Simultaneous spectrophotometric determination of the three structural isomers of cresol using multivariate regression methods

Marwa Said Moneeb[1*], Ismail Ibrahim Hewala[1], Hatem Abdelfattah Elmongy[2] and Abdel-Aziz Mohamed Wahbi[1]

[1]Department of Pharmaceutical Analytical Chemistry, Faculty of Pharmacy, University of Alexandria, Egypt
[2]Department of Pharmaceutical Analytical Chemistry, Faculty of Pharmacy, University of Damanhour, Egypt

Abstract: The simultaneous determination of the structural isomers of cresol was carried out using UV spectrophotometry by applying the principle component regression (PCR) and partial least squares (PLS) regression methods. Different concentration levels of cresol isomers were determined in their mixtures by construction of a partial factorial calibration design at four levels. Both multivariate calibration models were constructed using the correlation between the concentration and absorbance data matrices in the spectral region 283-305nm. The methods were validated by analyzing an independent validation set solutions of the same compounds. The methods were found to be accurate and precise as indicated by the mean % recovery (99.96-100.41%) and % relative standard deviation (0.15-0.72%), respectively. The methods were applied to the determination of cresol isomers in a topical veterinary preparation. The methods were proved to be applicable to the determination of the three cresol isomers without prior separation procedures, despite of the extensive spectral overlap of such compounds.

Keywords: Spectrophotometry, chemometrics, multivariate methods, isomers, cresol.

INTRODUCTION

Cresols, also known as methylphenols or hydroxytoluenes, are mixtures of the *o*-, *m*- and *p*- isomers that can be obtained from coal tar (fig. 1). It is official in the British pharmacopeia. Crude cresol of commercial grade contains approximately 20% of *o-cresol*, 40% of *m*-cresol, and 30% of *p*-cresol. Phenol is a common contaminant present in small amounts in cresols. Cresols are used as antiseptics, disinfectants and parasiticides in veterinary medicine (British Pharmacopoeia, 2010; OEHHA, 2000). It is known that the toxicity of cresol isomers varies markedly. A study on rat liver slices showed that the *p*-isomer is 5 to 10 times more toxic than the *m*- and the *o*-isomers (Thompson *et al*., 1994**).** Toxicity can be caused by inhalation, ingestion or skin contact (OEHHA, 2000).

Analysis of cresol isomers has been accomplished by gas chromatography (GC) with flame ionization detection (GC-FID) or coupled with mass spectrometry (GC-MS) (Cleghom *et al*., 1992; Williams *et al*., 1991). Capillary GC method has been reported for analysis of cresol isomers in urine (Bieniek, 1996). Co-electroosmotic capillary electrophoresis method has been used for their determination in combination with other phenolic compounds (Masselter *et al*., 1993; Masselter and Zemann 1995). Fluorimetric analysis of cresol isomers with a PLS multivariate calibration technique has been reported (Del Olmo *et al*., 1996). To the date there is no official method for the determination of cresol isomers.

Among the various methods available for the quantitation of drugs, spectrophotometry continues to be the most convenient analytical technique, due to its inherent simplicity, low cost and wide availability in most quality control laboratories. Literature review revealed that no spectrophotometric method has been reported for the simultaneous determination of cresol isomers.

Extensive UV spectral overlap is often a serious limitation to the application of univariate spectrophotometric methods. The UV absorption spectra of cresol isomers overlap extensively (fig. 2). Therefore the resolution of their ternary mixture is practically impossible by conventional spectrophotometry.

Recently, the multivariate regression methods, such as principal component regression (PCR) and partial least squares (PLS), have widely been used to resolve mixtures of two or more components spectrophotometrically without any prior separation steps (Kenneth *et al*., 1998; Kramer, 1998).

The aim of the present work is to develop spectrophotometric methods for the determination of cresol isomers in their mixtures and in a veterinary preparation using chemometric multivariate regression methods, principal component regression (PCR) and partial least squares (PLS). These methods can be considered useful alternatives to the reported sophisticated methods.

**Corresponding author:* e-mail: marwamoneeb@yahoo.com

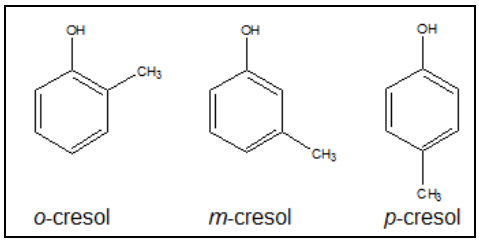

Fig. 1: The chemical structures of *o*-, *m*- and *p*-cresol.

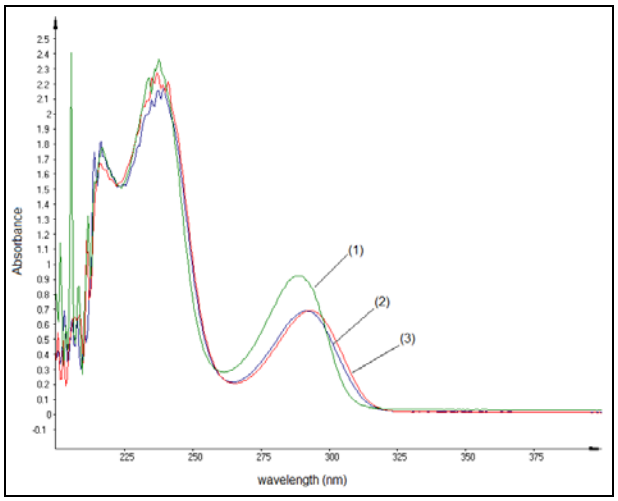

Fig. 2: The UV absorption spectra of solution containing 10µg mL^{-1} of (1) *o*-cresol, (2) *m*-cresol and (3) *p*-cresol in 0.1M NaOH-methanol (5:95% v/v).

MATERIALS AND METHODS

Materials and reagents

O-cresol (99%), *m*-cresol (99%), *p*-cresol (99%) and phenol (99%) were purchased from Sigma Aldrich, Egypt. Analytical grade methanol and sodium hydroxide were purchased from BDH, UK. Topical veterinary solution (Pharmachemical Pharmonia®) was purchased from Pharmachem, Australia. This aqueous solution is labeled to contain 30g L^{-1} cresol, in addition to sodium hydroxide, potassium hydroxide and oleic acid.

Instrumentation

Specrod-S600 Analytikjena UV-Vis photodiode array spectrophotometer interfaced to a personal computer loaded with WinAspect software for spectral data acquisition was used. The PCR and PLS analyses were performed using the Chemometric Toolbox 3.02 software (Kramer, 1995) for use with MATLAB 6.5.

Procedure

Preparation of standard solutions

100 mg of *o*-, *m*- and *p-cresol* was accurately weighed and transferred separately into 100mL volumetric flasks. Each isomer was dissolved in 5mL of 0.1M NaOH then diluted to volume with methanol (solution I). A 10mL aliquot of solution I was diluted to 100mL with the same solvent (solution II). Suitable dilutions of solution II with methanol were made to prepare solutions of the training set. The training set consisted of sixteen solutions containing *o*-, *m*-, and *p-cresol* in the concentration range of 0 to 15µg mL^{-1} (table 1). The training solutions have been prepared in the presence of phenol (3% of the total content of cresols in the solution). The validation set solutions were prepared similarly using independently prepared *o*-, *m*- and *p*-cresol stocks (table 1). These solutions were prepared in the concentration range of 0 to 12.5µg mL^{-1} (table 1).

Sample preparation

1mL aliquot of topical veterinary preparation (Pharmachemical Pharmonia®) was diluted to 100mL with NaOH-methanol (5:95% v/v). Six mL aliquot of this solution was diluted to 100mL with the same solvent to prepare the test solution.

Spectrophotometric measurement

The UV absorbances were recorded within the wavelength range 283-305 nm at 1 nm intervals.

RESULTS

Absorption spectra

The absorption spectra of *o*-, *m*- and *p-cresol* were separately recorded against 0.1M NaOH-methanol (5:95 % v/v) (fig. 2). It was found that the maximum absorption peaks (λ_{max}) of the three isomers were 289, 292 and 294 nm, respectively.

As could be perceived from fig. 2, the absorption spectra of the three isomers overlap extensively. Therefore, the resolution of their ternary mixture is practically impossible by conventional spectrophotometry. On the other hand, multivariate regression methods PCR and PLS offer an elegant complementary approach to solve such an analytical problem.

Construction of the multivariate models

Sixteen mixtures of the three isomers having variant proportions of them have been chosen as the training set (table 1). The training set was constructed using the 4-level partial factorial design (Brereton, 2003). This design involves $N=m \times l^p$ mixtures, where N is the number of mixtures, l is the number of concentration levels, $P \geq 2$ and $m \geq 1$ are integers. The levels selected were 0, 5, 10 and 15µg mL^{-1} for the three cresol isomers.

The multivariate regressions were computed with the PCR and PLS algorithms basing on the correlation for the absorbance and the matching concentration data matrices of the training set.

To generate a valid multivariate calibration, the constructed training set should contain all expected components in the unknown samples (Kramer, 1998).

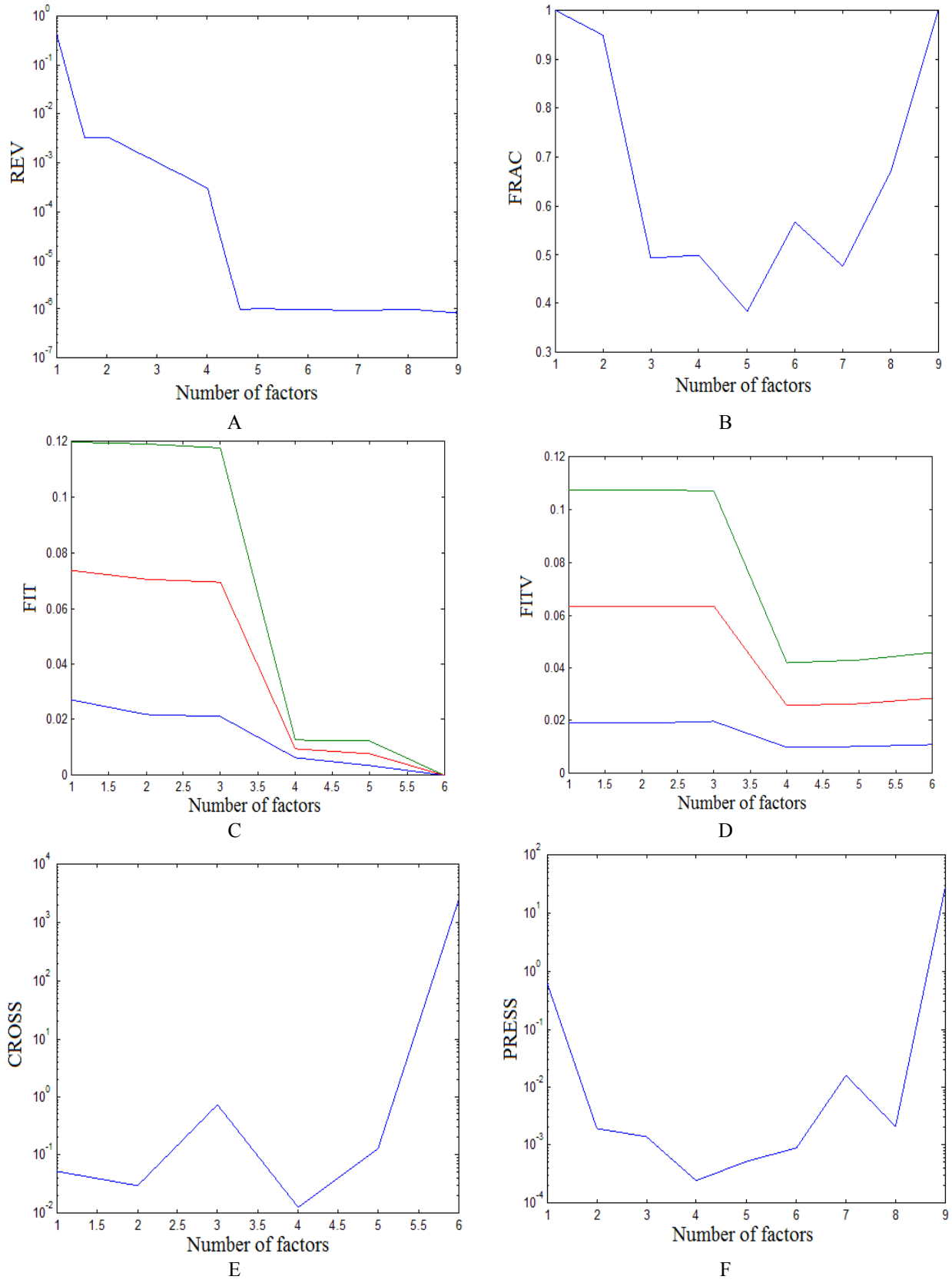

Fig. 3: Plots of (A) REV, (B) FRAC, (C) FIT, (D) FITV, (E) CROSS, (F) PRESS versus number of factors for the PCR model of *o*-, *m*- and *p*-cresol**.**

Table 1: Composition of the training and validation sets of *o*-, *m*- and *p*-cresol

Sample No.	Concentration μg mL^{-1}					
	Training set			Validation set		
	o-cresol	*m*-cresol	*p*-cresol	*o*-cresol	*m*-cresol	*p*-cresol
1	10	10	10	10	10	10
2	10	5	5	10	7.5	7.5
3	5	5	0	7.5	7.5	0
4	5	0	5	7.5	0	7.5
5	0	5	10	0	7.5	10
6	5	10	15	7.5	10	12.5
7	10	15	15	10	12.5	12.5
8	15	15	5	12.5	12.5	7.5
9	15	5	15	12.5	7.5	12.5
10	5	15	10	7.5	12.5	10
11	15	10	0	12.5	10	0
12	10	0	0	10	0	0
13	0	0	15	0	0	12.5
14	0	10	0	0	10	0
15	15	0	10	12.5	0	10
16	0	10	5	0	10	7.5

Table 2: Determination of *o*-, *m*- and *p*-cresol in the validation set mixtures

Sample No.	Percentage recovery					
	o-cresol		*m*-cresol		*p*-cresol	
	PCR	PLS	PCR	PLS	PCR	PLS
1	100.31	100.00	100.77	100.80	100.11	98.88
2	99.98	100.91	99.96	100.29	99.98	99.69
3	100.01	100.27	100.85	101.20	-	-
4	100.00	99.14	-	-	100.18	100.21
5	-	-	100.18	100.71	99.96	99.50
6	99.30	98.94	100.05	100.12	100.13	100.60
7	100.71	101.21	99.99	99.98	100.02	99.99
8	100.50	100.80	99.89	100.61	99.95	100.50
9	100.81	100.90	100.19	100.80	100.33	100.62
10	99.60	99.88	100.63	99.85	99.98	100.11
11	100.25	100.17	100.11	100.20	-	-
12	98.67	99.61	-	-	-	-
13	-	-	-	-	100.41	98.74
14	-	-	99.12	99.99	-	-
15	100.17	99.88	-	-	100.00	100.74
16	-	-	100.00	100.37	100.02	99.96
Mean	100.02	100.14	100.15	100.41	100.09	99.96
SD	0.60	0.72	0.46	0.41	0.15	0.66
RSD %	0.60	0.72	0.46	0.41	0.15	0.66

Phenol is a common contaminant of cresol mixtures (British Pharmacopoeia, 2010; OEHHA, 2000). Therefore, all the training solutions have been prepared in the presence of phenol. It is not necessary to include the concentration of the added phenol in the concentration matrix of the training set. This is because PLS and PCR are factor-based methods. They do not require that the concentration values for all of the existing components have to be provided. They can be used for very complex mixtures since only knowledge of constituents of interest is required (Kenneth *et al*., 1998; Kramer, 1998).

The validation of the developed calibrations was achieved by analyzing an independent set of validation mixtures. The validation set was constructed in a similar manner to the training set and its composition is shown in table 1.

Table 3: Statistical parameters of the validation solutions of *o*-, *m*- and *p*-cresol using PCR and PLS methods

Parameter	*o*-cresol		*m*-cresol		*p*-cresol	
	PCR	PLS	PCR	PLS	PCR	PLS
PRESS	0.0429	0.0587	0.0247	0.0295	0.0049	0.0593
SEP	0.0625	0.0731	0.0474	0.0518	0.0212	0.0734
MSEP	0.0036	0.0049	0.0021	0.0025	0.0004	0.0049
RMSEP	0.0598	0.0699	0.0453	0.0496	0.0203	0.0703
S^2	0.0039	0.005341	0.0022	0.0027	0.0004	0.0054
r	0.99988	0.99971	0.999951	0.999916	0.999987	0.999958
a	-0.017711	-0.090182	0.022842	0.074511	-0.030891	0.012067
Lower 95% [a]	-0.043933	-0.102661	-0.012155	-0.030152	-0.031994	-0.010541
Upper 95% [a]	0.053854	0.096193	0.077849	0.146011	0.025812	0.129485
b	1.011223	1.019711	0.998067	0.991333	1.001633	0.997442
Lower 95% [b]	0.996475	0.997769	0.989222	0.979836	0.996985	0.978518
Upper 95% [b]	1.025925	1.041631	1.006911	1.002831	1.006281	1.016282
S_a [c]	0.003872	0.005767	0.002326	0.003024	0.001222	0.004966
S_b [d]	0.006227	0.009275	0.00374	0.004862	0.001966	0.007985

[a]Lower and upper confidence limits for the intercept at 95% confidence level
[b]Lower and upper confidence limits for the slope at 95% confidence level
[c]standard error of the intercept
[d]standard error of the slope

Table 4: Determination of the three cresol isomers in veterinary preparation (Pharmachemical Pharmonia®) using PCR and PLS methods

	o-cresol		*m*-cresol		*p*-cresol	
	PCR	PLS	PCR	PLS	PCR	PLS
Mean found [a] (μg mL^{-1})	3.91	3.88	7.78	7.73	5.51	5.48
SD	0.06	0.08	0.08	0.11	0.03	0.01
RSD %	1.53	2.06	1.03	1.42	0.54	0.18

[a] Mean of five determinations.

The levels selected in the 4-level partial factorial design were 0, 7.5, 10 and 12.5μg mL^{-1} for the three cresol isomers.

The validation results presented in table 2 showed high accuracy and precision of the developed multivariate methods.

DISCUSSION

Selection of the optimum number of factors

Selecting the number of factors to be used in the calibration is a crucial step in factor-based methods. The factors having analytical information must be kept. Those containing only noise have to be discarded (Kenneth *et al*., 1998; Kramer, 1998).

The Chemometrics Toolbox 3.02 Software provides some indicator functions that can be applied to the determination of the most appropriate rank (number of factors). These include PCR-REV, PCR-FRAC, PCR-FIT, PCR-FITV, PCR-CROSS and PCR-PRESS for the PCR model (fig. 3), and PLS-REV, PLS-CROSS and PLS-PRESS for the PLS model (Kramer, 1995) (fig. 4).

The REV indicator calculates the reduced eigenvalues (REV) according to Malinowski's method (Kramer, 1998). Since eigenvalues are not calculated for PLS, PLS-REV creates its own "pseudo eigenvalues" for each absorbance and concentration factor by evaluating the amount of variance in the data modeled by each factor (Kramer, 1998; Kramer, 1995). Figs. 3a and 4a show that the fourth reduced eigenvalue is significantly higher than those of higher rank. Therefore in this case, the optimum system rank is four according to the REV indicator (Kramer, 1995).

PCR-FRAC is an empirical function based on the eigenvalues and it is related to PCR-REV and predates it (Kramer, 1998; Kramer 1995). According to the FRAC rule, the optimum system rank is one less than the rank where a minimum occurs in the plot (Kramer, 1995) (fig. 3b).

PCR-FIT determines how much error is present when a calibration matrix is used to predict the known concentrations of the training set as a function of the rank (number of factors) used in making the calibration. Fig. 3c shows that the errors drop at rank four and that they go

nearly to zero at the highest rank. This is because the first four factors contain all of the meaningful analytical information. All subsequent factors simply fit the residual noise better and better until all of the data are fit exactly when the complete set of factors is used (Kramer, 1995).

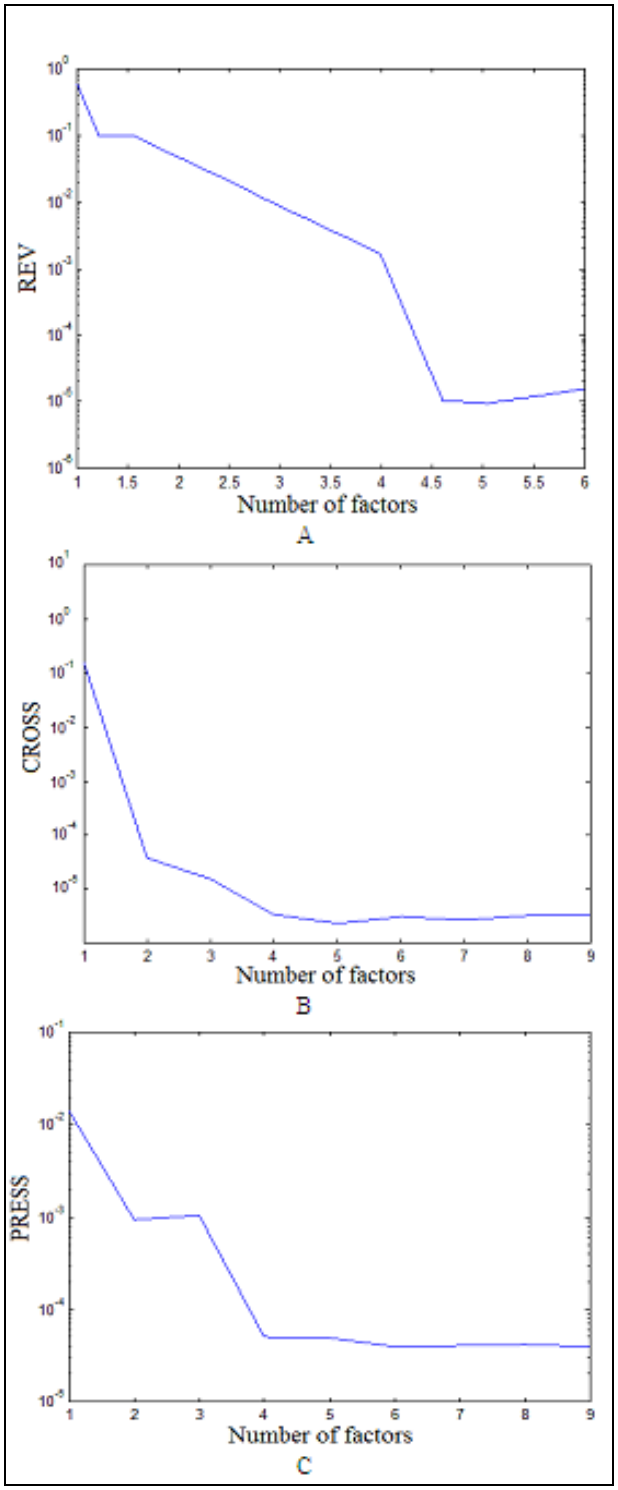

Fig. 4: Plots of (A) REV, (B) CROSS and (C) PRESS versus number of factors for the PLS model of *o*-, *m*- and *p*-cresol.

PCR-FITV works like PCR-FIT except it generates calibrations with the training set and checks the fit to the validation set. It is a more reliable test than PCR-FIT, but it requires validation data (Kramer, 1995).

The CROSS function performs a cross-validation procedure leaving out one sample at a time (Kramer, 1998; Espinosa-Mansila *et al*., 1995). It simulates a validation set by leaving out all possible combinations of one spectrum from the training set. The excluded spectrum is treated as an independent validation sample. The predicted residual error sum-of-squares PRESS is calculated for each developed calibration.

$$PRESS = \sum_{i=1}^{n} (C_i^{True} - C_i^{Predicted})^2$$

Where C_i^{True} is the actual concentration, $C_i^{Predicted}$ is the predicted one and n is the total number of validation mixtures. Then the PRESS for all the calibrations is examined and the one that provides the most appropriate results is chosen. The rank used in that calibration is the optimum one for the system (Kramer, 1998; Kramer, 1995). Figs. 3e and 4b show that for this set of data, errors are minimized when four factors are used.

The PRESS function generates a calibration for every possible rank. Each calibration is employed to estimate the concentrations of validation mixtures. Then for each calibration the PRESS value is determined. The calibration that gives the best results is selected. The rank used in that calibration is the most appropriate one for that system (Kramer, 1998). Figs. 3f and 4c show that for this set of data, errors are minimized when four factors are used.

The studied indicator functions demonstrated that a rank of four factors is the optimum system rank for both the PCR and PLS calibrations. The first three factors are suggested to be due to the investigated isomers as the main factors. The fourth factor is suggested to be due to the interference of phenol, which is the common impurity of cresol.

The constructed PCR and PLS models would span nearly all the data leaving only negligible residuals. The range of residuals not spanned with the four factors was found to be -10×10^{-3} to 8×10^{-3} for both PCR and PLS models.

The mean percentage recoveries (≈100%) and RSD % values (<1%) obtained for the determination of cresol isomers in validation samples using rank 4 indicated that the proposed multivariate models are of good accuracy and precision for the simultaneous determination of cresol isomers in their mixtures (table 2).

Method validation

The predictive capability of the proposed methods could be determined using some validation diagnostics. These

include the predicted residual error sum-of-squares (PRESS), the standard error of prediction (SEP), the mean squared error of prediction (MSEP), the root mean standard error of prediction (RMSEP), and the variance of prediction (s^2) (Kenneth *et al*., 1998; Kramer, 1998).

$$SEP = \left[\sum_{i=1}^{n}\left(C_i^{True} - C_i^{Predicted}\right)^2 / n-1\right]^{1/2}$$

$$MSEP = \sum_{i=1}^{n}\left(C_i^{True} - C_i^{Predicted}\right)^2 / n$$

$$RMSEP = \left[\sum_{i=1}^{n}\left(C_i^{True} - C_i^{Predicted}\right)^2 / n\right]^{1/2}$$

$$s^2 = \sum_{i=1}^{n}\left(C_i^{True} - C_i^{Predicted} - bias\right)^2 / n-1$$

The numerical values of the validation diagnostics for the proposed calibrations were found to be very small (table 3). This indicates the negligible error of prediction and the great predictive capability of the developed methods.

Further validation of the methods was accomplished using the predicted versus true concentration plot (Kenneth *et al*., 1998). Regression analysis for this linear relationship was computed for each multivariate model (table 3). For both the PCR and PLS calibrations, the 95% confidence interval of the intercept included the ideal value of zero and that of the slope included the ideal value of one. This indicates the absence of bias and the good fitness of the developed multivariate models (Miller and Miller, 2000).

Application

Determination of the three cresol isomers in a veterinary preparation

The multivariate models PCR and PLS have been applied to the simultaneous determination of *o*-, *m*- and *p*-cresol isomers in the aqueous veterinary solution Pharmachemical Pharmonia® as a representative example of cresol mixtures used in veterinary medicine. The results are compiled in table 4. The mean total cresol content was found to be ≈95.25% of the label claim. The results show that *o*-, *m*- and *p*-cresol are represented in the mixed cresol used in the preparation as ≈22.72%, 45.23% and 32.05%, respectively.

CONCLUSION

Techniques used to separate and quantify structural isomers are usually sophisticated and may be routinely not available in some quality control laboratories. In the UV region, minor differences between the extensively overlapped spectra of these isomers can lead to solve this problem using multivariate methods PCR and PLS. In this study simultaneous spectrophotometric determination of the three structural isomers of cresol (*o*-, *m*- and *p*-) has been achieved using multivariate regression methods PCR and PLS. These methods have been proved to be accurate and precise and were applied to the determination of cresol isomers in a veterinary solution. They were also time and cost saving as they required only direct dissolution of the samples in the specified solvent before measurement. No prior separation procedures are needed nor sophisticated instruments are required. The developed methods are suggested to be used in routine analysis of cresol isomers in their mixtures and preparations. Taking in consideration the difference in toxicity between the three isomers (Thompson *et al*., 1994), the determination of each isomer concentration in the veterinary preparations or in any other commercial cresol mixtures rather than the total cresol content would be of great importance.

REFERENCES

Bieniek G (1996). Simultaneous determination of phenol, cresol, xylenol isomers and naphthols in urine by capillary gas chromatography. *J. Chromatogr. B: Biomed. Sci. Appl*., **682**: 167-172.

Brereton RG (2003). Chemometrics: Data analysis for the laboratory and chemical plant, Wiley. Chichester, pp.69-76.

British Pharmacoepia (2010). The stationary office, London, pp.601-602.

Cleghom HP, Fellin P, Foster MG and Williams DT (1992). Determination of phenol, cresols and xylenols in workplace air using XAD-2 sorbent cartridges. *J. Toxicol. Environ. Chem*., **34**: 85-98.

Del Olmo M, Díez C, Molina A, De Orbe I and Vílchez JL (1996). Resolution of phenol, o-cresol, m-cresol and p-cresol mixtures by excitation fluorescence using partial least-squares (PLS) multivariate calibration. *Anal. Chim. Acta*., **335**: 23-33.

Espinosa-Mansilla A, Salinas F and Paya De Orbe I (1995). Simultaneous determination of sulfadiazine, doxycycline, furaltadone and trimethoprim by partial least squares multivariate calibration. *Anal. Chim. Acta*., **313**: 103-112.

Kenneth RB, Randy JP and Seasholtz MB (1998). Chemometrics: A Practical Guide, John Wiley and Sons, Inc., New York, pp.201, 204, 243-245, 278-290.

Kramer R (1995). Chemometrics TOOLBOX for use with MATLAB. The Math Works Inc., Natick. pp.1, 12, 5-25, 9, 30.

Kramer R (1998). Chemometric techniques for quantitative analysis. Marcel Dekker. Inc., New York, pp.1, 3, 7, 8, 13-16, 20, 23, 51-54, 68, 71-73, 79-81, 85-89, 99, 102, 104-110, 151.

Masselter SM and Zemann AJ (1995). Influence of buffer electrolyte pH on the migration behavior of phenolic compounds in co-electroosmotic capillary electrophoresis. *J. Chromatogr. A*., **693**: 359-365.

Masselter SM, Zemann AJ and Bobleter O (1993). Separation of cresols using coelectroosmotic capillary electrophoresis. *Electrophoresis*.,**14**: 36-39.

Miller JN and Miller JC (2000). Statistics and Chemometrics for Analytical Chemistry, 4th ed., Pearson Education Limited, England, pp.127-128.

Office of Environmental Health Hazard Assessment (OEHHA) (2000). Air toxics hot spots program risk assessment guidelines. Part III. Technical support document for the determination of noncancer chronic reference exposure levels Batch 2A. Available on-line at http://www.oehha.ca.gov.

Thompson DC, Perera K, Fisher R and Brendel K (1994). Cresol isomers: comparison of toxic potency in rat liver slices. *Toxicol. Appl. Pharmacol.*, **125**: 51-58.

Williams DT, Tran Q, Fellin P and Brice KA (1991). Evaluation of gas chromatography fourier transform infrared spectroscopy mass spectrometry for analysis of phenolic compounds. *J. Chromatogr. A.*, **549**: 297-312.