Available online at www.sciencedirect.com
SCIENCE DIRECT®
Talanta 63 (2004) 287–296
Talanta
www.elsevier.com/locate/talanta

Threshold bootstrap target factor analysis study of neodymium with pyridine 2,4 dicarboxylic acid N-oxide—an investigation of traceability

G. Meinrath [a,b,c], Z. Hnatejko [d], S. Lis [d,*]

[a] *RER Consultants, Schiessstattweg 3a, D-94032 Passau, Germany*
[b] *Technische Universität Bergakademie Freiberg, Institut für Anorganische Chemie, Leipziger Street 29, D-09596 Freiberg, Germany*
[c] *Technische Universität Bergakademie Freiberg, Institut für Geologie, Gustav-Zeuner Street 12, D-09596 Freiberg, Germany*
[d] *Department of Rare Earth, Faculty of Chemistry, A. Mickiewicz University in Poznañ, Grunwaldzka 6, 60780 Poznañ, Poland*

Received 11 April 2003; received in revised form 22 October 2003; accepted 29 October 2003

Abstract

Threshold bootstrap computer-assisted target factor analysis (TB CAT) has been applied to the evaluation of UV-Vis spectra of Nd(III) in aqueous solutions containing varying amounts of pyridine 2,4-dicarboxylic acid N-oxide (dipyr NO). Using the total uncertainty budget concept, probability distributions of the conditional formation constant(s) of the respective Nd(III) dipyr NO species have been evaluated. The experimental data are compatible with two different systems: a single Nd(III) dipyr NO species with $\log \beta_{11} = 2.8 \pm 0.11$ or, alternatively, a two species system with $\log \beta_{11} = 3.7 \pm 0.35$ and $\log \beta_{12} = 5.2 \pm 0.8$. The results seem to favor the single species interpretation. However, a discussion of the influence of spectral correlation indicates that the situation should be considered with caution. The results are compared to a previous TB CAT study on Nd(III) picolinic acid N-oxide interaction illustrating the advantage of arguing in terms of probability density functions. The pyridine N-oxides with substituted carboxylic acids are structurally poorly investigated. Hence, this study calls for more structural information on rare earth pyridine N-oxide derivates.

Keywords: Computer-assisted target factor analysis; Metrology; Thermodynamic data; Lanthanides; Correlation

1. Introduction

In many fields of science and technology exists that the reported value of a quantity is almost meaningless because the value and its uncertainty has been evaluated by methods and procedures that are either inadequate or inadequately reported. To amend the situation, the International Standardisation Organisation (ISO), in conjunction with several institutions including IUPAC, has issued the Guide to the Expression of Uncertainty in Measurement (GUM) [1] with the intention to establish a common basis for evaluation and communication of uncertainties. The GUM guidelines are fairly general. These guidelines establish a basic convention of reporting doubt in the value of a quantity relevant for any field which intends to play a role in the globalized world. Chemistry, being both a fundamental science and an important economic factor cannot avoid to take these guidelines serious. Chemical measurements are playing an increasingly important role in a globalized market place [2]. In response to the GUM, EURACHEM/CITAC has issued recommendations on the applications of GUM guidelines in chemical analysis [3]. The EURACHEM/CITAC recommendations mainly aim at well-defined procedures in analytical chemistry. For complex operations, e.g. the determination of thermodynamic formation constants, these recommendations need to be supplemented. Correlation effects within residuals and parameters, for instance, cannot be accounted for by the procedures proposed ref. [3]. It has been suggested previously to use computer-intensive resampling techniques to account for different types of correlation [4–7].

The study of metal ion interaction with organic ligands has been, still is and continues to be an active field of chemical research. A major reason for these continued interests is the fact that the interaction can be quantitatively expressed by formation quotients K'. These formation quotients are, at given conditions, fundamental constants of nature due their

* Corresponding author. Tel.: +48-61-8291-345; fax: +48-61-8658-008.
E-mail address: blis@amu.edu.pl (S. Lis).

0039-9140/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2003.10.045

relationship with the Gibbs Free Energy of Reaction, $\Delta G'_R$:

$$\Delta G'_R = -RT \ln K' \quad (1)$$

A large number of formation quotients has been reported in literature (e.g. [8–10]) and partly interpreted by linear free energy relationships [11–13]. This is, however, only one side of the medal.

The other side of the medal can be introduced by Hefter's statement that "it is an unfortunate, but only too well known, fact of solution chemistry that different techniques may yield different values of stability constants" [14]. Seen in a more general way, Hefter's statement directs to the doubt that is associated with any value of a thermodynamic constant reported in literature. Not only different methods yield different values for a quantity. Also repetitions of an experimental study will almost always result in a value that is numerically different from the previously reported one(s). In the majority of thermodynamic constants and formation quotients available in the literature, there is either no uncertainty attached or the meaning of the stated uncertainty is unclear. Such data are meaningless as soon as they are intended as a basis of important decisions [15–19]. Important decisions are decisions which affect third parties [20]. To satisfy quality assurance needs for thermodynamic data, references and standards, chains of comparisons and evaluations (traceability chain) and objective criteria for assessing the information content of experimental data are required.

Since values of thermodynamic quantities are derived from analytical procedures, the GUM guidelines also refer to these data. The details of their application, the prospects, and the possible limitations are a recent field of research. Our interest focuses on several aspects. First aspect is the relative importance of different uncertainty contributions, termed Type A and Type B uncertainties, respectively, in the GUM. Second aspect are the numerical and statistical properties of different algorithms and procedures commonly applied in the evaluation of thermodynamic data from analytical informations. In the contest of spectroscopic methods, we are interested in the contributions by several types of correlation, namely correlations in residuals and correlations in parameters (termed spectral correlation in the following). A further aspect are the implications regarding traceability and comparability of the evaluated informations.

Previous investigations are continued using a UV-Vis spectroscopic study of neodymium with pyridine 2,4 dicarboxylic acid N-oxide as a working example. Our main interest is on the numerical information about the interaction of the metal ion with the ligand available from spectral changes observed in 12 UV-Vis spectra as a function of varying ligand concentrations. In the course of our investigations, different metrological aspects of thermodynamic data have been investigated [5,6,21–23] from which the numerical technique of threshold bootstrap computer-assisted target factor analysis (TB CAT) developed. Factor analysis is a common chemometric technique to evaluate linear equations interpretable as additions of product terms [24]. In case of UV-Vis spectroscopic data, these data structures result from Bouguer–Lambert–Beer law

$$A = EC + N \quad (2)$$

where A is an $n \times m$ matrix of experimentally observed absorptions in m solutions at n wavelengths, E an $n \times p$ matrix of molar absorptions of p species at n wavelengths and C a $p \times m$ matrix holding concentrations of p species in m solutions. N is an $n \times m$ matrix of experimental residuals. These residuals are estimates of the errors. To resolve the experimental data matrix A into three matrices E, C and N is the task of factor analysis.

Target factor analysis is commonly a two-step procedure. In the first step, the abstract eigenvalues and eigenvectors of the absorption matrix A are obtained:

$$A = E^{\#} C^{\#} \quad (3)$$

The so-called 'abstract extinctions' (columns of eigenvector matrix $E^{\#}$) and 'abstract concentrations' (rows of eigenvector matrix $C^{\#}$) are void of all physical meaning. However, the factors (eigenvalues) are arranged in an hierarchical order: the factor explaining the highest amount of experimental variance is given in the first column and first row, respectively, while the column/row vector pair explaining the minute amount of experimental variance are found in the last column and row, respectively. The number of significant eigenvectors is termed data rank, while the remaining eigenvectors are considered as noise and commonly discarded from analysis [25]. There exist a variety of techniques to estimate the data rank, say, p.

Once a value for data rank p is available, the absorption matrix A will be recalculated by the first p column and row eigenvectors only, resulting in

$$A' = E^{\#} p C^{\#p^{-}} \quad (4)$$

Consequently

$$A - A' = N. \quad (5)$$

To obtain estimates for the physically meaningful quantities 'molar absorption' $\varepsilon(\lambda_j)$ and 'species concentration' c_i in each sample, the abstract eigenvectors must be transformed. This process is called target transformation according to

$$A' = E^{p} C^{p} = E^{\#p} T T^{-1} C^{\#p}. \quad (6)$$

The matrix T needs to be found be the experimenter. A large number of strategies are available (e.g. [26–30]) at various levels of sophistication.

While the majority of approaches estimates suitable factors (either columns of $E^{\#p}$ or rows of $C^{\#p}$), computer-assisted factor analysis takes another approach by directly estimating the rotation matrix T using numerical constraints like non-negativity of extinctions and absorptions and the spectral information of known component(s) like the uncoordinated metal ion. The first major reason to take this approach is the usually small dimensionality of T. If a chemical

system holds three species, T is 3×3 and only six elements in T have to be estimated [21,31]. The diagonal elements of matrix T are fixed to a non-zero value in order to avoid the trivial solution where all elements of T are zero. Because nothing can be a priori specified about the signs of the diagonal elements, the correct combination of positive and negative signs in the diagonal elements are found by testing all feasible permutations [21,31].

From an analysis of the distributions of data in noise matrix N some clues about the structure of residuals can be obtained. In classical factor analysis error theory (e.g. [24,32]) the errors in N are commonly assumed to be uncorrelated and identically and independently distributed (i.i.d.) with mean zero and standard deviation σ. In reality, however, spectral residuals are both non-i.i.d. and correlated [7]. While E and C are considered to be a property of a given chemical system, N is a random matrix. Hence, conclusions about E and C will include an element of randomness since $A = f(N)$. Consequently, the evaluated single component spectra in E and the respective species concentrations in C will vary randomly. Since correlation introduced bias into the evaluation procedure, the random distribution of E and C cannot be obtained from a statistical procedure based on the assumption of normality. In addition, since the information on the thermodynamic quantities has to be retrieved from the species concentrations in C which in turn depend on single component spectra properties like molar absorptions calculated in E, randomness in A also affects the formation quotients in a complicated manner. By classical error progression analysis, these multiple mutual correlations cannot be treated.

Alternatives are computer-intensive resampling techniques like bootstrap methods [33–35]. Combining computer-intensive resampling techniques with correlation-preserving resampling techniques like the threshold bootstrap [31,36] several numerical effects on an evaluated value of a quantity can be taken into account [4]. Finally, it is necessary to incorporate these statistical methods into an almost automatic algorithm for factor analysis which has been achieved in the computer-assisted target factor analysis (CAT) method resulting in the threshold bootstrap computer-assisted target factor analysis (TB CAT). There is a second major reason for the CAT approach: the algorithm must be able to proceed without human interference to perform a large number (about 1000–2500) of repeated cycles with slightly varying input data sets. Details of design and examples of application are given elsewhere and will not be repeated here [7,21,22,31].

Our interest in applying chemometric techniques is motivated by their potential role in metrology in chemistry. Hence, we are not focusing on the presentation of ever more complex spectroscopic systems. CAT, especially in its extension to TB CAT, is a tool to study the likely variability to be encountered in experimental data and its progression into the output of a complex data analysis procedure.

The TB CAT procedure allows application of metrological principles, especially evaluation of the total uncertainty budget [1,37], to the evaluation of quantities of chemical thermodynamics. Metrologically meaningful thermodynamic data in turn imply comparability, transparency and mutual recognition of measurement results in conjunction with cause-and-effect diagrams [3,21,38]. In turn, analysts have to accept that the reliability of a result has been usually overestimated in the past [39] and that statistics has to have a larger share in the evaluation of chemical data to avoid the unpleasant experience of disagreement in values reported from the study of the same quantity by different laboratories and techniques as mentioned by Hefter [14]. The large number of thermodynamic quantities with none or rather small uncertainties available in literature being not comparable among each other certainly must not considered as a highlight in chemistry [15,16,39,40].

2. Experimental and computational details

Neodymium perchlorate stock solution was prepared by dissolving Nd_2O_3 (spectroscopically pure, prepared in Department of Rare Earths, Faculty of Chemistry, Adam Mickiewicz University) in perchloric acid. Pyridine 2,4 dicarboxylic acid N-oxide was prepared as following. A mixture of pyridine dicarboxylic acid (5 g), potassium hydroxide (5 g KOH and 20 ml H_2O), acetic acid (150 ml) and 30% hydrogen peroxide (30 ml) was heated on a boiling water-bath for 6 h. Then a saturated solution of $Pb(OAc)_2$ (100 ml) was added. The resulting complex was filtered off and washed with hot water to pH 7. Free acid was liberated by passing hydrogen sulfide gas into a suspension of the complex in water (500 ml). The precipitated product was recrystallized from water or ethanol. Pyridine 2,4-dicarboxylic acid N-oxide was obtained with a 60% yield.

Working solutions were prepared by mixing the appropriate amount of aquoues solutions of pyridine 2,4-dicarboxylic acid N-oxide with the Nd(III) solutions. The Nd(III) concentration was kept fixed at 2×10^{-1} mol l^{-1}. Sodium perchlorate as a solid form (Aldrich, analar grade) was used in adjusting ionic strength. All solutions were prepared in volumetric flasks of 10 ml volume and had a final ionic strength of 0.1 M $NaClO_4$. The pH was measured with a combination glass electrode assembly (OSH-10-10 electrode and CP-315 meter, ELMETRON Co.) and the pH (5.2 ± 0.2) was kept constant by addition of either NaOH or $HClO_4$. The electrode system was calibrated by a two-point calibration.

The absorption spectra of Nd(III) complexes were obtained by means of a Shimadzu UV-2401 PC spectrophotometer. The spectra of sample solutions with varying molar ratios of the components (metal:ligand) were recorded at spectral slit width of 1.0 nm. Experiments were conducted at room temperature.

Nd(III)-pyridine 2,4-dicarboxylic acid N-oxide (dipyr NO) ligand interactions were studied by UV-Vis spectroscopy using the absorption band at 540–620 nm,

corresponding the hypersensitive transition $^4I_{9/2} \rightarrow {}^4G_{9/2}$ ($\lambda_{max} \sim 575$ nm). The Nd concentration was kept constant (2×10^{-3} mol l^{-1}). The concentration of pyridine 2,4-dicarboxylic acid N-oxide increased in the series of 12 experiments (as the molar ratio of the components (dipyr NO: Nd) from 0 to 8.0×10^{-3} mol l^{-1}. The UV-Vis spectra were recorded with 0.05 nm steps in digital form using quartz cuvettes of 5 cm path length. Four repeated scans were averaged to reduce the spectral noise.

The TB CAT algorithm has been implemented in a BASIC dialect. 2500 repetitions were performed to obtain the empirical probability densities. In each repetition, 2500 estimates for the respective formation constant(s) were obtained. The formation constants were weighted. The weight is the ratio between metal ion and the respective species. Thus, the influence of spectra where one of the components is available only in minor quantities is reduced. Such values are the more strongly influenced by small uncertainties and therefore less reliable. The 2500 values obtained for each formation constant in each resampling cycle are ranked and transformed into probability distribution functions [41].

The evaluation procedure follows the sequence given in Ref. [21]. The cause-and-effect diagram presented there is a basis of the following discussion. The Type B uncertainties considered in the evaluations are the uncertainty in the Nd(III) concentration with $u(\text{Nd}) = 7\%$, the uncertainty in the ligand concentration with $u(\text{dipyr NO}) = 5\%$ and an uncertainty in the volume of the additions of $u(\text{vol.}) = 2.5\%$. Other contributions are considered as negligible as discussed in detail previously [21].

3. Results and discussion

Complexes of lanthanide ions with pyridine carboxylic acid N-oxides have been a subject of several publications [42–47]. Systems bearing *N*-oxide groups furthermore form more stable complexes with lanthanides than the parent ligands. However, little has been published about the structure of the complexes [44,45]. Yoneda et al. studied the series of Ln(III) ions with picolinic acid N-oxide [42]. They concluded that picolinic acid N-oxide forms with Ln(III) chelates of six-membered rings and determined the stability constants of log β_{11} and log β_{12}, for Nd(III) the values are 2.91 and 5.06, respectively. Boyd et al. prepared and characterized by chemical analysis and UV and IR spectroscopy complexes of several Ln(III) ions (Ln = Ln, Pr, Sm, Dy, Ho Er and Yb) with picolinic acid N-oxide [43]. Studies indicated formation of cationic $[\text{Ln(HL)}_2\text{L}_2]\text{ClO}_4$ complexes in which this ligand is bidentate coordinating to the Ln(III) ions via N-oxide and carboxyl O atoms. Yan et al. described the synthesis and and X-ray crystal structure of the binuclear complex of 6-methylpicolinic (6-mepicno) acid N-oxide with La(III) [44]. In the complex of formula $\text{La}_2(\text{6-mepicno})_6 \cdot 6\text{H}_2\text{O}$ the coordination number of lanthanum is 9. Polymeric forms of Eu(III) and Er(III) complexes ions with coordination number of 8 with N-oxides of nicotinic and isonicotinic acids have been determined by Mao et al. [45]. Crystal structure and factors affecting the formation of coordination chains and dinuclear units of Eu(III) and Er(III) with nicotinic acid N-oxide of general formula $[\{\text{Ln(L)}_3(\text{H}_2\text{O})_2\}n] \cdot 4n\text{H}_2\text{O}$ and with isonicotinic acid N-oxide $[\{\text{Ln(L)}_2(\text{H}_2\text{O})_4\}_n](\text{NO}_3)_n \cdot n\text{H}_2\text{O}$ and $[\{\text{Ln}_2(\text{L})_4(\text{H}_2\text{O})_{10}\}](\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$, respectively, are presented. In our recent papers we described the synthesis of a series of pyridine carboxylic acid N-oxide ligands, L, and spectroscopic characterization of their Nd(III) and Eu(III) complexes and mixed ligand complexes. We have indicated the formation of LnL, LnL_2 and LnL_2X (where X is phen or byp) complexes in solutions and in the solid phase $[\text{LnL}_3(\text{H}_2\text{O})_2] \cdot 2\text{H}_2\text{O}$, using absorption, IR and luminescence spectroscopy, luminescence lifetime measurements and chemical analysis [46,47].

The interaction of Nd(III) with dipyr NO will be discussed on basis of the 12 experimental spectra shown in Fig. 1. The steady increase in absorption can be recognized readily as is the continuous variation with shape. A preliminary study on the likely amount of factors indicated two possibilities with either 2 or 3 components.

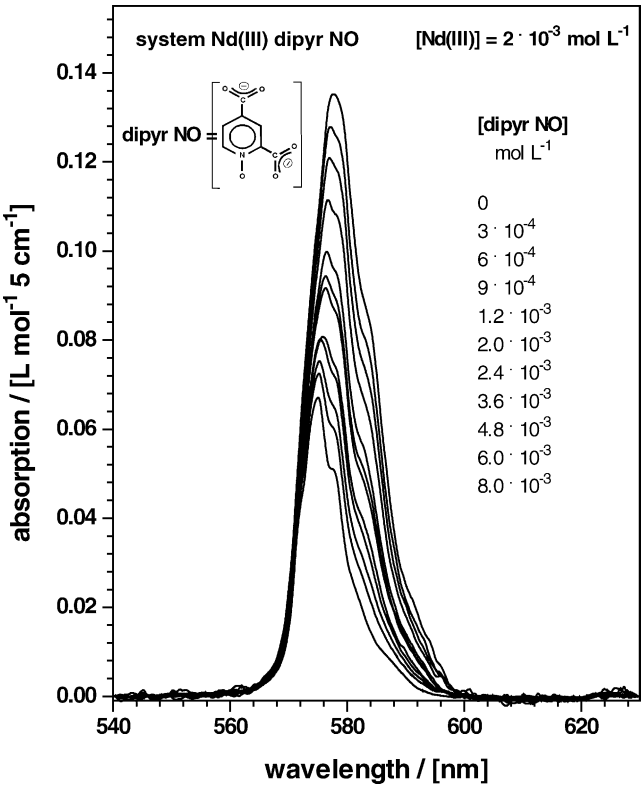

Fig. 1. Twelve experimental spectra of Nd(III) with varying concentrations of pyridine 2,4 dicarboxylic N-oxide (dipyr NO) at pH 5.2 ± 0.2 in 0.1N perchlorate medium at 22 ± 2 °C. The spectra are obtained by averaging four repeated scans.

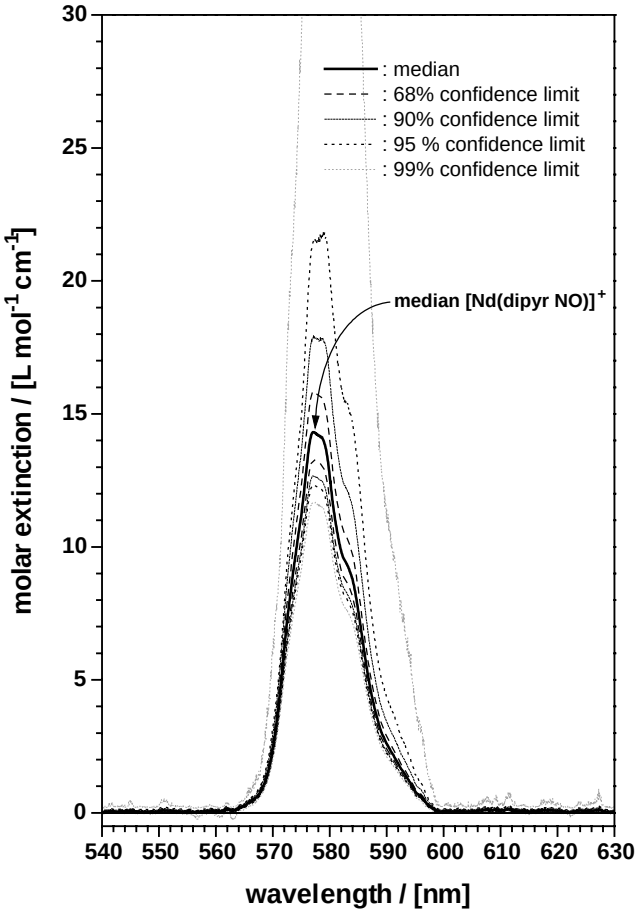

Fig. 2. Derived single component spectrum of species Nd(dipyr NO)$^{+}$ together with upper and lower confidence limits of 68, 90, 95 and 99% probability. The upper confidence limit shows a considerable spread and clearly shows a skewed distribution. The distribution of molar extinctions $\varepsilon(\lambda)$ is clearly skewed towards higher extinctions in the maximum of the spectral curve. The molar extinction in the median absorption maximum is $\varepsilon(577\,\text{nm}) = 14.2\,\text{l mol}^{-1}\,\text{cm}^{-1}$ with a lower 68% confidence limit of $\varepsilon(577\,nm) = 13.2\,\text{l mol}^{-1}\,\text{cm}^{-1}$ and an upper confidence limit of $\varepsilon(577\,\text{nm}) = 15.8\,\text{l mol}^{-1}\,\text{cm}^{-1}$.

3.1. Two-component model

In Fig. 2, as a result of the TB CAT analysis, the single component spectrum of a species Nd(dipyr NO)$^{+}$ is given together with four upper and lower confidence limits obtained from the empirical probability distribution after 2500 resampling cycles.

It is obvious that the distributed information about molar extinctions will result in distributed data for the calculated formation constant of the species Nd(dipyr NO)$^{2+}$. The empirical probability density of log $\beta(_{\text{NddipyrNO}})$ for the reaction

$$\text{Nd}^{3+} + \text{dipyr NO}^{2-} \leftrightarrow \text{Nd(dipyr NO)}^{+} \quad (3)$$

is given in Fig. 3. This density is obtained by ranking the 2500 values of log $\beta(_{\text{NddipyrNO}})$ obtained for each bootstrap cycle, evaluating the cumulative probability distribution and

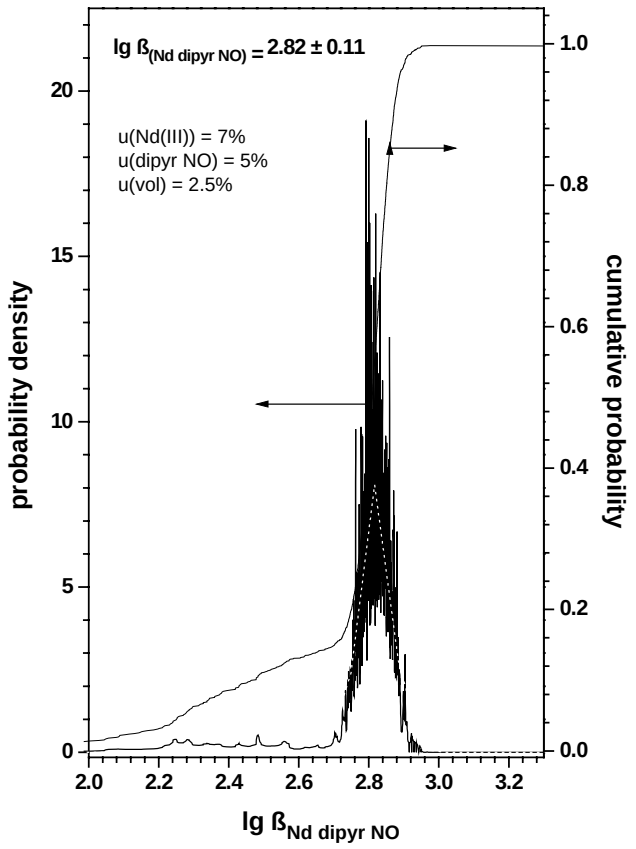

Fig. 3. Probability density (left scale) and cumulative distribution function (right scale) of the quantity log $\beta(_{\text{NddipyrNO}})$. The median is $\log \beta(_{\text{NddipyrNO}}) = 2.82$ with a 68% confidence region of ± 0.11. The curve is almost normally distributed.

taking the first derivative using a moving window regression method [41].

3.2. Three-component model

All information about the number of differently coordinated neodymium ions in solution is an indirect one. Factor analysis is principally capable to provide a clue to the likely number of different species in solution. A large number of such statistical test, e.g. empirical formula [24], eigen value analysis [25], F tests [48], correlation analysis [49] and canonical correlation [50], have been proposed. This—a small selection of possible references only—abundance of tests directs to the same dilemma: different techniques give different answers. CAT, as an un-supervised technique, studies the outcome of different assumptions about the likely number of species and indicates settings. There is no reason to discard some of the proposed interpretations. For the 12 Nd(III) dipyr NO solutions, a second possible interpretations has been suggested by the CAT algorithm where two coordinated species are available in solution. The single component spectra resulting from an analysis along this suggestion assuming the formation of a 1:1 Nd(III) dipyr NO species and a 1:2 Nd dipyr NO species, are given in Fig. 4a and b.

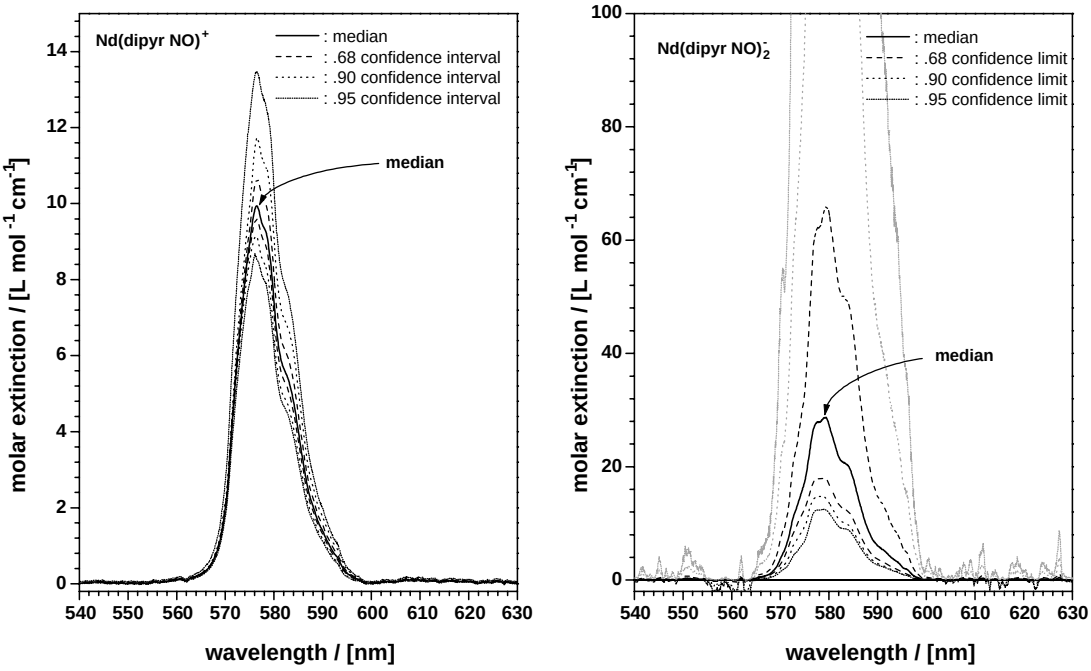

Fig. 4. Derived single component spectra of a species $Nd(dipyr\ NO)^{+}$ and $Nd(dipyr\ NO)_2^{-}$ together with the 68, 90 and 95% confidence regions. While the proposed spectrum for $Nd(dipyr\ NO)^{+}$ is quite narrow distributed, the proposed $Nd(dipyr\ NO)_2^{-}$ spectrum is strongly skewed towards high molar absorption.

The derived single component spectrum of $Nd(dipyr\ NO)^{+}$ is narrowly distributed. The spectral informations, however, are quite different. $Nd(dipyr\ NO)^{+}$ has a 68% confidence range of the maximum molar absorption between 9.9 and 10.7 $l\,mol^{-1}\,cm^{-1}$ with a median of 9.6 $l\,mol^{-1}\,cm^{-1}$.

The derived single component spectrum of $Nd(dipyr\ NO)_2^{-}$ has a much wider distribution in its calculated molar extinction with $18.0 < 28.7 < 66.0\,l\,mol^{-1}\,cm^{-1}$. The upper 95% confidence limit is as high as $500\,l\,mol^{-1}\,cm^{-1}$: a rather unusually large value for a rare earth aqueous species. However, it must be emphasized that the uncertainty in the molar extinctions takes into account all contributing uncertainties including the uncertainty in the likely chemical composition. It must not be confused with data obtained from evaluations where most uncertainty contributions are not considered at all.

In Fig. 5, a typical spectral deconvolution of an experimental spectrum by one set of the proposed single component spectra is shown using a solution with $2 \times 10^{-3}\,mol\,l^{-1}$ total concentration of pyridine 2,4-dicarboxylic acid N-oxide as an example. The 68% confidence region is obtained by classical least squares analysis. Fig. 5 shows that the alternative interpretation by two different coordinated species is not an unacceptable one—at least on the level of the spectral curves.

Fig. 6 presents the distributions of formation quotients from 2500 resampling cycles.

The distribution of log $\beta_{Nd(dipyrNO)}$ is comparatively narrow while the distribution of log log $\beta_{(Nd(dipyr\,NO)_2)}$ is more or less sloppy. Fig. 6 seems to argue against the

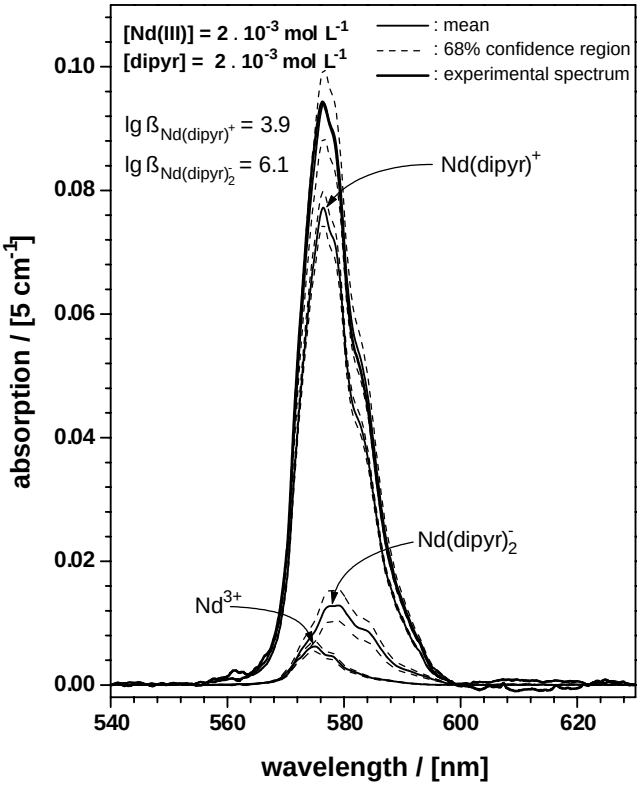

Fig. 5. Spectral deconvolution of an experimental spectrum of a solution holding pyridine 2,4 dicarboxylic acid N-oxide (dipyr NO) with $2 \times 10^{-3}\,mol\,l^{-1}$. From such an analysis, formation constants log $\beta_{(NddipyrNO)} = 3.9$ and log $\beta_{(Nd(dipyr\,NO)_2)} = 6.1$ result.

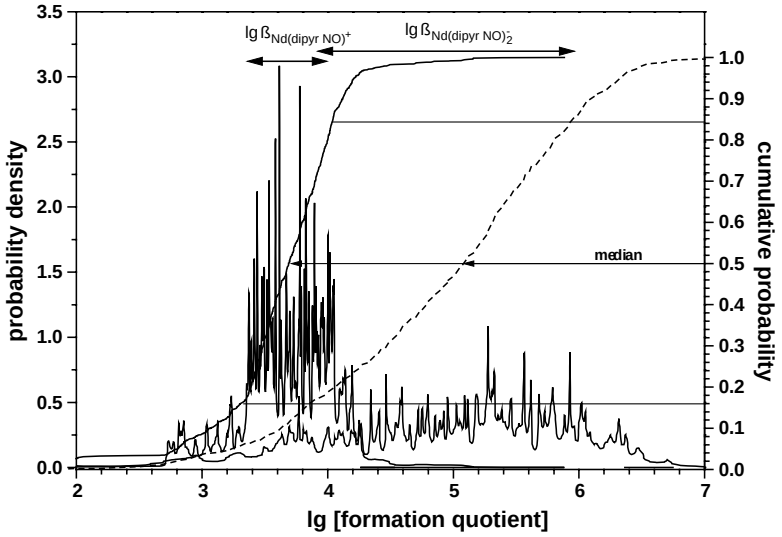

Fig. 6. Distribution of the formation quotients resulting from an interpretation of the experimental data by two coordinated species (three-component system). The cumulative distributions are given right, the probability densities left.

three-component interpretation. Systems with more factors are commonly expected to fit a given data set better than a system with less factors—and Fig. 6 is not in favour of such an interpretation.

There is, however, the experience that more factors do not only interpret a noisy data set better but also introduce considerable correlation. Spectral correlation is certainly one reason for the spread in the experimental results. Table 1 presents the correlations between the spectral curves for the both interpretations.

Each interpretation has highly correlating spectral curves. Correlation coefficients above ±0.80 are high and can be detrimental to a least-squares analysis [4,51]. To a large part, the strongly skewed distribution of the proposed molar extinction of a species Nd(dipyr NO)$_2^-$ is a consequence of the nuisance factor spectral correlation. The extreme value of $c = -0.97$ between the spectral curves of Nd(dipyr NO)$^+$ and Nd(dipyr NO)$_2^-$ results from the very close spectral shapes. Such a similarity in the spectral shape is not an unexpected for the 4f solution species [52,53] These small variations in the spectral curves are one of the interesting and challenging features of the 4f element aqueous solution chemistry. There is, however, up to now only qualitative experience what differences can in fact be resolved and what differences are dominated by even minor uncertainties.

Table 1
Correlations c of spectral curves for the three-component and the three-component interpretation of Nd(III) pyridine 2,4-dicarboxylic acid N-oxide interaction

Species	Nd^{3+}	Nd(dipyr NO)$^+$	Nd(dipyr NO)$_2^-$
Nd^{3+}	1	−0.90	
Nd(dipyr NO)$^+$	−0.90	1	
Nd^{3+}	1	−0.23	0
Nd(dipyr NO)$^+$	−0.23	1	−0.97
Nd(dipyr NO)$_2^-$	0	−0.97	1

There is a difference between the interpretation of the system by a two-component or a three-component system. This difference is in the magnitude of the first coordinated species Nd(dipyr NO)$^+$. This species has a formation quotient $\log \beta_{\mathrm{Nd(dipyr\,NO)^+}} = 2.8 \pm 0.11$ in the two-component interpretation, while in the three-component system the same quantity has a value of $\log \beta_{\mathrm{Nd(dipyr\,NO)^+}} = 3.7 \pm 0.35$. These differences are clearly significant as can be appreciated by comparing some values for formation quotients obtained under slightly different numerical regimes and similar compounds given in Fig. 7.

In Fig. 7, the probability density of $\log \beta_{(\mathrm{NdpicNO})}$ obtained for the Nd(III) interaction with picolinic acid N-oxide previously [21] is compared to the corresponding probability densities of formation constants in the N(III) dipyr NO system.

Two distributions are given for $\log \beta_{(\mathrm{NdpicNO})}$. One distribution presents the previously reported distribution where the $\log \beta$ values from all solutions has been given equal weight. Because equilibria where one or more components make only minor contributions are more sensitive to small random fluctuations, it is justified to put more weight to solutions that are well balanced. If such a weighing regime is applied, a more narrow and shifted distribution is observed. Hence, the decision to weigh the formation quotients resulted in more narrow and slightly shifted distribution but still overlapping with the previously reported one for the Nd(III) picolinic acid system [21]. Comparing the both distributions for $\log \beta_{\mathrm{Nd(picNO)}}$ and $\log \beta_{\mathrm{Nd(dipyrNO)}}$ shows that their difference is minor. It may be argued that the carboxylate group in 4-position may provide additional charge to stabilize the hexagonal ring with the coordinating Nd(III) ion. However, experimental evidence or support by

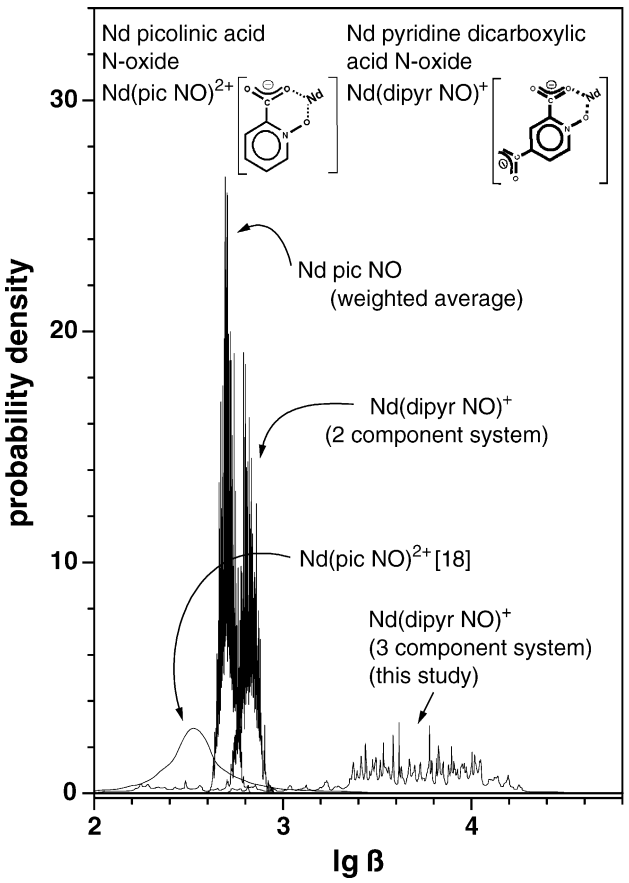

Fig. 7. A comparison of probability densities for the formation constants $\log \beta_{Nd(PicNO)}$ und $\log \beta(_{NddipyrNO})$. Two probability densities are given for $\log \beta_{Nd(PicNO)}$ where the difference is in the weighing of the various data. The other both densities correspond to the compound Nd(dipyr NO)$^{+}$ in a two-component and three-component, respectively. The structures of both ligands are presented in its basic and dibasic form, respectively. For picolinic acid N-oxide, a two component system holding Nd^{3+} and Nd(pic NO)$^{2+}$ has been found as the only feasible interpretation [21]. Please note that the integral over each distribution is unity.

crystallographic data is lacking, yet. This lack underscores the need for a thorough investigation of pyridine carboxylic acid N-oxide interaction with metal ions. The difference between the both distributions is less than the difference between the distributions in log $\beta_{Nd(picNO)}$ obtained from weighted and unweighted pooling of the calculated formation constants. The difference between the distributions from the two-component and the three-component interpretation of the Nd(III) dipyr NO system is clearly larger. If the system Nd(III) dipyr NO in fact should be forming a species Nd(dipyr NO)$_2^-$, it must be concluded that factor analysis is unable to resolve the system quantitatively. The reason for this inability is the high spectral correlation.

Factor analysis is also unable to provide inambiguous evidence about the number of species formed in the Nd dipyr NO system if the major contributions of measurement uncertainty are taken into account. Such a situation is not uncommon [22]. Comparison with studies of similar interactions is a possibility to arrive at some conclusions without making a personally biased decision. There is however the need to have experience about which magnitude of a difference is significant and where the resolution power of an experimental method is overestimated.

4. Conclusions

The few available quantitative studies of lanthanide interaction with carboxylic acid N-oxides have been interpreted without a thorough fundament of structural insight. The interpretation of chemical systems thus often had to rely on complicated mathematical and statistical procedures—while the statistical effects are often either underestimated or completely ignored [39]. Statisticians emphasize the variability of a result [51], while chemometricians emphasize the importance of such tools for looking behind doors otherwise closed [54,55].

This study was motivated by a third reason. We are interested in the reliability or its reciprocal quantity, the doubt, we have to associate with the conclusions forwarded on basis of the analysis of our studies. While the true values of a quantity can be found in nature, the doubt about our knowledge of such a value is a matter between human beings. A normative international guideline has been issued by the respective organizations, among those the IUPAC, and has been transferred into national norms by most industrialized countries [1,56]. The details of a transfer of such agreements into the field of analytical chemistry are still under discussion but this discussion is quite advanced [2,3,37–39]. Hence, there is a need for suitable methods and techniques that are able to contribute to the assessment of uncertainty in chemical measurement. Such methods and techniques can become quite sophisticated. In the mid run, however, such an approach will only be fruitful if it can achieve the major goal of comparability.

Traceability of values to common references, e.g. SI units, requires that all properties of the chemical and numerical procedures are thoroughly understood and quantified. It is, therefore, a difference whether the traceability of a analytical technique is investigated or whether a value of a chemical quantity has actually been traced to accepted references. Hence, the values reported in this study are not traced. The values reported are merely instruments to discuss the properties of the analytical technique. They show the high sensitivity of spectroscopic techniques to spectral correlation. Spectral correlation is an intrinsic property of an electronic transition in a chemical system. There is no way to reduce a given spectral correlation. Hence, the spectral correlation is an important quantity characterizing a spectroscopic system. Spectral correlation is likely to nullify the usefulness of spectral analysis for evaluation of thermodynamic data of those lanthanides where only transitions with coefficients of correlation > |0.8| are available [4].

Modern statistical tools have become available for chemists to explore their data conditional on the measurement

uncertainties and the epistemological uncertainties. These techniques are so-called computer-intensive statistical techniques. The bootstrap and the jackknife are prominent examples [57]. Computer-intensive statistical techniques replace mathematical sophistication by brute compouting power. However, these tools free the analyst from mathematical constraints and allow to concentrate on the relevant questions.

The present study is a mosaic piece in our investigation of metrological properties of numerical methods for extracting information from spectroscopic data in situations of strong spectral correlation. These studies have led to the development of threshold bootstrap computer-assisted target factor analysis (TB CAT) in combination with a metrological point of view. The concept of a total uncertainty budget plays a predominant role [21]. Consequently, the doubt to be associated with an interpretation of a data set is expressed by distribution curves of physical quantities. The distributions of different interpretations can be quantitatively compared.

Thus, TB CAT intends to contribute to metrology in chemistry where reasonably chosen uncertainties are expected to avoid the unfortunate situation highlighted by Hefter's statement that different methods provide different values for the same constant. The results forwarded from a study are probably less inambiguous—but at the same time also less discrepant with the outcome of other studies. Continuing work will show to what extent this expectation will fulfill.

Last but not least it should be emphasized that the observed limitations in the extraction of information from strongly correlated spectra cannot be overcome by including more spectra into the data analysis. If all relevant measurement uncertainty contributions are included into the analysis, meaningful information will be obtained in a rather limited concentration range only. If there are no independent methods used to assess the concentration of uncoordinated ligand (e.g. measurement of pH in case of the ligand OH^-), its concentration must be indirectly estimated based on the numerical analysis of the Nd(III) spectral signals. As the free ligand concentration decrease, the uncertainty increases drastically. Spectra taken in these concentration ranges will not provide useful information.

References

[1] ISO Guide to the Expression of Uncertainty in Measurement International Standardisation Organisation, Geneva, 1993.
[2] P. de Bièvre, P.D.P. Taylor, Fresenius J. Anal. Chem. 368 (2000) 567–573.
[3] EURACHEM/CITAC Quantifying Uncertainty in Chemical Analysis, second ed. EURACHEM 2000.
[4] Meinrath G., Lis S., Piskula, Z., Anal. Bioanal. Chem., 2004, in press.
[5] G. Meinrath, Fresenius J. Anal. Chem. 368 (2000) 574–584.
[6] G. Meinrath, Chemom. Intell. Lab. Syst. 51 (2000) 175–187.
[7] G. Meinrath, Anal. Chim. Acta 415 (2000) 105–115.
[8] J. Bjerrum, G. Schwarzenbach, Sillèn LG, 1957. Stability Constants of Metal Ion Complexes. Special Publication 6, R. Chem. Soc., London, UK.
[9] R.M. Smith, A.E. Martell, Critical Stability Constants, vol. 6. Plenum Press, New York, USA, 1989.
[10] Anderegg G., Critical Survey of Stability Constants of EDTA Complexes. IUPAC Chem Data Ser 14, Pergamon Press Oxford, UK, 1977.
[11] L.P. Hammett, Chem. Rev. 17 (1935) 125.
[12] B. Zilbergleyt, Z. Russ, J. Phys. Chem. (1983) 864–867.
[13] J.R. Zeevaart, N.V. Jarvis, W.K.A. Louw, G.E. Jackson, R. Mouton, J. Inorg. Biochem. 73 (1999) 265–272.
[14] G.T. Hefter, Critical Survey of Stability Constants and Related Thermodyamic Data of Fluoride Complexes in Aqueous Solution, IUPAC, Pergamon Press, 1984.
[15] O. Nitzsche, G. Meinrath, B. Merkel, J. Contam. Hydrol. 44 (2000) 223–237.
[16] G. Meinrath, S. Hurst, R. Gatzweiler, Fresenius J. Anal. Chem. 368 (2000) 561–566.
[17] C. Ekberg, G. Meinrath, B. Strömberg, J. Chem. Thermodyn. 35 (2002) 55–66.
[18] G. Meinrath, P.M. May, Mine Water Environ. 21 (2002) 24–35.
[19] C. Ekberg, A. Ödegaard-Jensen, G. Meinrath, Ljungskile-a Computer Program for Investigation of Uncertainties in Chemical Speciation. REPORT SKI 03:03, SKI Stockholm/S, 2003.
[20] P. de Bièvre, P.D.P. Taylor, Fresenius J. Anal. Chem. 368 (2000) 567–573.
[21] G. Meinrath, S. Lis, Anal. Bioanal. Chem. 372 (2002) 333–340.
[22] G. Meinrath, S. Lis, S. But, M. Elbanowski, Talanta 55 (2001) 371–386.
[23] G. Meinrath, Anal. Bioanal. Chem. 374 (2002) 796–805.
[24] E.R. Malinowski, Factor Analysis in Chemistry, second ed., Wiley, New York, USA, 1991.
[25] E.R. Malinowski, Anal. Chem. 49 (1977) 606–612.
[26] W. Windig, J. Guilment, Anal. Chem. 63 (1991) 1425–1432.
[27] H. Gampp, M. Maeder, Ch.J. Meyer, A.D. Zuberbühler, Talanta 32 (1985) 1133–1139.
[28] H. Cartwright, J. Chemometrics 1 (1987) 111–120.
[29] E.R. Malinowski, Anal. Chim. Acta 134 (1982) 129–137.
[30] F.J. Knorr, J.M. Harris, Anal. Chem. 53 (1981) 272–276.
[31] G. Meinrath, S. Lis, Fresenius J. Anal. Chem. 369 (2001) 124–133.
[32] Clifford A.A., Multivariate Error Analysis, Appl. Sci., London, UK, 1973.
[33] B. Efron, SIAM Rev. 21 (1979) 460–480.
[34] T. Roy, J. Chemomet. 8 (1994) 37–44.
[35] A.C. Davison, D.V. Hinkley, Bootstrap Methods and their Application. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, UK, 1997.
[36] D. Park, T.R. Willemain, Comp. Statist. Data Anal. 31 (1999) 187–202.
[37] G. Price, Accred. Qual. Assur. 1 (1996) 57–66.
[38] S.L.R. Ellison, V.J. Barwick, Analyst 123 (1998) 1387–1392.
[39] S. Ellison, W. Wegscheider, A. Williams, Anal. Chem. 69 (1997) 607A–613A.
[40] P. de Bièvre, Accred. Qual. Assur. 1 (1996) 91.
[41] G. Meinrath, C. Ekberg, A. Landgren, J.O. Liljenzin, Talanta 51 (2000) 231–246.
[42] H. Yoneda, G.R. Choppin, J.L. Bear, A.J. Graffeo, Inorg. Chem. 4 (1965) 244–246.
[43] S.A. Boyd, R.E. Kohrman, D.X. West, Inorg. Nucl. Chem. Lett. 12 (1976) 603–608.
[44] L. Yan, J. Liu, X. Wang, R. Yang, F. Song, Polyhedron 14 (1995) 3545–3548.
[45] J.-G. Mao, H.-J. Zhang, J.-Z. Ni, S.-B. Wang, T.C.W. Mak, Polyhedron 17 (1998) 3999–4009.
[46] S. Lis, Z. Hnatejko, P. Barczyñski, M. Elbanowski, J. Alloys Comp. 344 (2002) 70–75.

[47] S. Lis, Z. Hnatejko, S. But, A. Szyczewski, M. Elbanowski, Mol. Phys. 101 (2003) 977–981.
[48] E.R. Malinowski, J. Chemometrics 3 (1988) 49–60.
[49] D.N. Rutledge, A.S. Barros, Anal. Chim. Acta 454 (2002) 277–295.
[50] X.M. Tu, D.B. Burdick, D.W. Millican, L.B. McGown, Anal. Chem. 61 (1989) 2219–2224.
[51] N. Graeppi, D.H. Powell, G. Laurenczy, L. Zekany, A.E. Merbach, Inorg. Chim. Acta 235 (1995) 311–326.
[52] S. Lis, S. But, J. Incl. Phenom. 35 (1999) 225–231.
[53] I.E. Frank, J.H. Friedman, Technometrics 35 (1993) 109–135.
[54] S. Wold, Technometrics 35 (1993) 136–139.
[55] R. Hoffmann, V.J. Minkin, B.K. Carpenter, Bull. Soc. Chim. Fr. 133 (1996) 117–130.
[56] BIPM, Reconnaissance Mutuelle, Bureau International des Poids et Mésures. Sèvres, France, 1999.
[57] B. Efron, SIAM Rev. 21 (1979) 460–480.