HAL
open science

Inorganic salts diluted in water probed by Raman spectrometry: Data processing and performance evaluation

T.H. Kauffmann, M.D. Fontana

To cite this version:

T.H. Kauffmann, M.D. Fontana. Inorganic salts diluted in water probed by Raman spectrometry: Data processing and performance evaluation. Sensors and Actuators B: Chemical, 2015, 209, pp.154-161. 10.1016/j.snb.2014.11.108 . hal-01092027

HAL Id: hal-01092027

https://hal.science/hal-01092027v1

Submitted on 11 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.

*Manuscript
Click here to view linked References

Inorganic salts diluted in water probed by Raman spectrometry: Data processing and performance evaluation

T.H. Kauffmann[a,b] and M.D. Fontana[a,b]

[a]Université de Lorraine, Laboratoire Matériaux Optiques, Photonique et Systèmes, EA 4423 - 2, rue Edouard Belin, 57070 Metz, France.
[b]Supélec, Laboratoire Matériaux Optiques, Photonique et Systèmes, EA 4423 - 2, rue Edouard Belin, 57070 Metz, France.

Corresponding author: T.H. Kauffmann
E-mail address: thomas.kauffmann@univ-lorraine.fr
Tel.: +33 3 87 37 85 47

Abstract

The paper reports on a methodology for the determination, by means of Raman spectrometry, of the concentration of inorganic salts dissolved in aqueous solutions. This procedure is then applied to nitrate solutions with a varying content from 0 to 100 mM (mmol/l). For this we exploit the concentration dependence of the Raman peak lying at 1047 cm^{-1} specific to the NO_3^- anion. Different signal processing and normalization methods are used and compared to deduce the most reliable and robust calibration. Then cross-validations are done using the "leave-one-out" method to validate the prediction models. Several figures of merit such as sensitivity, signal to noise ratio, limits of detection and quantification are calculated to estimate the efficiency of our methodology for nitrate solutions. Various sources of uncertainties are considered and evaluated according to the ISO GUM and the standard uncertainty of the concentration is then calculated.

Keywords : Raman spectrometry; Salt solutions; Inorganic salts; Calibration; Normalization; Uncertainty.

1. Introduction

The detection and identification of chemicals in water by Raman spectroscopy (RS) has been demonstrated since several decades [1,2]. The advantages of RS compared with other techniques to detect anionic species are well known [1,3]. Indeed chromatography, mass spectrometry, colorimetry, ion selective electrodes, even if these techniques are sensitive, require sample preparation, or use of reagent, and/or frequent calibration [4]. The vibrational spectrum provides something like a fingerprint of one substance. However vibrational spectra of solute species by IR absorption is impossible owing to water absorption in IR range, so that only RS can be efficiently used. Furthermore Raman lines corresponding to substances are relatively narrow and generally distinct from the OH bands. In addition to remote sensing substances in tanks, lakes or rivers, portable Raman spectrometers are now inexpensive and robust [5].

The paper is devoted to the method allowing a quantitative and reliable exploitation of Raman spectra. It is well known that the Raman intensity is proportional to the species content [6,7], so that it can be used for the calibration of a substance although it is recognized that the main drawback of RS is the weakness of Raman intensity so that it is usually applied to large concentration. Nevertheless it is possible to measure concentration below 1 mM (mmol/L) using high performance laboratory spectrometers which are cumbersome and very expensive [8]. Furthermore, a long acquisition time of several minutes is generally needed and it is frequently necessary to apply several heavy treatments (pure water spectrum subtraction or deconvolution for example) in order to reach the micro molar level [9]. UV Raman resonance and SERS are also proposed to overcome the difficulties of normal RS. However, despite advantages, compared with normal RS, these techniques use more complex instrumentation, are more expensive, and specific to one or two species only [10]. The linearity of the response versus concentration could also be affected in

resonance conditions.

We show here that classical RS through portable spectrometer can provide reliable measurements of dissolved substances even for high dilutions. We consider as an example for all measurements and treatments the aqueous solutions of sodium nitrates. As they are used in fertilizers, aerosols and explosives a detection of nitrates and a reliable determination of their concentration is an important challenge [11,12,13].

This paper proposes in a first part a general method to achieve a reliable calibration of a substance diluted in water from a set of Raman spectra. At first, Raman signature specific to a substance and the concentration indicator are needed. Different data treatments are then described to evaluate the calibration methods. Several criteria revealing the quality of the quantitative analysis are presented. At last the methodology used to calculate the uncertainties coming from several error sources is described according to the Guide of Uncertainty and Measurement [14]. In a second part the suggested procedure with its different steps is applied to the determination of the concentration in sodium nitrate solutions. The results are presented and the figures of merit are calculated and discussed underlining the power of the method.

2. Methods for a quantitative exploitation of Raman data

2.1. Species Raman signature and concentration indicators

RS is well known analytical tool to detect inorganic and organic substances as well [1]. Thus database systems can be employed for the identification of species in the mixture. By contrast, the quantitative exploitation of Raman spectra is less frequent. It is mainly due to the difficulty to extract the absolute Raman intensity which is a complicated function of the set up characteristics (spectrometer, laser, detector…) and experimental conditions (temperature...). In fact it is possible to use the relative intensity to determine the species concentration C with a good reliability, if using appropriate procedure of data treatment. Different processes are presented and compared below.

Before any treatment, one (or two) Raman lines which are unambiguously attributed to a molecule or a substance have to be chosen [7]. These lines are generally intense and well defined. The position of this line (i.e. the frequency of one vibrational mode assigned to the line) and its width depend on the state of the substance and environment conditions [15]. Moreover the width is varying with spectrometer resolution. This line is the own signature of the substance to be analyzed.

Intensity at the maximum of this peak I_p is proportional to the substance content C, and therefore can be used to calibrate, and then to determine the content. Obviously the integrated intensity J_p (peak area) is proportional to C as well, so that both quantities I_p and J_p can be chosen as indicator parameters to derive the concentration C. As they depend on several internal (linked to the set up) and external (experimental conditions) parameters, I_p or J_p has to be related to C univocally. For this, I_p and J_p are required to be independent of these parameters. This is provided by the normalization of spectra, and thus transformation of I_p and J_p, so that only relative intensities are really exploited. Each use of I_p or J_p as concentration indicator has some advantages and drawbacks. The line position can be shifted by temperature change, or spectrometer drift causing a possible change in I_p value. The integrated intensity J_p is nearly unaffected by these small changes if the line width is not too small (several cm^{-1}) but requires additional calculation compared with I_p.

2.2. Data processing

2.2.1. Pre-treatments

To improve the quality of the Raman signal, some data treatments are often applied on the spectra [16]. The first process is generally a smoothing using several methods (moving average, Savitsky-Golay etc.) in order to diminish the spectral noise [17]. Then a baseline correction is usually done to overcome the background noise which may be due to the absorption of the solution for example.

2.2.2. Normalization

In order to achieve a calibration, a classical method in spectrometry consists in choosing a specific signature of the species, the intensity of which is concentration-dependent and to study its behavior. Thus, the intensity of a specific peak usually evolves with the species concentration and is therefore used for the calibration. The problem is that the intensity of a spectrum depends of several external parameters like the laser power, the integration time, the temperature or the optical probe, which give significant differences between series of spectra, independently of the sample [7]. Thereby if the spectra are not recorded under the same conditions they have to be normalized by reducing them on the same scale for comparison. For this, different methods of normalization exist which have all theirs advantages and drawbacks [17].

Within our method procedure we propose three kinds of normalization. One can consider the normalization to a reference peak the intensity I_{ref} which should be independent of the concentration and the indicator is therefore $I_p^{ref} = I_p / I_{ref}$ after the normalization process. In the second method the intensity can be compared to the intensity maximum I_{max} taken over the whole spectrum. We therefore have in this case $I_p^{max} = I_p / I_{max}$. The third method is the Standard Normal Variate (SNV) transformation which consists in subtracting the mean value of the spectrum I and dividing by the standard deviation s which leads to $I_p^{SNV} = (I_p - I) / s$.

Each method can be applied to the indicator J_p as well. We have respectively $J_p^{ref} = J_p / J_{ref}$ where J_{ref} is the area of the chosen reference, $J_p^{max} = J_p / J_{max}$ where J_{max} is the whole area under the spectrum and J_p^{SNV} which is the area J_p after the SNV transformation.

2.3. Criteria for quantitative analysis

2.3.1. Figures of merit

Several figures of merit (FOM) can be defined in order to estimate the performance of the methodology for the determination of the species concentration [18]. Some of FOM are linked to the signal to noise ratio *SNR* and thus to the instrument characteristics and the Raman efficiency of the substance. Related to *SNR* the limits of detection *LOD* and quantification *LOQ* are chosen and evaluated for a solution under study with a concentration *C*.

The signal to noise ratio is estimated by $SNR = (I_p - N) / \sigma_N$ where the signal is calculated as the difference between the intensity of the peak I_p and the baseline of the spectrum while the noise σ_N is given by the standard deviation of the intensity in a signal-free range.

LOD and *LOQ* are respectively estimated for a confidence level of 99% by $LOD = (3\times C) / SNR$ and $LOQ = (10\times C) / SNR$. *LOD* is the minimum detectable concentration which corresponds to a signal just above the noise level and *LOQ* is the limit from which we have a reliable concentration value. In practice *LOQ* is given by $LOQ = 3.3\times LOD$.

We consider also the sensibility *S* defined as the relative change in intensity (I_p or J_p) with the concentration and provided by the slope of the calibration.

2.3.2. Calibration model evaluation

The efficiency of a regression within a calibration model can be evaluated by different methods and criteria. Here, as usually in many studies [19,20], 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:

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

where n is the number of points used for the regression and y_i the mean of the y_i values.

2.3.3. Repeatability

The repeatability of the method is determined using several replicates made on one solution under the same experimental conditions. It is given here through the coefficient of variation $CV = \sigma / \mu$ where μ and σ are respectively the mean value and the standard deviation from a statistical study made on the predicted concentrations given by the model.

2.4. Uncertainty evaluation

According to the ISO GUM [14], the standard measurement uncertainty can be evaluated via two kinds of methods. The type A evaluation consists into a statistical study from several measurements of the same quantity obtained in assumed constant experimental conditions. The result is chosen as the average of the data and the standard uncertainty is given by the standard deviation of the mean. Uncertainty evaluation by means other than statistical analysis provides the type B contributions.

In the calculation of the whole uncertainty on the concentration, the squared of standard uncertainties obtained on the various sources of error have to be added. To deduce the standard uncertainty from the experimental uncertainty of type B, a rectangular distribution is conventionally assumed since no indication of the confidence, with which the error is estimated, is supplied by the manufacturer from the different used instruments [21].

The different sources of uncertainty are as follows. Firstly the uncertainty coming from the preparation of solutions used for the calibration is considered. Then the linear regression applied on the calibration points gives rise to an uncertainty evaluated by a type A method. Finally the uncertainty coming from the intensity measurement related to the instruments is taking into account.

2.4.1. Salt solutions uncertainty

Salt solutions used to achieve the calibration were prepared by dilution of a standard solution. The standard solution of a concentration C_0 was done according to $C_0 = m / V_0$ by weighing and adding a mass m of the solid substance into a flask containing a volume V_0 before filling with distilled water. The relative standard uncertainty of the standard solution concentration is therefore:

$$\left(\frac{u(C_0)}{C_0}\right)^2 = \left(\frac{u(m)}{m}\right)^2 + \left(\frac{u(V_0)}{V_0}\right)^2$$

Instrumental uncertainties Δm and ΔV_0 are given by the manufacturer so that the corresponding standard uncertainties are $u(m) = \Delta m / \sqrt{3}$ and $u(V_0) = \Delta V_0 / \sqrt{3}$ (type B evaluations).

Then the set of solutions were prepared by dilution of the standard solution and several concentrations are calculated according to $C = C_0.V / (V+V_{water})$ where V is the pipetted volume of the standard solution and V_{water} the volume of the added distilled water. Same pipettes of 5 mL were used for the standard solution and the water so that the uncertainty of the two volumes are the same and equal to $\Delta V = \Delta V_{water}$ i.e. a standard uncertainty $u(V) = \Delta V / \sqrt{3}$. Final volume $V+V_{water}$ is equal to 40 mL so that 8 pipetting of 5 mL were required to prepare a solution. Relative concentration uncertainties of the diluted solutions $u_{sol}(C)$ are then given by:

$$\left(\frac{u_{sol}(C)}{C}\right)^2 = \left(\frac{u(C_0)}{C_0}\right)^2 + 8\left(\frac{u(V)}{V}\right)^2$$

2.4.2. Calibration uncertainty

The confidence on the calibration depends on the reliability of the results derived from the fit of the plot $I = f(C)$ with a regression linear law $I_i = a.C_i + b$. For this we apply the type A method and calculate the standard deviation $s(I)$ between experimental data of I and the values as predicted by the law [22,23]:

$$s(I) = \frac{1}{\sqrt{n-2}} \sqrt{\sum_{i=1}^{n} (I_i - \hat{I}_i)^2}$$

where I_i and $\hat{I}_i$ are the experimental and predicted values respectively and n is the number of data used for the calibration. Then we calculate the standard uncertainty on the concentration C coming from the linear regression by the inverse calibration method with the formula [23,24]:

$$u_{calib}(C) = \frac{s(I)}{a} \sqrt{\frac{1}{N} + \frac{1}{n} + \frac{(I - \bar{I})^2}{a^2 \sum_{i=1}^{n} (C_i - \bar{C})^2}}$$

where I is the value coming from N replicates, I and C are the averages of the calibration values, C_i the different concentration values used in the calibration and a the slope of the calibration line. It can be noted that a corresponds to the sensitivity S of the probe to the sought substance. Moreover, one notes that the expression of $s(I) / a$ is the same as the RMSE used to evaluate the calibration regression.

2.4.3 Instrumental uncertainty

Here are concerned the uncertainty coming from the instruments and the chosen concentration indicator. Our indicator can be the intensity I_p or the area J_p of the signature peak after signal processing and normalization.

If the area J_p is taken instead of the intensity of the peak I_p, the peak area is approximated by a rectangle and calculated by $J_p \approx I_p \times dw_p$ where dw_p is the full width at half maximum (FWHM). The relative uncertainty on J_p is given by the calculation:

$$\left(\frac{u(J_p)}{J_p}\right)^2 = \left(\frac{u(I_p)}{I_p}\right)^2 + \left(\frac{u(dw_p)}{dw_p}\right)^2$$

The uncertainty on the intensity value $u(I)$ is given by the spectral noise estimated by the standard deviation of the intensity σ_n while the uncertainty $u(dw)$ of the width is estimated from the uncertainty of the line position by $u(dw) = \Delta dw / \sqrt{3}$ considering a type B evaluation. Δdw is provided by the CCD manufacturer as equal to ±1/2 pixel.

When normalization is applied on the indicator using a reference (I_{ref} or J_{ref}) relative uncertainties are convenient as well in the calculations. In this case we have $I_p^{ref} = I_p / I_{ref}$ or $J_p^{ref} = J_p / J_{ref}$ which give respectively:

$$\left(\frac{u(I_p^{\mathrm{ref}})}{I_p^{\mathrm{ref}}}\right)^2 = \left(\frac{u(I_p)}{I_p}\right)^2 + \left(\frac{u(I_{ref})}{I_{ref}}\right)^2$$

$$\left(\frac{u(J_p^{\mathrm{ref}})}{J_p^{\mathrm{ref}}}\right)^2 = \left(\frac{u(J_p)}{J_p}\right)^2 + \left(\frac{u(J_{ref})}{J_{ref}}\right)^2 = \left(\frac{u(I_p)}{I_p}\right)^2 + \left(\frac{u(dw_p)}{dw_p}\right)^2 + \left(\frac{u(I_{ref})}{I_{ref}}\right)^2 + \left(\frac{u(dw_{ref})}{dw_{ref}}\right)^2$$

The uncertainty $u_{instr}(C)$ of the concentration is then obtained by dividing $u(I_p^{ref})$ or $u(J_p^{ref})$ by the sensibility S (i.e. the slope a of the calibration).

2.4.4. Combined standard uncertainty

Finally the combined standard uncertainty of the concentration $u_c(C)$ is obtained by adding the squared value of each uncertainty coming from the different error sources according to $u_c(C) = \sqrt{u_{sol}^2(C) + u_{calib}^2(C) + u_{instr}^2(C)}$ and the expanded uncertainty is calculated by $U_c(C) = k.u_c(C)$ with a coverage factor $k = 2$ corresponding to a confidence level of 95%.

3. Results and discussion

The method described above with its different stages is applied to aqueous solution of nitrate.

3.1. Experimental

3.1.1. Salt solutions preparation

11 solutions of sodium nitrate $NaNO_3$ (purity >99%) were made by a dilution series of a standard solution preparing by weight. The concentration range was 0–100 mM (mmol/l) with a step of 10 mM.

3.1.2. Instrumentation

Raman spectra were recorded at room temperature with a RXN-1 Raman spectrometer (*Kaiser Optical Systems*) with an excitation laser line at 532 nm. The laser light was focused through a 50× objective (NA = 0.5) at a distance above 1 cm from the liquid sample contained in a quartz cell (contactless measurement) with a power of 70 mW. The Raman light was collected in backscattering configuration by the same optics 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*). The recording time was 2×30 sec and the spectral range was between 100 cm^{-1} and 4350 cm^{-1} within a spectral resolution about 1 cm^{-1}.

Data treatments such as baseline correction and normalization processes were performed by a self made program using the *R* language [25].

3.2. Nitrate signature

In the first step one chooses the specific signature of the nitrate in solution allowing its identification. Figure 1 shows the Raman spectrum of a sodium nitrate solution. The large peak located at 1047 cm^{-1} is present in all nitrate solutions and corresponds to the main vibrational mode of the nitrate anion NO_3^-. Its position was found to be independent of the monoatomic cation in aqueous solutions [15]. Besides the nitrate signature the spectrum exhibits the wide OH-stretching band (OH-SB) of the water molecule extending from 2800 to 3900 cm^{-1} and a weak band lying around 1635 cm^{-1} and corresponding to the OH-bending band (OH-BB). As shown in Figure 1 the intensity of the NO_3^- peak evolves monotonously with the nitrate concentration so that this peak can be considered for the quantitative analysis. A calibration can therefore be achieved using the

intensity I_p of the NO_3^- peak lying at 1047 cm^{-1} or the area J_p from 1000 to 1100 cm^{-1} after a linear correction of the baseline.

Fig1. Raman spectrum of a $NaNO_3$ solution of 100 mM showing the NO_3^- specific peak at 1047 cm^{-1}. The inset shows the peak evolution with the nitrate concentration in mM (mmol/l).

3.3. SNR, LOD and LOQ

The *SNR* is estimated for a nitrate solution of 10 mM as shown in Figure 2. The signal value is calculated as the difference between the NO_3^- peak intensity and the baseline and is equal to 1004 a.u. The spectral noise is given by the standard deviation of the intensity between 1200 and 1300 cm^{-1} and is equal to 30 a.u. We found therefore a *SNR* value of 33.5.

Finally the limit of detection is estimated by $LOD = 3\times10 / 33.5$ and found to be equal to 0.9 mM. The limit of quantification is then calculated by $LOQ = 3.3\times0.9$ and is therefore equal to 3 mM.

Fig2. Signal to noise ratio *SNR* evaluation considering the signal as the intensity of the NO_3^- peak at 1047 cm^{-1} and the noise as the standard deviation of the intensity between 1200 and 1300 cm^{-1}.

3.4. Comparison of normalization methods

The intensity of the signature peak is used for the concentration calibration but as it depends as well on many experimental conditions (recording time, laser power, CCD responses etc.) it must be normalized to be exploitable for the concentration determination. The different methods of normalization mentioned above are used and applied to both indicators I_p and J_p and then compared each to the others in order to obtain the best concentration calibration.

The first method is the normalization to the reference peak and instead the OH-SB often taken as reference for normalization in the literature [26,27] we prefer to consider $I_{OH\text{-}BB}$ of the OH-BB of the water spectrum. This band is lying about 1635 cm^{-1} and even if it is less intense than the OH-SB it has as advantage to be not (or very slightly) affected by the nitrate introduction as shown in Figure 3. The second method takes the intensity as reference as the maximum of the spectrum, here $I_{OH\text{-}SB}$ of the OH-SB. The last method uses the intensity of the NO_3^- peak after a SNV transformation applied on the whole spectrum.

If the area J_p under the NO_3^- peak is chosen as indicator, the first method takes the area under the OH-BB from 1500 to 1800 cm^{-1} and the second normalization method is to divide J_p by the whole area of spectrum. As the OH-SB can be also used for comparison with the different normalization processes the area from 2600 to 3900 cm^{-1} is calculated as well. Finally within the third method the SNV transformation is applied to the spectrum before calculating J_p.

Fig3. Raman spectrum in $NaNO_3$ solution of 0, 30, 50 and 100 mM showing the NO_3^- specific peak at 1047 cm^{-1} evolving with the nitrate concentration and the unaffected OH bending band at 1635 cm^{-1}.

In order to improve the quality of the calibration a linear baseline correction is applied <u>after</u> the SNV transformation. Figure 4 shows clearly the contribution of a baseline correction after SNV normalization. Here the regression coefficient R^2 found for the calibration using the area of the NO_3^- peak increases from 0.812 to 0.899 after a SNV transformation and reaches 0.992 when a baseline correction is applied in addition. Baseline corrections are also applied on the NO_3^- peak and on the OH bending and stretching bands for the other normalization methods <u>before</u> the ratio calculation.

Fig4. NO_3^- peak for several nitrate solutions from 0 to 100 mM: (a) Raw data, (b) after SNV normalization applied on the whole spectral range and (c) after baseline correction from 1000 to 1100 cm^{-1}.

After normalization and linear baseline correction, the different intensity or area values I_p and J_p of the NO_3^- peak are plotted in function of the concentration in order to obtain the calibration lines

shown in Figure 5. Then a linear regression is applied to evaluate the regression coefficient R^2 and the sensitivity S. Results are summarized in Table 1.

Table 1. Calibration results compared for different normalization methods. Results without normalization process (raw spectra) are also given for comparison.

Normalization methods	R^2	Slope S
Indicator I_p		
Raw spectra I_p	0.9537	1.2×10^{2}
OH stretching band $I_p^{OH\text{-}SB}$ (i.e. I_p^{max})	0.9984	5.9×10^{-4}
OH bending band $I_p^{OH\text{-}BB}$	0.9990	3.1×10^{-2}
SNV I_p^{SNV}	0.9986	2.3×10^{-3}
Indicator J_p		
Raw spectra J_p	0.9571	1.6×10^{3}
Whole Area J_p^{max}	0.9951	1.3×10^{-5}
OH stretching band $J_p^{OH\text{-}SB}$	0.9926	1.8×10^{-5}
OH bending band $J_p^{OH\text{-}BB}$	0.9980	3.8×10^{-3}
SNV J_p^{SNV}	0.9924	3.1×10^{-2}

Firstly one can see that applying normalization significantly improves the regression coefficient in each case whatever are the indicator and the normalization method. Thus R^2 increases from 0.95 to over 0.99 within all normalization methods.

All normalization methods give rise to R^2 values very close each to the other. By contrast they yield very different values of sensitivity S. As a consequence S can be considered as the criterion in the comparison between the different normalization methods.

In the first case when I_p is the indicator, the normalization to the reference peak i.e. the OH-BB gives the highest sensitivity equal to 3.1×10^{-2} mM^{-1} more than one order of magnitude larger than S provided by the SNV transformation. The normalization by the spectrum maximum i.e. the intense OH-SB gives the worst sensitivity (52 times lower).

If the area J_p is taken as indicator, the sensitivity is clearly better when the SNV normalization is applied and has the same value (3.1×10^{-2} mM^{-1}) as in the best situation of the first case. The normalization method with the OH-BB taken as reference gives a sensitivity 8 times lower. Other methods yield a bad sensitivity, nearly 3 orders of magnitude less than the SNV transformation.

Fig5. Calibration lines using (a) I_p and (b) J_p as indicator for different normalization methods.

The SNV transformation provides a good calibration (R^2 and sensitivity) when the area J_p is taking as indicator and is a simple and easy method to be applied.
The choice of a reference band gives also good results in calibration (R^2) but the sensitivity is smaller particularly if the chosen band is wide. Moreover this method can be less convenient because it requires searching for a band reference which is independent of concentration and requiring several processing (smoothing, baseline etc.).

The normalization by reference to the spectrum maximum, here the intense OH-SB has to completely excluded, not only because it gives a very poor sensitivity (or very small calibration slope) but also since it can lead to an erroneous determination of C. Indeed there are no *a priori* reasons for which the maximum could be independent of C. In our case OH-SB is affected by the introduction of nitrate, and other substances as well.

3.5. Prediction model

The methodologies implemented to establish the calibration laws according to the indicator I_p or J_p can be summarized as follows:

(a) the successive treatments as follows are to applied to the indicator I_p: Raw spectra ; Smoothing (optional) ; Baseline corrections on the NO_3^- spectral range (1000–1100 cm^{-1}) and on the OH-BB (1500–1800 cm^{-1}) ; Normalization $I_p^{OH\text{-}BB} = I_p / I_{OH\text{-}BB}$; Plot $I_p^{OH\text{-}BB}$ vs C ; Linear fit ; Calibration law.

(b) the successive treatments as follows are to applied to the indicator J_p: Raw spectra ; Smoothing (optional) ; SNV transformation ; Baseline correction on the NO_3^- spectral range (1000–1100 cm^{-1}) ; Area calculus J_p^{SNV} ; Plot J_p^{SNV} vs C ; Linear fit ; Calibration law.

The calibration models C_{pred} vs C are then achieved by reversing the calibration laws in order to predict the concentration values. Results of calibration give regression coefficients R^2 equal to 0.999 and 0.992 for I and J respectively, root mean square errors *RMSEC* of 1.1 mM and 3.1 mM and zero *BIAS*.

The models are then checked by cross-validation 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. The cross-validation gives a good coefficient of regression R^2 equal to 0.998, a *RMSEV* of 1.4 mM and a zero *BIAS* using the intensity $I_p^{OH\text{-}BB}$. Results are worse using the area J_p^{SNV} with R^2 equal to 0.984, a *RMSEV* of 4.1 mM and a weak *BIAS* of 0.5 mM. Calibration and validation models are given in Figure 6.

Fig6. (a) Calibration and (b) cross-validation lines and results using the OH bending band normalization for I_p and the SNV transformation for J_p. R^2, *RMSE* and *BIAS* are given in both cases.

The repeatability is finally determined using 30 replicates of a solution of $NaNO_3$. We found a coefficient of variation equal to 4.5 % with $I_p^{OH\text{-}BB}$ as indicator and 3.7 % if J_p^{SNV} is chosen. Comparison results of the two best normalization methods are summarized in Table 2.

Table 2. Calibration and validation results, sensitivity and repeatability for $NaNO_3$ solutions according the best normalization methods.

Accuracy	Calibration		$I_p^{OH\text{-}BB}$	J_p^{SNV}
	R^2	unitless	0.999	0.998
	RMSEC	mM	1.1	3.1
	BIAS	mM	0	0
	Validation		$I_p^{OH\text{-}BB}$	J_p^{SNV}
	R^2	unitless	0.998	0.984
	RMSEV	mM	1.4	4.1
	BIAS	mM	0	0.5
Sensitivity	S	10^{-2}/mM	3.1	3.1
Repeatability	*CV* for 30 replicates of $NaNO_3$	%	4.5	3.7

3.6. Uncertainty calculation

Here the uncertainties of the concentration are calculated for a solution of C = 10 mM for the two possible choices of indicator.

3.6.1. Solution preparation

Salt solutions used to achieve the calibration were prepared by dilution of a standard solution of C_0 = 100 mM of sodium nitrate (molar mass of 85 g/mol). The standard solution was done according to $C_0 = m / V_0$ by weighing and adding m = 4.25 g of $NaNO_3$ into a flask containing a volume V_0 = 500 mL before filling with distilled water.

Uncertainties due to the weighing machine and the flask are given by the manufacturer and equal respectively to $\Delta m = \pm 0.01$ g and $\Delta V_0 = \pm 0.5$ mL so that the corresponding standard uncertainties are $u(m) = 5.77\times10^{-3}$ g and $u(V_0) = 2.89\times10^{-1}$ mL and the uncertainty of the pipetted volumes are equal to $\Delta V = \pm0.03$ mL which gives a standard uncertainty $u(V) = 1.73\times10^{-2}$ mL.

The standard uncertainty of the standard solution is found to be equal to $u(C_0) = 0.15$ mM which gives for a solution of 10 mM an uncertainty $u(C_{sol}) = 0.1$ mM.

3.6.2. Calibration

For the method using the intensity $I_p^{OH\text{-}BB}$, the standard deviation of the calibration between experimental and predicted data is found to be $s(I_p^{OH\text{-}BB}) = 3.42\times10^{-2}$. With a calibration slope $a = 3.11\times10^{-2}$, $N = 2$ replicates, $n = 11$ calibration values, $I_p^{OH\text{-}BB} = 1.63$, $C = 50$ mM and with an intensity value $I_p^{OH\text{-}BB} = 4.14\times10^{-1}$ for $C = 10$ mM, we find a concentration uncertainty coming from the calibration of $u_{calib}(C) = 0.9$ mM.

If J_p^{SNV} is chosen as indicator, we find a standard deviation of the calibration $s(J_p^{SNV}) = 9.62\times10^{-2}$ and with a calibration slope $a = 3.14\times10^{-2}$, the same numbers of replicates and calibration values ($N = 2$ and $n = 11$), $J_p^{SNV} = 1.78$, $C = 50$ mM and with an area value $J_p^{SNV} = 5.84\times10^{-1}$ for $C = 10$ mM, we find a concentration uncertainty coming from the calibration of $u_{calib}(C) = 2.6$ mM.

3.6.3. Instrumentation

To calculate the uncertainty coming from the instruments, we assume first that the uncertainty on the intensity is given by the spectral noise and can be estimated by the standard deviation σ_n used previously in the *SNR* calculation so that $u(I_p) = u\ (I_{OH\text{-}SB}) = 30$ a.u. Then for a solution of 10 mM we find a peak intensity $I_p = 1124$ a.u. and an intensity of $I_{OH\text{-}SB} = 2715$ a.u. for the OH-BB used here as reference giving $I_p^{OH\text{-}BB} = 4.14\times10^{-1}$. The calculation gives rise to an instrumental uncertainty of $u_{instr}(C) = 0.4$ mM.

For the area method, the spectral noise is estimated by the standard deviation σ_n of the intensity between 1200 and 1300 cm^{-1} but here after the SNV normalization and the uncertainty is found to be $u(J_p^{SNV}) = 8.9\times10^{-4}$. The CCD manufacturer provides an uncertainty of ±1/2 pixel which represents for our spectrometer an uncertainty of $\Delta dw = \pm1$ cm^{-1} (2048 pixels for a spectral range of 4250 cm^{-1}) and assuming a rectangular distribution for the uncertainty on the line position we have $u(dw_p) = 1/\sqrt{3}$. Therefore with the values of $I_p^{SNV} = 3.29\times10^{-2}$, with $dw_p = 12$ cm^{-1} for the peak width and with $J_p^{SNV} = 5.84\times10^{-1}$ we find in this case an instrumental uncertainty of $u_{instr}(C) = 1$ mM.

3.6.4. Combined standard uncertainty

Finally we calculate a combined standard uncertainty of the concentration $u_c(C) = 1$ mM with $I_p^{OH\text{-}BB}$ as indicator (i.e. an expanded uncertainty $U_c(C) = 2$mM) and $u_c(C) = 2.8$ mM if J_p^{SNV} is chosen

(i.e. $U_c(C)$ = 5.6 mM). Table 3 summarizes the various sources of uncertainty in both methods.

Table 3. Uncertainties calculated for a nitrate solution of 10 mM according the two possible indicators

Source of uncertainty	Standard uncertainty	
	$I_p^{OH\text{-}BB}$	J_p^{SNV}
Solution preparation $u_{sol}(C)$	0.1 mM	0.1 mM
Calibration $u_{calib}(C)$	0.9 mM	2.6 mM
Instrumentation $u_{instr}(C)$	0.4 mM	1.0 mM
Combined standard uncertainty $u_c(C)$	1 mM	2.8 mM
Expanded uncertainty $U_c(C)$ (k = 2)	2 mM	5.6 mM

4. Conclusion

We propose a general procedure for the quick determination, by means of a portable Raman probe, of the concentration of an inorganic salt diluted in aqueous solution. We underline the effects of the normalization method and the indicator on the abilities and performances of the Raman sensor. Sensibility and accuracy are shown to be particularly dependent on these choices.

The described method may is here applied on nitrates but may be used for other inorganic salts (sulfates, phosphates, carbonates, ...) giving a specific line, signature of the sought anion. Nitrate solutions from 0 to 100 mM were chosen as example and investigated through the Raman peak located at 1047 cm^{-1} as the specific signature of the NO_3^- anion. Several spectral processing and normalization methods were compared and discussed in order to get the best concentration calibration. Two methods give highest sensitivity equal to 3.1×10^{-2} mM^{-1} in both cases and good calibration parameters (R^2, *RMSE* and *BIAS*) according to the indicator chosen to calibrate the concentration variations. Thus SNV (Standard Normal Variate) transformation is better when the area under the peak is considered while a normalization by the OH bending band intensity taking as reference is preferable if the intensity at the maximum of the signature peak is used as indicator. In each case a cross-validation using the "leave-one-method" was done to validate the models.

Repeatability of the methods and some figures of merit are calculated for a nitrate solution of 10 mM. Limits of detection and quantification were therefore found to be 0.9 mM and 3 mM respectively and the repeatability of the both methods is less than 5%. The efficiency of our probe and technique can be pointed out if comparing the *LOD* value with the required water drinkability limit (50 mg/l or 0.8 mM).

Finally the uncertainty of the concentration is estimated according to the ISO GUM taking into account the various sources of uncertainties (solution preparation, calibration, instruments) and are found to be equal to 2 mM and 5.6 mM with the peak intensity and the area respectively taken as indicators.

References

[1] K. Nakamoto, Infrared and Raman Spectra of Inorganic and Coordination Compounds, J. Wiley and sons (2009).
[2] D.E. Irish and A.R. Davis, Interactions in aqueous alkali metal solutions, Canadian Journal of Chemistry 46 (1968) 943-951.
[3] L. A. Lyon et al., Raman Spectroscopy, Analytical Chemistry, 70 (1998) 341R.
[4] M.A. Giambra, Application of ion chromatography to qualitative and quantitative determination of the main inorganic ionic components of samples from a production process of potassium sulphate, Analytica Chimica Acta, 530 (2005) 41-48.
[5] I. R. Lewis and H. G. M. Edwards, Handbook of Raman Spectroscopy (2001).
[6] R. Loudon, The Raman effect in crystals, Advances in Physics, 13 (1964) 423-482.
[7] M.D. Fontana et al., Raman spectroscopic sensors for inorganic salts, RSC Books, Spectroscopic Properties of Inorganic and Organometallic Compounds, 44 (2013) 40-67.
[8] W.W. Rudolph and G.T. Hefter, Quantitative analysis in alkaline aluminate solutions by Raman spectroscopy,

Analytical Methods, 1 (2009) 132-138.
[9] W.W. Rudolph and G. Irmer, Analysis of ionic species in natural mineral waters studied by Raman spectroscopy, ICORS 2004 abstract.
[10] T. Ling et al., Reflectance based optical fibre sensor for ammonium ion using solid-state Riegler's reagent, Sensors and Actuators B: Chemical, 173 (2012) 614-619.
[11] J.A. Camargo et al., Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates, Chemosphere 58 (2005) 1255–1267.
[12] L. Monteagudo et al., River eutrophication: Irrigated vs. non-irrigated agriculture through different spatial scales, Water Research 46 (2012) 2759-2771.
[13] M.J. Moorcroft et al., Detection and determination of nitrate and nitrite: a review, Talanta 54 (2001) 785-803.
[14] Guide to the Expression of Uncertainty in Measurement, ISO, Geneva, (1995) ISBN 92-67-10188-9.
[15] K. Ben Mabrouk et al., Raman study of cation effect on sulfate vibration modes in solid state and in aqueous solutions, Journal of Raman Spectroscopy, 44 (2013) 11 1603-1608.
[16] P. Lasch, Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging, Chemometrics and Intelligent Laboratory Systems, 117 (2012) 100–114.
[17] J. Palacky et al., SVD-based method for intensity normalization, background correction and solvent subtraction in Raman spectroscopy exploiting the properties of water stretching vibrations, Journal of Raman Spectroscopy, 42 (2011), 1528–1539.
[18] I. Taverniers et al., Trends in quality in the analytical laboratory. II. Analytical method validation and quality assurance, Trends in Analytical Chemistry, 23 (2004) 8 535-552.
[19] G. Kang et al., Quantitative analysis of mixed hydrofluoric and nitric acids using Raman spectroscopy with partial least squares regression, Talanta, 81 (2010) 1413-1417.
[20] F. Estienne et al., Multivariate calibration with Raman spectroscopic data: a case study, Analytica Chimica Acta, 424 (2000) 185–201.
[21] T.H. Kauffmann and M.D. Fontana, Optical sensor of salt concentration: uncertainty evaluation, Sensors and Actuators B: Chemical, 161 (2012) 1 21-27.
[22] A. Rosengard, Probabilités et statistiques en recherche scientifique, Dunod, (1972) pp. 178–194.
[23] M. Feinberg, Validation interne des méthodes d'analyse, Techniques de l'ingénieur, (2001) 224–228.
[24] T. Lu, C. Chen, Uncertainty evaluation of humidity sensors calibrated by saturated salt solutions, Measurement, 40 (2007) 591–599.
[25] R Core Team (2012). R : A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http ://www.R-project.org/.
[26] R. Li et al., Hydrogen bonded structure of water and aqueous solutions of sodium halides: a Raman spectroscopic study, Journal of Molecular Structure, 707 (2004) 83-88.
[27] Q. Sun and C. Qin, Chemical Geology, 283 (2011) 274-278.

Figure1
Click here to download high resolution image

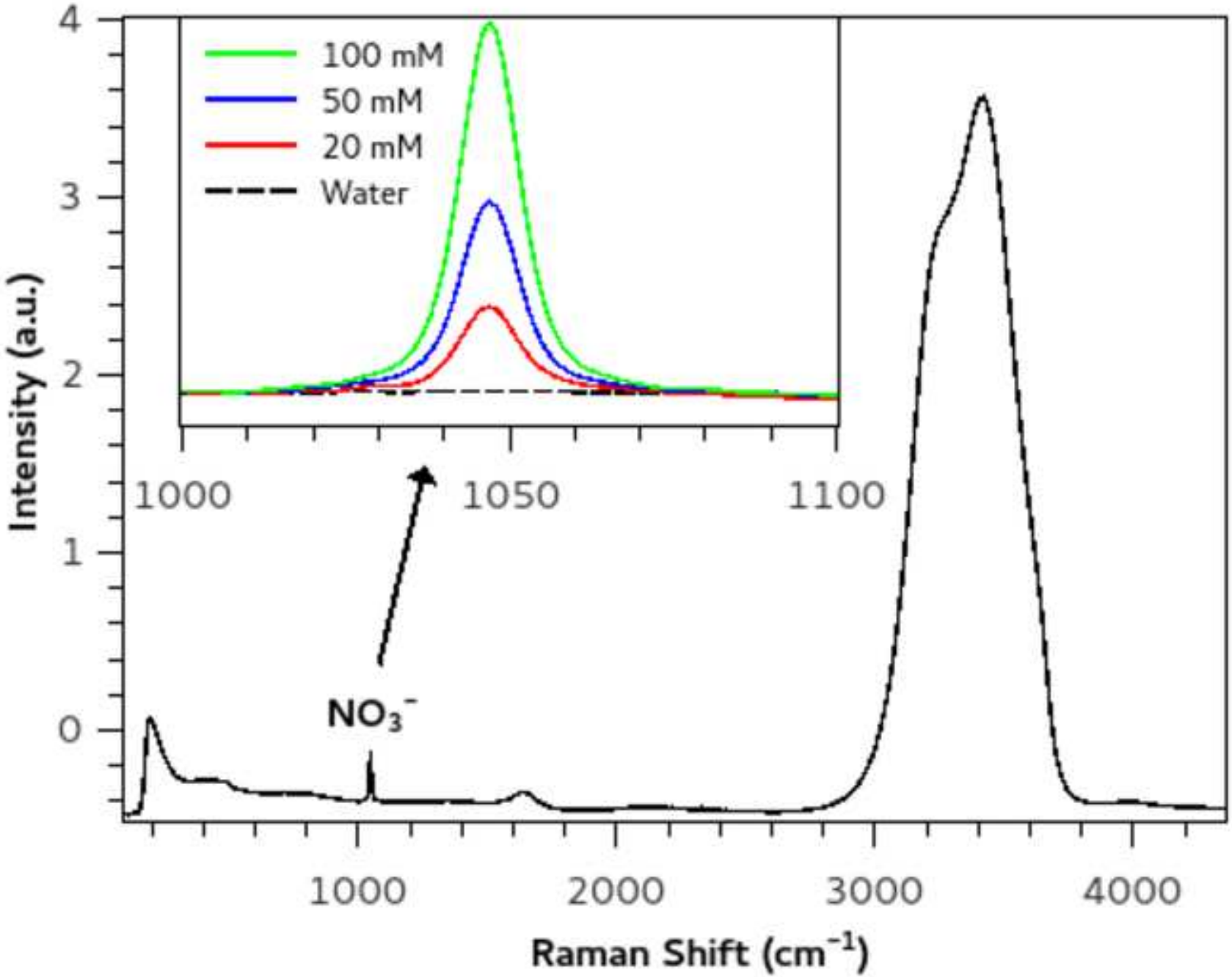

Figure2
Click here to download high resolution image

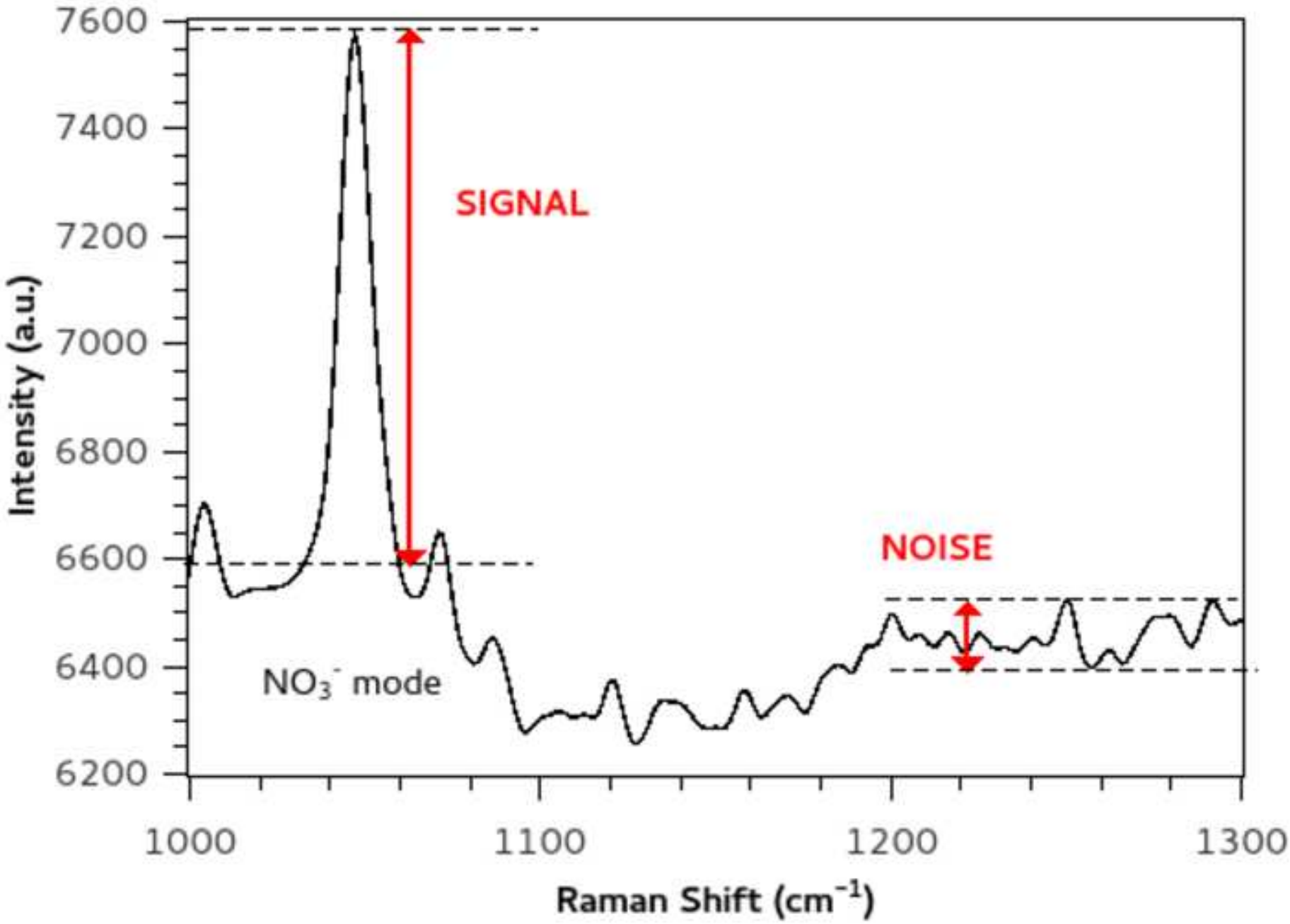

Figure3
Click here to download high resolution image

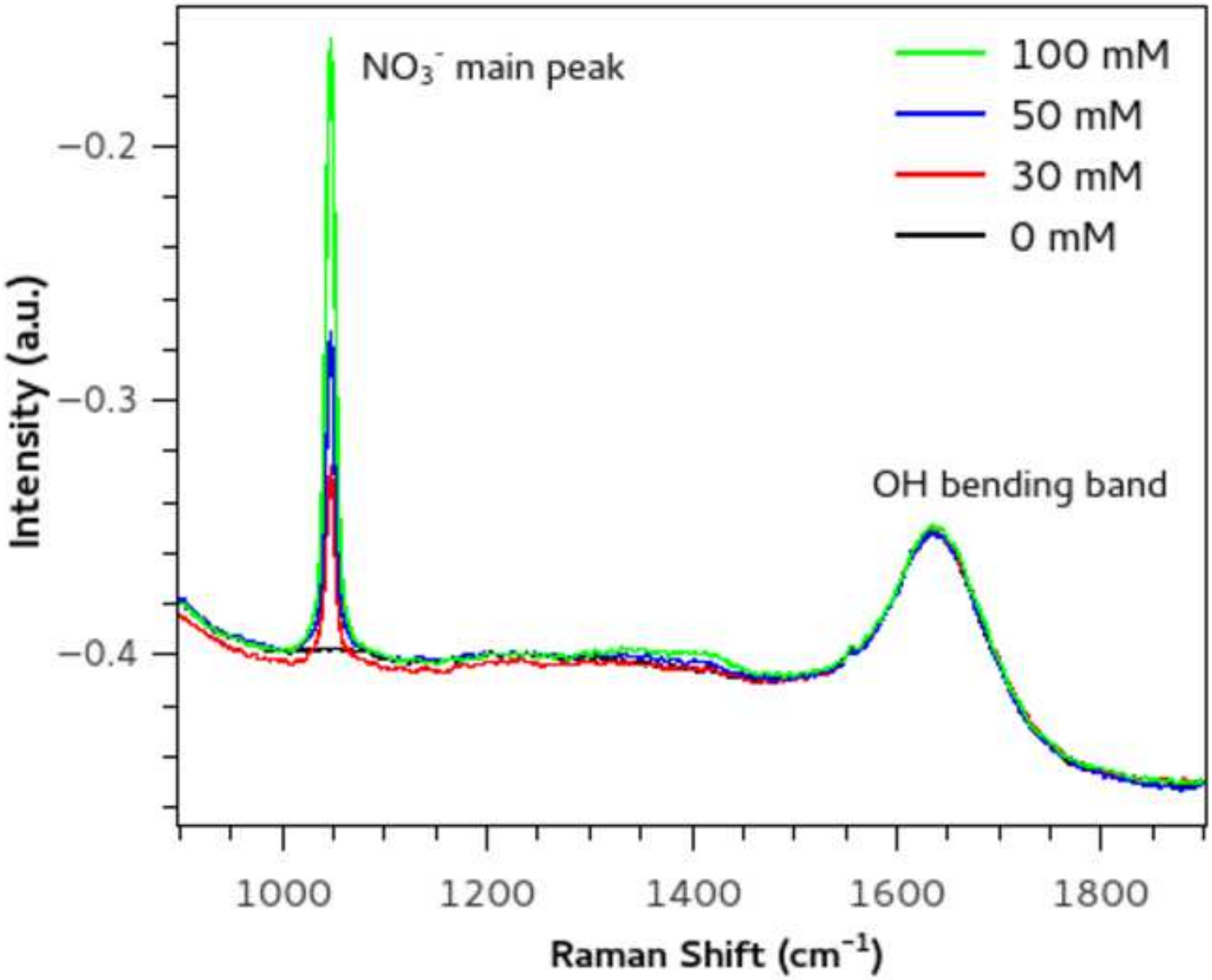

Figure4

Click here to download high resolution image

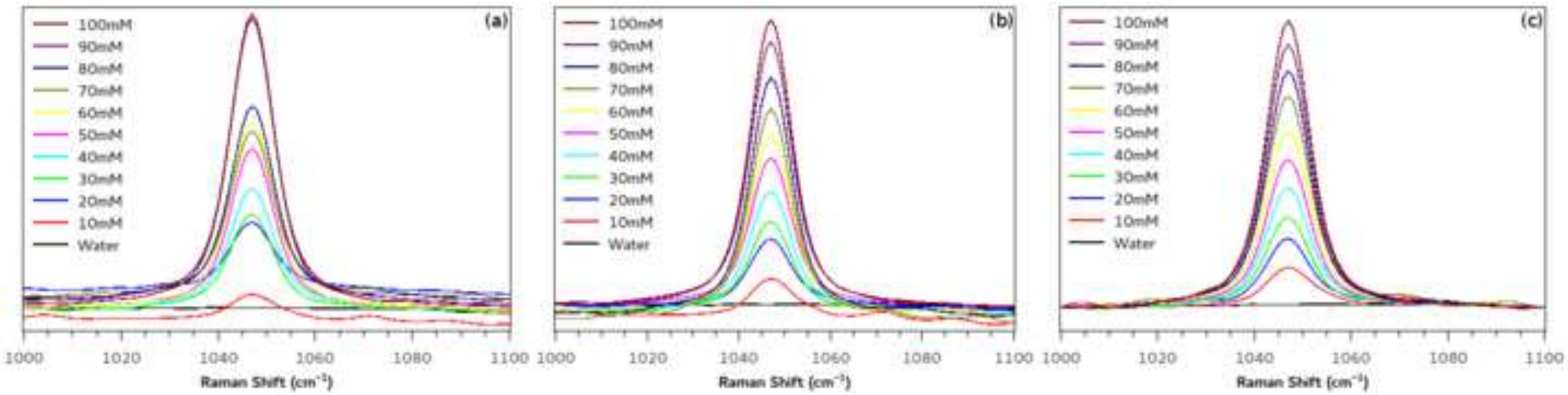

Figure5
Click here to download high resolution image

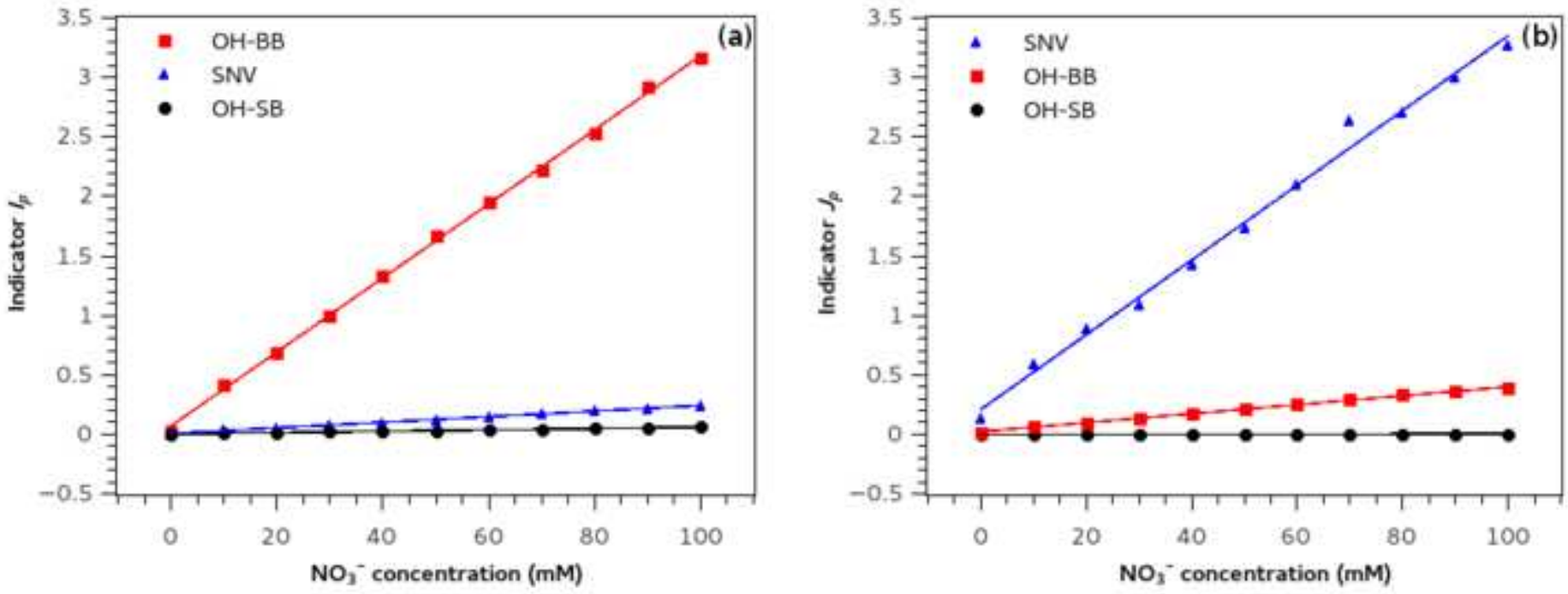
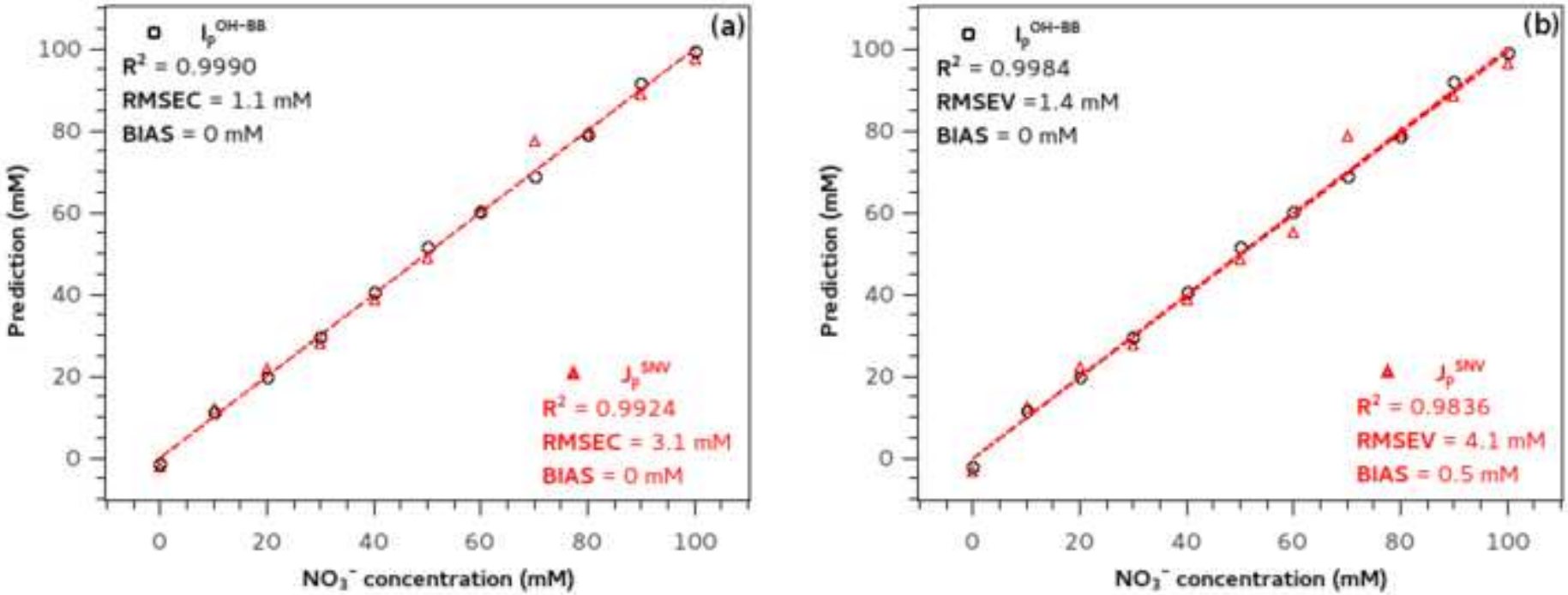
Figure6
Click here to download high resolution image