calibration model.

3.6. Artificial neural network

ANN is typically organized in layers where these layers are made up of a number of interconnected nodes which contain an activation function. Input vectors are presented to the network via the input layer which communicates to one or more "hidden layers" where the actual processing is done via a system of weighted "connections". Most ANNs contain some form of "learning rule" which modifies the weights of the connections according to input patterns that it is presented with. There are many different kinds of learning rules used by neural networks, in this work, Back-Propagation Neural Networks (BP-ANN) was applied. In BP-ANN, "learning" is a supervised process that occurs with each cycle of "epoch" (i.e., each time the network is presented with a new input pattern) through a forward activation flow of inputs and the backwards error propagation of weight adjustment. There are many variations of the BP algorithm. The simplest implementation of BP learning updates the network weights and biases in the direction in which performance function decreases most rapidly, the negative of the gradient. An iteration of this algorithm can be written as

$$X_{k+1} = X_k - \delta_k g_k, \tag{2}$$

where X_k is a vector of current weights and biases, g_k is the current gradient, and δ_k is the learning rate. In this work, gradient descent with momentum is applied and the performance function was the Mean of the Sum of Squares Error (MSSE), the average squared error between the network outputs and the actual output. For the basic gradient descent algorithm, the weights and biases are moved in the direction of the negative gradient of the performance function. Gradient descent with momentum often provides faster convergence because momentum allows a network to respond not only to the local gradient but also to recent trends in the error surface. Momentum can also help the network to overcome a shallow local minimum in the error surface and settle down at or near the global minimum. Momentum can be added to back-propagation learning by making weight changes equal to the sum of a fraction of the last weight change and the new change suggested by the back-propagation rule. The magnitude of the momentum constant is allowed to be any number between 0 and 1. When the momentum constant is 0, the weight change is based solely on the gradient. When the momentum constant is 1, the new weight change is set to equal the last weight change and the gradient is simply ignored. The performance of the network was also tested by reducing the dimension of the input vectors before the training process. The selected model for ANN was 2–1 input–output pair model by Tan-sigmoid (called Tansig) as a transfer function in each layer. On the other hand selected factors for ANN model were 2–1 input–output pair containing Tansig as a transfer function for each model. This selected parameter was applied for both species in this work. As a result of using these models for quantitative analysis, great modification was demonstrated in the result (Table 1). Results of calibration step for PLS and ANN has been detailed in Table 2.

Table 2

Prediction of SMX and TMP concentration (g per 100 ml) for calibration samples by PLS and ANN algorithm

TMP			SMX		
Actual	PLS	ANN	Actual	PLS	ANN
0.30	0.29	0.30	0.50	0.48	0.49
0.40	0.43	0.41	1.00	0.99	0.99
0.50	0.48	0.49	1.25	1.26	1.25
0.60	0.61	0.61	1.50	1.52	1.49
0.70	0.70	0.71	1.75	1.76	1.75
0.80	0.77	0.79	2.25	2.26	2.26
0.90	0.89	0.89	2.50	2.49	2.50
1.00	1.00	1.01	2.75	2.70	2.73
1.10	1.09	1.10			
1.20	1.20	1.20			
1.30	1.29	1.29			
1.400	1.43	1.41			

Table 3

Comparing the results of PLS–FTIR and ANN–FTIR algorithms with UV-VIS reference method in determination of SMX and TMP[a]

	Actual[b]	UV-Vis	FTIR–PLS	FTIR–ANN
SMX	0.300	0.300 ± 0.016	0.296 ± 0.011	0.299 ± 0.001
	0.310	0.330 ± 0.016	0.333 ± 0.011	0.323 ± 0.001
	0.350	0.340 ± 0.016	0.344 ± 0.011	0.349 ± 0.001
TMP	0.800	0.850 ± 0.009	0.843 ± 0.009	0.825 ± 0.005
	0.800	0.740 ± 0.009	0.746 ± 0.009	0.796 ± 0.005
	1.000	1.140 ± 0.009	1.120 ± 0.009	1.020 ± 0.005

Notes: [a]SD is for $n = 6$ for both methods; [b]concentrations are in g per 100 ml.

3.7. Comparing the determination of SMX and TMP in real samples by PLS1 and ANN

SMX (8×10^{-3} and 10×10^{-3} g ml^{-1}) and TMP (0.20×10^{-4} g ml^{-1}) solutions were prepared of Alborz Daru Co. suspension product, being predicted by PLS and ANN model. Using this solution, standard addition method was employed to provide 3.0×10^{-3}, 3.1×10^{-3} and 3.5×10^{-3} g ml^{-1} solutions. Finally, the results were compared with the result obtained from the UV-Vis method (Table 3). Generally an acceptable concordance was found between the results and it can be observed that the ATR–FTIR method could provide an accurate determination method for TMP and SMX in Co-Trimoxazole® oral suspension. Also prediction by ANN shows great improvement in results compared to those of PLS.

4. Conclusion

While the components present in a sample are interfering, the determination process is not desirable performed by univariate calibration. Thus it is necessary to apply a multivariate calibration technique. ATR–FTIR spectrometry in combination with chemometrics, enables to develop a rapid and precise method for determination of SMX and TMP in Co-Trimoxazole® oral suspension. The results obtained in this research illustrated that this procedure is successfully applicable in quantitative analysis of SMX

and TMP in pharmaceutical industry scale. Finally, it was concluded that ANN provides better statistical results in comparison with PLS.

References

[1] M. Aboudan, A.F. Shoukry and Y.M. Issa, *Anal. Lett.* **29** (1996), 19–28.
[2] J.J. Berzas, J. Rodríguez and G. Castañeda, *Anal. Chim. Acta* **349** (1997), 303–311.
[3] R.F. Bevill, Sulfonamides, in: *Veterinary Pharmacology and Therapeutics*, N.H. Booth and L. McDonald, eds, Iowa Univ. Press, Iowa City, IA, USA, 1998, pp. 785–795.
[4] R.M. Carvalho, C. Mello and L.T. Kubota, *Anal. Chim. Acta* **420** (2000), 109–121.
[5] A.M. Fransisco-Salinas, I. Duran-Meras and M.D. Moreno, *Anal. Lett.* **27** (1994), 1893–1906.
[6] M. Fresta, P.M. Furneri, E. Mezzasalma, V.M. Nicolosi and G. Pugeisi, *Antimicrob. Agents Chemother.* **40** (1996), 2865–2873.
[7] P. Issopoulos, *Fresenius' J. Anal. Chem.* **354** (1996), 378–380.
[8] M.E.M. Hassouna, *Anal. Lett.* **30** (1997), 2341–2352.
[9] C.T. Hung and D.G. Perrier, *J. Liquid Chromatogr.* **8** (1985), 521–536.
[10] L.M. Lin, *Acta Pharmacol. Sin.* **26** (1991), 858–864.
[11] New Mexico AIDS InfoNet, Trimethoprim/sulfamethoxazole, Fact Sheet No. 535, 2001.
[12] R.M.E. Richards, R.B. Taylor and Z.Y. Zhu, *J. Pharm. Pharmacol.* **48** (1996), 981–984.
[13] J.W. Turczan, *Anal. Chim. Acta* **92** (1977), 123–128.
[14] K. van der Steuijt and P.J. Sonneveld, *Chromatogr.* **422** (1987), 328–333.
[15] A. Wang, F. Gong, H. Li and Y. Fang, *Anal. Chim. Acta* **386** (1999), 265.
[16] S. Wold, *Chemom. Intell. Lab. Syst.* **14** (1992), 71–84.