A Comparative Chemometric Study for Quantitative Determination of Duloxetine Hydrochloride in presence of its Toxic Impurity 1-Naphthol

Basma H. Anwar[1], Nessreen S. Abdelhamid[1] , Maimana A. Magdy[1] , Ibrahim A. Naguib[1]*

1 Pharmaceutical Analytical Chemistry Department, Faculty of Pharmacy, Beni-Suef University, Alshaheed Shehata Ahmad Hegazy St., 62514, Beni-Suef, Egypt

Abstract: Three multivariate chemometric models, called classical least squares (CLS), partial least-squares (PLS) and linear support vector regression (SVR), are developed for the quantitative determination of Duloxetine Hydrochloride (DUL) in presence of its toxic impurity (1-naphthol) in raw materials and pharmaceutical dosage form by using UV spectral data. The three methods are compared among each other and the advantages and disadvantages are discussed. For good results, a two-factor, a four-level experimental design was used, resulting in a training set of 16 mixtures containing different ratios of each component. The test set consisting of nine mixtures was necessary to test the ability of the proposed methods to predict DUL in presence of its impurity 1-naphthol. The results show the success of the three developed methods to determine DUL in presence of small levels of its toxic impurity with good accuracy and selectivity. The results of the dosage form were compared statistically to that of the reported HPLC method, with no significant difference in accuracy and precision, indicating that the suggested calibration models are able and suitable for routine analysis of the drug in bulk and pharmaceutical dosage forms. Compared to the CLS and PLS models, the SVR model gives the best results regarding to the accuracy with a lower prediction error and better generalization ability. However, the CLS and PLS models are found to be simpler and faster in usage and management.

Keywords: Duloxetine hydrochloride;1-naphthol; classical least squares; partial least squares; linear support vector regression; chemometry; spectrophotometry

i. Introduction

Duloxetine hydrochloride (DUL); shown in **Fig.1**, which is chemically identified as (+)-(S)-N-methyl-γ-(1-naphthyloxy)-2-thiophenepropylamine hydrochloride[1] is used for treating depression and anxiety. It also helps to relieve nerve pain in peripheral neuropathy in diabetic patients. In addition, it helps to improve the appetite, mood, energy level, and sleep. Moreover, it relieves nervousness. Duloxetine is a serotonin-norepinephrine reuptake inhibitor. It works by restoring the balance of certain natural substances like serotonin and norepinephrine in the brain[2]. British Pharmacopoeia stated that 1-naphthol (naphthalen-1-ol); shown in **Fig.1**, is a potential impurity of DUL[1]. Frequent exposure to 1-naphthol causes potential hepatic toxicity to humans.1-Naphthol is metabolized to glucuronic acid and sulphate esters by hepatocytes causing bleeding on the surface of the hepatocytes, which together with a dose-dependent decrease in

intracellular glutathione (GSH) speed up the onset of cytotoxicity[3]. 1-Naphthol also has potential toxicity and is harmful to the freshwater fish[4]. Accordingly, determination of DUL in presence of 1-naphthol has high importance due to the potential toxicity of 1-naphthol on human and aquatic life. Different methods were reported for the analysis of DUL, including UV-spectrophotometry [5, 6], HPLC [7-9], TLC [10, 11] and ion selective electrode [12] . There are two HPLC chromatographic methods for detection of DUL in presence of its common impurities including 1-naphthol were reported [13, 14]. However, there are no chemometric analytical methods for the determination of DUL in its dosage form without interference from its toxic impurity 1-naphthol. The presented work aims to determine DUL quantitatively in the presence of its toxic impurity, and to compare between the classical least squares (CLS), partial least-squares (PLS) and linear support vector regression (SVR) chemometric models, finding the advantages and the limitations of each model. The three chemometric models were applied to UV spectrophotometric data to determine DUL simultaneously in the presence of small amounts of 1-naphthol according to the guidelines of the International Conference on Harmonization (ICH) [15]. Hence, chemometric methods could be a suitable tool for data analysis from such cheap and simple spectrophotometric equipment. The chosen models provide high accuracy and precision for the quantitative analysis of DUL in a pharmaceutical dosage form compared to that of the reported HPLC method[9].

ii. Experimental

1. *Instruments*

A double beam UV-visible spectrophotometer (SHIMADZU, Japan) model UV-1601 PC with quartz cell of 1cm pathlength was connected to IBM compatible computer. The software was UVPC personal spectroscopy software version 3.7.

2. *Material and reagents*

a) Pure samples

- Duloxetine Hydrochloride (DUL) was kindly supplied by Eva pharm Company for pharmaceuticals and chemicals, Giza, Egypt. Its purity was found to be 99.91 %.
- 1-naphthol was bought from El Nasr Company for Pharmaceutical Industries, Cairo, Egypt. Its purity was found to be 99.39 %.

b) Pharmaceutical dosage form

CYMBATEX® Capsules are containing delayed release pellets batch No. 605324. Each capsule is labeled to contain 60 mg of DUL. The capsules are manufactured by Eva Pharma company, Egypt.

c) Chemicals and reagents

All chemicals and solvents used in this work were of analytical grade and were used without additional purification. Methanol and Ethanol HPLC grade are acquired from Fischer, UK.

3. *Standard solutions (stock and working)*

a) Stock standard solutions (1 mg/mL)

An amount equal to 0.1 gm of each of DUL and 1-naphthol were accurately weighed into two separate 100-mL volumetric flasks, 50 mL of methanol was added to each flask, shaken to dissolve then the volume was completed with methanol to the mark.

b) *Working standard solutions (100 µg/mL)*

A volume equal to 10 mL of each of DUL and 1-naphthol stock standard solutions were accurately transferred into two separate 100-mL volumetric flasks, then the volume was completed with ethanol to the mark, which is one of the green solvents to the environment.

iii. Procedure

1. *Linearity*

UV spectra for different aliquots of DUL ranging between (1 -15) µg/mL were recorded from 221 to 338 nm. DUL exhibited linearity between (2-14) µg/mL at its λ_{max} at 229 nm. The overlapped spectra of 6 µg/mL DUL and 3.8 µg/mL 1-naphthol are shown in Figure 2.

2. *Experimental Design*

a) Calibration set:

A multilevel, multifactor calibration design composed of four levels and two factors was designed using four concentration levels coded as +2, +1, −1, +1, where −1 is the central level for each of the components to be analyzed, including the drug DUL and its impurity 1-naphthol. The design purpose is to distribute the mixture space well; which means that there are four mixtures for each compound at each concentration level, resulting in 16 mixtures for the training set [16]. The central level of the design was 10 µg/mL for DUL and 0.2 µg/mL for 1-naphthol. Selection of the concentration for the levels of DUL was based on its calibration ranges, while the concentration levels of the impurity 1-naphthol were calculated according to the ICH requirements for impurity levels. Because minimal levels are allowed, we involved the impurity in levels from 1 to 5% of the drug based on a molar basis to cover a wide range of possibilities. **Table.1** represents the concentration design matrix.

b) Test set:

To test the validity and sensitivity of the proposed multivariate chemometric models, independent test set mixtures were obtained by repeating the preparation of four mixtures of the training set (1, 5, 9, and 11) and the preparation of four independent mixtures other than the training set mixtures but within the concentration space of the design, as shown in **Table.1**. The two-dimensional scores plot for the first two principal components of the concentration matrix was obtained to confirm the orthogonality, symmetry, and rotatability of the training set mixtures (presented as circles), as shown in **Fig.3**. Mean centering of the data seems to be the best preliminary processing to obtain the best results.

c) Analysis of Cymbatex capsules:

The contents of 10 Cymbatex capsules were weighed, emptied, finely powdered and mixed well. An amount equivalent to 60 mg DUL was accurately weighed and transferred into a 100 mL volumetric flask to which 75 mL methanol was added. The prepared solution was sonicated for 30 min, cooled, and

the volume was completed to the mark with methanol to obtain 1 mg/mL stock solution. The stock solution was filtered and diluted with ethanol to obtain 100 µg/mL working solution. Lastly, 1.2 mL of DUL working solution were diluted to 10 mL with ethanol. The average of three corresponding spectra was recorded. This experiment was repeated three times, and the obtained spectra were analyzed by the suggested chemometric models.

d) Software:

Codes for CLS and PLS (PLS1 algorithm) [17], bootstrap, and grid search were written using MATLAB 7.5.0 (R2007b). The codes for the SVR algorithm were downloaded from the Internet Web site http://onlinesvr.altervista.org/. All calculations were done using an Intel® Core™ 2 Duo CPU with 2.20 GHz and 3.00 GB of RAM under Microsoft Windows 7.

e) Chemometric Methods:

Multivariate calibration methods purpose to establish a relationship between the spectral data and the concentrations of the proposed compounds. Various designs have been used in this study such as CLS, PJS and linear SVR which were mentioned in many previous works [17, 18]

Classical least squares (CLS)

As mentioned by by Haaland and Thomas[19]. The classical least-squares (CLS) model is typically written as

$$\boldsymbol{X} = \boldsymbol{CK} + \boldsymbol{E_x} \qquad \text{..... eq(1)}$$

Where $\boldsymbol{X}$ is an I×J matrix of the UV spectra for the J variables (wavelengths) and I samples. $\boldsymbol{C}$ is a matrix of dimensions I×M of the concentration values for the M components. K is the M × J matrix of the pure component spectra, and E_x is the matrix of the residual error.

Partial least squares regression (PLSR)

Mathematically in PLSR, the predictor matrix (X) and the response vector (c) are decomposed using a given number of PLS components (latent variables) [19-22], according to equations:

$$\mathbf{X = T + P.\ E} \qquad \text{..... eq(2)}$$

$$\mathbf{c = T + q.\ f} \qquad \text{..... eq(3)}$$

T and P are the scores and loadings for X, q is the loading vector for c, and E and f are the residuals for X and c, respectively.

Optimization of the number of latent variables for the PLS model

Cross-validation method using leave one out (LOO-CV) was used to select the optimum number of PLS components which was the one that gives the lowest value of RMSECV. The basis of the CV method was discovered in details by Haaland and Thomas [19] and it was used in previous works [20]..

Support vector regression (SVR)

Consider a data set X (I×J) with an output vector c. The objective is to find a multivariate regression function f(x) based on X to predict a desired output property (e.g., the concentration of a chemical compound) from a sample (e.g., a spectrum). The complete SVR equations are fully derived in[21, 22] and the summary equation is given by

$$f(x) = \sum_{ij=1}^{N}(\alpha_i - \alpha_i^*)\ (\emptyset\ (x_i).\emptyset\ (x_j)) + b \qquad \ldots\ldots\ \text{eq(4)}$$

where α_i and α_i^* are the Lagrange multipliers satisfying the constraint $0 < \alpha_i, \alpha_i^* <$ C.C is an additional parameter called the penalty error or regularisation constant which determines the trade-off between the training error and model simplicity.

Optimization of the number of linear SVR model parameters

ε-insensitive loss function was applied and used this study to optimize the SVR model. The basics of this function was explained in details by Gunn and Parrella[23, 24]. The primary range of values was (0.01–1) for ε and (30–1000) for C.

iv. Results and discussion

Parameters' Optimization Results

1) For the PLS model: LOO-CV method was used to find of the optimum number of PLS components to perform the calibration model on the training set. The results show that the optimum number of PLS components was 2 for DUL, as shown in **Fig.4**.
2) For SVR model: the results of the grid search that gave the lowest RMSECV **(Eq.5)** were ε = 0.15 and C = 420 for DUL.

Data Analysis Results

The presented work aimed firstly to determine DUL quantitatively in the presence of its toxic impurity 1-naphthol using three popular chemometric methods named CLS, PLS and linear SVR. The three multivariate models were able to use the UV data and overcome the overlapping spectra of the components which was shown in **Fig.2**. Secondly, this work includes a comparative study of the three chemometric methods through the analysis of DUL in the presence of its toxic impurity.

All the CLS, PLS and linear SVR models were successfully able to detect the concentrations of DUL in the training set and the test set, indicated by a high recovery percentage with low SD, as presented in **Table.2**. The RMSEP is a parameter used to assess the predictive abilities of the three models (**Table.2**). RMSEP

comparative plots between CLS, PLS and linear SVR for the prediction of test set samples are shown in Fig.5. The comparison proves the generalization characters and model transferability of the presented chemometric models, especially for linear SVR.

Various comparison parameters are included in our study, such as root mean square error of calibration (RMSEC), RMSEP, calculation and computational procedures, and optimization steps. First, for the RMSEC, which provides the auto-predictive error value, it was noticed that linear SVR shows the lowest RMSEC compared to CLS and PLS, indicating the highest accuracy of SVR followed by PLS then CLS.

In addition, the corresponding SD of the PLS model is smaller than that of CLS and SVR, indicating the highest precision of PLS followed by SVR then CLS. Moreover, the comparative bar plot in Fig.5 shows that compared to CLS and PLS, linear SVR gives the lowest RMSEP, reflecting the highest capability to handle future samples and the highest generalization capability. Furthermore, concerning the calculation and computational procedures, CLS and PLS are simpler than SVR, because the latter needs additional steps in calculation and more time is consumed for optimization.

In this study, 4 fold CV was used for optimizing the SVR parameters to avoid overfitting by predicting small subsets of data (as used in the LOO-CV technique), so that the robustness and its generalization ability of the model would be increased. Thus performing SVR is better and preferable than with CLS and PLS [25]. The linear SVR model in this study was used rather than the nonlinear SVR model as the later require optimization of large number of parameters then that of the linear SVR, so this makes linear SVR able to save time and human resources.

From the above discussion and the literature, the three models shows a better chance to save time, money and use fewer equipment if compared with the reported HPLC method [9]. They also give satisfied results with good prediction ability.

Application of the Proposed Methods to the Pharmaceutical Formulation

The suggested chemometric methods were applied for analysis of DUL in Cymbatex capsules, and satisfactory results with good recoveries were obtained. These results were statistically compared to the results obtained from applying the reported HPLC method [9] using t- and F-tests. The values obtained are less than the theoretical ones, indicating no significant difference between the two proposed methods and the reference HPLC method with respect to accuracy and precision (**Table.3**).

v. *Conclusion*

In general, the presented work is proposed to determine DUL quantitatively in presence of its toxic impurity 1-naphthol. Moreover, it aimed to provide a comparative study among CLS, PLS and linear SVR chemometric models, finding the advantages and limitations of each model. The results obtained are an

encouragement for performing smart chemometric approaches, especially linear SVR, for the quantitative determination of different pharmaceutical products with cheap and simple equipment like the UV spectrophotometer, even though if the number of interfering components is large and the resulted spectra are overlapped with many interfering species. However, CLS and PLS are much simpler and faster and would be more suitable for routine analysis of such simple mixtures. Moreover, the three models can save time, money and equipment rather than the reported HPLC methods with good results and detection ability.

vi. *Refferences*

1. *The British pharmacopeia.* Her Majesty's Offeice, Londin, 2013.
2. *Martindale-The complete drug reference (thirty-first-edition).* pharmaceutical press, 2012.
3. Mary D'arcy Doherty, G.M.C.a.M.T.S., *Mechanisms of toxic injury to isolated hepatocytes by l-naphthol.* Biochemical Pharmacology, 1984. **33**(4): p. 543-549.
4. K. S. Tilak, D.M.R., A. Priyamvada Devi and A. S. Murty, *Toxicity of carbaryl and 1-naphthol to four species of freshwater fish.* Journal of biosciences, 1981. **3**(4): p. 457–461.
5. Mohammad Yunoos, D.G.S., B.pragati Kumar, Shahul Hameed and Azmath Hussain, *Simple UV Spectrophotometric Determination of Duloxetine Hydrochloride in Bulk and in Pharmaceutical Formulations.* E-Journal of Chemistry, 2009. **7**(3): p. 785-788.
6. M. M. Kamila, N.M., L. K. Ghosh, *A validated UV spectrophotometric method for determination of duloxetine hydrochloride.* Pharmazie, 2006. **62**(2007): p. 414–415.
7. Usmangani K. Chhalotiya, K.K.B., Dimal A. Shah, Sunil L. Baldania, *Development and Validation of a Stability-Indicating RP-HPLC Method for Duloxetine Hydrochloride in its Bulk and Tablet Dosage Form.* scientia pharmaceutica 2010. **2010**(78): p. 857–868.
8. V.R. Sinha, A., R. Kumria, and J.R. Bhinge, *Stress Degradation Studies on Duloxetine Hydrochloride and Development of an RP-HPLCMethod for its Determination in Capsule Formulation.* Journal of Chromatographic Science, 2009. **47**.
9. Chusena Narasimharaju Bhimanadhuni, D.R.G.K., Andhra Pradesh, Chintha Srinivas, *Development and validation of RP-HPLC method for determination of Duloxetine hydrochloride in bulk and dosage form.* International Current Pharmaceutical Journal, 2012. **1**(5): p. 98-102.
10. Sheikh Shahnawaz, A.S., Mir Tariq Masroor and Vandana Arora, *Stability-Indicating HPTLCMethod for Determination of DuloxetineHydrochloride in Bulk Drug and Tablet Formulation.* Chromatography Research International, 2011. **2011**.
11. Suneela S. Dhaneshwar, P.D., M. Patil, G. Vadnerkar, and S. R. Dhaneshwar, *Development and validation of a HPTLC method for Estimation of Duloxetine Hydrochloride in Bulk Drug and in Tablet Dosage Form.* Indian Journal Of Pharmaceutical Sciences, 2008. **70**(2): p. 233–236.
12. Reda A. Ammar, H.O., Abdulrhman Al-Warthan, *Quantitative Determination of Duloxetine Hydrochloride in Pharmaceuticals and Urine Using Prepared Ion Selective Membrane Electrode.* International Journal of ELECTROCHEMICAL SCIENCE, 2012. **7** (2012): p. 4369 - 4380.
13. N.V.V.S.S. Raman, K.A.H., A.V.S.S. Prasad, K. Ratnakar Reddy, K. Ramakrishna, *Determination of duloxetine hydrochloride in the presence of process and degradation impurities by a validated stability-indicating RP-LC method.* Journal of Pharmaceutical and Biomedical Analysis, 2009. **51**(2010): p. 994–997.
14. Navneet Kumar, D.S., Bhupendrasinh Vaghela, P.Sunil Reddy & Anuj Raghav, *Development and validation of a stability indicating rp -lc method for the estimation of process related impurities and degradation products of duloxetine hydrochloride and phthalic acid in duloxetine hydrochloride delayed release capsules.* Journal of Liquid Chromatography & Related Technologies, 2012. **35**(7): p. 867-881.
15. *ICH, Q2 (R1) Validation of Analytical Procedures: Text and Methodology.* ICH Harmonized Tripartite Guideline,, 2005.
16. Brereton, R.G., *Multilevel Multifactor Designs for Multivariate Calibration.* The Analyst, 1997. **122**: p. 1521–1529.
17. Gasteiger, J., *Handbook of Chemometrics.* WILEVECH verlag GmbH & Co., Weinheim, Germany, 2003. **3**.
18. S. WOLD, A.R., H. WOLD AND W. J. DUNN, *The Collinearity Problem In Linear Regression.The Partial Least Squares (Pls) Approach To Generalized Inverses.* Society for Industrial and Applied Mathematics, 1984. **5**(3): p. 735-743.
19. Haaland, D.M., & Thomas, *E. V.* Anal. Chem., 1988. **60**: p. 1193-1202.
20. Brereton, R.G., John Wiley & Sons, *Chemometrics: Data Analysis for the Laboratory and Chemical Plant.* Chichester, United Kingdom, 2003.
21. Shawe-Taylor, N.C.a.J., *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods.* Cambridge University Press,Cambridge, UK, 2000.
22. J.A.K. Suykens, T.V.G.a.J.D.B., *Least Squares Support Vector Machines.* J. World Scientific, , Singapore, 1999.
23. Gunn, S.R., *Support Vector Machines for Classification and Regression,*

Technical Report. University of Southampton, UK, 1998.

24. Parrella, F., *Online Support Vector Regression.* Thesis in Information Science, 2007.

25. Naguib, I.A., *Stability indicating analysis of bisacodyl by partial least squares regression, spectral residual augmented classical least squares and support vector regression chemometric models: A comparative study.* Bulletin of Faculty of Pharmacy, Cairo University, 2011. **49**: p. 91-100.

Table.1: the 3 level 2 factor experimental design of 16 training set mixtures with 9 test set mixtures shown as concentrations of the mixture components in µg/mL

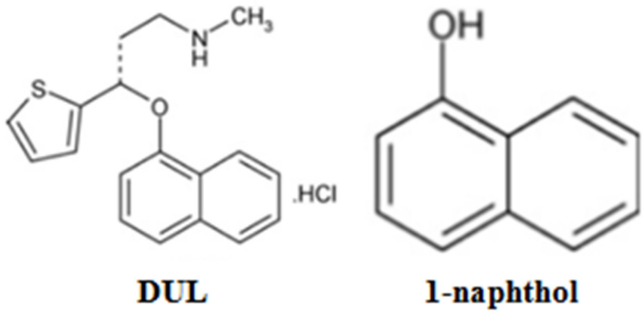
Training set		Test set	
DUL	1-naphthol	DUL	1-naphthol
10	0.2	10	0.2
10	0.1	10	0.1
9	0.1	12	0.1
9	0.4	13	0.1
12	0.1	13	0.1
9	0.2	11	0.1
10	0.5	12	0.3
13	0.5	13	0.3
13	0.1	11	0.2
9	0.5	-	-
13	0.2	-	-
10	0.4	-	-
12	0.4	-	-
12	0.5	-	-
13	0.4	-	-
12	0.2	-	-

Table.2: Analysis results for the prediction of training set (auto-prediction) and independent test set of DUL in presence of its impurity 1-naphthol by CLS, PLS and linear SVR chemometric methods

Taken (µg/mL)	Training set						Taken (µg/mL)	Test set					
	CLS		PLS		Linear SVR			CLS		PLS		Linear SVR	
	Found (µg/mL)	% R	Found (µg/mL)	% R	Found (µg/mL)	% R		Found (µg/mL)	% R	Found (µg/mL)	% R	Found (µg/mL)	% R
10	10.27	102.69	10.18	101.81	10.17	101.67	**10**	10.30	102.96	10.32	103.23	10.32	103.19
10	9.93	99.30	10.03	100.34	10.01	100.09	**10**	10.02	100.23	10.11	101.07	10.10	101.04
9	8.82	98.00	8.86	98.42	8.85	98.33	**12**	12.09	100.73	12.08	100.63	12.05	100.41
9	8.97	99.65	9.16	101.74	9.15	101.67	**13**	13.08	100.58	12.97	99.76	12.91	99.30
12	12.03	100.26	11.99	99.91	11.94	99.53	**13**	13.10	100.81	13.51	103.89	13.41	103.18
9	8.78	97.57	8.85	98.39	8.85	98.33	**11**	10.97	99.75	10.92	99.30	10.87	98.83
10	10.07	100.66	10.03	100.34	10.08	100.82	**12**	12.11	100.88	12.06	100.54	12.04	100.35
13	13.08	100.59	13.00	99.97	13.01	100.09	**13**	13.12	100.94	13.09	100.66	13.05	100.41
13	13.23	101.78	13.15	101.16	13.10	100.79	**11**	10.95	99.51	10.32	103.23	10.32	103.19
9	8.91	99.03	8.96	99.61	9.01	100.07	**Mean (%)**		100.86		101.14		100.84
13	12.94	99.54	12.90	99.19	12.86	98.91	**SD**		0.94		1.61		1.61
10	10.05	100.52	10.12	101.22	10.15	101.45	**RMSEP**		0.1372		0.1215		0.1176
12	12.05	100.39	11.96	99.64	11.96	99.65							
12	11.96	99.64	11.91	99.22	11.91	99.26							
13	13.02	100.16	13.02	100.18	13.01	100.08							
12	11.80	98.31	11.88	98.98	11.85	98.75							
Mean (%)		99.88		100.01		99.97							
SD		1.32		1.07		1.10							
RMSEC		0.13716		0.1215		0.10892							

Table.3: Statistical analysis of the three proposed CLS, PLS and linear SVR chemometric methods and the reported HPLC method for determination of DUL in pharmaceutical formulation

Parameters	CLS	PLS	Linear SVR	Reported HPLC method[9]
Mean (µg)	108.79	109.63	109.79	109.17
SD	0.57	0.69	0.76	0.42
Variance	0.327	0.482	0.585	0.175
N	6	6	6	6
Student's t-test (2.228)[a]	1.323	1.378	1.751	-
F-test (5.050)[a]	1.868	2.750	3.338	-



DUL 1-naphthol

Figure.1: chemical structure of DUL and its impurity 1-naphthol

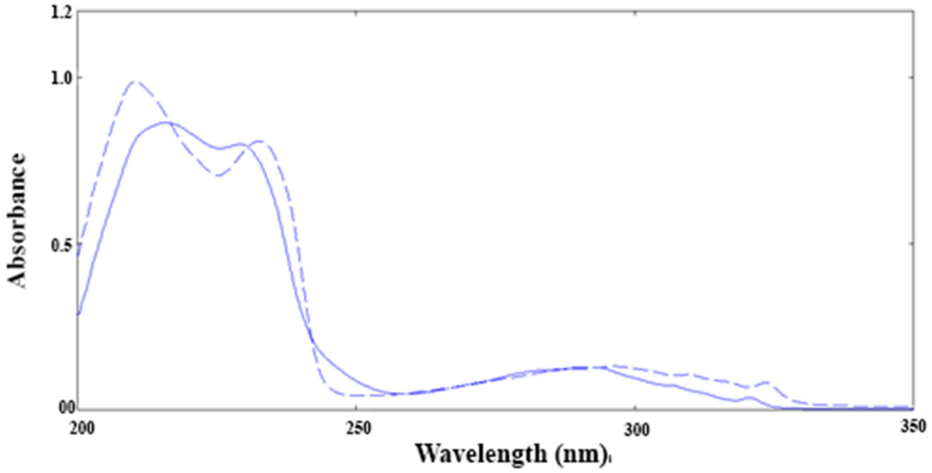


Figure.2: UV spectra of 6 µg/mL of DUL (—) and 3 µg/mL of 1-naphthol (- - -) using ethanol as a solvent

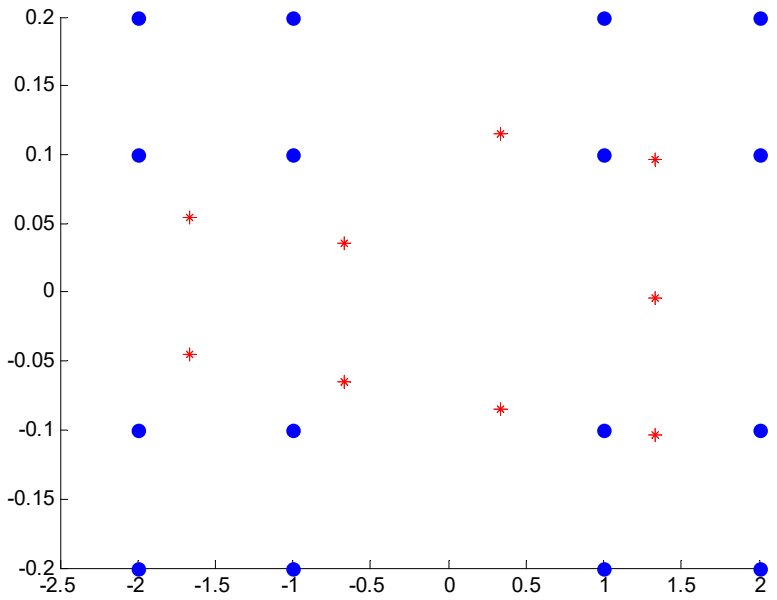


Figure.3: two-dimensional scores plot for the mean centered 16 training set samples (circles) and 8 test set samples (stars) of concentration matrices of the 3-level-2-factor experimental design.

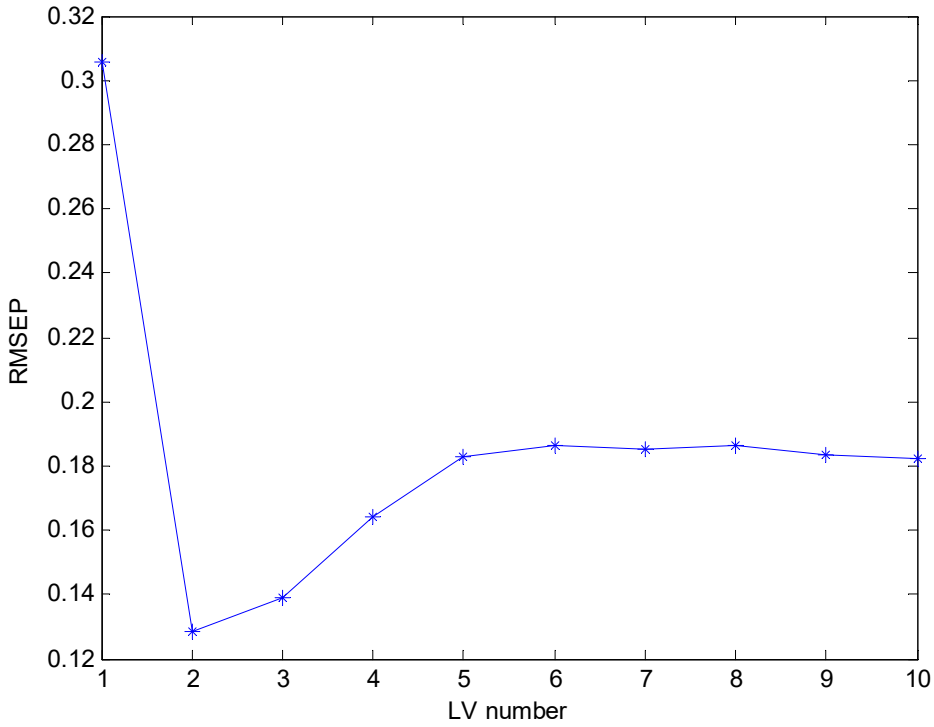


Figure.4: selection of the optimum number of PLS components (latent variables(LVs)) via plotting the number of PLS components versus the corresponding root mean square error of prediction (RMSEP) by using bootstrap technique

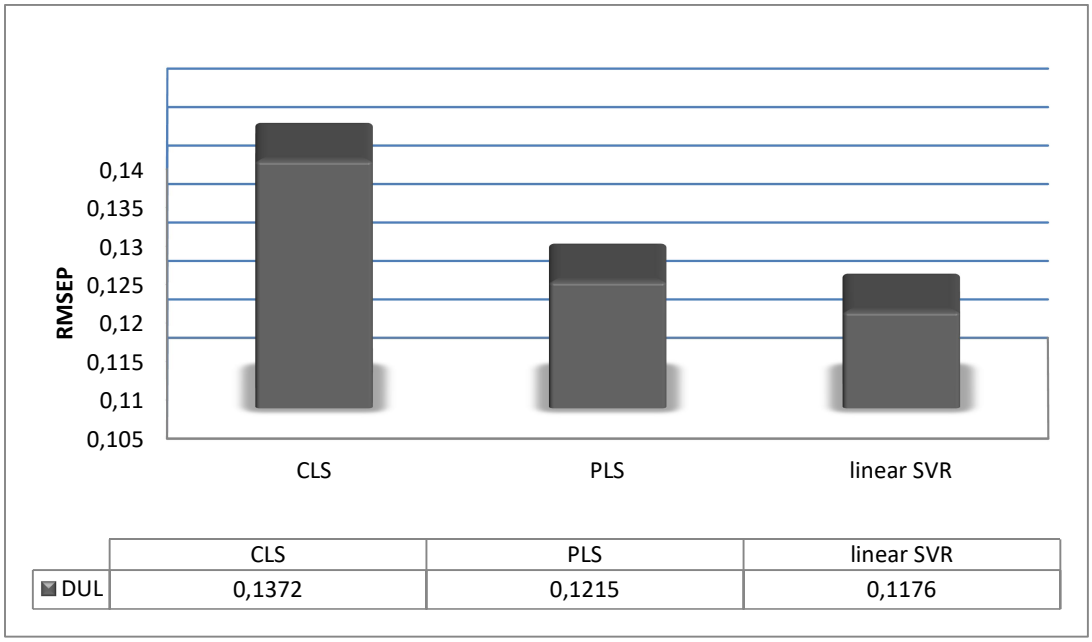


Figure.5: RMSEP plots for the prediction of the independent test set samples for DUL using CLS, PLS and linear SVR models.