PAPER

Check for updates

Resolution and quantification challenge of modern chemometric models in the determination of anti-migraine tablets containing ergotamine, caffeine, acetaminophen, and metoclopramide†

Mahmoud M. Elkhoudary, [ID] *[ab] Randa A. Abdel Salam[b] and Ghada M. Hadad[b]

This study is a comparison between the performance of five multivariate models in the determination of the unique mixture of ergotamine (ERG), metoclopramide (MET), caffeine (CAF), and paracetamol (PAR) in laboratory-prepared mixtures and in pharmaceutical formulations. Two supervised learning machine methods—artificial neural networks (PC-ANN) preceded by principle component analysis and support vector regression (SVR)—were compared with a spectral residual augmented classical least squares (SRACLS) method, multicurve resolution alternating least squares (MCR-ALS) method, and principle component based method; partial least squares (PLS). The results showed the superiority of linear learning machine methods in handling extremely noisy and complex spectral data, especially during the determination of the challenging mixture under study. ERG (the component with a close to undetectable concentration and with the lowest ratio in the studied dosage form) was only determined using three chemometric models, with root mean squared error of prediction (RMSEP) for the proposed models of 0.0879, 0.0694, and 0.0250 for PLS, SVR and PC-ANN, respectively. In addition, the results suggest that ANN is the method of choice for the determination of mixtures with extreme conditions; for example, components with a very low contribution in the overall spectra, components with narrow informative range, and extremely nonlinear spectral data.

1. Introduction

Headache disorders are some of the most frequently reported symptoms and have been associated with impaired quality of life, increased incidence of depression, musculoskeletal pain, and disability.[1] Epidemiologic studies have found that 57% of males and 76% of females have one or more headache attacks per month.[2] A migraine headache can cause intense throbbing or a pulsing sensation in one area of the head and is commonly accompanied by nausea, vomiting, and extreme sensitivity to light and sound.[3] Some migraines are preceded or accompanied by sensory warning symptoms (aura), such as flashes of light, blind spots, or tingling in the arm or leg.[3] The prevalence of migraine headaches in the Arab world is similar to the range reported worldwide.[4] The prevalence of migraine headache in six Arab countries (Egypt, Saudi Arabia, Tunisia, Oman, Jordan, and Qatar) ranges from 3% to 12%.[4]

Medications can help in reducing the frequency and severity of migraines. Pharmaceutical companies offer different kinds of analgesic and nonsteroidal anti-inflammatory drug mixtures (NSAIDs), with or without ergot alkaloids and caffeine (CAF) for acute headache therapy. Antiemetics such as metoclopramide (MET) and domperidone are often included in compound anti-migraine preparations. In addition, attacks not responding to simple analgesics or NSAIDs may be treated with specific anti-migraine drugs such as selective serotonin (5-HT1) agonists (*e.g.*, sumatriptan) or ergot derivatives (*e.g.*, ergotamine [ERG] and dihydroergotamine). As ERG itself can also exacerbate nausea and vomiting, concurrent administration of MET or domperidone, may be required.[2] Among the most commonly used antimigraine polypill combinations are those containing ERG, CAF, paracetamol (PAR), and MET. The structures of these drugs are provided in Fig. 1.

ERG [(5/S)-12-hydroxy-2-methyl-5-benzylergotaman-3,6,18-trione tartrate][5] is an alkaloid derived from ergot. It has strong vasoconstrictor effects, and may have a partial agonist action at serotonin (5-HT) receptors. ERG is used in the treatment of migraines and cluster headaches.[6]

MET [4-amino-5-chloro-*N*-[2-(diethylamino)ethyl]-2-methoxy benzamide][5] is used for its prokinetic and antiemetic properties. MET is used to combat nausea and vomiting associated with various migraine-induced gastrointestinal disorders.[6]

[a]Pharmaceutical Chemistry Department, Faculty of Pharmacy, University of Tabuk, Tabuk, Kingdom of Saudi Arabia. E-mail: melkhoudary@ut.edu.sa; Tel: +966 593503202

[b]Pharmaceutical Analytical Chemistry Department, Faculty of Pharmacy, Suez Canal University, Ismailia, Egypt. Fax: +20 643561877; Tel: +20 1110777922

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7ra00257b
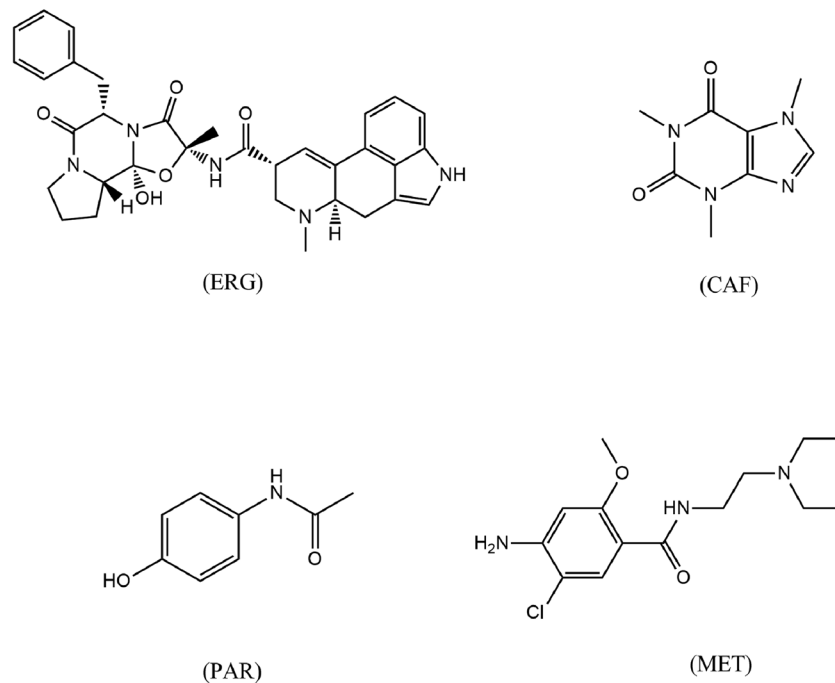
Fig. 1 Chemical structures of ERG, MET, CAF and PAR.

PAR [N-(4-hydroxyphenyl) acetamide] has analgesic and antipyretic activities with weak anti-inflammatory properties.[5]

CAF [1,3,7-trimethyl-3,7-dihydro-1H-purine-2,6-dione][5] is a methyl xanthine that acts by inhibiting the enzyme phosphodiesterase with an antagonistic effect at adenosine receptors. It acts as a stimulant of the central nervous system (CNS) and is sometimes given with ERG in formulations for the treatment of migraine.[6]

Several analytical procedures have been described for the simultaneous determination of mixtures containing PAR and CAF, including spectrophotometry,[7] chemometrics,[8] high-performance thin-layer chromatography (HPTLC),[9] high-performance liquid chromatography (HPLC),[10–12] and gas chromatography.[13] The United States Pharmacopeia (USP)[14] describes the determination of ERG and CAF tablets by HPLC with fluorimetric detection for ERG at $\lambda_{em} = 435$ nm and $\lambda_{ex} = 325$ nm, and ultraviolet (UV) detection for CAF at 254 nm. Similarly, various analytical procedures have been described for the analysis of mixtures containing ERG and CAF. These methods include spectrophotometry,[15] thin-layer chromatography (TLC),[16] and HPTLC.[17] Moreover, PAR and MET have been determined in combination using spectrophotometry.[18]

A detailed literature survey revealed that two methods have been used for the determination of ERG, MET, CAF, and PAR simultaneously using capillary electrophoresis (CE) in their dosage form[19] and HPLC with UV and fluorescence detection.[20] In contrast, the chemometric models proposed in this publication are based on the simple and widely available UV spectrophotometer. Furthermore, the introduced models offer a simple, accurate, time saving, and ecofriendly alternative to the commonly used chromatographic techniques, which suggests that these methods may be valid candidates for the

quality control analysis of these compounds in combined pharmaceutical tablets. In the current work, we developed five multivariate methods that can detect and quantitate MET, CAF, and PAR in raw material or pharmaceutical formulations. Additionally, one model was able to detect and quantitate ERG as well as the other three components.

Chromatographic methods are well known for their superior abilities in the separation of complex matrices and quantitation of their components at trace levels. However, these methods require unique expensive equipment and costly chemicals that are used in large quantities and can harm the environment. Attempts to develop methods that use fast, simple, and less expensive instruments are attracting considerable attention currently. The recent implementation of chemometric techniques during the manipulation of raw spectroscopic data allows mining of more useful information, and suggests the application of these methods in more complex situations.

This work aimed primarily to challenge the commonly used chemometric models[21,22] in dealing with UV data of mixtures that cannot be detected and quantified with univariate and most multivariate spectrophotometric methods. The proposed models were partial least squares (PLS), multivariate curve resolution with alternating least squares (MCR-ALS), spectral residual augmented classical least squares (SRACLS), support vector regression (SVR), and artificial neural networks preceded by principle component analysis of raw data (PC-ANN). The proposed mixture for this challenge was ERG, MET, CAF, and PAR in a ratio of 1 : 5 : 50 : 325, respectively. This mixture has the unique property of having considerable variability in the ratios of its components in its commercial dosage form, creating a constraint that nearly all the compounds fall outside the linearity range while others are in ratios that are nearly

undetectable. The ability of multivariate models to deal with this kind of problem is variable based on how each model manipulates the raw data to find reasonable predictions.

The number-one priority for each model was to accurately and selectively determine ERG (with the lowest ratio of 1), which has relatively low ratio in comparison with PAR (with the highest ratio of 325) in the presence of all other compounds. Furthermore, the work offers a comparison between five multivariate models with different mathematical backgrounds in terms of qualitative and quantitative abilities in order to qualify the multivariate model that can best deal with mixtures that have this kind of nonlinearity and complexity.

2. Experimental

2.1. Instruments

The measurements were performed using a UV-visible spectrophotometer (Agilent Technologies Cary 5000 UV-VIS-NIR Spectrophotometer), equipped with a quartz 1 cm cuvette. Spectra were recorded in the range 200.0–400.0 nm at 1.0 nm intervals. The spectra were collected using Cary Win UV software (Agilent, USA).

2.2. Material and reagents

All experiments were performed using pharmaceutical-grade authentic standards of ERG (Egyptian International Pharmaceutical Industries Company, 10th of Ramadan City, Egypt), MET (Tabuk Pharmaceuticals, Saudi Arabia), CAF, and PAR (ACROS, Belgium), and all standards were certified to have a purity of 99.3–99.7% (w/w) on a dried basis.

The studied dosage form was Metograine® tablets from EVA Pharma, Egypt (batch no. 009264), which was labeled to contain 1 mg ERG, 50 mg CAF, 325 mg PAR, and 5 mg MET per tablet and was purchased from the local market.

Analytical grade methanol was purchased from Fisher Scientific Ltd.

2.3. Stock and standard solutions

Stock solutions for ERG, MET, CAF, and PAR were each prepared by accurately weighing pure powder (25 mg) of each, which was then transferred into 50 mL volumetric flasks and diluted to the mark with methanol. Dissolution was achieved with the help of an ultrasonic bath for approximately 15 min. The calibration-standard working solutions were prepared by dilution of the stock standard solutions with methanol to reach the proposed linearity concentration ranges of 9–50, 2.5–15, 1.8–20, and 5–70 μg mL^{-1} for ERG, MET, CAF, and PAR, respectively.

The preparations were carried out based on the intensity of the absorbance of each component spectra and the overall contribution of each component in the final mixture to allow adjustment of the final proportions to 10 mL. All stock and working standard solutions were kept away from light in order to avoid photodegradation and stored refrigerated at 4 °C; the solutions showed stability for 3 days under these conditions.

2.4. Experimental design

2.4.1. Calibration and test sets. To construct calibration set, a multilevel multifactor design was performed.[23] A calibration set of 25 samples was prepared for calibration. A four-factor calibration design with five levels was used in concentrations of 0.4–0.6, 2–3, 20–30, and 130–195 μg mL^{-1} for ERG, MET, CAF, and PAR, respectively. The studied dosage form contained ERG, MET, CAF, and PAR in a ratio of 0.5 : 2.5 : 25 : 165, respectively. The concentration levels were coded −2, −1, 0, +1, and +2, where the level coded 0 is the central level for the design. Each concentration level was based on the component level in the studied dosage form. Mean centering proved to be the preprocessing method of choice based on optimum results for the five developed models.

The 2D plot for scores of the first two PCs of the mean centered concentration matrix was obtained to confirm that training samples cover the mixture space fairly and ensure orthogonality, symmetry and rotatability,[23] as indicated in Fig. 2. The developed models were tested for validity and predictive ability, with a set of test mixtures of 12 samples that falls inside the concentration space of the design in Fig. 2.

Table 1 represents the concentration design matrix for both calibration and test sets. The range used was 280–380 nm, and 101 data points were used in data modelling in the case of PLS, MCR-ALS, SRACLS, and SVR models. In contrast, for ANN, the first four principle component scores for ERG, MET, CAF, and PAR were used instead of spectral data in the data modelling.

2.4.2. Analysis of Metograine® tablets. Twenty Metograine® tablets were weighed and ground separately to form a homogenous powder. An accurately weighed portion equivalent to 1 mg ERG, 5 mg MET, 50 mg CAF, and 325 mg PAR, according to the labeled claims, was separately dissolved in methanol via ultrasonication for 15 min. The solution was cooled then transferred into a 100 mL volumetric flask and then diluted to volume
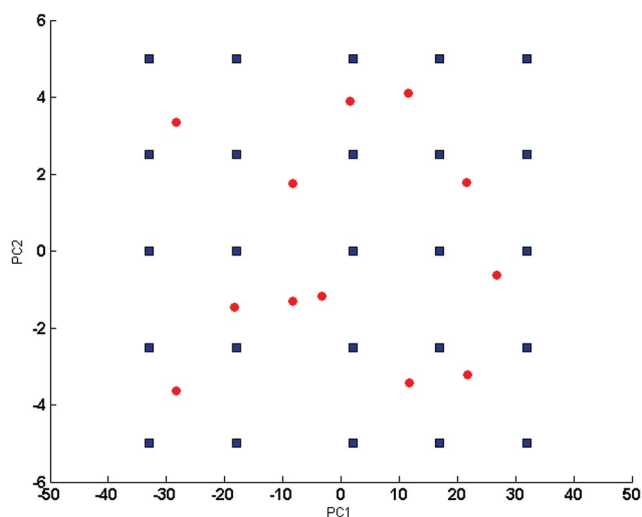

Fig. 2 2D scores plot for the mean centered 25 training samples (■) and 12 test samples (●) concentration matrix of the 5 level 4 components experimental design.

Table 1 Five-level 4-factor experimental design of 25 calibration set mixtures together with 16 test set mixtures shown as concentrations of the mixture components in µg mL^{-1a}

Mixture no.	Calibration set				Test set			
	ERG	MET	CAF	PAR	ERG	MET	CAF	PAR
1	0.50	2.50	25.00	165.00	0.45	3.00	25.00	160.00
2	0.50	2.00	20.00	195.00	0.45	2.25	25.00	165.00
3	0.40	2.00	30.00	145.00	0.50	2.50	20.00	170.00
4	0.40	3.00	22.50	195.00	0.50	3.00	27.50	180.00
5	0.60	2.25	30.00	165.00	0.55	2.10	25.00	195.00
6	0.45	3.00	25.00	145.00	0.40	3.00	25.00	150.00
7	0.60	2.50	22.50	145.00	0.60	2.50	22.50	190.00
8	0.50	2.25	22.50	180.00	0.40	2.75	20.00	140.00
9	0.45	2.25	27.50	195.00	0.55	2.50	27.50	190.00
10	0.45	2.75	30.00	180.00	0.40	2.00	22.00	160.00
11	0.55	3.00	27.50	165.00	0.60	2.75	27.00	140.00
12	0.60	2.75	25.00	195.00	0.50	2.25	20.00	180.00
13	0.55	2.50	30.00	195.00				
14	0.50	3.00	30.00	130.00				
15	0.60	3.00	20.00	180.00				
16	0.60	2.00	27.50	130.00				
17	0.40	2.75	20.00	165.00				
18	0.55	2.00	25.00	180.00				
19	0.40	2.50	27.50	180.00				
20	0.50	2.75	27.50	145.00				
21	0.55	2.75	22.50	130.00				
22	0.55	2.25	20.00	145.00				
23	0.45	2.00	22.50	165.00				
24	0.40	2.25	25.00	130.00				
25	0.45	2.50	20.00	130.00				

^{a} CAF, caffeine; ERG, ergotamine; MET, metoclopramide; PAR, paracetamol.

with methanol. The solutions were filtered through a 0.45 µm disposable filter paper before use. Finally, an aliquot of the prepared solution was diluted to 10 mL and then measured. The average of three corresponding spectra was recorded. The experiment was repeated six times and the obtained spectra were analyzed by the chemometric model that showed best performance.

2.4.3. Software. Data analysis was carried out using Matlab® 7.10.0.499 (R2010a), Eigenvector PLS_Toolbox 7.3. The MCR-ALS GUI was downloaded from the website http://www.mcrals.info/, the Neural Network Toolbox implemented in Matlab®, the codes for SRACLS were written in lab and codes for SVR algorithm were downloaded from the website http://onlinesvr.altervista.org/.

3. Chemometric methods

3.1. PLS

Multiple linear regression (MLR), principal component regression (PCR), and partial least squares regression (PLS) are examples of well-established and widely used models for first-order data calibration.

PLS depends on the decomposition of the predictor matrix X and the concentration vector c simultaneously using a given number of latent variables (LVs).[24] Cross validation (CV)[25] is applied to predict the optimum number of PLS latent variables. For each number of latent variables, the performance of the model was evaluated based on the root mean square error of CV (RMSECV) (Table 2, eqn (1) and (2)).

3.2. SRACLS

SRACLS is a modification to the classical least-squares (CLS) model that was devised to overcome the disadvantages of CLS and give a comparable performance to PLS[26] (Table 2, eqn (3)).

3.3. MCR-ALS

MCR techniques are applied to decompose bilinear signals to their pure contributions. They have the advantage over all first-order calibration methods that they have the possibility of recovering individual pure spectra of the components contributing in the analyzed mixtures, the property that is limited in all other methodologies and compensated for by efficient mathematical algorithms to overcome non-selectivity and overlapping in the measured signal.[27] In addition, these techniques are considered a multi-wavelength representation of the Beer–Lambert law (Table 2, eqn (4)–(6)).

3.4. Supervised machine learning methods

Machine learning methods, such as SVR and ANN[28,29] are more often implemented in analysis processes these days because of improvements in the performance of personal computers and the considerable capabilities of these models in dealing with complicated cases where the data contain noise, sources of nonlinearity as weighing errors, uncalibrated glassware and concentrations out of the Beer–Lambert calibration range.[26,30–32]

3.4.1. Artificial neural networks (ANN). An ANN is composed of artificial neurons connected by weights. The network parameters are modified until the network output matches the target. Many input/target pairs are used to ensure good training of the network and reliable results.[33] A feed-forward back propagation ANN model was the learning method adopted in this study. Back-propagation ANN has many benefits in signal processing, data reduction, and prediction of spectra.[34,35]

Data reduction of the inputs data is a common step to reduce the computing time and to filter the noise by selecting only data relevant to the analyzed components.[32,36] In PC-ANN, data are compressed into scores that best describe the data signal, then they are used as input data.

3.4.2. Support vector regression (SVR). The objective of SVR is to find a multivariate regression function $f(x)$ based on signal (UV absorbance) data set matrix X to predict a desired output property (concentration). The complete SVR equations are fully derived in ref. 37 and 38, and the summary equation is given in Table 2, eqn (7).

There are many parameters that need to be adjusted for successful SVR predictions such as C (regularization constant), ε (insensitive loss function by Vapnik[28,39]), and σ (kernel width parameter). The kernel used in this work is the Gaussian Radial Basis Function (RBF) (Table 2, eqn (8)).

Table 2 Complete list of equations used for developing and optimizing the five compared chemometric models[a]

Equation no.	Equation form	Variables	Application model
1	$\text{RMSECV} = \sqrt{\text{PRESS}/n}$	n	PLS, SVR
2	$\text{PRESS} = \sum (Y_{\text{pred}} - Y_{\text{true}})^2$	$Y_{\text{pred}}, Y_{\text{true}}$	PLS, SVR
3	$X = CK + E$	X, C, K, E	SRACLS
4	$X = CS^{\text{T}} + E$	X, C, S^{T}, E	SRACLS
5	$\% \ \text{lof} = 100x \sqrt{\dfrac{\sum e_{ij}^2}{\sum d_{ij}^2}}$	$\% \ \text{lof}, e_{ij}^2, d_{ij}^2$	MCR-ALS
6	$R^2 = 100x \left(1 - \dfrac{\sum e_{ij}^2}{\sum d_{ij}^2} \right)$	R^2, e_{ij}^2, d_{ij}^2	MCR-ALS
7	$f(x) = \sum_{i,j=1}^{N} (\alpha_i - \alpha_i^*) \langle \varnothing(x_i) \cdot \varnothing(x_j) \rangle + b$	X, α_i, α_i^*	SVR
8	$K(x_i, x_i) = \exp\left(\dfrac{-\|x_i - x_j\|^2}{2\sigma^2} \right)$	K, σ	SVR
9	$\text{RMSEP} = \sqrt{\dfrac{\sum_{i=1}^{n} (\hat{c}_i - c_i - \text{bias})^2}{n-1}}$	$c_i, \hat{c}_i, n$	PLS, SRACLS, MCR-ALS, SVR, PC-ANN
10	$\text{Bias} = \dfrac{\sum_{i=1}^{n} \hat{c}_i - c_i}{n}$	$c_i, \hat{c}_i, n$	PLS, SRACLS, MCR-ALS, SVR, PC-ANN

[a] n, is the number of training samples; Y_{pred} and Y_{true} are predicted and true concentrations in µg mL^{-1}, respectively; $X(i,j)$ is a matrix of the UV absorbance spectra for the j variables (wavelengths in our case) and the i samples; $C(i,n)$ is a matrix of the concentration profiles for the n components; $K(n,j)$ is a matrix of the pure component signals (spectra at unit concentrations in our case); E is the residual error matrix; $S^{\text{T}}(n,j)$ is the matrix of pure spectra of n components at j measured wavelengths; $\%$ lof is the percentage lack of fit; R^2 is the explained variance; e_{ij}^2 and d_{ij}^2 are the elements of the residual matrix (E) and spectral matrix (X), respectively; α_i and α_i^* are the Lagrange multipliers that satisfy the constraint $0 \leq \alpha_i$ and $\alpha_i^* \leq C$; σ is the kernel width parameter; c_i is the reference concentration; $\hat{c}_i$ is the calculated concentration.

Radial SVR function (non-linear) is used for modelling, where the optimum values for ε, C, and σ were obtained by running a grid search based on leave one-out-CV to give the lowest RMSECV. The tested values for ε was (0.01–1.01), for C (30–990), and for σ (0.1–10.1). The grid search was performed in two stages, the first using a wide grid followed by a fine search. With every set of SVR parameters, the SVR model was built on the 24 samples and leaving only one sample. This is followed by predicting the RMSECV for the samples that have been removed, and then the average of RMSECV after all samples have been removed is computed (Table 2, eqn (1) and (2)).
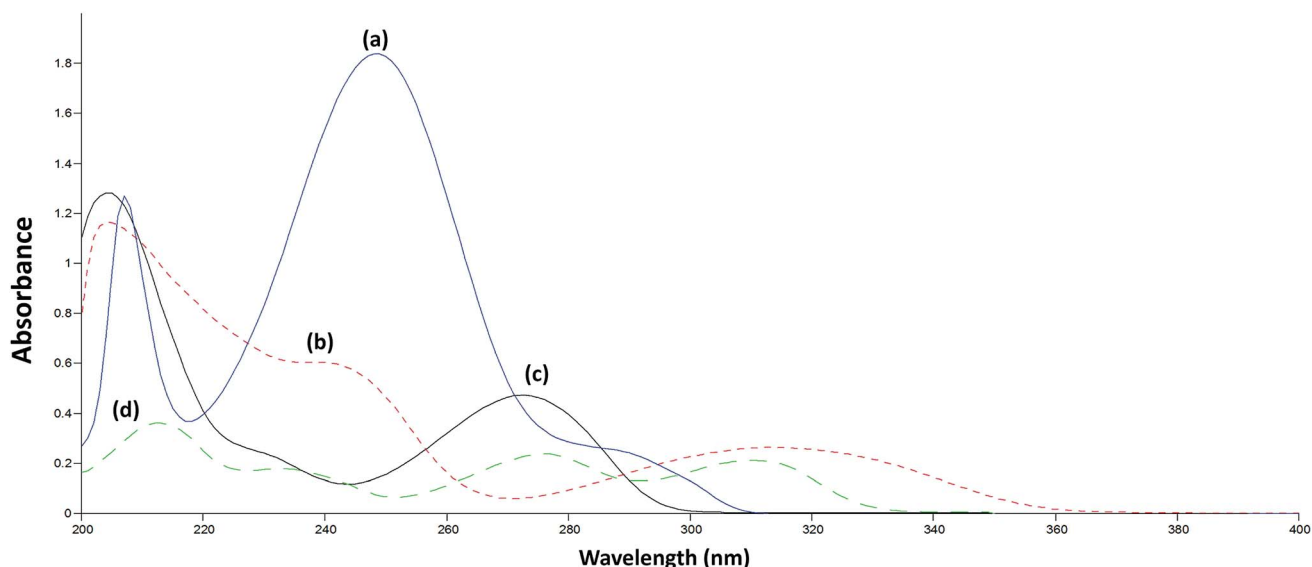

Fig. 3 Overlay absorption spectra of 20 µg mL^{-1} of (a) PAR, 20 µg mL^{-1} of (b) ERG, 10 µg mL^{-1} of (c) CAF and 5 µg mL^{-1} of (d) MET.

Table 3 Comparison of different error estimates of the proposed calibration methods[a]

Method	PLS				SRACLS				MCR-ALS				SVR				PC-ANN			
	ERG	MET	CAF	PAR	ERG	MET	CAF	PAR	ERG	MET	CAF	PAR	ERG	MET	CAF	PAR	ERG	MET	CAF	PAR
RMSEC ($\mu g\ mL^{-1}$)	0.0658	0.0271	0.2119	2.1271	—	0.0908	0.4217	2.1437	—	0.0439	1.7018	2.1589	0.0505	0.0180	0.2217	1.1110	0.0202	0.0439	0.3492	2.1589
RMSECV ($\mu g\ mL^{-1}$)	0.0766	0.0311	0.2669	2.6168	—	0.0804	0.4310	2.5833	—	0.2766	0.2670	2.4423	0.0499	0.0217	0.2098	2.0068	0.0202	0.0439	1.7018	2.1589
RMSEP ($\mu g\ mL^{-1}$)	0.0879	0.3648	2.0732	6.4209	—	0.2950	2.1171	5.9514	—	0.2029	2.5266	5.0593	0.0694	0.2893	2.0891	5.8506	0.0250	0.1543	0.8865	4.0885

PLS

True ($ng\ mL^{-1}$)				ERG		MET		CAF		PAR	
ERG	MET	CAF	PAR	Found	R%	Found	R%	Found	R%	Found	R%
0.45	3.00	25.00	160.00	0.47	105.13	2.89	96.41	25.93	103.72	164.90	103.06
0.45	2.25	25.00	165.00	0.47	105.11	2.11	93.91	26.04	104.14	167.39	101.45
0.50	2.50	20.00	170.00	0.48	95.72	2.47	98.84	21.64	108.21	173.57	102.10
0.50	3.00	27.50	180.00	0.43	86.17	2.95	98.40	29.39	106.87	185.46	103.03
0.55	2.10	25.00	195.00	0.45	81.43	2.14	101.97	27.41	109.65	198.05	101.56
0.40	3.00	25.00	150.00	0.47	116.44	3.07	102.43	27.84	111.37	155.32	103.55
0.60	2.50	22.50	190.00	0.50	83.24	2.94	117.74	24.61	109.38	200.15	105.34
0.40	2.75	20.00	140.00	0.55	138.05	3.11	113.09	22.07	110.33	148.00	105.72
0.55	2.50	27.50	190.00	0.50	91.76	3.00	120.11	29.60	107.64	198.86	104.66
0.40	2.00	22.00	160.00	0.56	141.16	2.46	123.14	24.57	111.70	169.38	105.86
0.60	2.75	27.00	140.00	0.57	94.17	3.36	122.21	29.10	107.77	145.67	104.05
0.50	2.25	20.00	180.00	0.59	117.91	2.87	127.75	22.30	111.52	184.38	102.44
Mean (%)					104.69		108.53		109.67		103.57
S.D.					20.13		12.18		2.68		1.57

SRACLS

ERG		MET		CAF		PAR	
Found	R%	Found	R%	Found	R%	Found	R%
—	—	2.95	98.43	25.85	103.40	164.01	102.50
—	—	2.18	96.67	26.27	105.08	167.33	101.41
—	—	2.48	99.19	21.83	109.13	174.24	102.49
—	—	3.00	99.96	29.61	107.67	183.43	101.91
—	—	2.11	100.52	28.07	112.27	197.59	101.33
—	—	3.17	105.79	27.55	110.22	153.54	102.36
—	—	2.79	111.63	24.92	110.76	199.84	105.18
—	—	3.08	112.09	21.54	107.70	148.13	105.81
—	—	2.87	114.94	29.86	108.59	197.26	103.82
—	—	2.36	118.03	24.55	111.61	169.84	106.15
—	—	3.35	121.76	28.43	105.30	143.85	102.75
—	—	2.68	119.01	22.31	111.57	185.23	102.91
—		108.17		108.61		103.22	
—		9.11		2.87		1.66	

[a] CAF, caffeine; ERG, ergotamine; MET, metoclopramide; PAR, paracetamol; PLS, partial least squares; RMSECV, root mean square error of CV; RMSEP, relative mean squared error of prediction; SRACLS, spectral residual augmented classical least squares method.

4. Results

4.1. Parameter optimization

Optimization of multivariate models is a very important step that benefits from analyst experiences to achieve optimum performance of each model and avoid common multivariate problems such as overfitting. In this work, each compound was calibrated individually in each model. The individual calibration gave results that are superior to global calibration due to the fact that not all the analyzed compounds can be calibrated with all tested models. This was expected in our study mixture due to the abnormal variability in its component ratios.

Parameters of each tested model were studied and then optimized to achieve optimum performance. For PLS, appropriate selection of the number of LVs to be used to construct the model is the key to achieve correct quantitation. The method developed by Haaland and Thomas[24] was used to determine the optimum number of factors, which involves selecting that model including the smallest number of factors that result in an insignificant difference between the corresponding RMSECV and the minimum RMSECV. The optimum number of PLS latent variables by the leave-one-out CV technique was 4 for all analyzed compounds. For SRACLS, the optimum number of loadings P_{new} used for augmentation of $\hat{K}$ matrix was determined with the help of leave-one-out CV and found to be 10, 8, and 3 for MET, CAF and PAR, respectively. For SVR, RBF kernel demonstrated better results due to the assumed non-linearity of the analyzed components. The grid search showed that the lowest RMSECV resulted in $\varepsilon = 0.06$, 0.01, 0.21 and 0.21, $C = 60$, 990, 990 and 990 and $\sigma = 0.1$, 5.1, 4.1 and 2.6 for ERG, MET, CAF and PAR, respectively.

For MCR-ALS, different constraints were studied and optimized, and satisfactory results were observed when applying non-negativity constraints (nnl) for spectral and concentration matrices, with a low lack of fit percentage (% lof) of 1.9743 and a high variance captured by the model (R^2) of 99.96. The least number of iterations that could explain our models was 2, 6, and 11 for MET, CAF, and PAR, respectively.

For ANN, a low number of nodes was used to avoid high computer processing time, overfitting, and noise incorporation in data modelling.[40,41] Accordingly, the input matrix was reduced using principal component analysis (PCA) from 101 points to only four principal components for all analyzed compounds. These components contained only the scores that best describe the analyzed components. A single hidden layer was sufficient to solve the mixture's components under study, where more hidden layers may cause overfitting.[34] Different parameters were optimized using a trial-and-error method to achieve the best network architecture. The number of hidden neurons was optimized by testing different numbers of neurons and comparing the relative mean squared error of prediction (RMSEP) values, where 20, 17, 16, and 20 neurons for ERG, MET, CAF, and PAR, respectively, achieved the best performance. The transfer functions that performed best in ANN modelling were tansig–purelin, due to the assumed non-linearity of ERG, MET, and CAF. However, purelin–purelin performed better during

modeling of PAR. The training stopped when the RMSEP of the training set decreased and that of the test set increased. In addition, several learning functions were studied and their performance was nearly the same (*i.e.*, there was no decrease in RMSEP). A Levenberg–Marquardt training algorithm (TRAINLM) was thus used, as it saves time. In addition, the learning rate was set to 0.1 while the coefficients of decrease and increase of learning were set to 0.001.

4.2. Spectral analysis

A comprehensive comparison was made between the original and the reconstructed spectra of the calibration matrix in order to select the most suitable range of wavelengths. The spectra were recorded over the range 200.0–400.0 nm. The spectra of ERG, MET, CAF, and PAR have similarities in the informative spectral range (200.0–380.0 nm) (Fig. 3). However, the wavelengths preceding 280.0 nm were excluded as they were highly influenced by the high absorbance characteristics of PAR (the highest ratio component), and that was demonstrated by high noise in the calibration matrix, lower precision, and masking the contribution of the other components available in the
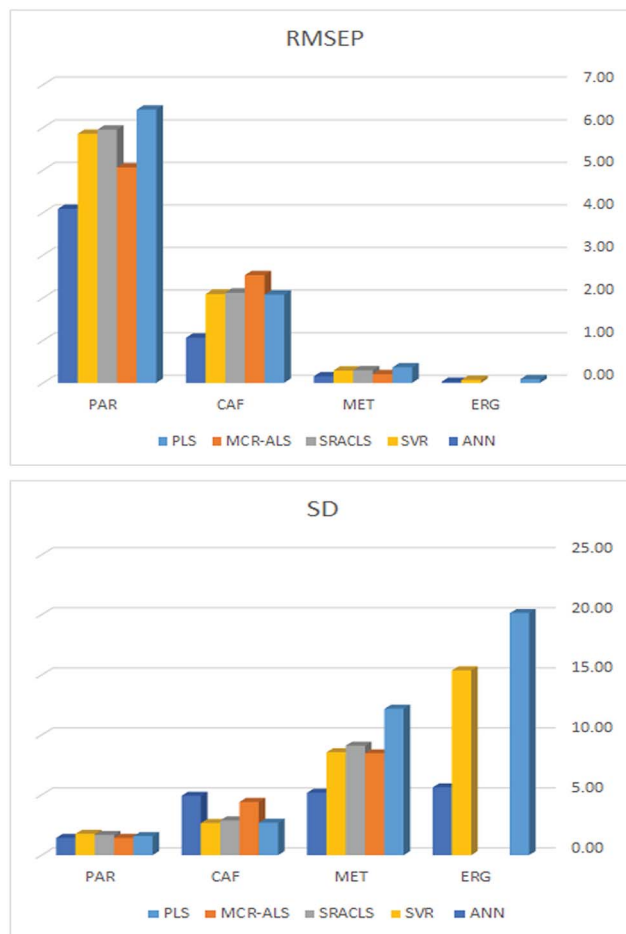

Fig. 4 Bar plots for comparison of the RMSEP and SD values obtained by application of the five proposed methods for the analysis of the independent test set.

mixture. In addition, wavelengths above 380.0 nm were also excluded due to poor spectral characteristics for all the analyzed compounds at their measured concentration levels.

The performance of the proposed models was assessed by recovery percentage and RMSEP for predictive ability and SD for intermediate precision (Table 3, Fig. 4). The proposed models were found able to resolve ERG, MET, CAF, and PAR in the calibration and test sets (Tables S1–S3,† 4 and 5). In addition, all proposed models could successfully predict the concentrations of MET, CAF, and PAR in the training and test sets with high intermediate precision, except for MET where PLS had lower precision compared with the other four models (Tables S1–S3,† 4 and 5). For ERG (the compound with the lowest ratio), PC-ANN, SVR, and PLS could be ordered increasingly according to their quantitative ability and precision (Tables 4 and 5, Fig. 4). Only PC-ANN showed remarkable superiority over all other models in quantifying, with success and high precision, the independent test set of all compounds. In addition, PC-ANN was applicable for the analysis of ERG, MET, CAF, and PAR in the tested dosage form (Table 6), where the tablet excipients did not interfere with the spectra of ERG, MET, CAF, and PAR. Error estimates for the proposed models are presented in Table 3, where the RMSEP value calculated for the independent test set (Table 2, eqn (9) and (10)) was used as an estimate of efficiency for the PC-ANN model over other proposed models.

4.3. Analysis of pharmaceutical products

The most qualified model (PC-ANN) was successfully applied for the determination of ERG, MET, CAF, and PAR in Metograine® tablets. Six replicate determinations were made. Satisfactory results were obtained for each compound in a good agreement with the labeled claims (Table 6).

The mean percentage recovery of declared contents ($n = 6$) was found to be 99.47, 101.14, 100.07, and 101.63 for ERG, MET, CAF, and PAR, respectively, in Metograine® tablets.

5. Discussion

Five multivariate calibration models were pushed to resolve and quantify a unique mixture as a guide for the pharmaceutical drug analysis community and to help them choose a suitable method for the analysis of mixtures with extreme nonlinear problems.

The mixture chosen for this study was unique in that it possesses inherent criteria resulting in extreme nonlinearity in its spectral data. First, the ratios of the components of this mixture are highly variable (0.5 : 2.5 : 25 : 165 for ERG, MET, CAF, and PAR, respectively), where determination of ERG (with the lowest ratio) and PAR (with the highest ratio) simultaneously with MET and CAF is nearly impossible using univariate, and most multivariate, approaches. The other problem that assures the uniqueness of this mixture is that all components have a rule in the overall nonlinearity of the spectral data for the analyzed mixtures due to design constraints. MET has a small contribution in the overall nonlinearity as only −2 level of the calibration design is below the proposed linearity range.

Table 4 Analysis results for the prediction of the independent set by PLS, SRACLS, MCR-ALS, and SVR models[a]

Method				MCR-ALS								SVR							
True (ng mL⁻¹)				ERG		MET		CAF		PAR		ERG		MET		CAF		PAR	
ERG	MET	CAF	PAR	Found	R%	Found	R%	Found	R%	Found	R%	Found	R%	Found	R%	Found	R%	Found	R%
0.45	3.00	25.00	160.00	—	—	2.98	99.34	25.87	103.50	163.97	102.48	0.50	111.03	2.95	98.50	25.94	103.75	164.43	102.77
0.45	2.25	25.00	165.00	—	—	2.10	93.55	25.89	103.56	168.25	101.97	0.50	110.97	2.18	97.09	26.03	104.11	167.07	101.25
0.50	2.50	20.00	170.00	—	—	2.40	95.84	21.53	107.65	173.29	101.93	0.50	99.68	2.54	101.63	21.62	108.09	172.29	101.35
0.50	3.00	27.50	180.00	—	—	2.86	95.40	29.40	106.89	183.76	102.09	0.50	99.89	3.02	100.57	29.59	107.61	183.75	102.08
0.55	2.10	25.00	195.00	—	—	2.12	100.90	27.23	108.92	198.44	101.77	0.50	90.88	2.20	104.99	27.69	110.77	195.58	100.30
0.40	3.00	22.50	150.00	—	—	3.13	104.44	27.87	111.47	153.86	102.57	0.50	124.42	3.13	104.41	27.83	111.32	154.97	103.32
0.60	2.50	20.00	190.00	—	—	2.89	115.78	24.74	109.96	198.65	104.55	0.51	85.50	2.88	115.12	24.81	110.27	199.20	104.84
0.40	2.75	27.50	140.00	—	—	2.89	105.17	22.51	112.57	146.25	104.47	0.50	125.89	3.06	111.39	22.03	110.16	146.45	104.61
0.55	2.50	22.00	190.00	—	—	2.56	102.47	30.28	110.10	196.37	103.35	0.50	90.89	2.84	113.78	29.58	107.58	196.51	103.43
0.40	2.00	27.00	160.00	—	—	1.95	97.66	25.35	115.22	168.36	105.23	0.50	125.02	2.38	118.99	24.45	111.15	169.98	106.24
0.60	2.75	20.00	140.00	—	—	2.47	89.65	30.48	112.91	141.83	101.31	0.50	83.31	3.23	117.63	28.90	107.05	146.70	104.79
0.50	2.25	20.00	180.00	—	—	1.85	82.13	23.67	118.37	181.19	100.66	0.50	99.96	2.73	121.24	22.17	110.87	185.40	103.00
			Mean (%)		—		98.53		110.09		102.70		103.95		108.78		108.56		103.16
			S.D.		—		8.47		4.42		1.41		15.38		8.55		2.65		1.75

[a] CAF, caffeine; ERG, ergotamine; MCR-ALS, multicurve resolution alternating least squares; MET, metoclopramide; PAR, paracetamol; SVR, support vector regression.

CAF has higher contribution due to only +2 level of the calibration design being at the border of linearity range and all other design levels were outside this range. PAR and ERG have the highest contribution in model nonlinearity as PAR levels incorporated in the design are much higher than the proposed linearity range, while ERG levels incorporated in the design are much lower than the proposed linearity range and nearly undetectable.

A combination of the problems described above produces a unique mixture with severe nonlinearity characteristics. While most of the published work is dealing with nonlinearity problems focused mainly on mixtures with minor nonlinearity problems (weighing errors, instrumental noise, and uncalibrated glassware, for example), our work makes a resolution and quantification challenge for five different chemometric models with different mathematical backgrounds to deal with a mixture that represents an extreme in nonlinearity characteristics. The result of this comparison can guide analysts to suitable models that can be used in the determination of dosage forms with similar problems in future instances.

Each multivariate model used in this study has different characteristics. PLS is considered the standard chemometric model for many applications and is well known for its simpler computation and conception compared with other more complex models such as SVR and ANN.[42] SRACLS and MCR-ALS have the ability to extract pure spectra of each component, which increases the resolution power of these models. Usually, ANN and SVR models have advantages over other multivariate models because they can perform nonlinear regression. The comparison was made on the basis of recovery percentage, standard deviations, and RMSEP.

On application of the PLS model, only four LVs were used to describe the four analyzed compounds, despite having the nonlinearity problems and concentration ratio variability mentioned above. Furthermore, it was not expected that PLS would perform better than the other models (SRACLS and MCR-ALS) in the determination of the challenging component ERG. This can be explained by the spectral characteristics of each compound; ERG and MET (compounds with the lowest contribution) show absorbance over the entire analyzed spectral range (280.0–380.0 nm) while CAF and PAR show absorbance only over a specific portion of the analyzed spectral range (280.0–350.0 nm) (Fig. 3). The region of the spectral range in which only the small concentration compounds (ERG and MET) show absorbance characteristics may have helped the PLS model to discriminate these compounds without the need for more LVs. In addition, it was observed that PLS has the lowest performance among the studied multivariate methods due to the simplicity of its calculations and its inferior ability to deal with nonlinearity in spectral data (Fig. 4).

For MCR-ALS, it was observed that the model could resolve the spectra of ERG, MET, CAF, and PAR efficiently, where it was able to account for 99.96% of the variance in the analyzed spectra; this was beneficial in the qualitative confirmation of their presence. The quantitative performance of the model was higher than the PLS model, especially in the determination of PAR despite failure to quantitate ERG (Fig. 4).

For SRACLS, the number of **k**-loading vectors used for describing each component decreased with increasing concentration ratio for each compound. The quantitative performance of the model was higher than the PLS model in the determination of MET, CAF, and PAR despite failure to quantitate ERG (Fig. 4).

Learning machine methods (*e.g.* ANN and SVR) were applied to test their well-known merits over other models in the manipulation of spectral data with extreme nonlinearity, especially for ERG for which the other models studied performed badly in their prediction. In general, PC-ANN and SVR performed better than any other tested model for the determination of ERG, MET, CAF, and PAR. PC-ANN had superior

Table 5 Analysis results for the prediction of the independent set by PC-ANN model[a]

Method				PC-ANN							
ERG	MET	CAF	PAR	ERG		MET		CAF		PAR	
True (ng mL^{-1})				Found	R%	Found	R%	Found	R%	Found	R%
0.45	3.00	25.00	160.00	0.50	111.64	3.01	100.44	23.79	95.16	163.21	102.01
0.45	2.25	25.00	165.00	0.47	104.28	2.06	91.77	25.15	100.59	166.13	100.68
0.50	2.50	20.00	170.00	0.49	97.33	2.66	106.25	18.81	94.07	170.43	100.25
0.50	3.00	27.50	180.00	0.49	97.67	3.14	104.56	27.84	101.24	185.19	102.89
0.55	2.10	25.00	195.00	0.52	94.24	2.23	106.06	24.80	99.19	196.92	100.98
0.40	3.00	25.00	150.00	0.44	110.34	3.13	104.46	26.63	106.53	154.48	102.99
0.60	2.50	22.50	190.00	0.59	98.10	2.87	114.83	22.24	98.85	196.44	103.39
0.40	2.75	20.00	140.00	0.42	105.89	2.81	102.23	20.11	100.53	143.58	102.56
0.55	2.50	27.50	190.00	0.54	97.31	2.57	102.80	28.92	105.16	196.61	103.48
0.40	2.00	22.00	160.00	0.40	101.23	2.06	103.00	23.09	104.94	165.31	103.32
0.60	2.75	27.00	140.00	0.58	96.55	2.88	104.80	27.18	100.68	143.09	102.21
0.50	2.25	20.00	180.00	0.50	99.96	2.33	103.67	19.36	96.82	178.47	99.15
		Mean (%)		101.21		103.74		100.31		101.99	
		S.D.		5.62		5.18		2.88		1.41	

[a] ANN, artificial neural networks; CAF, caffeine; ERG, ergotamine; MET, metoclopramide; PAR, paracetamol.

Table 6 Analysis of the Metograine® tablets with the PC-ANN model[a]

Dosage form	Metograine® tablets			
Method	PC-ANN			
Drug	ERG	MET	CAF	PAR
R%	105.25	97.19	94.07	102.01
	103.04	106.06	101.24	100.68
	104.28	104.46	99.19	100.25
	97.33	94.07	106.53	102.89
	92.67	102.23	98.85	100.98
	94.24	102.80	100.53	102.99
Mean (%)	99.47	104.60	100.07	101.63
S.D.	4.96	5.33	3.69	1.06

[a] ANN, artificial neural networks; CAF, caffeine; ERG, ergotamine; MET, metoclopramide; PAR, paracetamol.

performance over SVR, especially in predictions of ERG (Fig. 4). This may be on account of merits of ANN algorithms and choosing PCs that best represent the relevant spectral data with minimal contribution of noise.

To summarize, multivariate models can deal with nonlinearity problems in spectral data with different levels of efficiency. The PC-ANN model has merits over other models in successfully extracting useful information from extreme and complex nonlinear spectral data that helps in resolution and quantitation of compounds under investigation. SRACLS and MCR-ALS have limited ability in dealing with severe nonlinearity in spectral data, although they introduce useful information on the qualitative composition of the analyzed mixtures. The PLS model has some intrinsic abilities that allow it to resolve and quantitate the analyzed compounds.

The PC-ANN method can be applied in the quality control of ERG, MET, CAF, and PAR in laboratory-prepared mixtures and pharmaceutical formulations, with the advantage of spectrophotometric methods for the quantitative determination of samples with high sensitivity, minimum sample preparations, minimum laboratory consumption, and low cost materials.

6. Conclusion

The present work introduced a comparative study of five different multivariate calibration models; PLS, SRACLS, MCR-ALS, SVR, and PC-ANN. SD and RMSEP values in independent test sets reveal that all models could resolve and predict MET, CAF, and PAR with different levels of accuracy and precision, while PC-ANN gave the best results, particularly in the prediction of EGR (a nearly undetectable compound with the lowest concentration ratio). The SVR and PC-ANN methods showed better performance and superior quantification power of spectral data with extreme complexity and nonlinearity characteristics. The PLS model has an inherent merit in dealing with nonlinear data, allowing the model to resolve the four analyzed compounds and quantitate them with moderate accuracy and precision.

The proposed chemometric models make use of the simplicity of spectrophotometric methods and widen the area of application of these methods in cases where severe nonlinearity in spectral data is expected. In addition, these methods avoid the need for tedious treatments and the complications of time- and cost-consuming chromatographic techniques. Furthermore, the proposed PC-ANN model can be used for routine quality-control analysis of ERG, MET, CAF, and PAR in their dosage form without interference from added excipients.

References

1 A. Molarius and Å. Tegelberg, *Headache: The Journal of Head and Face Pain*, 2006, **46**, 73–81.
2 J. O'Sullivan and J. McCabe, *J. Am. Assoc. Nurse Anesth.*, 2006, **74**, 61–69.
3 J. Olesen, *Headache: The Journal of Head and Face Pain*, 1978, **18**, 268–271.
4 H. T. Benamer, D. Deleu and D. Grosset, *J. Headache Pain*, 2010, **11**, 1–3.
5 *The British Pharmacopeia*, Her Majesty's Stationary Office, London, 2013.
6 S. C. Sweetman, *Martindale: The complete drug reference*, Pharmaceutical Press, Great Britain, 33rd edn, 2002.
7 V. Vichare, P. Mujgond, V. Tambe and S. Dhole, *Int. J. PharmTech Res.*, 2010, **2**, 2512–2516.
8 M. Khoshayand, H. Abdollahi, M. Shariatpanahi, A. Saadatfard and A. Mohammadi, *Spectrochim. Acta, Part A*, 2008, **70**, 491–499.
9 J. K. Inamadugu, R. Damaramadugu, R. Mullangi and V. Ponneri, *Biomed. Chromatogr.*, 2010, **24**, 1006–1014.
10 D. Šatínský, I. Neto, P. Solich, H. Sklenářová, M. Conceicao, B. Montenegro and A. N. Araújo, *J. Sep. Sci*, 2004, **27**, 529–536.
11 H. Li, C. Zhang, J. Wang, Y. Jiang, J. P. Fawcett and J. Gu, *J. Pharm. Biomed. Anal.*, 2010, **51**, 716–722.
12 Y. Zhang, N. Mehrotra, N. R. Budha, M. L. Christensen and B. Meibohm, *Clin. Chim. Acta*, 2008, **398**, 105–112.
13 T. Harsono, M. Yuwono and G. Indrayanto, *J. AOAC Int.*, 2005, **88**, 1093–1098.
14 *The United States Pharmacopeia, 32nd edition, The National Formulary, 27th edn*, United States Pharmacopeial Convention, Washington, DC, 2009.
15 A. Halim, *Die Pharmazie*, 1981, **36**, 157.
16 M. Amin and W. Sepp, *J. Chromatogr. A*, 1976, **118**, 225–232.
17 M. Aranda and G. Morlock, *J. Chromatogr. Sci.*, 2007, **45**, 251–255.
18 S. Wadher, P. Pathankar, M. Puranik, R. Ganjiwale and P. Yeole, *Indian J. Pharm. Sci.*, 2008, **70**, 393.
19 M. A. Sultan, H. M. Maher, N. Z. Alzoman, M. M. Alshehri, M. S. Rizk, M. S. Elshahed and I. V. Olah, *J. Chromatogr. Sci.*, 2012, **51**, 502–510.
20 F. Ibrahim and M. Wahba, *Sep. Sci. Technol.*, 2014, **49**, 2228–2240.
21 A. E. Gindy and G. M. Hadad, *J. AOAC Int.*, 2012, **95**, 609–623.
22 M. Wesolowski and B. Suchacz, *J. AOAC Int.*, 2012, **95**, 652–668.
23 R. G. Brereton, *Analyst*, 1997, **122**, 1521–1529.

24 D. M. Haaland and E. V. Thomas, *Anal. Chem.*, 1988, **60**, 1193–1202.

25 M. Stone, *J. R. Stat. Soc.*, 1974, **36**, 111–147.

26 I. A. Naguib, *Bull. Fac. Pharm.*, 2011, **49**, 91–100.

27 J. Jaumot, R. Gargallo, A. de Juan and R. Tauler, *Chemom. Intell. Lab. Syst.*, 2005, **76**, 101–110.

28 F. Parrella, *degree of Information Science*, University of Genoa, 2007.

29 M. M. Elkhoudary, I. A. Naguib, R. A. A. Salam and G. M. Hadad, *J. Fluoresc.*, 2017, 1–12.

30 A. Abbaspour and L. Baramakeh, *Spectrochim. Acta, Part A*, 2006, **64**, 477–482.

31 Y. Dou, H. Mi, L. Zhao and Y. Ren, *Anal. Biochem.*, 2006, **351**, 174–180.

32 M. M. Elkhoudary, R. A. Abdel Salam and G. M. Hadad, *Spectrochim. Acta, Part A*, 2014, **130**, 222–229.

33 H. Demuth, M. Beale and M. Hagan, *MATLAB User's Guide, version 4.0: Neural network toolbox*, MathWorks Inc., Natick, MA, USA, 2005.

34 A. R. Khanchi, M. K. Mahani, M. Hajihosseini, M. G. Maragheh, M. Chaloosi and F. Bani, *Food Chem.*, 2007, **103**, 1062–1068.

35 H. W. Darwish, M. I. Attia, A. S. Abdelhameed, A. M. Alanazi and A. H. Bakheit, *Molecules*, 2013, **18**, 974–996.

36 A. Afkhami, M. Abbasi-Tarighat and H. Khanmohammadi, *Talanta*, 2009, **77**, 995–1001.

37 L. A. Baumes, J. M. Serra, P. Serna and A. Corma, *J. Comb. Chem.*, 2006, **8**, 583–596.

38 U. Thissen, B. Ustun, W. J. Melssen and L. M. Buydens, *Anal. Chem.*, 2004, **76**, 3099–3105.

39 S. R. Gunn, *Support Vector Machines for Classification and Regression*, University of Southampton, UK, 1998.

40 M. M. Elkhoudary, R. A. A. Salam and G. M. Hadad, *Spectrochim. Acta, Part A*, 2014, **130**, 222–229.

41 A. M. Yehia and H. M. Mohamed, *Spectrochim. Acta, Part A*, 2016, **152**, 491–500.

42 T. Mehmood, K. H. Liland, L. Snipen and S. Sæbø, *Chemom. Intell. Lab. Syst.*, 2012, **118**, 62–69.