HAL
open science

Simultaneous quantification of ionic solutions by Raman spectrometry and chemometric analysis

T.H. Kauffmann, M.D. Fontana

To cite this version:

T.H. Kauffmann, M.D. Fontana. Simultaneous quantification of ionic solutions by Raman spectrometry and chemometric analysis. IMEKO XXI World Congress, Aug 2015, Prague, Czech Republic. pp.1782-1786. hal-01227871

HAL Id: hal-01227871

https://centralesupelec.hal.science/hal-01227871

Submitted on 12 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

XXI IMEKO World Congress "Measurement in Research and Industry"
August 30 - September 4, 2015, Prague, Czech Republic

SIMULTANEOUS QUANTIFICATION OF IONIC SOLUTIONS BY RAMAN SPECTROMETRY AND CHEMOMETRIC ANALYSIS

Thomas H. Kauffmann [1,2], *Marc D. Fontana* [1,2]

[1] University of Lorraine, Laboratoire Matériaux Optiques, Photonique et Systèmes, Metz, France
[2] CentraleSupélec, Laboratoire Matériaux Optiques, Photonique et Systèmes, Metz, France

Abstract **-** We report on the detection by Raman spectrometry of several pollutants dissolved in water such as nitrates, sulfates or chlorides. Unlike current water probes often based on conductivity our Raman probe measures simultaneously the concentration of each species in a few seconds only. We show how chemometrics methods based on statistical analysis like the principal components analysis and the partial least square regression can be applied for the spectra treatment.

Keywords: Raman spectrometry, salt solutions, chemometrics, PCA, PLS regression

1. INTRODUCTION

Raman spectrometry (RS) is a well known optical technique generally used to study the vibration of molecular bonds[1]. Based on the inelastic diffusion of the light by a material, RS provides several information about the sample which can be a solid, a liquid or a gas. RS is proposed here as new probe for the simultaneous detection of several salts diluted in water and the determination of theirs concentrations. The different parameters related to a specific Raman line (i.e. position, intensity, full width at half maximum) give information about the sample[2, 3]. Thus, the line positions related to the molecular composition are used to identify species and the peaks intensities give rise to the concentrations. RS is particularly adapted for liquid analysis, allowing fast and in-situ measurements (no sampling) and associated with chemometrics methods[4] provides an efficient probe for water monitoring.

2. MATERIAL AND METHODS

2.1. Raman spectrometry of salt solutions

Raman spectrometry consists in exciting a material with a laser and collecting the light diffused by it. Indeed, although the majority of the scattered light does not exchange energy with the material i.e. no change in wavelength (Rayleigh scattering), a small part of the light interacts with the sample and has therefore a different wavelength (inelastic scattering). This is the Raman scattering which evidences vibration modes of the molecules. A Raman spectrum therefore provides a fingerprint of the molecular bonds of a material and the study of its evolution or its modification provides information about the sample (bonds, concentration, temperature, mechanical stress etc.). Here Raman spectrometry is used as a probe for the detection and the measurement of salt diluted in water with all the advantages of this technique: use of contacless or immersion probe, no sample preparation, in-situ and fast measurement.

Analysis of Raman spectra of salts diluted in water shows a modification of the water spectrum by the anions present in the solution (the OH stretching band in particular) and characteristic peaks[5]. As shown in Fig.1 some of these ions have a specific signature in the wavenumber range w around 1000 cm^{-1} ($\sim$ 30 THz). For example, the line located at 981 cm^{-1} is specific to the sulfate anion SO_4^{2-} and the line at 1048 cm^{-1} to the nitrate anion NO_3^- . In addition to being indicators of the type of ions present in solution, these peaks evolve in intensity I with the concentration of the species so that a calibration can be usually done for each anion[6, 7]. Moreover, other salts such as halides (chlorides for example) do not exhibit specific lines (not active in Raman) but just slightly disturb the wide OH stretching band of the water spectrum lying between 2700 and 3900 cm^{-1} so that the determination of the concentration is more difficult in this case[8, 9] especially since the OH band is also affected by the salts giving direct signatures.

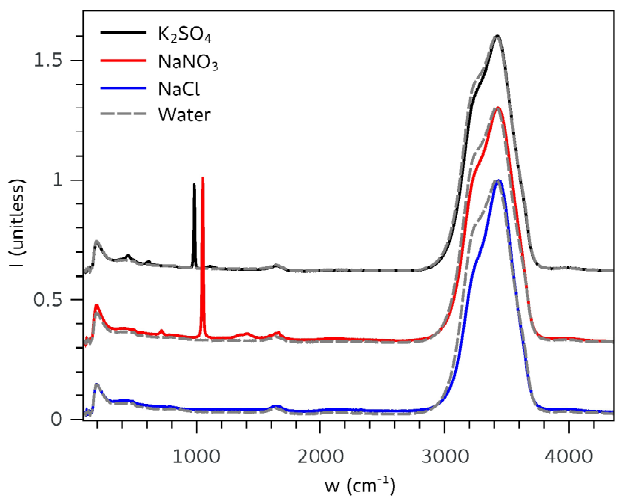

Fig. 1. Normalized Raman spectra (deliberately shifted) of salt solutions showing caracteristic lines of NO_3^- and SO_4^{2-} anions around 1000 cm^{-1} and the wide OH stretching band of the water affected by the anions. Water spectrum is given in dash line for comparison.

2.2. Principal Components Analysis

Chemometric analysis includes all the statistical and mathematical methods for extracting information from a spectrum[4, 10]. Generally based on the analysis of a large number of data, several methods are used to rank the spectra in order to sort the corresponding species. Others methods are used to link the spectra evolution with species concentration to establish a calibration law. Chemometric analysis is widely used in IR spectroscopy[11] and more and more in Raman studies[12, 13]. Here we use the principal components analysis (PCA) which allows to highlight similarities between spectra according the principal components. PCA is used to detect the presence of several anions in salt solutions, especially halides.

The principle of a PCA is to reduce the number of spectral variables using an orthogonal transformation and turn them into uncorrelated variables called principal components (PC) ranked in order of largest possible variance. The interest of a PCA is that the number of principal components is very small (generally less than 10) compared to the large number of original variables. In this study, Raman spectra have 2048 spectral variables i.e. intensities, which are reduced by PCA to scores and loadings. Scores are the coordinates of each spectrum in the different dimensions (principal components) and loadings are the representation of these dimensions according to the wavenumbers. The distribution of these scores in the space of dimensions is used to group individuals (spectra) according to common principal components and thus to link them to a physical or chemical parameter. This "scores plot" allows highlighting similarities between spectra according principal components. It is a very useful tool to class or group spectra and therefore the related species.

2.3. Partial Least Square Regression

Partial Least Square Regression (PLSR) is a regression method developed by H. Wold in the 1960s[14] widely used in spectrometric application (IR or Raman) for quantitative analysis in industry. PLS regression works in the same way as the PCA but take into account the response values e.g. species concentrations and their variations, in the determination of the calibration model.

The efficiency of a regression within a calibration model can be evaluated by different methods and criteria. Here, as usually in many studies[15, 16], three parameters are calculated in order to estimate the power of the regression: the regression coefficient (R^2), the root mean square error (RMSE) and the bias (BIAS) which are detailed hereafter.

The regression coefficient corresponds to the spreading predictions parallel to the ideal line. R^2 is usually given for a regression but this parameter although necessary is not sufficient. Thus, the RMSE evaluates the differences between values $\hat{y}_i$ predicted by the model and the observed value y_i and has the same unit as the predicted value. The BIAS provides the offset on the prediction. The true values are underestimated or overestimated if the BIAS is negative or positive respectively.

These three criteria can be calculated by the following expressions:

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-1}} \tag{2}$$

$$BIAS = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \tag{3}$$

where n is the number of points used for the regression and $\bar{y_i}$ the mean of the y_i values.

2.4. Experimentation

Several salted solutions of nitrate, sulfate, and halide (fluoride, chloride, bromide and iodide) at different concentrations from 10 to 100 mmol/L were prepared by weight. Spectra were recorded in 5 seconds at room temperature with a RXN-1 Raman spectrometer (Kaiser Optical Systems) which can be used for *in situ* measurements. An excitation laser line at 532 nm was used and the laser light was focused through a 10x objective at a distance above 1 cm from the liquid sample contained in a quartz cell. The Raman light is collected in backscattering configuration by the same optic, filtered to get rid of the Rayleigh line and is diffracted through a grating on a CCD camera (Andor) using two lines of 1024 pixels and cooled at -40 °C. Acquisition was done by the iCRaman software (Mettler Toledo) and the spectral range starts at 100 cm^{-1} and ends at 4350 cm^{-1} within a spectral resolution of 2 cm^{-1}. Signal processing (SNV normalization) and chemometric analysis (PCA and PLSR) were performed according to a home-made program using the R language[17].

3. RESULTS

3.1. Ion identification

First, a PCA was applied on the set of Raman spectra consisting of nitrate, sulfate, chloride and iodide and the scores plot resulting of the projection of the different spectra according to the two first principal components PC_1 and PC_2 is given in Fig.2. One can see that a PCA allows classing the spectra in several clusters according to the type of anion: NO_3^-, SO_4^{2-}, Cl^- and I^-. Because of the intense peak about 1000 cm^{-1} nitrate and sulfate spectra can be clearly separated. Halides which do not have direct signatures but disturb the OH stretching band are grouped together.

However, it is possible to differentiate chlorides and iodides into this group. For each anion, the corresponding spectra are ranked in the order of concentrations from 10 to 100 mmol/L.

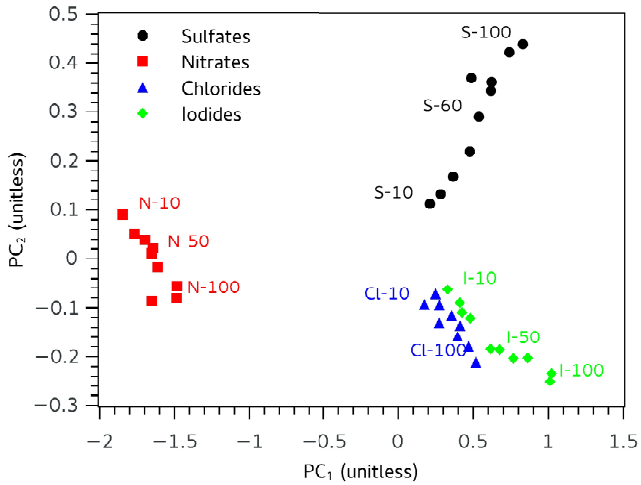

Fig. 2. Scores plot resulting from a PCA on Raman spectra of nitrates, sulfates, chlorides and iodides solutions.

3.2. Ion quantification

Then, a PLS regression was applied on the set of spectra in order to build prediction models for these 4 anions. In order to evaluate the quality of the model established for each anion, a cross validation was done using the leave-one-out method. One removes the first of the n samples and performs the calibration model with the $n-1$ remaining samples to predict the concentration of the removed sample. This sample is then set back in the calibration set while the second one is removed for prediction. The same operation is repeated for all samples.

10 PC were used to establish the calibration for the halides (Cl^- and I^-) and 8 PC were used for NO_3^- and SO_4^{2-}. The number of PC *nPC* needed for each anion in the PLSR is given by the evolution of the RMSE in validation according to the number of PC (see Fig.3). Although the RMSE in calibration always decreases with number of PC, the RMSE in validation passes through a minimum before increasing (over-fitting phenomenon).

Results of the PLS regression showing values of the regression coefficient R^2, the root mean square error RMSE and the bias are given in Tab.1 for each anion in calibration and validation.

Nitrate and sulfate spectra give a RMSE in validation less than 1 mmol/L (i.e. $< 1\%$), no bias and a value of $R^2 > 0.999$. Halides give higher RMSE (2.5 mmol/L for I^- and 6.9 mmol/L for Cl^-) and lower R^2 especially for the chlorides. However, these results are very good and allows quantification of chlorides with an error less than 250 mg/L (limit of drinkability).

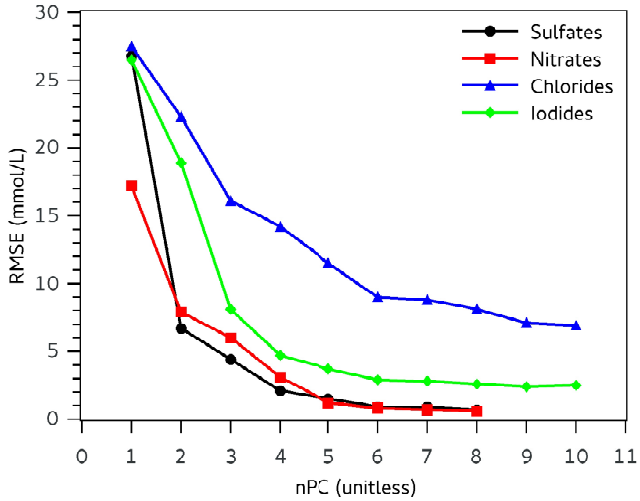

Fig. 3. Scree plot showing the optimal number of principal components to choose for each anion in the PLS regression.

Table 1. Results of PLS regression made on nitrates, sulfates, chlorides and iodides solutions. RMSE and BIAS are expressed in mmol/l.

	Calibration			Validation		
Anion	R^2	*RMSE*	*BIAS*	R^2	*RMSE*	*BIAS*
NO_3^-	0.9999	0.2	0	0.9995	0.6	0
SO_4^{2-}	0.9999	0.2	0	0.9994	0.7	0
Cl^-	0.9987	1	0	0.9369	6.9	-0.3
I^-	0.9998	0.4	0	0.9925	2.5	0

3.3. Halides analysis

Finally, this method was applied on a set of halides spectra which do not give specific signature in the Raman spectrum and are therefore more difficult to detect. Fig.4 illustrate the slight influence of the halides on the OH stretching band of the water spectrum for two concentration values of 50 and 100 mmol/L. Here, only the spectral range between 2600 and 3900 cm^{-1} corresponding to the OH stretching band was used in the PCA calculation.

The scores plot given in Fig.5 allows to differentiate the several halides which are ranked according the first component in the order of their mass. PLS regression was also used for the halides analysis as the first attempt to create a quantification model. 7 components were used for F^-, 6 for Cl^-, 5 for Br^- and 4 for I^-. Results of the PLS are given in Tab.2.

Although the calibration results are correct, those found in validation are not as good as previously. We

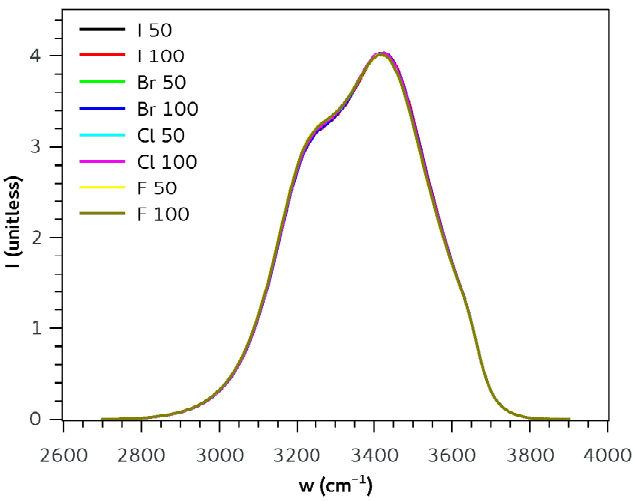

Fig. 4. Slight influence of halide solutions on the OH stretching band of the water spectrum for 50 and 100 mmol/L.

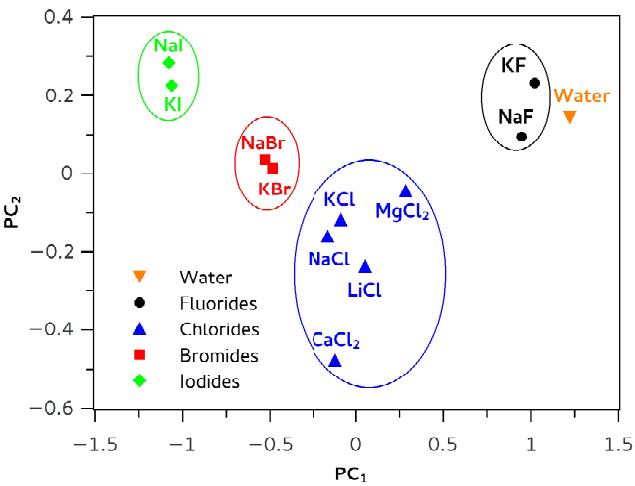

Fig. 5. Scores plot resulting from a PCA on Raman spectra of halides solutions.

found here a R^2 less than 0.9 for each anion and a RMSE between 9.2 and 22.7 mmol/L. However, although the method still needs to be improved, these first results on halides which have little influence on the water spectrum for these concentration range (10-100 mmol/L), show the possibility of quantification and are encouraging for the future.

4. CONCLUSIONS

Raman spectrometry can be a very useful tool for water analysis. Thus this optical technique is able to quickly and simultaneously detect and measure several substances dissolved in water like nitrates, sulfates, and halides sought for water quality control. We have shown that chemometric analysis can differentiate anions in solutions making it easier to detect the presence of species that do not have characteristic peaks but slightly influence the spectrum of water like the chlorides. PLS regressions applied on several salt spectra give rise to good calibration models allowing predicting the concentration of each anion. The quantification method needs to be improved in the case of halides analysis for low concentrations. This method will be applied for the simultaneous prediction of anion concentration in salt mixture solutions.

Table 2. Results of PLS regression made on fluorides, chlorides, bromides and iodides solutions. RMSE and BIAS are expressed in mmol/l.

	Calibration			Validation		
Anion	R^2	*RMSE*	*BIAS*	R^2	*RMSE*	*BIAS*
F^-	0.9973	1.4	0	0.7730	13.2	-0.2
Cl^-	0.9874	3.1	0	0.6369	16.9	0.7
Br^-	0.9323	7.2	0	0.3318	22.7	-0.6
I^-	0.9677	5	0	0.8869	9.2	0.1

REFERENCES

[1] L.A. Lyon et al., Raman spectroscopy, *Analytical Chemistry*, 70, 341R-361R, 1998.

[2] Ph. Colomban, Imagerie Raman de matriaux et dispositifs htrognes, *Techniques de l'Ingnieur*, 04, RE 5 1-13, 2002.

[3] J. Barbillat et al., Spectromtrie Raman, *Techniques de l'Ingnieur*, P2865, 1-31, 1999.

[4] S. Wold and M. Sjostrom, Chemometrics, present and future success, *Chemometrics and Intelligent Laboratory Systems*, 44, 3-14, 1998.

[5] K. Ben Mabrouk, T.H. Kauffmann, H. Aroui, M.D. Fontana, Raman study of cation effect on sulfate vibration modes in solid state and in aqueous solutions, *Journal of Raman Spectroscopy*, 44(11), 1603-1608, 2013.

[6] M.D. Fontana, K. Ben Mabrouk, T.H. Kauffmann, Raman spectroscopic sensors for inorganic salts, *Spectroscopic Properties of Inorganic and Organometallic Compounds*, J. Yarwood, R. Douthwaite, S. Duckett, editor, 44, 40-67, RSC Publishing, 2013.

[7] T.H. Kauffmann and M.D. Fontana, Inorganic salts diluted in water probed by Raman spectrometry: Data processing and performance evaluation, Sensors and Actuators B: Chemical, 209, 154-161, 2015.

[8] S.A. Burikov, T.A. Dolenko, V.V. Fadeev, A.V. Sugonyaev, New opportunities in the determination of inorganic compounds in water by the method of laser Raman spectroscopy, *Laser Physics*, 15(8), 1175-1179, 2005.

[9] T.H. Kauffmann, M.D. Fontana, Optical sensor of salt concentration: Uncertainty evaluation, *Sensors and Actuators B: Chemical*, 161, 21-27, 2012.

[10] S. Wold, Chemometrics; what do we mean with it, and what do we want from it?, *Chemometrics and Intelligent Laboratory Systems*, 30, 109-115, 1995.

[11] C. Muehlethaler, G. Massonnet, P. Esseiva, The application of chemometrics on Infrared and Raman spectra as a tool for the forensic analysis of paints, *Forensic Science International*, 209n 173-182, 2011.

[12] J.B. Cooper, Chemometric analysis of Raman spectroscopic data for process control applications, *Chemometrics and Intelligent Laboratory Systems*, 46, 231-247, 2009.

[13] O. Svensson, M. Josefson, F.W. Langkilde, Reaction monitoring using Raman spectroscopy and chemometrics, *Chemometrics and Intelligent Laboratory Systems*, 49, 49-66, 1999.

[14] H. Wold, Nonlinear estimation by iterative least squares procedures, *Research Papers in Statistics*, 411444, 1966.

[15] G. Kang et al., Quantitative analysis of mixed hydrofluoric and nitric acids using Raman spectroscopy with partial least squares regression, *Talanta*, 81, 1413-1417, 2010.

[16] F. Estienne et al., Multivariate calibration with Raman spectroscopic data: a case study, *Analytica Chimica Acta*, 424, 185201, 2000.

[17] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2012, ISBN 3-900051-07-0, http://www.R-project.org/.