



*Anal. Bioanal. Chem. Res., Vol. 3, No. 1, 65-72, June 2016.*

## Determination of $^{137}\text{Ba}$ Isotope Abundances in Water Samples by Inductively Coupled Plasma-optical Emission Spectrometry Combined with Least-squares Support Vector Machine Regression

E. Zolfonoun<sup>a,\*</sup>, S.M.R. Pakzad<sup>a</sup> and M. Salahinejad<sup>b</sup>

<sup>a</sup>NFCRS, Nuclear Science and Technology Research Institute, Tehran, Iran

<sup>b</sup>Environmental Laboratory, NSTRI, Tehran, Iran

(Received 15 November 2015, Accepted 7 April 2016)

A simple and rapid method for the determination of  $^{137}\text{Ba}$  isotope abundances in water samples by inductively coupled plasma-optical emission spectrometry (ICP-OES) coupled with least-squares support vector machine regression (LS-SVM) is reported. By evaluation of emission lines of barium, it was found that the emission line at 493.408 nm provides the best results for the determination of  $^{137}\text{Ba}$  abundances. After recording the emission spectra in the range of 493.362-493.467 nm, quantification of  $^{137}\text{Ba}$  abundances was performed with the aid of LS-SVM algorithm. The obtained results revealed that using LS-SVM as a nonlinear modeling approach improves the predictive quality of the developed models compared with partial least squares (PLS) method. The calculated results proved that the combination of ICP-OES and LS-SVM is a suitable and low cost technique for the determination of  $^{137}\text{Ba}$  abundances. Performance of the proposed method was examined through measuring  $^{137}\text{Ba}$  abundances in synthetic mixtures and water samples.

**Keywords:**  $^{137}\text{Ba}$ , ICP-OES, LS-SVM, PLS

### INTRODUCTION

Barium is a silvery-white metal which exists in nature only in ores containing mixtures of elements. It combines with other chemicals such as sulfur or carbon and oxygen to form barium compounds. Barium compounds are used by the oil and gas industries to make drilling muds. They are also used to make paint, bricks, ceramics, glass, and rubber [1]. Naturally occurring barium consists of six stable isotopes,  $^{138}\text{Ba}$  (71.7%),  $^{137}\text{Ba}$  (11.2%),  $^{136}\text{Ba}$  (7.9%),  $^{135}\text{Ba}$  (6.6%),  $^{134}\text{Ba}$  (2.4%) and  $^{132}\text{Ba}$  (0.1%). The determination of the isotopic composition of a particular element is a very useful technique to provide information about the geographic, chemical, and biological origins of substances [2,3]. Since the isotopes of a particular element have identical chemical properties despite slightly different masses, they may enter into chemical reactions at different rates which can induce local variations in isotopic

composition. The measurement of these variations can reveal information about the origin and/or the processes that have affected the material analyzed. Stable isotope data can be relevant to a range of research areas in geology, hydrology, food authentication, forensics, medicine and environmental sciences [4,5].

The conventional methods for determination of isotopic composition of elements are thermal ionization mass spectrometry (TIMS) and inductively coupled plasma mass spectrometry (ICP-MS). TIMS is a magnetic sector mass spectrometry technique that is capable of making very precise measurements of isotope ratios of elements that can be ionized thermally, usually by passing a current through a thin metal ribbon under vacuum. However, time-consuming sample preparation, analyte purification, and sample determination processes may limit the routine application of TIMS for isotope ratio measurements [6]. ICP-MS is another method for the determination of isotopic composition of elements that uses an argon inductively coupled plasma (ICP) as an ionization device interfaced to a

\*Corresponding author. E-mail: e.zolfonoun@gmail.com

low-resolution quadrupole mass analyzer [7-9]. Although ICP-MS is an attractive technique for determination of isotope ratios, due to its rapidity and selectivity, high instrumental and operational costs make it difficult to facilitate this method in routine analysis.

The isotopic composition of some elements can be determined by spectroscopic techniques, such as atomic emission or atomic absorption [10,11]. These determinations are based on the isotopic shifts of different isotopes of a given element in the emission or absorption lines. The determination of isotope ratios by atomic absorption or atomic emission measurements have been reported for Li [12], Pb [13], U [14,15], B [16,17] and Hg [18].

Inductively coupled plasma-optical emission spectrometry (ICP-OES) is one of the most powerful and popular analytical tools for the determination of trace elements in a variety of different sample matrices. Compared with the mass spectrometric methods, ICP-OES is a simple, fast and less expensive method. However, due to the small isotopic shifts, the emission spectra of isotopes cannot be resolved by conventional ICP-OES spectrometers. This prevents the quantitative determination of the isotopes by direct emission measurements [16].

Under computer-controlled instrumentation, multivariate chemometrics methods play a very important role in the multicomponent analysis of the mixtures by spectroscopic techniques [19-21]. The application of quantitative chemometrics to multivariate chemical data is becoming more widespread owing to the availability of digitized spectroscopic data and commercial software for laboratory computers.

Recently, multivariate calibration methods have been employed to determine the isotopic composition of elements by various atomic spectrometry techniques such as laser-induced breakdown spectroscopy and inductively coupled plasma-optical emission spectrometry. Zolfonoun *et al.* proposed the application of ICP-OES for the determination of boron isotopic ratio using multivariate curve resolution with alternating least squares (MCR-ALS) algorithm [16]. Chan *et al.* used PLS calibration for the determination of  $^{235}\text{U}$  and  $^{238}\text{U}$  using laser induced breakdown spectroscopy (LIBS) [22]. Zolfonoun utilized ICP-OES combined with derivative spectroscopy, MCR-ALS and PLS methods, for

the determination of  $^6\text{Li}$  and  $^7\text{Li}$  abundances [23]. Also, Khayatzadeh Mahani *et al.* applied PLS calibration for the quantification of  $^{235}\text{U}$  abundance by using ICP-OES technique [10].

Support vector machine (SVM), introduced by Vapnik [24], is a valuable tool for solving pattern recognition and classification problems. The SVM algorithm can be applied to regression problems by the introduction of an alternative loss function. Due to its advantages and remarkable generalization performance over other methods, SVM has attracted attention and gained extensive application (Vapnik 1998) [25]. Least-square-support vector machine (LS-SVM) is an alternate formulation of SVM described by Suykens and Vandewalle [26]. The LS-SVM can resolve linear and non-linear multivariate calibration problems in a relatively fast manner [27]. Recently, multicomponent determinations based on the LS-SVM application in spectrophotometric and voltammetric data have been published [28-30].

The aim of this work is to develop a rapid and inexpensive method for the quantitative determination of  $^{137}\text{Ba}$  abundances by ICP-OES. Since conventional ICP-OES spectrometers are not able to directly determine the isotopic composition of elements due to the limitation in spectral resolution, LS-SVM algorithm as an efficient multivariate calibration method was employed for the quantification of  $^{137}\text{Ba}$  abundances.

## THEORY

### Theory of Least-square Support Vector Regression

In principle, LS-SVM always fits a linear relation ( $y = wx + b$ ) between the independent variable or regressor ( $x$ ) and the dependent variable ( $y$ ). The best relation is the one that minimizes the cost function ( $Q$ ) containing a penalized regression ( $e_i$ ) error term:

$$Q = \frac{1}{2}W^T W + \frac{1}{2}\gamma \sum_{i=1}^n e_i^2$$

subject to:

$$y_i = W^T \phi x_i + b + e_i \quad i = 1, \dots, n$$

where  $\phi$  denotes the feature mapping parameter. The first

part of this cost function is a weight decay which is used to regularize weight sizes and penalize large weights. Due to this regularization, the weights converge towards similar values. Large weights deteriorate the generalization ability of the LS-SVM method because they can cause excessive variance. The second part of cost function is the regression error for all training data. The relative weight of this part as compared to the first part indicated by the parameter  $\gamma$ , which has to be optimized by the user. Similar to other multivariate statistical models, performance of the LS-SVMs for quantitative studies depends on the combination of several parameters. The attainment of the kernel function is cumbersome and depends on each case. The radial basis function (RBF),  $\exp(-(\|x_i - x_j\|^2)/2\sigma^2)$ , is by far the most popular choice of the kernel types because of its localized and finite responses across the entire range of the real  $x$ -axis. In this kernel function,  $\sigma^2$  is the width of the Gaussian function which should be optimized by the user, to obtain the support vectors. For  $\sigma$  of the RBF kernel it should be stressed that it is very important to perform a careful model selection of the tuning parameters, in combination with the regularization constant  $\gamma$ , in order to achieve a good generalization model.

### Partial Least Squares

Partial least squares regression (PLSR) has become the most frequently used method for the multivariate calibration because high-performance calibration models are obtained, while the software is not only available but also easily implemented. PLS is a linear modeling technique where information in the matrix of predictors  $D$  is projected onto a small number of underlying ("latent") variables called PLS components, referred to as latent variables. The matrix of responses  $Y$  is simultaneously used in estimating the "latent" variables in  $D$  that will be the most relevant for predicting the  $Y$  variables. The number of significant factors for the PLS algorithm was determined using the cross-validation method. With cross-validation, one sample was kept out (leave one out) of the calibration and used for prediction. The process was repeated so that each of the samples was kept out once.

### Statistical Parameters

The quality of each model was assessed by applying the

$k$ -fold cross-validation procedure. In  $k$ -fold cross-validation, the data is first partitioned into  $k$  equally (or nearly equally) sized segments or folds. The regression model will then be trained and tested  $k$  times. Each time the model is built using  $(k-1)$  folds as the training sample and the remaining single fold is retained for testing.

For the evaluation of the performance of multivariate calibration models, the root mean square error (RMSE) can be used:

$$RMSE = \left[ \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]^{0.5}$$

The square of the correlation coefficient ( $R^2$ ), indicating the fitting quality of all the data to a straight line, is calculated for the checking of each calibration, and is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

where  $\hat{y}_i$  is the estimated value of the  $i_{th}$  object and  $y_i$  is the corresponding reference value of this object,  $\bar{y}$  is the mean of reference values and  $n$  is the total number of objects in the corresponding set.

## EXPERIMENTAL

### Reagents

All reagents used were of analytical grade and were used as supplied. <sup>137</sup>Ba and naturally abundant barium standard solutions (100  $\mu\text{g ml}^{-1}$  as Ba) were prepared by dissolving appropriate amounts of barium-<sup>137</sup>Ba carbonate (91.7% <sup>137</sup>Ba) (ISOFLEX) and barium carbonate (Sigma-Aldrich) in 1 M hydrochloric acid, respectively. Working solutions were prepared by adequate dilution with ultrapure water.

### Instrumentation and Software

All the measurements were carried out using a Perkin Elmer (Optima 7300 DV) simultaneous inductively coupled plasma-optical emission spectrometer coupled to a concentric nebulizer and equipped with a charge coupled device (CCD) detector. Operational conditions were

**Table 1.** Operating Parameters for the ICP-OES

Plasma gas	Argon
Plasma gas flow rate	15 l min <sup>-1</sup>
Auxiliary gas flow rate	0.5 l min <sup>-1</sup>
Frequency of RF generator	40 MHz
RF generator power	1.2 kW
Nebulizer gas flow rate	0.8 l min <sup>-1</sup>
Sample flow rate	0.8 ml min <sup>-1</sup>
Integration time	3 s
Viewing mode	Radial
Viewing height	15 mm
Spectral resolution	15 pm

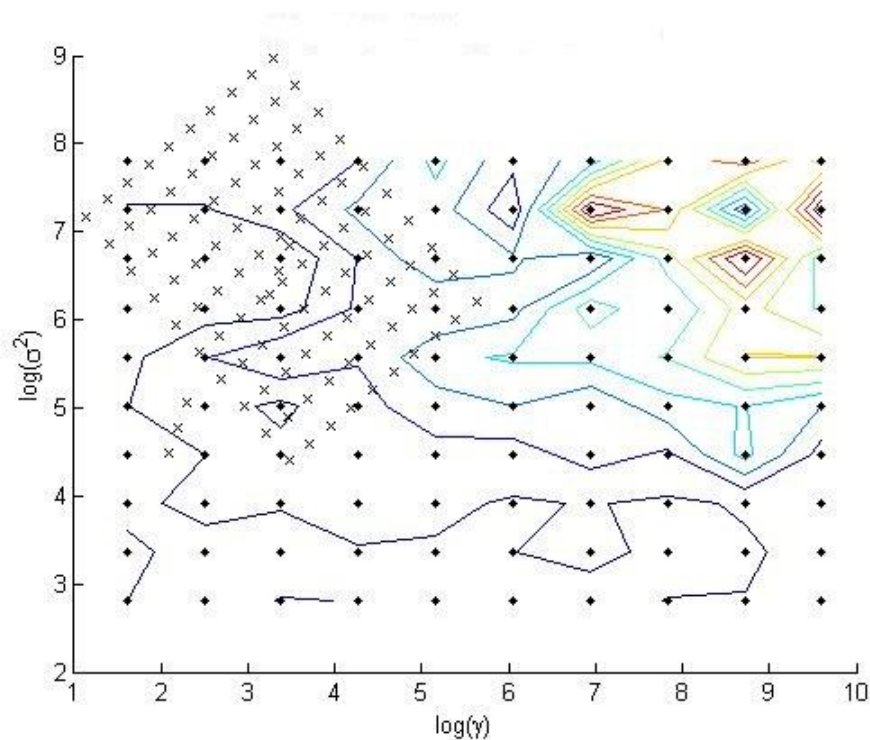
optimized and summarized in Table 1. The data were processed using MATLAB software version 7.7. The LS-SVM optimization and model results were obtained using the LS-SVM lab toolbox version 1.5 (Matlab toolbox for least squares support vector machines) [25].

## RESULTS AND DISCUSSION

Barium has a number of emission lines for determination by ICP-OES method. However some of these emission lines have very low sensitivity for determination of barium. Therefore, in this study five emission lines (493.408, 455.403, 413.243, 233.527 and 230.425 nm) were selected to evaluate the ability of ICP-OES method to determine <sup>137</sup>Ba abundance.

### LS-SVM Determination of <sup>137</sup>Ba Abundances

Determination of <sup>137</sup>Ba abundance by ICP-OES coupled



**Fig. 1.** Optimization of the LS-SVM algorithm for <sup>137</sup>Ba abundance analysis, based on the emission line at 493.408 nm.

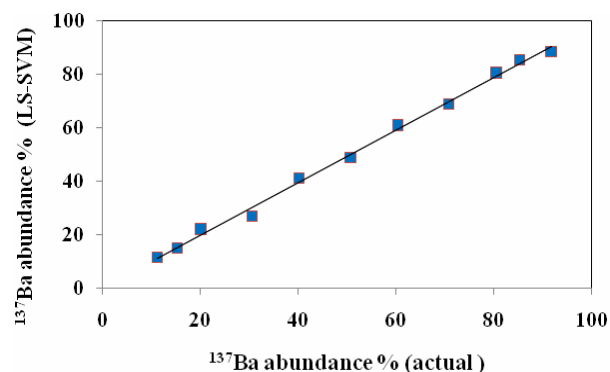
with LS-SVM method involves constructing calibration and validation sets. A training set of 11 samples was taken ( $X_{(11 \times 10)}$ ). The  $^{137}\text{Ba}$  abundances in the calibration set were in the range of 11.23-91.7% ( $y_{(11 \times 1)}$ ). For constructing calibration samples, the concentration of barium was kept constant at  $5 \mu\text{g ml}^{-1}$ . For the validation set, 6 samples were prepared. The LS-SVM was applied first to calibration samples for the analysis of  $^{137}\text{Ba}$  abundances. In this study, LS-SVM was performed with radial basis function (RBF) as a kernel function because of its good general performance and limited parameters. Thus  $\gamma$  (the relative weight of the regression error) and  $\sigma^2$  (the kernel parameter of the RBF kernel) need to be optimized. To determine the optimal parameters, a grid search was performed based on leave-one-out cross-validation on the original training set for all parameter combinations of  $\gamma$  and  $\sigma^2$ . Since because the grid search has been performed over just two parameters, a contour plot of the optimization error can be visualized easily (Fig. 1). From Fig. 1, the optimal parameter settings can now be selected from a smooth subarea with a low cross-validation RMSE.

The quality of each model was assessed by applying the 5-fold cross-validation procedure. The statistical parameters for determination of  $^{137}\text{Ba}$  abundances are presented in Table 2.

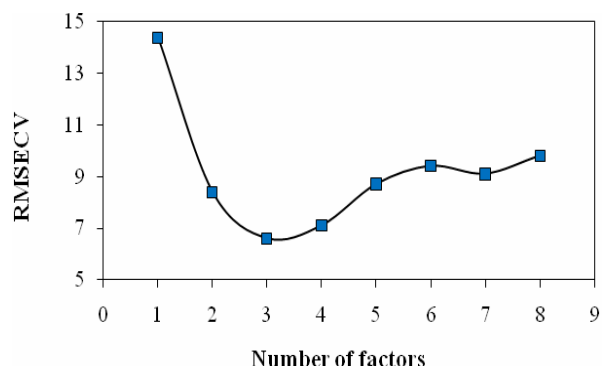
As seen in this table, the statistical parameters of LS-SVM model for the emission line at 493.408 nm are superior to those of the other emission lines. The best model corresponds to the emission line at 493.408 nm with the highest  $R^2$  value of 0.994 for the test set and the lowest RMSEP value (3.360). The value of the determination coefficient of the 5-fold cross-validation for the model obtained with the emission line at 493.408 nm ( $R^2_{\text{CV}} = 0.968$ ) is slightly higher than that concerning the emission line at 455.403 nm ( $R^2_{\text{CV}} = 0.950$ ). Therefore, the emission line at 493.408 nm was selected for determination of  $^{137}\text{Ba}$  abundance. In Fig. 2, the plot of the predicted  $^{137}\text{Ba}$  abundances by LS-SVM model versus actual values is represented. The agreement between observed and predicted values and high correlation coefficient, confirms the high prediction ability of LS-SVM modeling.

### Comparison with PLS Method

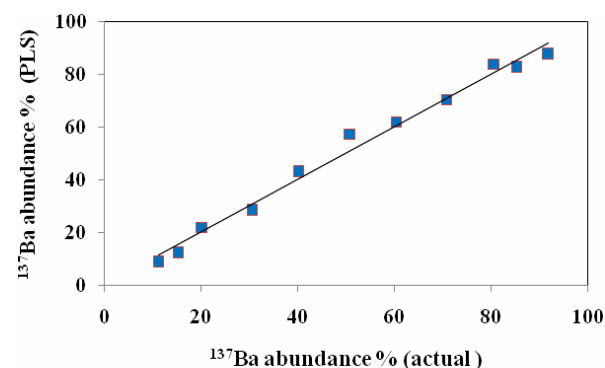
In order to determine the optimum number of factors



**Fig. 2.** Plot of the estimated  $^{137}\text{Ba}$  abundances by the LS-SVM method vs. the actual values.



**Fig. 3.** Optimization of the number of principal components for the PLS model.



**Fig. 4.** Plot of the estimated  $^{137}\text{Ba}$  abundances by the PLS method vs. the actual values.

(latent variables) for the partial least squares calibration model, the leave-one-out cross-validation procedure was applied. The most convenient PLS model that resulted in the best fitness contained three latent variables (Fig. 3). In Fig. 4, the plot of the estimated  $^{137}\text{Ba}$  abundances versus actual values are represented. The statistical parameters calculated for the PLS model are presented in Table 3. As seen, the statistical parameters of LS-SVM models are superior to that of PLS.

Although PLS method assumes a linear relationship between the independent and dependent variables, small deviation from linearity is acceptable and can be readily suppressed by including additional modelling factors.

However, in the presence of substantial nonlinearity, PLS tends to give large prediction errors and calls for more suitable models. Intrinsically non-linear calibration methods

are applicable in the later cases [28]. LS-SVM is a nonparametric and flexible modelling method which is capable of fitting complex and nonlinear relationships between species and predictors and therefore produces superior model compared to the model developed by PLS approach.

### Determination of $^{137}\text{Ba}$ Abundances in Synthetic Samples

The predictive ability of the calibration model was assessed using six synthetic samples with various  $^{137}\text{Ba}$  abundances. The results obtained by applying LS-SVM based calibration model to synthetic samples are listed in Table 4. Table 2 shows  $R^2$  and RMSEP for the test set. The results demonstrated that the LS-SVM method serves as a good model for the quantification of  $^{137}\text{Ba}$  abundances.

**Table 2.** Statistical Parameters for the LS-SVM Model

Wavelength	Spectral range	$R^2$	RMSEC	$R^2_{CV}$	RMSECV	$R^2_{\text{Test set}}$	RMSEP
493.408	493.362-493.467	0.994	2.332	0.968	5.502	0.977	3.360
455.403	455.357-455.470	0.988	3.227	0.950	6.686	0.932	5.650
413.243	413.021-413.121	0.850	11.531	0.472	20.860	0.723	9.697
233.527	233.507-233.539	0.782	13.781	0.578	16.978	0.782	13.781
230.425	230.409-230.441	0.764	14.426	0.235	24.368	0.784	13.301

**Table 3.** Statistical Parameters for the PLS Model

Wavelength	Spectral range	$R^2$	RMSEC	$R^2_{CV}$	RMSECV	$R^2_{\text{Test set}}$	RMSEP
493.408	493.362-493.467	0.982	3.736	0.948	6.487	0.952	4.411
455.403	455.357-455.470	0.974	4.411	0.917	8.201	0.924	8.058
413.243	413.021-413.121	0.792	12.667	0.550	20.404	0.749	12.486
233.527	233.507-233.539	0.841	11.083	0.664	16.583	0.787	11.477
230.425	230.409-230.441	0.753	13.831	0.444	20.879	0.714	13.919

### Accuracy and Capability of the Method

In order to evaluate the applicability of the proposed method, spiked water samples with  $^{137}\text{Ba}$  (well water, river water and sea water) were analyzed for the quantification  $^{137}\text{Ba}$  abundances. In addition, the predicted results using ICP-OES combined with the LS-SVM method were compared with the ICP-MS method. Comparisons of the prediction capability of the proposed method and the ICP-MS method for determination of  $^{137}\text{Ba}$  abundances are presented in Table 5. The results demonstrated that the proposed method is suitable for determination of  $^{137}\text{Ba}$  abundances in real water samples.

### CONCLUSIONS

In the present study, the application of ICP-OES coupled with computational methods for determination of  $^{137}\text{Ba}$  abundances is proposed. The LS-SVM and PLS methods were employed for the quantification of  $^{137}\text{Ba}$  abundances. The results obtained for determination of  $^{137}\text{Ba}$  abundances demonstrated that the LS-SVM method performs somewhat better than the PLS method. The proposed method offers a simple and a reasonable substitute for expensive mass spectrometric methods for the quantification of barium isotopes. This study shows that complex problems that may

**Table 4.** Results of the Measuring  $^{137}\text{Ba}$  Abundances in Validation set

Sample number	$^{137}\text{Ba}$ abundance% (Actual)	$^{137}\text{Ba}$ abundance% (LS-SVM)	Relative error (%)
1	25.71	21.69 (1.12) <sup>a</sup>	-15.66
2	35.37	33.90 (0.69)	-4.16
3	45.03	49.66 (0.78)	10.28
4	55.49	58.60 (1.02)	5.60
5	65.14	69.41 (0.73)	6.55
6	75.61	75.22 (0.83)	-0.51

<sup>a</sup>Values in parentheses are R.S.D.s based on the three replicate analyses.

**Table 5.** Determination of  $^{137}\text{Ba}$  Abundance% in Spiked Water Samples (n = 3)

Sample	Total barium concentration ( $\text{mg l}^{-1}$ )	$^{137}\text{Ba}$ abundance%		Significant error (t-test, 95% confidence level)
		Proposed method	ICP-MS	
Well water	0.008	$40.04 \pm 1.02^a$	$42.06 \pm 0.75$	No
River water	0.030	$55.36 \pm 0.77$	$53.35 \pm 1.05$	No
Caspian seawater	0.070	$61.96 \pm 1.32$	$62.34 \pm 0.55$	No

<sup>a</sup>Standard deviation.

arise when quantifying one or several analytes can be addressed using suitable chemometric techniques.

## REFERENCES

- [1] CRC Handbook of Chemistry and Physics, 87<sup>th</sup> ed., 2006.
- [2] P. Desmoulins, Stable Isotopes, Application-Production, Gif-Sur-Yvette, France, 1994.
- [3] Z. Muccio, G.P. Jackson, Analyst 134 (2009) 213.
- [4] L. Simon, C. Lecuyer, C. Marechal, N. Coltice, Chem. Geol. 225 (2006) 61.
- [5] K. Allmen, M.E. Böttcher, E. Samankassou, T.F. Nägler, Chem. Geol. 277 (2010) 70.
- [6] A. Sonoda, Y. Makita, K. Ooi, T. Hirotsu, J. Nucl. Sci. Technol. 39 (2002) 295.
- [7] D. Beauchemin, Anal. Chem. 78 (2006) 4111.
- [8] S. Misra, P.N. Froelich, J. Anal. At. Spectrom. 24 (2009) 1524.
- [9] J.S. Becker, H. Sela, J. Dobrowolska, M. Zoriy, J.S. Becker, Int. J. Mass Spectrom. 270 (2008) 1.
- [10] M. Khayatzaadeh Mahani, A.R. Khanchi, M. Heidari, A. Ahmadi, J. Anal. At. Spectrom. 25 (2010) 1659.
- [11] S. Thangavel, S.V. Rao, K. Dash, J. Arunachalam, Spectrochim. Acta Part B 61 (2006) 314.
- [12] J.A. Wheat, Appl. Spectrosc. 25 (1971) 328.
- [13] P.L. Larkins, Spectrochim. Acta Part B 39 (1984) 1365.
- [14] J.A. Goleb, Anal. Chem. 35 (1963) 1978.
- [15] M. Krachler, D.H. Wegen, J. Anal. At. Spectrom. 27 (2012) 335.
- [16] E. Zolfonoun, S.J. Ahmadi, Spectrochim. Acta Part B 81 (2013) 64.
- [17] X. Mao, A.A. Bol'shakov, D.L. Perry, O. Sorkhabi, R.E. Russo, Spectrochim. Acta Part B 66 (2011) 604.
- [18] A.A. Ganeyev, S.E. Sholupov, Spectrochim. Acta Part B 47 (1992) 1325.
- [19] J. Moros, S. Garrigues, M. Guardia, Trends Anal. Chem. 29 (2010) 578.
- [20] A. Rouhollahi, H. Tavakoli, S. Nayebi, J. Ghasemi, M. Noroozi, M. Hashemi, Iran. J. Chem. Chem. Eng. 26 (2007) 41.
- [21] J.B. Ghasemi, E. Zolfonoun, Spectrochim. Acta Part A 115 (2013) 357.
- [22] G.C.Y. Chan, X. Mao, I. Choi, A. Sarkar, O.P. Lama, D.K. Shuh, R.E. Russo, Spectrochim. Acta part B 89 (2013) 40.
- [23] E. Zolfonoun, J. Anal. At. Spectrom. 30 (2015) 2003.
- [24] C. Cortes, V. Vapnik, Mach. Learn. 20 (1995) 273.
- [25] V. Vapnik, Statistical Learning Theory, John Wiley, New York, 1998.
- [26] J.A.K. Suykens, J. Vandewalle, Neural Process. Lett. 9 (1999) 293.
- [27] J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, Least-Squares Support Vector Machines, World Scientifics, Singapore, 2002.
- [28] A. Niazi, J. Ghasemi, M. Zendehtdel, Talanta 74 (2007) 247.
- [29] A.A. Ensafi, F. Hasanpour, T. Khayamian, A. Mokhtari, M. Taei, Spectrochim. Acta A 75 (2010) 867.
- [30] J.B. Ghasemi, E. Zolfonoun, Environ. Monit. Assess. 184 (2012) 3971.