

Quantitative analysis and resolution of pharmaceuticals in the environment using multivariate curve resolution-alternating least squares (MCR-ALS)

AHMED MOSTAFA*
HEBA SHAABAN

Department of Pharmaceutical
Chemistry, College of Clinical Pharmacy
Imam Abdulrahman Bin Faisal
University, King Faisal Road
P.O. Box 1982, Dammam 31441
Saudi Arabia

The study presents the application of multivariate curve resolution alternating least squares (MCR-ALS) with a correlation constraint for simultaneous resolution and quantification of ketoprofen, naproxen, paracetamol and caffeine as target analytes and triclosan as an interfering component in different water samples using UV-Vis spectrophotometric data. A multivariate regression model using the partial least squares regression (PLSR) algorithm was developed and calculated. The MCR-ALS results were compared with the PLSR obtained results. Both models were validated on external sample sets and were applied to the analysis of real water samples. Both models showed comparable and satisfactory results with the relative error of prediction of real water samples in the range of 1.70–9.75 % and 1.64–9.43 % for MCR-ALS and PLSR, resp. The obtained results show the potential of MCR-ALS with correlation constraint to be applied for the determination of different pharmaceuticals in complex environmental matrices.

Keywords: multivariate curve resolution alternating least squares, correlation constraint, partial least squares, multivariate calibration, environmental analysis, UV-Vis spectrophotometry

Accepted September 29, 2018
Published online October 15, 2018

Different classes of pharmaceuticals and personal care products are becoming a source of environmental contamination. They have been reported to occur in different water systems such as surface water, groundwater and wastewater (1–3). Non-steroidal anti-inflammatory drugs (NSAIDs) such as naproxen, ketoprofen, ibuprofen and paracetamol, antidepressants, antibiotics and antiepileptics are examples of some of the pharmaceuticals that are commonly detected in water environments (4–6). Caffeine is also detected in different environmental compartments, because a significant amount of caffeine could be disposed of household wastes from unconsumed caffeine-containing beverages or could leave the human body unchanged *via* urine or feces (7). In addition, caffeine

* Correspondence; e-mail: ammostafa@iau.edu.sa; ammostaf@uwaterloo.ca

is commonly used in combination with analgesics and dietary supplements (8). Triclosan is a broad-spectrum antimicrobial agent that is widely used in different personal care products and has been detected in wastewater (9). Continual discharge of pharmaceuticals into the environment might lead to adverse effects on humans and wildlife. For instance, accumulation of NSAIDs such as diclofenac in water may cause harmful renal effects in humans (10). There is, therefore, a great demand to monitor these emerging contaminants in such complex environmental matrices.

Liquid chromatography (LC) (10) and gas chromatography (GC) (11) are the most commonly employed analytical techniques for the analysis of pharmaceuticals in environmental samples. Despite the high sensitivity and selectivity of the chromatographic techniques used, they require tedious and time-consuming procedures of sample preparation. Moreover, these techniques require expensive solvents and sophisticated instruments.

UV-Vis spectrophotometry is a simple and cost-effective technique. However, the critical challenge of this technique is the occurrence of overlapped spectra and matrix interferences (12). In such situations, multivariate calibration models could be effectively used since they provide results from mathematical resolution of the overlapped spectra that are comparable to those obtained by chromatographic techniques.

Multivariate curve resolution (MCR) is one of the most commonly used multivariate calibration models. Multivariate curve resolution-alternating least squares (MCR-ALS) is a mathematical algorithm proposed by Tauler in 1995 (13). It decomposes the data matrix into a bilinear model producing pure spectral and concentration profiles of the mixture components. The algorithm has been successfully used to analyze data coming from different sources such as IR (14), chromatography (15), hyperspectral imaging (16), nuclear magnetic resonance (17) and X-ray fluorescence (18). Parastar and Shaye (19) applied the algorithm to spectrophotometric determination of some pharmaceuticals in different water samples. Estrogens in natural water were determined using the same algorithm applied to LC with diode array detection (20). For more MCR-ALS applications, readers are referred to refs. 21 and 22.

In this work, MCR-ALS with correlation constraint has been applied to overcome the complex matrix effect and to quantitatively determine selected pharmaceuticals in different aqueous media using first-order UV spectrophotometric data. In addition, a partial least squares regression (PLSR) method was also developed and optimized. Optimized methods were statistically compared to a reported HPLC method (23). To the best of our knowledge, this is the first study to use MCR-ALS for simultaneous determination of ketoprofen (KTP), naproxen (NPX), paracetamol (PAR) and caffeine (CAF) in the presence of triclosan (TRC) as interfering species in different environmental water samples.

EXPERIMENTAL

Chemicals and reagents

Pure standards of KTP, NPX, PAR, CAF and TRC were purchased from Sigma-Aldrich (now Merck KGaA, Germany) and certified to contain $\geq 98, 99, 99, 99$ and 99% API, resp., as per official methods (26). Hydrochloric acid and methanol were obtained from Merck (Germany). Ultra-pure water (18.2 M Ω) was purified using a Pure Lab Ultra water system (ELGA, High Wycombe, UK) and used for all sample preparations.

Instrumentation and software

A Hach UV-Visible spectrophotometer, model DR6000 (Hach, USA) with 1-cm quartz cell was used for spectrophotometric measurements. The wavelength range was 200–400 nm, bandwidth was 1 nm and the wavelength scanning speed was 2800 nm min⁻¹. Spectra were exported as Excel files. Microsoft Excel 2010 was used for plotting the acquired spectra. PLS Toolbox software version 8.5 (R8.5.2) (Eigenvector Research, Inc., Manson, WA, USA) was used for PLSR analysis. MCR-ALS calculations were performed using MCR-ALS GUI 2.0 software for use with Matlab 2015a (24), freely available (25).

Standard solutions and real samples

Stock solutions of the four analytes and TRC were prepared separately in methanol, at a concentration of 100 µg mL⁻¹ and stored in dark at 4 °C. Working standard solutions were freshly prepared by appropriate dilution in 0.05 mol L⁻¹ HCl. A set of 25 calibration solutions of the four analytes and TRC was prepared in the concentration range of 1.0–11.0 µg mL⁻¹. A validation set of additional 7 samples containing the four analytes and TRC was similarly prepared.

Sea and well waters were collected from Al Khobar city, Saudi Arabia. Wastewater samples were collected from the Dammam wastewater treatment plant (Dammam, Saudi Arabia). All samples were collected in amber glass bottles with TeflonTM lined caps and then transferred to the laboratory in ice boxes. Water samples were filtered through 0.45-µm Nylon Acrodisc membrane filters (Gelman Sciences Inc., USA). No further treatment was applied to any sample. The samples were spiked with standard solutions of KTP, NPX, PAR, CAF and TRC in the concentration range of 1.0–11.0 µg mL⁻¹. A test set of 20 samples was prepared. All samples were prepared in duplicate. All samples were prepared in 0.05 mol L⁻¹ HCl with no pH adjustment.

A multilevel multifactor design (27) was used to build the calibration model. A five-factor, five-level design was used in the concentration range of 1.0–11.0 µg mL⁻¹ for the four analytes and TRC. The selected design provided factors that are mutually orthogonal and span each other's calibration space symmetrically. Concentrations of the validation set were selected randomly within the calibration range and the test set was fixed at concentration levels of 4.0, 5.0, 6.0, 8.0 and 9.0 µg mL⁻¹ for each analyte. Table I shows the calibration matrix along with the validation and test sets.

Chemometric methods

Multivariate curve resolution alternating least squares (MCR-ALS). – MCR extracts relevant information of pure components in a mixture through bilinear model decomposition of the data matrix. This model can be expressed in Eq. 1:

$$D = CS^T + E \quad (1)$$

where D is the experimental data matrix containing the measured spectra. Columns of matrix C contain the concentration profiles of all analytes and S^T is the matrix of the corresponding pure spectra. E is the matrix associated to experimental error and represents the data that is not explained by the model (residuals).

Table I. Concentration matrix used for the preparation of calibration, validation and test sets of KTP, NPX, PAR, CAF and TRC

No.	Calibration set						Validation & test sets					
	Concentration ($\mu\text{g mL}^{-1}$)						Concentration ($\mu\text{g mL}^{-1}$)					
	KTP	NPX	PAR	CAF	TRC	Matrix	KTP	NPX	PAR	CAF	TRC	Matrix
1	6.0	6.0	6.0	6.0	6.0	-	10.0	3.0	10.0	3.0	10.0	-
2	6.0	1.0	1.0	11.0	3.5	-	10.0	10.0	3.0	3.0	10.0	-
3	1.0	1.0	11.0	3.5	11.0	-	5.0	10.0	10.0	3.0	3.0	-
4	1.0	11.0	3.5	11.0	6.0	-	3.0	3.0	10.0	10.0	3.0	-
5	11.0	3.5	11.0	6.0	3.5	-	3.0	3.0	3.0	3.0	3.0	-
6	3.5	11.0	6.0	3.5	3.5	-	3.0	10.0	3.0	3.0	10.0	-
7	11.0	6.0	3.5	3.5	8.5	-	5.0	5.0	5.0	10.0	5.0	-
8	6.0	3.5	3.5	8.5	11.0	-	4.0	4.0	4.0	4.0	4.0	Sea water
9	3.5	3.5	8.5	11.0	8.5	-	5.0	5.0	5.0	5.0	5.0	Sea water
10	3.5	8.5	11.0	8.5	6.0	-	6.0	6.0	6.0	6.0	6.0	Sea water
11	8.5	11.0	8.5	6.0	11.0	-	8.0	8.0	8.0	8.0	8.0	Sea water
12	11.0	8.5	6.0	11.0	11.0	-	9.0	9.0	9.0	9.0	9.0	Sea water
13	8.5	6.0	11.0	11.0	1.0	-	4.0	4.0	4.0	4.0	4.0	Tap water
14	6.0	11.0	11.0	1.0	8.5	-	5.0	5.0	5.0	5.0	5.0	Tap water
15	11.0	11.0	1.0	8.5	1.0	-	6.0	6.0	6.0	6.0	6.0	Tap water
16	11.0	1.0	8.5	1.0	6.0	-	8.0	8.0	8.0	8.0	8.0	Tap water
17	1.0	8.5	1.0	6.0	8.5	-	9.0	9.0	9.0	9.0	9.0	Tap water
18	8.5	1.0	6.0	8.5	8.5	-	4.0	4.0	4.0	4.0	4.0	Waste water
19	1.0	6.0	8.5	8.5	3.5	-	5.0	5.0	5.0	5.0	5.0	Waste water
20	6.0	8.5	8.5	3.5	1.0	-	6.0	6.0	6.0	6.0	6.0	Waste water
21	8.5	8.5	3.5	1.0	3.5	-	8.0	8.0	8.0	8.0	8.0	Waste water
22	8.5	3.5	1.0	3.5	6.0	-	9.0	9.0	9.0	9.0	9.0	Waste water
23	3.5	1.0	3.5	6.0	1.0	-	4.0	4.0	4.0	4.0	4.0	Well water
24	1.0	3.5	6.0	1.0	1.0	-	5.0	5.0	5.0	5.0	5.0	Well water
25	3.5	6.0	1.0	1.0	11.0	-	6.0	6.0	6.0	6.0	6.0	Well water
							8.0	8.0	8.0	8.0	8.0	Well water
							9.0	9.0	9.0	9.0	9.0	Well water

Information about estimation of the possible number of components, non-negativity and correlation constraints have been described elsewhere (24, 28).

Quality of the MCR-ALS model can be assessed using the percentage of lack of fit (Eq. 2):

$$\text{lack of fit}(\%) = 100 \sqrt{\frac{\sum_{i,j} e_{ij}^2}{\sum_{i,j} d_{ij}^2}} \quad (2)$$

where d_{ij} is an element of the data matrix D and e_{ij} is the related residual (difference between experimental data input and model reproduced data).

Partial least-squares regression (PLSR). – PLSR decomposes the spectral data matrix D and concentration vector c simultaneously into scores T and loadings P and q according to Eqs. 3 and 4:

$$D = TP^T + E \quad (3)$$

$$c = Tq + f \quad (4)$$

where T is the score matrix, P^T and q are the matrix and vector loadings describing the variance in D and c , resp. E and f are the experimental errors (residuals) in D and c . This decomposition looks for a set of components called factors or latent vectors in such a way that the first few factors explain as much as possible of the covariance between D and c . This is then followed by a regression step where the decomposition of D is used to predict c (29). In this work, data was mean-centered before decomposition.

Validation of models

In order to evaluate the quality of prediction of the developed MCR-ALS and PLSR models, a set of external validation samples were used. From the results of these samples, some figures of merit were calculated according to the following equations to describe the validation results.

Root mean square error of prediction (RMSEP):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (c_i - \hat{c}_i)^2}{n}} \quad (5)$$

bias:

$$\text{bias} = \frac{\sum_{i=1}^n (c_i - \hat{c}_i)}{n} \quad (6)$$

standard error of prediction (SEP):

$$SEP = \sqrt{\frac{\sum_{i=1}^n (c_i - \hat{c}_i - \text{bias})^2}{n-1}} \quad (7)$$

relative percentage error in the concentration predictions (RE, %):

$$RE(\%) = 100 \sqrt{\frac{\sum_{i=1}^n (c_i - \hat{c}_i)^2}{\sum_{i=1}^n c_i^2}} \quad (8)$$

where c_i and \hat{c}_i are the known and predicted analyte concentration in sample i , resp., and n is the total number of samples used in the validation set.

In addition, the slope, intercept and correlation coefficient were calculated for a linear regression fit found between the known and predicted concentrations.

RESULTS AND DISCUSSION

Multivariate calibration results

Fig. 1 shows the pure UV absorption spectra of the four analytes KTP, NPX, PAR, CAF along with the interfering component TRC at a concentration of $10 \mu\text{g mL}^{-1}$. As can be observed, the spectra are highly overlapped. In addition, the TRC spectrum shows a strong overlap with target analytes along the wavelength range of 200–340 nm. As a result, quantitative resolution of target analytes is impossible using any univariate or conventional method. Therefore, MCR-ALS and PLSR were proposed to resolve these complex mixtures with the interfering component. NPX shows very strong absorption in the range preceding 240 nm, which led to noisy spectra of the assured calibration mixtures in this range. Therefore, wavelengths of 200–240 nm were excluded. Moreover, wavelengths of 340–400 nm

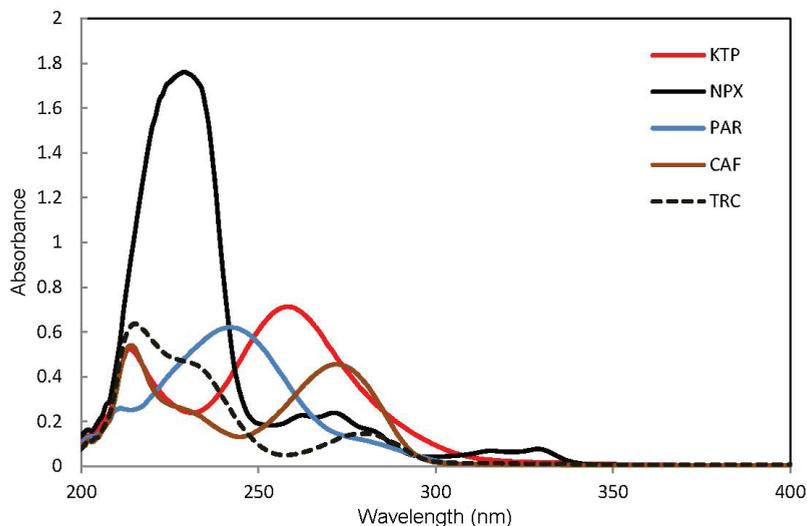


Fig. 1. Pure UV-Vis spectra of $10.0 \mu\text{g mL}^{-1}$ of ketoprofen (KTP), naproxen (NPX), paracetamol (PAR), caffeine (CAF) and triclosan (TRC).

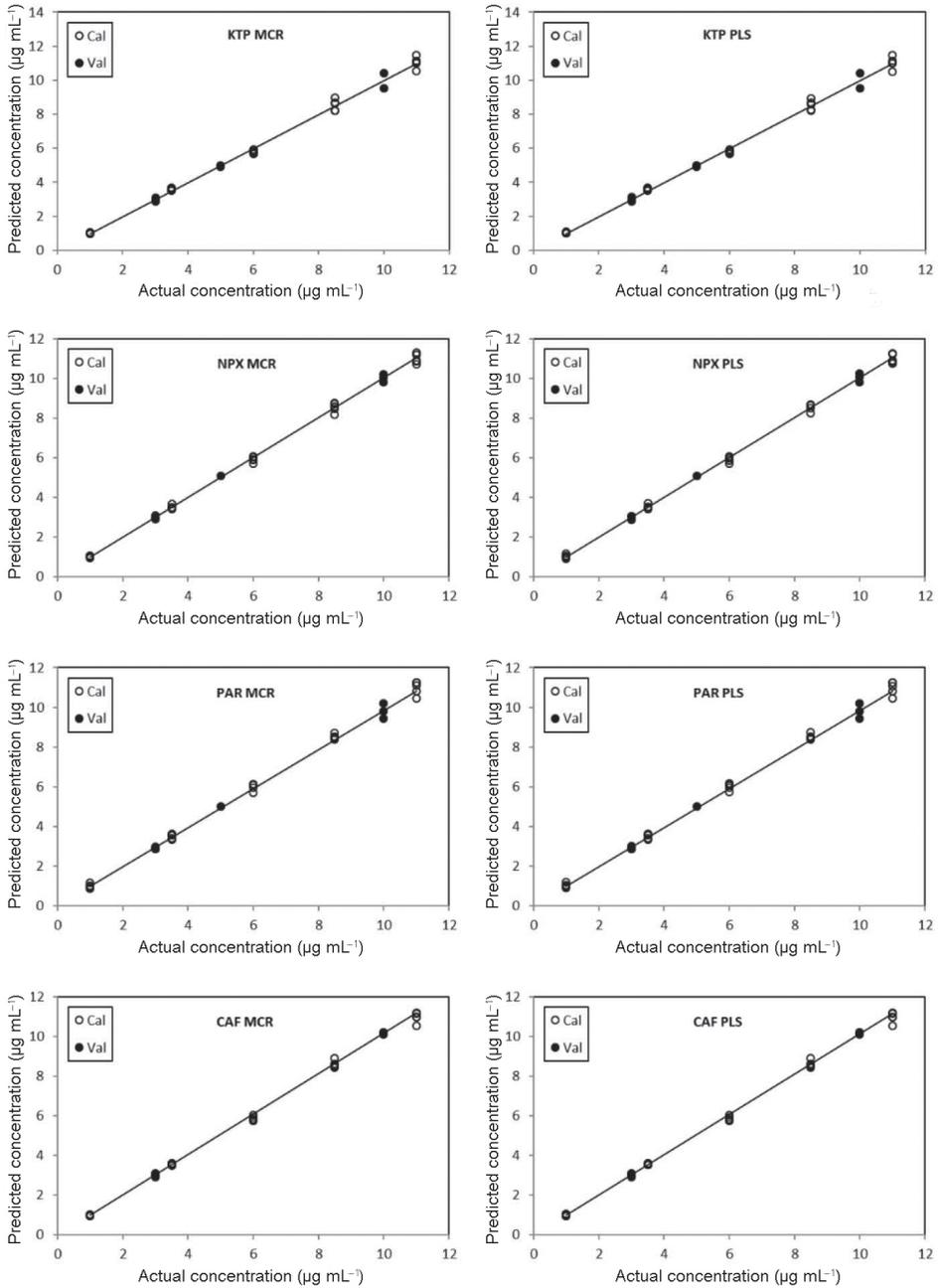


Fig. 2. Plot of actual analyte concentrations *vs.* MCR-ALS and PLS predicted values.

were excluded as well due to poor absorption of the target analytes at the concentration levels measured. The wavelength range of 240–340 nm was therefore selected for quantitative determination of the target analytes using MCR-ALS and PLSR models.

Five major components were initially determined in the data matrix before starting the MCR-ALS using singular value decomposition (SVD). Pure spectra of target analytes were used as initial estimates to check the MCR-ALS resolution and to decrease the rotational ambiguity effects. Initial spectral profile estimates were calculated using the simple-to-use interactive self-modeling mixture analysis (SIMPLISMA) (30). MCR-ALS was applied to the 25-mixture calibration data set using non-negativity constraints in spectral and concentration profiles. A fast non-negativity constrained least squares algorithm (FNNLS, fast NNLS) (31) was used. In addition, correlation constraint was applied (24), and the variable containing quantitative information of target analytes was selected (*i.e.*, concentration profile of the four target analytes in the data matrix). The convergence criterion was set at 0.1 % and the maximum number of iterations was 100. However, no more than 10 iterations were required to achieve convergence in all the tested samples. The regression model was developed.

Fig. 2 shows the scatter plot of resolved MCR-ALS concentration values *vs.* the actual concentrations. MCR-ALS recovered spectral profiles of the four target analytes and the interference component as well (Fig. 3). Correlation coefficients higher than 0.998 were achieved for all components. Table II presents the figures of merit of the regression model of the calibration set. The results show excellent correlation coefficients ($R > 0.998$) and low relative error RE = 3.15, 2.28, 2.52 and 2.30 % for KTP, NPX, PAR and CAF, resp. NPX showed the best statistical results of the calibration set. This was mainly due to the UV absorption band in the wavelength range of 310–340 nm, with almost no interference from either the interfering component (TRC) or the other three analytes (Fig. 1).

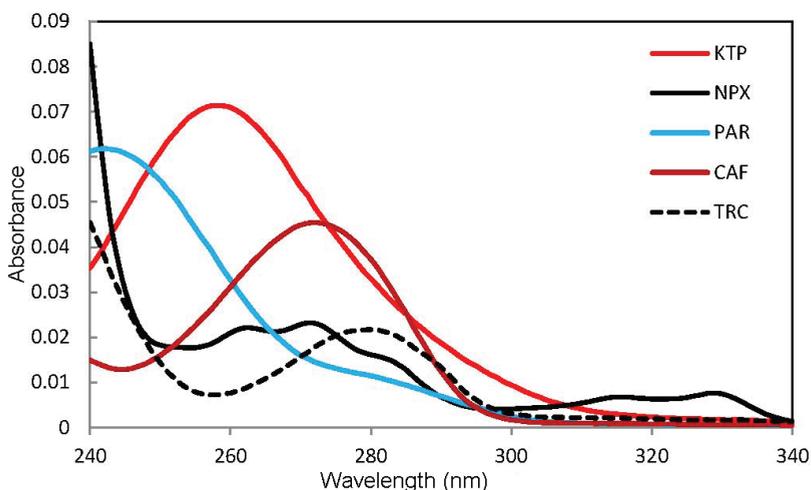


Fig. 3. MCR-ALS resolved spectral profiles of the four target analytes (KTP, NPX, PAR and CAF) and the interfering component (TRC).

Table II. Figures of merit of MCR-ALS and PLSR models for the calibration set of KTP, NPX, PAR and CAF

Parameters	MCR					PLSR				
	KTP	NPX	PAR	CAF	CAF	KTP	NPX	PAR	CAF	CAF
Calibration range ($\mu\text{g mL}^{-1}$)	1.0-11.0	1.0-11.0	1.0-11.0	1.0-11.0	1.0-11.0	1.0-11.0	1.0-11.0	1.0-11.0	1.0-11.0	1.0-11.0
Slope (b)	1.0000	1.0000	1.0000	1.0000	1.0000	0.9962	0.9982	0.9976	0.9979	0.9979
Standard error of slope	1.29×10^{-2}	9.37×10^{-3}	1.04×10^{-2}	9.46×10^{-3}	9.46×10^{-3}	1.28×10^{-2}	6.10×10^{-2}	1.03×10^{-2}	9.44×10^{-3}	1.23×10^{-2}
Intercept (a)	-3.55×10^{-2}	2.31×10^{-3}	-6.22×10^{-3}	1.55×10^{-2}	1.55×10^{-2}	2.27×10^{-2}	1.06×10^{-2}	1.47×10^{-2}	1.23×10^{-2}	1.23×10^{-2}
Standard error of intercept	9.01×10^{-2}	6.52×10^{-2}	7.22×10^{-2}	6.59×10^{-2}	6.59×10^{-2}	8.92×10^{-2}	8.76×10^{-3}	7.17×10^{-2}	6.58×10^{-2}	6.58×10^{-2}
RMSECV	21.9×10^{-2}	15.7×10^{-2}	17.6×10^{-2}	16.1×10^{-2}	16.1×10^{-2}	21.8×10^{-2}	14.9×10^{-2}	17.5×10^{-2}	16.0×10^{-2}	16.0×10^{-2}
SEP	21.5×10^{-2}	15.6×10^{-2}	17.2×10^{-2}	15.7×10^{-2}	15.7×10^{-2}	21.3×10^{-2}	6×10^{-2}	17.1×10^{-2}	15.7×10^{-2}	15.7×10^{-2}
Bias	5.9×10^{-15}	3.73×10^{-16}	9.25×10^{-15}	1.41×10^{-14}	1.41×10^{-14}	-4.0×10^{-13}	3.64×10^{-12}	-5.4×10^{-12}	6.0×10^{-13}	6.0×10^{-13}
RE (%)	3.15	2.28	2.52	2.30	2.30	3.13	2.13	2.51	2.30	2.30
Correlation coefficient (R)	0.9981	0.9990	0.9988	0.9989	0.9989	0.9981	0.9991	0.9988	0.9989	0.9989

Table III. Figures of merit of MCR-ALS and PLSR models for the validation set of KTP, NPX, PAR and CAF

Parameters	MCR-ALS					PLSR				
	KTP	NPX	PAR	CAF	CAF	KTP	NPX	PAR	CAF	CAF
Accuracy (Mean \pm SD) ^a	100.8 ± 2.10	100 ± 2.29	98.1 ± 2.62	100.8 ± 2.21	100.8 ± 2.21	101.0 ± 2.36	100.2 ± 2.63	98.1 ± 2.90	101.0 ± 2.21	101.0 ± 2.21
Repeatability (RSD, %) ^b	1.98	1.52	2.31	1.87	1.87	2.12	1.75	2.47	2.01	2.01
Intermediate precision (RSD, %) ^c	2.21	1.60	2.39	2.15	2.15	2.50	1.91	2.50	2.33	2.33
RMSEP	0.169	0.124	0.210	0.106	0.106	0.166	0.135	0.218	0.104	0.104
SEP	0.156	0.115	0.195	0.098	0.098	0.153	0.138	0.224	0.106	0.106
Bias	-6.01×10^{-2}	-1.51×10^{-2}	0.102	-5.66×10^{-2}	-5.66×10^{-2}	-6.21×10^{-2}	-2.12×10^{-2}	0.102	-5.85×10^{-2}	-5.85×10^{-2}
RE (%)	2.68	1.75	2.97	1.80	1.80	2.63	1.91	3.08	1.76	1.76
Correlation coefficient (R)	0.9989	0.9993	0.9985	0.9998	0.9998	0.9989	0.9992	0.9983	0.9989	0.9989

^a Mean and standard deviation for 7 determinations.

^b Intra-day relative standard deviation ($n = 3$), average of three different concentrations repeated three times within a day.

^c Inter-day relative standard deviation ($n = 3$), average of three different concentrations repeated three times on three different days.

The same calibration data matrix used for MCR-ALS analysis was also used to develop the PLSR model for comparison. The data was divided into two input matrices: one as the calibration matrix and the other as the test matrix as required by the software. Different preprocessing methods such as auto-scaling and mean-centering were tested. Mean-centering was selected due to its higher sensitivity and good results. The LOO-CV was used to determine the number of PLSR latent factors based on the root mean square error of cross validation (RMSECV). The optimum number of latent factors was selected. The selected model was the one with the smallest number of factors with RMSECV, which was not significantly greater than that for the model that yielded the lowest RMSECV. The model showed five factors. In contrast to MCR-ALS, PLSR is not able to provide an estimation of pure spectra of the analytes in the mixture. Table II shows the figures of merit of the PLSR calibration model compared to the MCR-ALS model. The results show that both models are comparable. PLSR showed excellent correlation coefficients between 0.9981 and 0.9991 and low relative errors between 2.13 and 3.13 %. In summary, MCR-ALS and PLSR models showed comparable and satisfactory calibration results where there was no matrix effect. However, it is important to test the prediction ability of the developed models on external data sets (*i.e.*, validation and test sets). Therefore, the models were applied for the prediction of validation and test sets with different concentrations within the calibration range.

Validation set results

The developed models were applied for prediction of the concentrations of KTP, NPX, PAR and CAF in an external validation set of 7 synthetic mixtures with different concentrations (Table I). The MCR-ALS algorithm was performed using the same constraints applied for the calibration set. Table III shows the figures of merit for predictions of the validation set in terms of percent recovery, RMSEP, SEP, RE (%), bias and *R*. Percent recoveries are satisfactory, ranging from 98.1 ± 2.6 to 100.8 ± 2.2 %. The model showed excellent correlation coefficients, *R*, between 0.9985 and 0.9998 and low relative errors RE between 1.75 and 2.68 %. Fig. 2 shows the regression plots of the MCR-ALS predicted analyte concentrations *versus* actual concentrations.

The validation set was mean-centered and the PLSR model was applied to it using 5 latent factors. Regression plots of the predicted *versus* actual concentrations showed high correlation coefficients between 0.9983 and 0.9998 (Fig. 2 and Table III). Table III shows the obtained figures of merit of external validation of the MCR-ALS model compared to the PLSR model. Both models showed comparable and satisfactory results.

Real samples

To test the ability of the developed models to quantify target analytes in real samples and overcome the interference from background constituents, four different water samples (sea, tap, waste and well water) were spiked with target analytes and analyzed using the developed models. Table I shows the test samples spiked with appropriate amounts of analytes, in duplicate, at five different concentrations (4.0, 5.0, 6.0, 8.0 and 9.0 $\mu\text{g mL}^{-1}$).

The same MCR-ALS constraints and PLSR pre-processing of calibration and validation sets were applied to the set of real samples. Table IV shows the statistical figures of merit of the real sample set of both models. The MCR-ALS results obtained revealed slightly higher relative errors RE (%) than those obtained for the synthetic mixtures. This was

Table IV. Figures of merit of MCR-ALS and PLSR models for prediction of test sets (spiked sea, tap, waste and well water)^a

Test	Parameters	MCR-ALS					PLSR				
		KTP	NPX	PAR	CAF	KTP	NPX	PAR	CAF		
Sea water	RMSEP	0.353	7.55×10^{-2}	0.200	0.250	0.357	0.106	0.190	0.243		
	SEP	0.315	6.76×10^{-2}	0.179	0.224	0.320	9.51×10^{-2}	0.170	0.217		
	Bias	-0.257	-3.99×10^{-2}	0.152	0.180	-0.264	-5.81×10^{-2}	0.144	0.174		
	RE (%)	8.99	1.93	5.10	6.37	9.11	2.71	4.85	6.19		
	Correlation coefficient (R)	0.9924	0.9981	0.9982	0.9991	0.9922	0.9980	0.9981	0.9990		
Tap water	RMSEP	0.139	8.23×10^{-2}	6.68×10^{-2}	0.125	0.140	9.43×10^{-2}	6.44×10^{-2}	0.124		
	SEP	0.124	7.36×10^{-2}	5.97×10^{-2}	0.111	0.125	8.43×10^{-2}	5.76×10^{-2}	0.111		
	Bias	-6.47×10^{-2}	2.54×10^{-2}	-1.49×10^{-2}	-2.93×10^{-2}	-7.36×10^{-2}	8.05×10^{-3}	-2.11×10^{-2}	-3.15×10^{-2}		
	RE (%)	3.54	2.10	1.70	3.18	3.56	2.40	1.64	3.17		
	Correlation coefficient (R)	0.9980	0.9978	0.9990	0.9982	0.9980	0.9975	0.9990	0.9982		
Waste water	RMSEP	0.335	8.82×10^{-2}	0.383	0.376	0.342	8.97×10^{-2}	0.370	0.366		
	SEP	0.299	7.89×10^{-2}	0.342	0.336	0.306	8.03×10^{-2}	0.331	0.327		
	Bias	-0.258	2.72×10^{-2}	0.294	0.290	-0.263	8.44×10^{-3}	0.285	0.283		
	RE (%)	8.53	2.25	9.75	9.57	8.72	2.29	9.43	9.33		
	Correlation coefficient (R)	0.9927	0.9944	0.9979	0.9975	0.9927	0.9942	0.9978	0.9976		
Well water	RMSEP	0.222	0.102	0.330	0.330	0.232	0.112	0.321	0.323		
	SEP	0.199	9.16×10^{-2}	0.295	0.295	0.207	0.100	0.287	0.289		
	Bias	-0.145	-4.42×10^{-3}	0.236	0.249	-0.152	-2.42×10^{-2}	0.228	0.243		
	RE (%)	5.67	2.61	8.40	8.42	5.91	2.86	8.18	8.24		
	Correlation coefficient (R)	0.9973	0.9983	0.9922	0.9974	0.9973	0.9981	0.9920	0.9974		

^aConcentration levels of 4.0, 5.0, 6.0, 8.0 and 9.0 µg mL⁻¹.

Table V. Statistical comparison between the results of MCR-ALS and PLSR models and the reported HPLC method for prediction of test sets (spiked sea, tap, waste and well water)^b

Test	Parameter	MCR-ALS				PLSR				HPLC (ref. 23)			
		KTP	NPX	PAR	CAF	KTP	NPX	PAR	CAF	KTP	NPX	PAR	CAF
Sea water	Recovery (%)	104.4	101.9	95.6	95.5	104.4	102	95.6	95.6	100.8	101.1	97.6	100.9
	SD ^b	5.67	2	1.96	4.3	6.03	2.5	1.78	4.16	2.57	1.80	1.02	3.58
	t-test	2.01	0.97	1.99	1.97	1.89	1.29	2.11	1.96	(2.31) ^c	(2.31) ^c	(2.31) ^c	(2.31) ^c
	F-value	4.88	1.23	3.69	1.44	5.32	1.93	3.03	1.35	(6.39) ^c	(6.39) ^c	(6.39) ^c	(6.39) ^c
Tap water	Recovery (%)	101.4	98.7	99.7	100.9	101.4	98.8	99.7	100.9	100.7	99.8	99.7	100.1
	SD ^b	2.19	2.12	1.51	2.38	2.21	2.83	1.58	2.35	1.92	2.03	1.18	2.37
	t-test	1.10	1.83	0.01	0.50	0.97	1.02	0.07	0.51	(2.31) ^c	(2.31) ^c	(2.31) ^c	(2.31) ^c
	F-value	1.30	1.10	1.62	1.01	1.33	1.95	1.78	1.02	(6.39) ^c	(6.39) ^c	(6.39) ^c	(6.39) ^c
Waste water	Recovery (%)	107.7	99.6	92.3	93.6	107.7	99.6	92.4	93.7	102.1	100.0	98.2	99.3
	SD ^b	4.02	2.94	4.62	5.45	4.33	3.02	4.25	5.26	3.09	1.95	4.02	4.52
	t-test	2.14	0.26	2.03	1.42	2.03	0.21	2.09	1.41	(2.31) ^c	(2.31) ^c	(2.31) ^c	(2.31) ^c
	F-value	1.69	2.26	1.32	1.45	1.96	2.39	1.12	1.36	(6.39) ^c	(6.39) ^c	(6.39) ^c	(6.39) ^c
b	Recovery (%)	104.2	99.5	94.6	95	104.2	99.6	94.6	95.1	100.9	100.1	100.1	98.8
	SD ^b	3.19	2.66	4.18	4.46	3.47	3.34	4.01	4.32	2.97	2.37	3.95	3.82
	t-test	1.37	0.32	1.93	1.14	1.32	0.25	1.90	1.12	(2.31) ^c	(2.31) ^c	(2.31) ^c	(2.31) ^c
	F-value	1.15	1.25	1.12	1.37	1.37	1.98	1.03	1.28	(6.39) ^c	(6.39) ^c	(6.39) ^c	(6.39) ^c

^a Concentrations: 4.0, 5.0, 6.0, 8.0 and 9.0 µg mL⁻¹.

^b Mean and standard deviation for 5 determinations.

^c Theoretical values for t and F at (p = 0.05).

mainly due to the presence of the interfering matrix in real samples. Waste water samples showed the highest relative errors between 2.25–9.75 %. Nevertheless, NPX relative error was not that high with RE of 2.25 %, which was due to its characteristic absorption band in the wavelength range of 310–340 nm. Very similar results were obtained by the application of the PLSR model, as shown in Table IV. Results of MCR-ALS and PLSR models were compared with the reported HPLC method (23) based on Student's *t*-test and *F*-ratio at 95 % confidence level and showed no significant difference regarding accuracy and precision (Table V).

To sum up, MCR-ALS showed prediction ability comparable to the well-established PLSR. However, MCR-ALS has the advantage of recovering pure spectra of the analytes of interest as well as the interfering components (Fig. 3), thus allowing their possible identification and/or confirmation.

Method validation

The proposed methods were further validated in terms of linearity and range, as shown in Table II. Selectivity, accuracy, precision and repeatability of the proposed methods were assessed by the analysis of the external validation set, as shown in Table III. Satisfactory validation results were obtained for both models and showed that the models were accurate, precise, robust and specific over the specified range. Moreover, variation of hydrochloric acid strength by ± 0.02 mol L⁻¹ and temperature by ± 2 °C had no significant effects on the developed models.

CONCLUSIONS

MCR-ALS with the correlation constraint has been proven as an effective and accurate tool for the spectrophotometric determination of different pharmaceuticals in natural water samples when the sample matrix effect exists and in the presence of other interfering components. The predictive capability of MCR-ALS was comparable to that of the PLSR model. However, MCR-ALS has the advantage of recovering qualitative information about the analytes of interest and interfering species.

Acknowledgments. – This work was supported by the Deanship of Scientific Research, Imam Abdulrahman Bin Faisal University, Kingdom of Saudi Arabia (Grant No. 2015-082).

REFERENCES

1. S. D. Richardson and T. A. Ternes, Water analysis: emerging contaminants and current issues, *Anal. Chem.* 83 (2011) 4614–4648; <https://doi.org/10.1021/ac200915r>
2. K. Kümmerer, *Pharmaceuticals in the Environment: Sources, Fate, Effects and Risks*, Springer Science & Business Media, Heidelberg 2008, pp. 521.
3. D. S. Aga, *Fate of Pharmaceuticals in the Environment and in Water Treatment Systems*, CRC Press, Boca Raton (FL) 2007.
4. J. Rivera-Utrilla, M. Sánchez-Polo, M. Á. Ferro-García, G. Prados-Joya and R. Ocampo-Pérez, Pharmaceuticals as emerging contaminants and their removal from water. A review, *Chemosphere*, 93 (2013) 1268–1287; <https://doi.org/10.1016/j.chemosphere.2013.07.059>

5. T. A. Ternes, Occurrence of drugs in German sewage treatment plants and rivers 1, *Water Res.* **32** (1998) 3245–3260; [https://doi.org/10.1016/S0043-1354\(98\)00099-2](https://doi.org/10.1016/S0043-1354(98)00099-2)
6. M. D. Celiz, J. Tso and D. S. Aga, Pharmaceutical metabolites in the environment: analytical challenges and ecological risks, *Environ. Toxicol. Chem.* **28** (2009) 2473–2484; <https://doi.org/10.1897/09-173.1>
7. H. Shaaban, High speed hydrophilic interaction liquid chromatographic method for simultaneous determination of selected pharmaceuticals in wastewater using a cyano-bonded silica column, *J. Liq. Chromatogr. Relat. Technol.* **41** (2018) 180–187; <https://doi.org/10.1080/10826076.2018.1429282>
8. E. Gracia-Lor, N. I. Rousis, E. Zuccato, R. Bade, J. A. Baz-Lomba, E. Castrignanò, A. Causanilles, F. Hernández, B. Kasprzyk-Hordern and J. Kinyua, Estimation of caffeine intake from analysis of caffeine metabolites in wastewater, *Sci. Total Environ.* **609** (2017) 1582–1588; <https://doi.org/10.1016/j.scitotenv.2017.07.258>
9. F. Tohidi and Z. Cai, Fate and mass balance of triclosan and its degradation products: comparison of three different types of wastewater treatments and aerobic/anaerobic sludge digestion, *J. Hazard. Mater.* **323** (2017) 329–340; <https://doi.org/10.1016/j.jhazmat.2016.04.034>
10. H. Shaaban and T. Górecki, High temperature-high efficiency liquid chromatography using sub-2 μm coupled columns for the analysis of selected non-steroidal anti-inflammatory drugs and veterinary antibiotics in environmental samples, *Anal. Chim. Acta* **702** (2011) 136–143; <https://doi.org/10.1016/j.aca.2011.06.040>
11. K. Kotnik, T. Kosjek, U. Krajnc and E. Heath, Trace analysis of benzophenone-derived compounds in surface waters and sediments using solid-phase extraction and microwave-assisted extraction followed by gas chromatography-mass spectrometry, *Anal. Bioanal. Chem.* **406** (2014) 3179–3190; <https://doi.org/10.1007/s00216-014-7749-0>
12. A. El-Gindy, S. Emara and A. Mostafa, UV partial least-squares calibration and liquid chromatographic methods for direct quantitation of levofloxacin in urine, *J. AOAC Int.* **90** (2007) 1258–1265; <https://doi.org/10.1039/c0ay00662a>
13. R. Tauler, Multivariate curve resolution applied to second order data, *Chemom. Intel. Lab. Syst.* **30** (1995) 133–146; [https://doi.org/10.1016/0169-7439\(95\)00047-X](https://doi.org/10.1016/0169-7439(95)00047-X)
14. W. Chen, X.-Y. Liu, B.-C. Huang, L.-F. Wang, H.-Q. Yu and B. Mizaikoff, Probing membrane fouling via infrared attenuated total reflection mapping coupled with multivariate curve resolution, *Chemphyschem* **17** (2016) 358–363; <https://doi.org/10.1002/cphc.201500932>
15. M. Navarro-Reig, J. Jaumot, A. Baglai, G. Vivó-Truyols, P. J. Schoenmakers and R. Tauler, Untargeted comprehensive two-dimensional liquid chromatography coupled with high-resolution mass spectrometry analysis of rice metabolome using multivariate curve resolution, *Anal. Chem.* **89** (2017) 7675–7683; <https://doi.org/10.1021/acs.analchem.7b01648>
16. D. A. Forchetti and R. J. Poppi, Use of NIR hyperspectral imaging and multivariate curve resolution (MCR) for detection and quantification of adulterants in milk powder, *LWT-Food Sci. Technol.* **76** (2017) 337–343; <https://doi.org/10.1016/j.lwt.2016.06.046>
17. F. Puig-Castellví, I. Alfonso and R. Tauler, Untargeted assignment and automatic integration of ^1H NMR metabolomic datasets using a multivariate curve resolution approach, *Anal. Chim. Acta* **964** (2017) 55–66; <https://doi.org/10.1016/j.aca.2017.02.010>
18. J. B. Ghasemi, M. K. Rofouei and N. Amiri, Multivariate curve resolution alternating least squares in the quantitative determination of sulfur using overlapped S ($K\alpha$)–Mo ($L\alpha$) emission peaks by wavelength dispersive X-ray fluorescence spectrometry, *X-Ray Spectrom.* **44** (2015) 75–80; <https://doi.org/10.1021/acs.analchem.6b03116>
19. H. Parastar and H. Shaye, Comparative study of partial least squares and multivariate curve resolution for simultaneous spectrophotometric determination of pharmaceuticals in environmental samples, *RSC Adv.* **5** (2015) 70017–70024; <https://doi.org/10.1039/C5RA10658C>

20. R. L. Pérez and G. M. Escandar, Liquid chromatography with diode array detection and multivariate curve resolution for the selective and sensitive quantification of estrogens in natural waters, *Anal. Chim. Acta* **835** (2014) 19–28; <https://doi.org/10.1016/j.aca.2014.05.015>
21. C. Ruckebusch and L. Blanchet, Multivariate curve resolution: a review of advanced and tailored applications and challenges, *Anal. Chim. Acta* **765** (2013) 28–36; <https://doi.org/10.1016/j.aca.2012.12.028>
22. M. Garrido, F. Rius and M. Larrechi, Multivariate curve resolution-alternating least squares (MCR-ALS) applied to spectroscopic data from monitoring chemical reactions processes, *Anal. Bioanal. Chem.* **390** (2008) 2059–2066; <https://doi.org/10.1007/s00216-008-1955-6>
23. J. Santos, I. Aparicio, E. Alonso and M. Callejón, Simultaneous determination of pharmaceutically active compounds in wastewater samples by solid phase extraction and high-performance liquid chromatography with diode array and fluorescence detectors, *Anal. Chim. Acta* **550** (2005) 116–122; <https://doi.org/10.1016/j.aca.2005.06.064>
24. J. Jaumot, A. de Juan and R. Tauler, MCR-ALS GUI 2.0: New features and applications, *Chemom. Intel. Lab. Syst.* **140** (2015) 1–12; <https://doi.org/10.1016/j.chemolab.2014.10.003>
25. Multivariate Curve Resolution Homepage; <http://www.mcrals.info>; last access date Sept 4, 2018
26. *USP 29, NF 24*, USP Convention, Rockville (MD) USA, 2005; <http://www.pharmacopeia.cn/usp.asp>; last access date Sept 4, 2018
27. R. G. Brereton, Multilevel multifactor designs for multivariate calibration, *Analyst* **122** (1997) 1521–1529; <https://doi.org/10.1039/a703654j>
28. T. Azzouz and R. Tauler, Application of multivariate curve resolution alternating least squares (MCR-ALS) to the quantitative analysis of pharmaceutical and agricultural samples, *Talanta* **74** (2008) 1201–1210; <https://doi.org/10.1016/j.talanta.2007.08.024>
29. A. R. de Carvalho, M. del Nogal Sánchez, J. Wattoom and R. G. Brereton, Comparison of PLS and kinetic models for a second-order reaction as monitored using ultraviolet visible and mid-infrared spectroscopy, *Talanta* **68** (2006) 1190–1200; <https://doi.org/10.1016/j.talanta.2005.07.053>
30. W. Windig and J. Guilment, Interactive self-modeling mixture analysis, *Anal. Chem.* **63** (1991) 1425–1432; <https://doi.org/10.1021/ac00014a016>
31. R. Bro and S. De Jong, A fast non-negativity-constrained least squares algorithm, *J. Chemom.* **11** (1997) 393–401; [https://doi.org/10.1002/\(SICI\)1099-128X\(199709/10\)11:53.0.CO;2-L](https://doi.org/10.1002/(SICI)1099-128X(199709/10)11:53.0.CO;2-L)