



Spectrophotometric determination of Rare Earth Elements in aqueous nitric acid solutions for process control[☆]

Oxana Ye. Rodionova^a, Tatyana I. Tikhomirova^b, Alexey L. Pomerantsev^{a,c,*}

^a *Semenov Institute of Chemical Physics RAS, Kosygin Str. 4, 119991 Moscow, Russia*

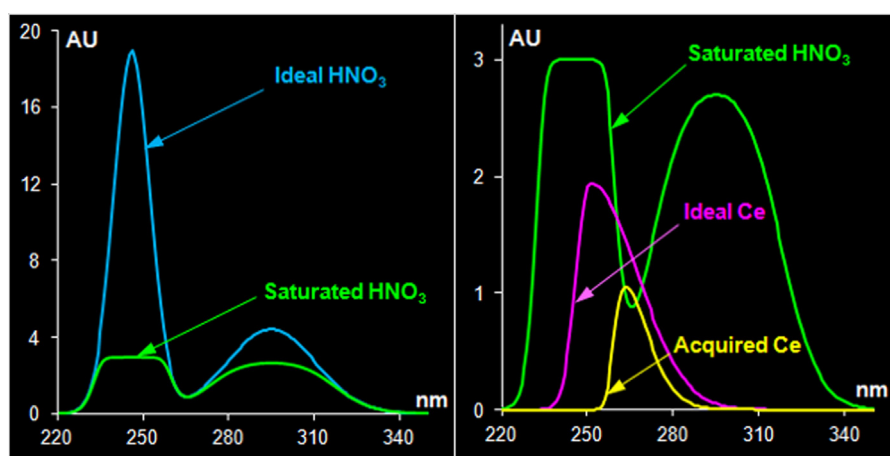
^b *Department of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia*

^c *Institute of Natural and Technical Systems RAS, Kurortny pr. 99/18, 354024 Sochi, Russia*

HIGHLIGHTS

- Spectrophotometric determination of Rare Earth mixtures in HNO₃ is studied.
- Reliable models for quantification of Nd and Pr are constructed.
- Models for semi-quantitative analysis of Ce are established.
- Models' performances are analyzed in out-of-control cases.
- The predictability of the PLS and MCR-ALS models is compared.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 9 December 2014

Received in revised form 7 February 2015

Accepted 11 February 2015

Available online 14 February 2015

Keywords:

Rare Earth elements

Ultraviolet–visible (UV–vis)

spectrophotometry

Partial least squares (PLS)

Multivariate curve resolution–alternating

least squares (MCR–ALS)

Out-of control behavior

ABSTRACT

Noninvasive analytical control is of special interest for the complicated and hazardous production processes. On-line monitoring provides a unique opportunity to determine critical concentrations rapidly and without serious risks to operating personnel and the environment. Models for quantitative determination of concentrations of Rare Earth Elements in complex mixtures in nitric acid serve for these purposes. Here, the feasibility of simultaneous determination of cerium, praseodymium, and neodymium using the whole UV–vis spectroscopic range, together with chemometric data processing, is studied. The predictability of two chemometric techniques, partial least squares regression and correlation constrained multivariate curve resolution–alternating least squares are compared. Models' performances are analyzed in out-of-control cases.

© 2015 Elsevier B.V. All rights reserved.

[☆] This work was first presented at the 14th Conference on Chemometrics in Analytical Chemistry, Richmond, USA, 2014.

* Corresponding author at: Semenov Institute of Chemical Physics RAS, Kosygin Str. 4, 119991, Moscow, Russia. Tel.: +7 495 9397483; fax: +7 495 9397483.

E-mail address: forecast@chph.ras.ru (A.L. Pomerantsev).

1. Introduction

Intensive spectrophotometric studies of Rare Earth Elements (REEs) took place more than 60 years ago [1–7]. Most of the extinction profiles we can refer to now were obtained back in these

years. For each REE the efforts were focused on revealing ‘the bands that are completely isolated from background absorption due to other ions’ [1]. However, REEs typically have an abundant UV–vis spectrum with many peaks, due to which a classic univariate approach was not very successful. Afterwards the direct UV–vis spectroscopy lost its popularity and was replaced with more selective methods based on chromogenic reagents [8]. The interest in spectrophotometry of REEs has recently been revived in connection with process analytical technology (PAT) and multivariate approach. Noninvasive analytical control is of special interest in the nuclear industry, in the field of spent nuclear fuel reprocessing [9,10]. On-line control of these processes provides a unique opportunity to determine the indispensable concentrations rapidly and without serious risk of radiation pollution [10,11].

PAT rapidly penetrates into various branches of industry. The PAT based applications imply combining the on-line methods of measurements with chemometric data processing [12]. Such a way of data acquisition essentially limits the available analytical means as no dilution or concentration is possible, no exogenous reagents can be applied, etc. Moreover, the determination of components in complex mixtures becomes rather difficult, as elements with very high and very low concentrations (extinctions) should be quantified simultaneously. For these reasons, many interesting and appropriate laboratory scaled PAT solutions have not been implemented in practice. An additional obstacle is posed by the variability of process parameters, which are often out of control in comparison with laboratory scaled models. These circumstances require the development of special chemometric methods. When it comes to calibration models the following issues should be taken into account.

1. The concentrations in real life production could be occasionally out of the range studied in the laboratory. Therefore, the model should be, to some extent, stable in case of a prediction out of the explored concentration range.
2. The absorbance of some analytes can be very high in complex mixtures. As a result, the corresponding peaks become distorted and flattened.

The aim of the study is an elaboration of the calibration models for simultaneous quantification of REEs – cerium, neodymium, and praseodymium, in the nitric acid solutions using spectrophotometry and chemometric modeling. As the purpose of the models is an on-line process control, they should be stable when it comes to issues listed above. The selection of instruments and experimental set-up aims to reproduce the real process analyzer facilities. Practically, this means that the spectral resolution and wavelength accuracy are subpar, the optical path length is fixed, no dilution is used, and no extra reagents are employed.

Two standard calibration techniques are employed in this study. The first is the partial least squares (PLS) regression as it is the most popular and well-established multivariate calibration technique [13]. Another approach considered is the correlation constrained multivariate curve resolution–alternating least squares (MCR–ALS) [14,15]. These methods are compared to select an appropriate solution for REE concentrations determination.

2. Theory

A calibration model relates two parts of data, namely, the $(I \times J)$ matrix $\mathbf{X}$ that contains spectra of I samples recorded for J wavelengths, and the $(I \times N)$ matrix $\mathbf{C}$ of components' concentrations. N stands for the number of components. Direct calibration is based on the principle of linearity, also known as the Lambert–Beer law,

$$\mathbf{X} = \mathbf{CS}^t \quad (1)$$

In this equation, $\mathbf{S}$ is the $(J \times N)$ matrix of pure spectra, also called the *sensitivity matrix*. If matrix $\mathbf{S}$ is known, the concentrations are calculated by an equation

$$\mathbf{C} = \mathbf{XB}, \quad (2)$$

where $\mathbf{B} = \mathbf{S}(\mathbf{S}^t\mathbf{S})^{-1}$ is the $(J \times N)$ matrix. However, in practice, matrix $\mathbf{S}$ is usually unknown and should be reconstructed from a training subset of data.

A basic model of the inverse calibration is given in a similar form (Eq. (2)), where the targeted matrix $\mathbf{C}$ (concentrations) is expressed directly in terms of the known spectra matrix $\mathbf{X}$. Despite the fact that such a representation of calibration equation contradicts the fundamental relation given in Eq. (1), this approach provides a better quality of modeling [13,16]. In this case, $\mathbf{B}$ is the $(J \times N)$ matrix of coefficients, which are estimated using the training subset of data.

The obtained model (i.e. matrix $\mathbf{B}$) is then applied to another subset $\mathbf{X}_{\text{new}}$ that can be a validation, or test, or new data set. The predicted concentrations

$$\mathbf{C}_{\text{new}} = \mathbf{X}_{\text{new}}\mathbf{B} \quad (3)$$

can be compared with the known reference values, or directly used for the process monitoring.

Two methods employed in this paper are the correlation constrained MCR–ALS (direct calibration) and PLS (inverse calibration). The assumption of a bi-linear relation between the experimental data and the components' concentrations/pure spectra serves as the basis of both methods.

2.1. PLS

PLS is a well-known method of inverse calibration described elsewhere in literature [17]. The method yields the prediction of an analyte concentration without reconstructing the pure spectra of the mixture components. The procedure develops a model by means of gradual increase of its complexity, i.e. the number of latent variables (LVs). The optimal complexity is selected comparing the behavior of calibration and validation errors against the increasing number of LVs.

In the present study each chemical component is calibrated individually, this means that the PLS1 technique is used. Let K be the selected number of LVs, then for each REE the following formula is applied

$$\mathbf{c} = \mathbf{Tq} + \mathbf{f} \quad (4)$$

where $\mathbf{c}$ is the $(I \times 1)$ vector of concentrations for one chemical component, $\mathbf{T}$ is the $(I \times K)$ matrix of the PLS X-scores, $\mathbf{q}$ is the $(K \times 1)$ vector of PLS Q-loadings, and $\mathbf{f}$ is the $(I \times 1)$ vector of residuals.

2.2. Correlation constrained MCR–ALS

Another well known method is MCR–ALS [18]. This is a direct calibration approach that resolves spectroscopic data, and, afterwards, it uses the obtained spectroscopic profiles for the quantitative calibration. MCR–ALS is based on the equation

$$\mathbf{X} = \mathbf{CS}^t + \mathbf{E} \quad (5)$$

Here $\mathbf{S}$ is the $(J \times N)$ matrix of pure spectra, N is the number of components in the system, and $\mathbf{E}$ is the $(I \times J)$ matrix, which contains variations not explained by the model.

The procedure starts with evaluation of the initial estimate of matrix $\mathbf{S}$, or matrix $\mathbf{C}$ [19]. In calibration, it is natural to use the known matrix $\mathbf{C}_{\text{ref}}$ for the initial step. Afterwards, matrices $\mathbf{C}$ and $\mathbf{S}$ are found by the ALS method that consequently minimizes the sum

of the squared residuals

$$\| \mathbf{X} - \mathbf{C}\mathbf{S}^t \|^2. \quad (6)$$

The algorithm consists of two steps, the C-type step, and the S-type step, which are repeated until convergence. At the C-type step, the value of $\mathbf{S} \equiv \mathbf{S}_{\text{hat}}$ is fixed, and $\mathbf{C}$ matrix is calculated using the unconstrained least squares (LS) estimator

$$\mathbf{C}_{\text{in}} = \mathbf{X}\mathbf{S}_{\text{hat}}(\mathbf{S}_{\text{hat}}^t\mathbf{S}_{\text{hat}})^{-1} \quad (7)$$

$\mathbf{C}_{\text{in}}$ is transformed into $\mathbf{C}_{\text{hat}}$ to incorporate the constraints for the concentration profile.

For the S-type step, the value of $\mathbf{C} \equiv \mathbf{C}_{\text{hat}}$ is fixed and matrix $\mathbf{S}$ is found by a similar formula

$$\mathbf{S}_{\text{in}} = \mathbf{X}^t\mathbf{C}_{\text{hat}}(\mathbf{C}_{\text{hat}}^t\mathbf{C}_{\text{hat}})^{-1}. \quad (8)$$

Afterwards, $\mathbf{S}_{\text{in}}$ is transformed into $\mathbf{S}_{\text{hat}}$ to incorporate some constraints for the spectral profile. The calculations, given in Eqs. (7) and (8) and transformations subject to corresponding constraints are repeated till convergence.

To give a physicochemical meaning to the LS estimates $\mathbf{C}_{\text{in}}$ and $\mathbf{S}_{\text{in}}$, necessary constraints are applied at each step. For example, there are natural non-negativity constraints that force the concentrations and spectra to be equal or greater than zero. In calibration problems, additional correlation constraints [14,15] are applied. They are based on regressions that relate the known reference concentration matrix, $\mathbf{C}_{\text{ref}}$, with matrix $\mathbf{C}_{\text{in}}$ obtained in the ALS procedure by Eq. (7).

$$\mathbf{C}_{\text{ref}} = \mathbf{C}_{\text{in}}\mathbf{A} + \mathbf{G}. \quad (9)$$

Here $\mathbf{A}$ is the $(N \times N)$ matrix of slopes, and $\mathbf{G}$ is the $(I \times N)$ matrix of intercepts [14]. In a simple case, a univariate regression $\mathbf{c}_{n,\text{ref}} = a_n\mathbf{c}_{n,\text{in}} + g_n$ is developed for each component concentration vector $\mathbf{c}_n$ [20]. Then, $\mathbf{A} = \text{diag}(a_1, \dots, a_n)$, and $\mathbf{G} = (g_1\mathbf{1}, \dots, g_n\mathbf{1})$, where $\mathbf{1}$ is the $(N \times 1)$ vector of units.

The estimated regression matrices $\mathbf{A}$ and $\mathbf{G}$ are used to obtain the adjusted concentration matrix $\mathbf{C}_{\text{hat}}$ by the following formula

$$\mathbf{C}_{\text{hat}} = \mathbf{C}_{\text{in}}\mathbf{A} + \mathbf{G}. \quad (10)$$

In this study we make an attempt to substitute the linear correlation given by Eq. (9) by a non-linear transformation in an effort to account for a non-linear relationship between $\mathbf{c}_{\text{ref}}$ and $\mathbf{c}_{\text{in}}$ for calibration of cerium. Eq. (9) is modified as follows

$$\mathbf{c}_{\text{ref}} = f(\mathbf{c}_{\text{in}}, \mathbf{a}) \quad (11)$$

Function $f(\cdot)$ as well as the unknown parameters $\mathbf{a}$ are selected to fit the $\mathbf{c}_{\text{ref}}/\mathbf{c}_{\text{in}}$ dependence in the best way.

In MCR-ALS, the prediction is performed by the following formula

$$\mathbf{C}_{\text{new}} = \mathbf{X}_{\text{new}}\mathbf{S}_{\text{hat}}(\mathbf{S}_{\text{hat}}^t\mathbf{S}_{\text{hat}})^{-1}\mathbf{A} + \mathbf{G} \quad (12)$$

where $\mathbf{X}_{\text{new}}$ is a matrix of new (spectral) data, and matrices $\mathbf{A}$, $\mathbf{G}$, and $\mathbf{S}_{\text{hat}}$ have been obtained at the calibration stage.

For the beginning of the MCR-ALS procedure it is necessary to determine the number of chemical components, i.e. the N value. In the current study the number of chemical components is known, so N is equal to 3. Of course an effective rank of the $\mathbf{X}$ -matrix may be higher, e.g. if the interaction of components are presented in the analytes. In our case, the PCA of $\mathbf{X}$ data also returns $N=3$.

2.3. Figures of merit

We use the root mean square errors (RMSE) of calibration, RMSEC, and prediction, RMSEP, which are calculated by a formula

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^I (c_i - \hat{c}_i)^2}{I}} \quad (13)$$

where c_i are the reference concentration values, $\hat{c}_i$ are the estimated values, and I is the number of samples. RMSEC characterizes the quality of calibration. RMSEP not only assesses the quality of prediction but for the PLS model, also helps to select the optimal number of LVs.

RMSE is not an appropriate measure in case of a mixture calibration, because the component concentrations can be presented in very different scales. For this reason the relative error, RE, is applied. Presented in percents (RE%), it is calculated as follows

$$\text{RE}(\%) = 100 \sqrt{\frac{\sum_{i=1}^I (c_i - \hat{c}_i)^2}{\sum_{i=1}^I c_i^2}} \quad (14)$$

where c_i are the reference concentration values, $\hat{c}_i$ are the estimated values, and I is the number of samples in the calibration, or validation, or test sets.

The 'predicted vs. reference' plots are constructed for the model performance evaluation. The slope, offset, and squared correlation coefficient characterize the model quality and help the comparison of various methods. Figures of merit are calculated separately for each REE.

2.4. Computing

All calculations, data-pretreatment, PLS modeling, MCR-ALS procedure were performed using the standard Excel functions, and Chemometrics Add-in [16].

3. Experimental

3.1. Sample preparation

The stock solutions of 50.00 g L⁻¹ of neodymium(III), praseodymium(III), and cerium(III) were made up using the nitrate hexahydrate oxides Nd(NO₃)₃·6H₂O, Pr(NO₃)₃·6H₂O, and Ce(NO₃)₃·6H₂O by Merck (extra pure), and 0.7 M HNO₃ by Merck. Working solutions for the calibration and validation data sets were prepared by aliquot dilution and mixing of these stock solutions.

The samples for the test set have the same concentration of the REEs dissolved in nitric acid of various concentrations. The 4.0 M HNO₃ was made by diluting of 33 mL of HNO₃ in 100 mL of water using a volumetric flask; other nitric acid solutions were prepared by dilutions of this solution. Double-distilled water was used. The test set samples were prepared using the accurate weighing of each REE, which were then mixed and dissolved in the appropriate HNO₃ solution.

3.2. Instrumentation and experimental measurements

The samples are subjected to the UV-vis spectroscopy in the transmittance mode, and then converted to the absorbance units. Spectra are collected with a wavelength increment of 1 nm among consecutive measurements over the range of 200–1000 nm using spectrophotometer UNICO SQ-2800 with photometric range of 0.01–3 AU. The acquisition time is 8 min. A 10 mm path length quartz cuvette is used.

3.3. Data subsets

The goal of the study is to construct calibration models for quantification of each REE in the presence of the other REEs. The

concentration ranges of REEs were defined by the customer and are presented in Table 1.

Two level full factor design was used as a basement for the investigation of the range of interest. Eight solutions were prepared, which are samples Nos. 1–6 in the calibration set and samples Nos. 1 and 2 in the validation set (see Table 2). In addition three samples with average concentrations for all three REEs, i.e. an intermediate point, were added. These are samples 7 and 8 from the calibration set and sample 3 from the validation set. Solutions with very low (No. 12 from the calibration set and No. 6 from the validation set) and very high (Nos. 9–11 from the calibration set and Nos. 7 and 8 from the validation set) concentrations were added for an assessment of method performance. For the evaluation of spectral behavior dependant on changes in concentrations for each REE, one component solution with two concentration levels was prepared. These are samples Nos. 9–14 in the validation set. The remaining two samples in validation set (Nos. 4 and 5) were chosen for better data interpretation.

As a result the complete concentration ranges vary for cerium from 3×10^{-4} to 5 g L^{-1} , for neodymium from 4×10^{-2} to 30 g L^{-1} and for praseodymium from 1×10^{-4} to 15 g L^{-1} . All these samples were divided into calibration and validation sets. These two subsets are prepared with the constant nitric acid concentration that equals 0.7 M.

In addition, a special test subset contains samples having the same composition of the REEs dissolved in nitric acid of various concentrations, which are 0.1 M, 0.4 M, 1.8 M, 3.0 M, and 4.0 M. This special subset is aimed at assessment of the model stability, i.e. at verification of the calibration performance under unusual conditions that differ from the calibration ones.

Additionally, the background spectra of HNO_3 at concentrations of 0.1 M, 0.4 M, 0.7 M, 1.8 M, 3.0 M, and 4.0 M are acquired, and then subtracted from sample spectra that have the corresponding nitric acid concentrations. The baseline offset method is applied to the raw data in order to correct shift and slope. The resulting spectra are shown in Fig. 1.

The validation subset is selected in a special way. The concentration range of praseodymium, [0,10], is completely inside its calibration range, [0,15]. The concentration range of neodymium, [0,30], is twice larger than in the calibration subset, [0,15]. And

for cerium, both subsets share the same concentration range, [0,5]. In this way, we provide for all variants that can be expected in practical applications, namely, a new sample is located inside, outside, or on the border of the calibration concentration range. This design allows us to verify the steadiness of the model against unexpected variations in the monitoring process concentrations that could go out of the concentration range studied in the laboratory. It is worth noting that all mixtures, except three samples used to control the experiment reproducibility, have different REE compositions, though the concentrations of one/two components may be equal (factorial design). The first samples in the calibration and validation sets can be considered as an example. In these mixtures concentrations of Pr and Ce are equal, but the concentration of Nd is different. Such a design is used not only to quantify a specific component but also to assess the influence of other components in determination of analyte concentration in a mixture.

The test subset is designed in another way. It is known that industrial processes might have variations in the nitric acid media concentration. Table 2 shows that both the calibration and validation subsets contain samples with different REE concentrations dissolved in the same HNO_3 solvent, 0.7 M. On the contrary, the test subset includes samples that have the same REE concentrations, but the acid concentration varies. This helps us to evaluate another model feature that is stability against alternating nitric acid concentration.

Subtracting the background spectrum of nitric acid from the raw spectra requires the quantification of its concentration in the mixture. Actually, HNO_3 does not absorb in the range of 300–1000 nm. Thus, it does not influence on the Pr and Nd determination. At the same time, nitric acid has huge peaks in the range 200–300 nm, and, in this way, hampers calibration of Ce. There exist many ways of determining HNO_3 concentration, but in a real application this step may reduce the speed of the ongoing process.

4. Results

4.1. Quantitative determination of praseodymium and neodymium

The MCR-ALS procedure is utilized with the non-negativity constraints both for concentrations and spectral profiles, and correlation constraints for the concentration profiles. Neodymium and praseodymium are reliably predicted with an MCR-ALS model (Table 3, MCR-ALS columns).

A sharp increase of validation error for neodymium is explained by the fact that the concentration range in the validation subset is two times greater than that for calibration. However, considering

Table 1
Predefined REE's concentrations.

	HNO_3 (M)	Pr (g L^{-1})	Nd (g L^{-1})	Ce (g L^{-1})
Min	0.7	0.30	0.60	0.03
Max	0.7	10.00	12.00	0.60

Table 2
Samples used for analysis.

	Calibration set				Validation set				Test set			
	HNO_3 (M)	Pr (g L^{-1})	Nd (g L^{-1})	Ce (g L^{-1})	HNO_3 (M)	Pr (g L^{-1})	Nd (g L^{-1})	Ce (g L^{-1})	HNO_3 (M)	Pr (g L^{-1})	Nd (g L^{-1})	Ce (g L^{-1})
1	0.7	10.00	0.60	0.60	0.7	10.00	12.00	0.60	0.1	3.00	2.00	0.30
2	0.7	0.30	12.00	0.60	0.7	0.30	0.60	0.03	0.4	3.00	2.00	0.30
3	0.7	0.30	0.60	0.60	0.7	3.00	2.00	0.30	1.8	3.00	2.00	0.30
4	0.7	10.00	12.00	0.03	0.7	8.00	10.00	0.50	3.0	3.00	2.00	0.30
5	0.7	0.30	12.00	0.03	0.7	1.00	3.50	2.00	4.0	3.00	2.00	0.30
6	0.7	10.00	0.60	0.03	0.7	0.01	0.04	0.002				
7	0.7	3.00	2.00	0.30	0.7	1.00	30.00	5.00				
8	0.7	3.00	2.00	0.30	0.7	1.00	30.00	2.00				
9	0.7	15.00	5.00	5.00	0.7	0.00	0.00	2.00				
10	0.7	15.00	5.00	2.00	0.7	0.00	0.00	0.50				
11	0.7	5.00	0.30	5.00	0.7	1.00	0.00	0.00				
12	0.7	1×10^{-4}	4×10^{-4}	3×10^{-4}	0.7	0.25	0.00	0.00				
13					0.7	0.00	3.50	0.00				
14					0.7	0.00	0.875	0.00				

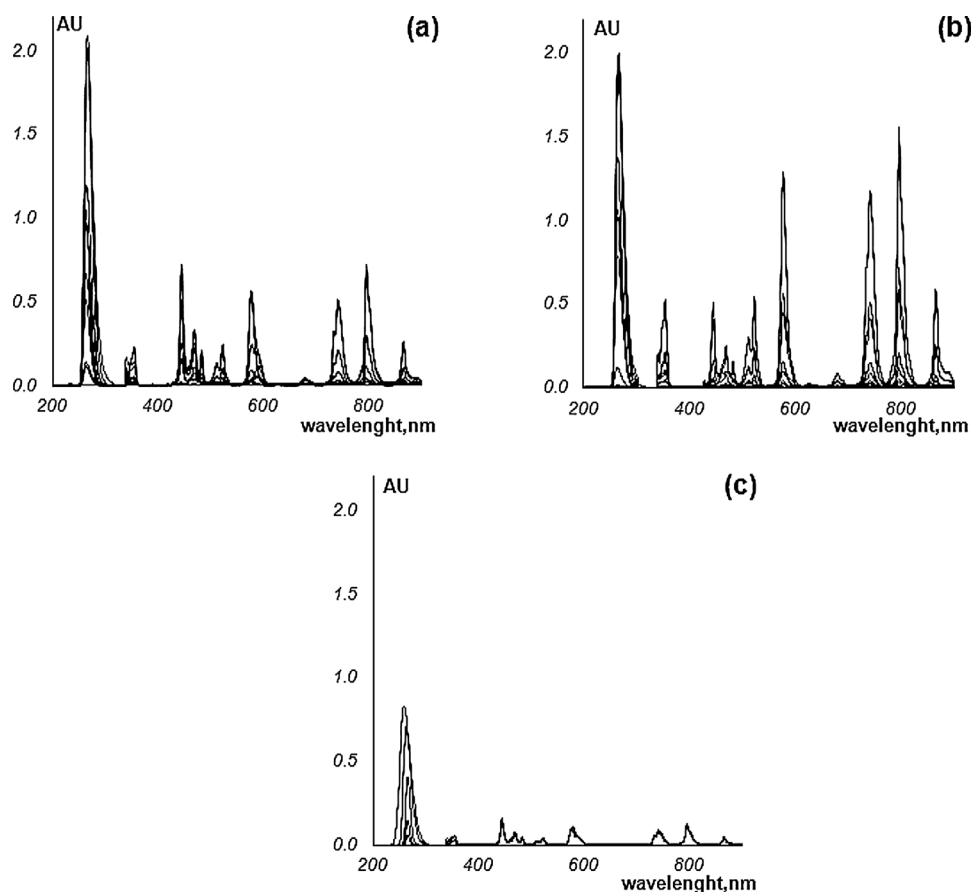


Fig. 1. Absorbance spectra: training subset (a), validation subset (b), and test subset (c).

validation errors in the percentage scale, we see that they are not so large. Consider, for example, the residuals of Nd. Sample No. 6 from the calibration set (see Table 2) has the residual $0.60 - 0.48 = 0.12$ that is 20% of the reference value. On the other side, sample No. 8 from the validation set (see Table 2) is predicted with the residual $30.00 - 27.35 = 2.65$ that is 9% of the reference value. If samples Nos. 7 and 8 with extra high Nd concentration are excluded from the validation set, then RMSEPs in both methods will be compatible to RMSECs, namely $RMSEP_{PLS} = RMSEP_{MCR-ALS} = 0.08$.

The pure spectra of praseodymium (a) and neodymium (b) were recovered using MCR-ALS and recalculated to the molar extinction coefficients. They are presented in Fig. 2.

Table 3

Results of quantitative determination of Nd and Pr by PLS (4 LVs) and MCR-ALS (3 components).

	Pr		Nd	
	PLS	MCR-ALS	PLS	MCR-ALS
Calibration				
RMSEC	0.13	0.20	0.08	0.16
Slope	1.00	1.00	1.00	1.00
Offset	0.00	0.01	0.00	0.00
R^2	1.00	1.00	1.00	1.00
RE (%)	2	3	1	2
Validation				
RMSEP	0.33	0.25	1.03	1.05
Slope	1.04	1.03	0.91	0.91
Offset	-0.02	-0.10	0.17	0.16
R^2	0.99	0.99	1.00	1.00
RE (%)	9	7	8	9

Additionally, some reference values found in literature are shown. Black solid curves represent the extinction coefficients' profiles found in Ref. [21]. Squares (1) stand for extinction coefficients given in Ref. [1] for each REE in 0.45 M HNO_3 . Triangles (2) correspond to the values presented in Ref. [2]. Dots (3) demonstrate the neodymium extinction coefficient [3] obtained in 0.03 M HNO_3 solution.

Praseodymium and neodymium have a number of sharp absorption peaks in nitric acid. The main praseodymium peaks can be seen around 445 nm, 467 nm, 481 nm, and 589 nm. The main neodymium peaks are located around 353 nm, 520 nm, 578 nm, 741 nm, 798 nm, and 866 nm. In general, extinction coefficients for praseodymium and neodymium are consistent with results presented in Ref. [21]. However, the pure spectra found by MCR should be taken with a care due to a well known scaling ambiguity [19].

Cerium has only one peak at 253 nm and therefore it is not shown graphically. Our estimation for the peak height is $100 \text{ M}^{-1} \text{ cm}^{-1}$. The extinction of cerium in sulfuric acid given in Ref. [4] is seven times larger. It is known that cerium [2] as well as nitric acid [22] have no absorption bands in the visible range and therefore the latter two components cannot interfere in quantification of praseodymium and neodymium in a mixture. At the same time the extinction coefficients of these REEs are rather low in comparison with cerium. The highest absorption peak of praseodymium, around 445 nm, has an extinction coefficient of $6.8 \text{ M}^{-1} \text{ cm}^{-1}$ while for the highest neodymium peak, around 798 nm, the extinction coefficient equals $8.6 \text{ M}^{-1} \text{ cm}^{-1}$. At the same time, a unique peak of cerium has the extension coefficient that is at least 10 times larger. This should be taken into account in selection of the optical pathlength for the spectra acquisition. Namely, too short a path is

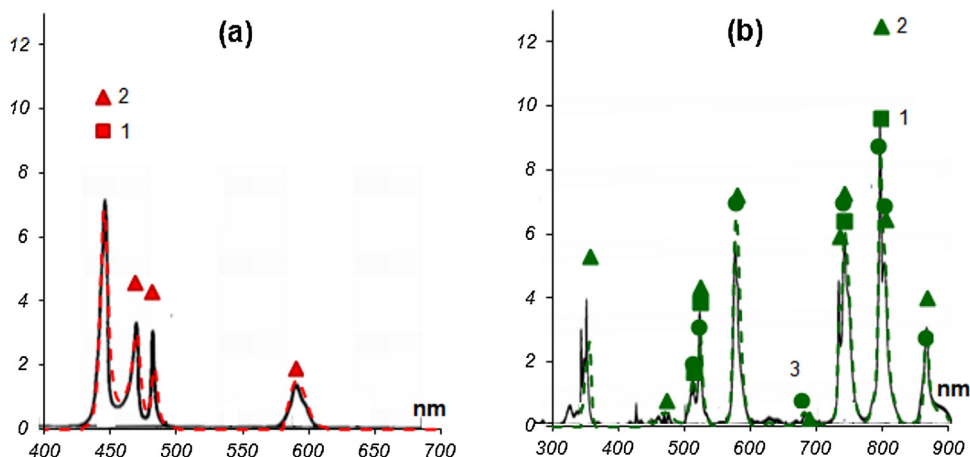


Fig. 2. Extinction coefficients, $M^{-1} cm^{-1}$, left plot (a) is for Pr, right plot (b) is for Nd. Dashed lines represent coefficients recovered by MCR-ALS. Coefficients found in literature: black solid lines are from Ref. [21], squares (1, ■) are from Ref. [1], triangles (2, ▲) are from Ref. [2], and dots (3, ●) are from Ref. [3].

not suitable for praseodymium and neodymium determination, whereas too long a path distorts the cerium peak.

The acquired spectra of mixtures are well-fitted by the recovered spectra in the visual range, see for example Fig. 3.

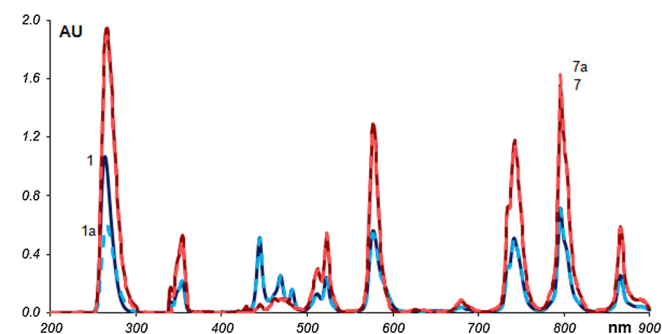


Fig. 3. Results of MCR-ALS. Acquired (solid, 1 and 7) and recovered (dashed, 1a and 7a) spectra for the first (blue, 1) and seventh (red, 7) validation samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For comparison of calibration/prediction results, individual PLS models for Nd and Pr are established. In both models, four LVs clearly manifest themselves as the minimum of RMSEP=0.34, at RMSEC=0.13 for praseodymium and RMSEC=0.08, RMSEP=1.03 for neodymium. Predicted vs. reference plots (Fig. 4) affirm that both methods provide very similar results. It is worth noting that zero, or close to zero concentrations, which sometimes are predicted by PLS as negative values, are estimated by MCR-ALS as zero, due to the non-negativity constraint applied to the concentration profiles (Fig. 4a).

The advantage of MCR-ALS is its well-known ability to present pure spectra of the components. The drawback is a complicated calculation procedure when compared to regression methods. The PLS calibration provides information mostly about the concentration profile. We applied the MCR-ALS for calibration with the intention to establish a model that would be more stable in prediction of the out-of-control cases. In the investigated case, where the neodymium concentration in the validation set is twice higher than that in the training set, no outperformance of MCR-ALS over PLS is noticed. Of course, this is not a general conclusion, but this can be explained by the fact that bi-linearity, as the basic principle both for MCR-ALS and PLS, is obeyed within those investigated conditions.

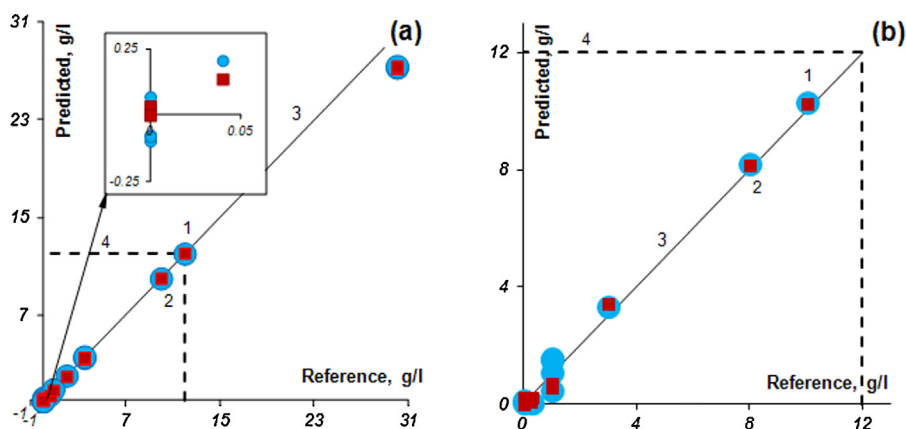


Fig. 4. Prediction of neodymium (a) and praseodymium (b). Dots (1, ●) stand for PLS, squares (2, ■) for MCR-ALS, line (3) is the target line, dashed line (4) represents calibration range.

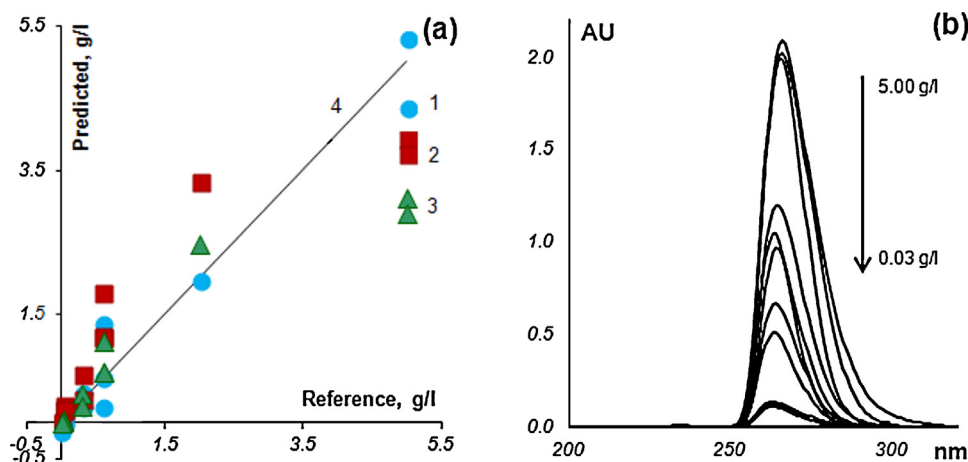


Fig. 5. Cerium determination. (a) Calibration: dots (1, ●) mark PLS, squares (2, ■) mark MCR-ALS, triangles (3, ▲) stand for MCR-ALS-NC, (4) target line; (b) acquired calibration spectra.

4.2. Quantitative analysis of cerium

Spectrophotometric determination of cerium is a much more challenging task.

Both calibration models, constructed with the help of MCR-ALS and the PLS regression in the range of 200–320 nm demonstrate low quality and obvious non-linearity in the predicted vs. reference plot (see Fig. 5a). The PLS method suggests application of three LVs. A summary of the results is presented in Table 4. The reasons of such a poor modeling become clear after a more detailed analysis of the acquired spectra.

It is known that cerium has no absorbance bands in the visual and near-infrared range, but has a large absorption peak [4] near 253 nm with an extinction coefficient of about $700 \text{ M}^{-1} \text{ cm}^{-1}$. Thus, measurements in the UV range provide a unique chance to determine cerium in nitric acid using the in-line spectrophotometry. However, the calibration results manifest a violation of the Lambert–Beer law. To understand the reason, we should note that the aqueous solution of HNO_3 has two partly overlapping peaks around 247 nm and 295 nm. The detailed description of the behavior of flattened HNO_3 spectra and quantitative determination of nitric acid concentration in aqueous solutions is described in Ref. [22]. In case we had an ‘ideal’ instrument, the spectrum of 0.7 M nitric acid would look like the dashed curve (1) in Fig. 6a.

The peak height at 247 nm is 18.8 AU, and at 295 nm the peak height is 4.4 AU. Due to a detector limitation we acquire a distorted spectrum (curve 2) flattened at the level of 3 AU. Taking this into account, we can explain the non-linear behavior of the cerium absorbance against the concentration of the component. Let us consider a spectrum of 0.4 g L^{-1} of cerium acquired with 0.7 M HNO_3 as a spectral background. Instead of an ‘ideal’ spectrum

Table 4
Results of quantitative determination of Ce by PLS (3 LVs), MCR-ALS (3 components), MCR-ALS-NC (3 components).

	PLS	MCR-ALS	MCR-ALS-NC
Calibration			
RMSEC	0.33	0.76	0.84
Slope	0.97	0.73	0.60
Offset	0.04	0.51	0.26
R^2	0.97	0.84	0.88
RE (%)	15	35	39
Validation			
RMSEP	0.70	0.89	0.67
Slope	1.09	0.88	0.69
Offset	0.26	0.58	0.34
R^2	0.86	0.73	0.77
RE (%)	43	54	41

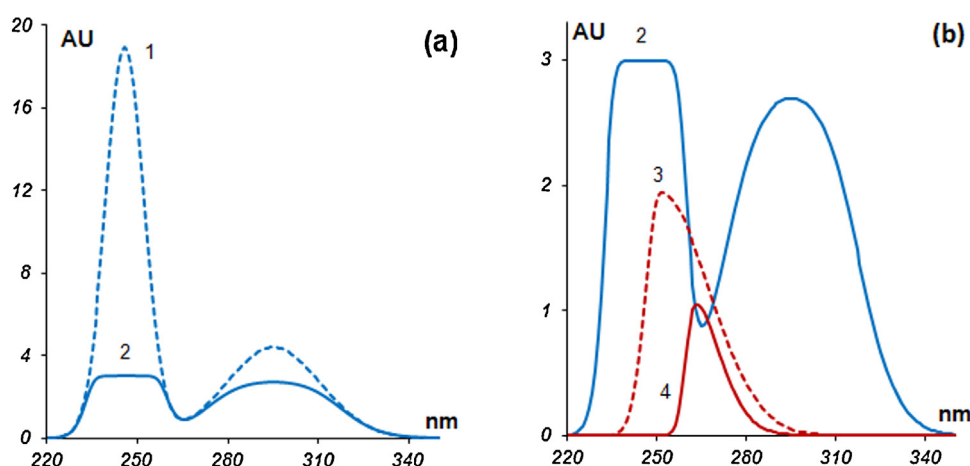


Fig. 6. Spectra of HNO_3 and Ce in the UV region: ideal (1) and saturated (2) spectra of 0.7 M HNO_3 , ideal (3) and acquired (4) spectra of Ce (0.4 g L^{-1} in 0.7 M HNO_3).

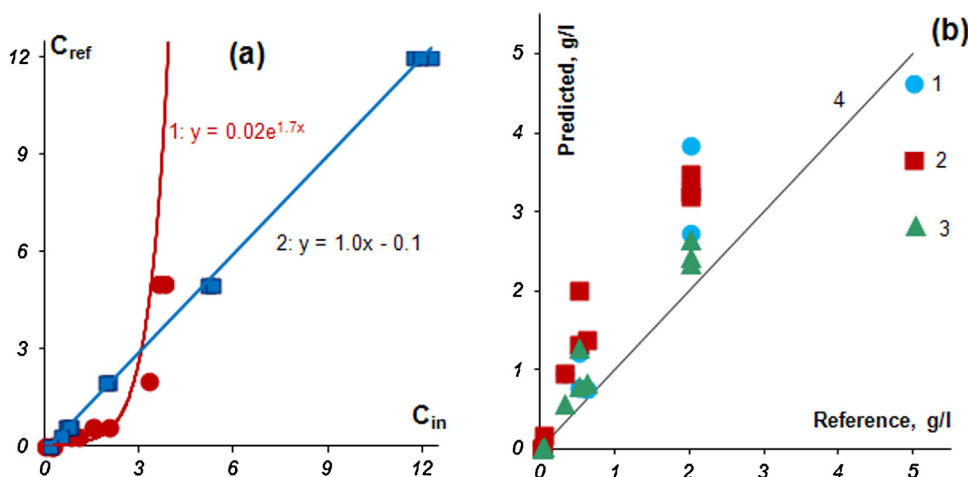


Fig. 7. Cerium determination. (a) Correlation constraint in MCR-ALS method. Dots (1, ●) stand for Ce, squares (2, ■) for Nd. (b) Validation set, dots (1, ●) mark PLS, squares (2, ■) mark MCR-ALS, triangles (3, ▲) stand for MCR-ALS-NC.

(curve 3) with a peak at 253 nm, we acquire a distorted spectrum (curve 4) with a peak shifted to 268 nm. The latter curve is the difference between the saturated spectrum of Ce in HNO_3 and the saturated spectrum of HNO_3 as a background. This distortion explains a non-linear pattern presented in Fig. 5, as well as the cerium peak shortage shown in Fig. 2.

In the MCR-ALS calibration, a non-linear dependence between C_{in} and C_{ref} values is clearly seen in Fig. 7a.

The conventional MCR-ALS method applied to the calibration of cerium produces point set 1, which should, but does not fit in with a linear regression. For comparison, a similar point set 2, obtained by calibration of neodymium, is also shown. This set follows a linear regression quite well. In case of cerium, it seems natural to replace the linear dependence given by Eq. (9), with a nonlinear, e.g. with polynomial, or exponential regression (Eq. (11)) that is shown by curve 1 in Fig. 7a. This technique (labeled further as MCR-ALS-NC) is applied for cerium calibration. The results are shown in Fig. 5a by triangles marked as 3. Figures of merit are presented in Table 4, last row. It shows that the calibration results are worse but the validation is a little bit better due to a more precise prediction in the range from 0 to 2 g L^{-1} . In general, the application of a non-linear constraint does not provide an essential improvement (see Fig. 7b). The obtained results would be better to use for qualitative or semi-quantitative analysis rather than for the determination of cerium concentration.

5. Discussion

As discussed above, for practical implementation, it is important to assess the stability of the models in out-of-control cases.

For this purpose we use a special test set, in which the samples have different concentrations of HNO_3 (see last part of Table 2) but identical concentrations of REEs. The results are presented in Fig. 8. It is known that certain spectral bands of the solution of trivalent lanthanides are sensitive to the changes in the acid concentrations, i.e. the peak of praseodymium at 575 nm. These bands are called “hypersensitive” [2]. The other characteristic bands are not so sensitive. As a result, the prediction of praseodymium and neodymium test samples is less accurate than for validation samples but still satisfactory. This is due to the employment of the spectra in the whole visual range, but not in several characteristic bands alone [9]. PLS works slightly better for neodymium, whereas MCR-ALS predicts praseodymium a little bit better. We can

conclude that changes in nitric acid concentration in the range 0.1–4 M do not have an essential influence on quantitative determination of these REEs.

Determination of cerium is of the other kind. The application of PLS for prediction of the test samples shows the worst result among the three methods. Results of the MCR-ALS prediction are better but still unsatisfactory. This outcome can be explained by the fact that both methods, PLS and MCR-ALS are designed to work in the frame of bi-linearity. Our attempt to adapt the MCR model by

Table 5
Test set. Performance of various methods.

Methods	Pr	Nd	Ce
PLS	15%	7%	210%
MCR-ALS	11%	7%	142%
MCR-ALS-NC	–	–	84%

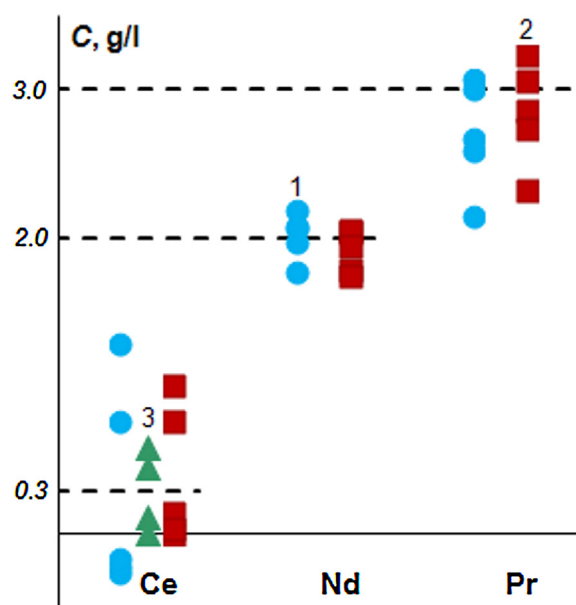


Fig. 8. Test set prediction. Dots (1, ●) stand for PLS, squares (2, ■) for MCR-ALS, and triangles (3, ▲) for MCR-ALS-NC (Ce only). The black dotted lines refer to the reference concentration values.

application of the non-linear constraint had a modest success. The summarized results for the test set prediction are presented in Table 5.

6. Conclusions

The study confirms that spectrophotometry together with chemometric data processing is an appropriate tool for the on-line control of REEs in complex mixtures. Chemometric modeling is more or less complicated depending on the absorbance properties of the specific REE. Established models are stable to some extent for out-of-control cases.

Models for the quantitative determination of neodymium and praseodymium in ternary mixtures with cerium in nitric acid are developed using the MCR-ALS and PLS methods. The predictive ability of the models assessed with the help of the specially designed validation and test sets is very similar. The anticipation that MCR-ALS will outperform PLS for the extreme cases of prediction out of concentration range, or for the changes in the nitric acid concentration, does not justify our expectations. This can be explained by the fact that bi-linearity between the experimental data and the concentrations of the components/pure spectra that is in the background of both methods, is preserved for the analysis of extreme cases. At the same time the MCR approach in addition to Pr and Nd quantification provides information regarding the pure spectra. Thus, both techniques can be reliably used for the prediction of neodymium and praseodymium in nitric acid by means of the in-line spectrophotometry. There have been several studies where the performance of PLS and MCR-ALS for calibration of complex mixtures are compared [14,15,18]. The result is that the predictions made by both of these methods are comparable, which agrees with our results.

Application of spectrophotometry for the on-line prediction of cerium concentrations in nitric acid is a hard problem, due to the high absorptivity of nitric acid that overlaps with cerium absorption peaks. Both methods provide poor prediction. Our exercise in damage limitation by introducing a non-linear constraint into MCR-ALS gave very modest results. Thus, in case of cerium, we yielded partly negative result, which we consider worth reporting. The developed models may only be used for a semi-quantitative analysis. A solution to the HNO_3/Ce problem may be sought by various methods. One of the directions is to change the experimental set-up. Another possible approach is the application of various non-linear calibration methods.

We consider the recently proposed non-linear multivariate curve resolution (NL-MCR-ALS) method [22] as the most prospective approach for the analysis of the cerium/nitric acid mixtures. The preliminary results show that this technique

provides a good calibration accuracy; it is stable to some extent for out-of-control cases.

References

- [1] A.W. Wylie, Spectrophotometric determination of praseodymium, neodymium and samarium, *J. Soc. Chem. Ind.* 69 (1950) 143–147.
- [2] W.T. Carnell, The absorption and fluorescence spectra of Rare Earth ions in solution, in: K.A. Gschneider Jr., L. Eyring (Eds.), *Handbook on the Physics and Chemistry of Rare Earths*, North-Holland, Amsterdam, NL, 1979.
- [3] J. Foos, A.S. Kertes, M. Peleg, Spectra of neodymium(III) in nitrate–water and nitrate–chloride melts, *J. Inorg. Nucl. Chem.* 36 (1974) 837–840.
- [4] H.L. Greenhaus, A.M. Feibush, L. Gordon, Ultraviolet spectrophotometric determination of cerium(III), *Anal. Chem.* 29 (10) (1957) 1531–1534.
- [5] C.J. Rodden, Spectrophotometric determination of praseodymium, neodymium and samarium, *J. Res. Natl. Bur. Stand.* 26 (6) (1941) 557–570.
- [6] J. Hoogschagen, A.P. Snoek, C.J. Gorter, The absorption of light in aqueous solutions of neodymium salts, *Physica* 10 (8) (1943) 693–698.
- [7] F.W. Paul, Absorption spectra of cerium, neodymium and samarium, *Phys. Rev.* 49 (1936) 156–162.
- [8] O. Stoyanov, N.A. Chivireva, I.V. Stoyanova, E.V. Timukhin, V.P. Antonovich, Indirect photometric determination of europium(II) in some inorganic materials containing aliovalent europium species, *J. Anal. Chem.* 66 (2011) 470–475.
- [9] S.A. Bryan, T.G. Levitskaia, A.M. Johnsen, C.R. Orto, J.M. Peterson, Spectroscopic monitoring of spent nuclear fuel reprocessing streams: an evaluation of spent fuel solutions via Raman, visible, and near-infrared spectroscopy, *Radiochim. Acta* 99 (2011) 563–571.
- [10] D. Kirsanov, V. Babain, M. Agafonova-Moroz, A. Lumpov, A. Legin, Combination of optical spectroscopy and chemometric techniques—a possible way for on-line monitoring of spent nuclear fuel (SNF) reprocessing, *Radiochim. Acta* 100 (2011) 185–188.
- [11] I.D. Nickson, C. Boxall, A. Jackson, G.O.H. Whillock, A spectrophotometric study of cerium IV and chromium VI species in nuclear fuel reprocessing process streams, *IOP Conf. Ser. Mater. Sci. Eng.* 9 (2010) 012011.
- [12] A.L. Pomerantsev, O. Ye Rodionova, Process analytical technology: a critical view of the chemometricians, *J. Chemom.* 26 (2012) 299–310.
- [13] T. Naes, T. Isaksson, *Multivariate Calibration and Classification*, Wiley, Chichester, 2002.
- [14] R.R. de Oliveira, K.M. Gomes de Lima, R. Tauler, A. de Juan, Application of correlation constrained multivariate curve resolution alternating least-squares methods for determination of compounds of interest in biodiesel blends using NIR and UV–visible spectroscopic data, *Talanta* 125 (2014) 233–241.
- [15] M.C. Antunes, J.E.J. Simao, A.C. Duarte, R. Tauler, Multivariate curve resolution of overlapping voltammetric peaks: quantitative analysis of binary and quaternary metal mixtures, *Analyst* 127 (6) (2002) 809–817.
- [16] A.L. Pomerantsev, *Chemometrics in Excel*, John Wiley & Sons, Hoboken, NJ, 2014.
- [17] S. Wold, M. Sjostroma, L. Eriksson, PLS-regression: a basic tool of chemometrics, *Chemom. Intell. Lab. Syst.* 58 (2001) 109–130.
- [18] A. de Juan, R. Tauler, Multivariate curve resolution (MCR) from 2000: progress in concepts and applications, *Crit. Rev. Anal. Chem.* 36 (2006) 163–176.
- [19] D. Juan, S.C. Rutan, M. Maeder, R. Tauler, Multivariate curve resolution chapters, in: D. Brown, R. Tauler, B. Walczak (Eds.), *Comprehensive Chemometrics*, vol. 2, Elsevier, Amsterdam, 2009 325–344, 473–506.
- [20] T. Azzouz, 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.
- [21] D.L. Ryabchikov, V.A. Ryabukhin, *The Analytical Chemistry of Rare-Earth Elements and Yttrium*, Nauka Publishers, Moscow, 1966 (in Russian).
- [22] A.L. Pomerantsev, Y.V. Zontov, O. Ye. Rodionova, Nonlinear multivariate curve resolution alternating least squares (NL-MCR-ALS), *J. Chemom.* 28 (2014) 740–748.