Postprint of: Bruździak P., Vapor correction of FTIR spectra – A simple automatic least squares approach, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, Vol. 223 (2019), 117373, DOI: 10.1016/j.saa.2019.117373

Vapor correction of FTIR spectra – a simple automatic least squares approach

Piotr Bruździak*

Gdańsk University of Technology, Department of Physical Chemistry,
Narutowicza 11-12, 80-233 Gdańsk, Poland

Abstract

FTIR spectroscopy is one of the best techniques to study intermolecular interactions. However, such an application requires high quality spectra with as little noise as possible, which are often difficult to obtain. One of the main sources of unwanted interference is water vapor. Here a robust method is proposed for automatic, fast and reliable vapor correction of FTIR spectra. The presented least squares approach of vapor subtraction using many vapor spectra and a special residual function provides a much better correction. It does not rely on the researcher's experience, no coefficients are arbitrarily chosen or tweaked, thus such results are more trustworthy and accurate.

Keywords: FTIR Spectroscopy, Vapor Correction, Least Squares

---

1. Introduction

FTIR spectroscopy is an invaluable tool facilitating a non-invasive testing of various types of samples. It is often used in a simple way to identify samples, confirm the presence of various functional groups or the degree of

---

*piotr.bruzdziak@pg.edu.pl

conversion of substrates. Such a use of this kind of spectroscopy often requires neither excessive measurement parameters nor a sophisticated treatment of spectral data and the confirmation of the presence or absence of characteristic absorption bands can be made even with a significant noise.

However, it is not the case when the FTIR spectroscopy is used to study interactions in solutions, mainly aqueous, or changes in the secondary structure of proteins. Such a kind of experiments requires high resolutions, a large number of independent scans, and yet the changes in the shape of measured spectra in a series are usually small and obscured by noise and other interfering factors. One of the main of such factors is water vapor present even in a well purged spectrometers, especially when semi-open ATR accessories are used. Many strategies exist to diminish its influence on spectra: purging with a dry inert gas or dried air, air-evacuated vacuum spectrometers, software-based automatic vapor subtraction algorithms, advanced model-based algorithms [1], special methods of data acquisition [2], or a simple by hand subtraction performed by a researcher.

In the case of protein solution spectra measurements, the amide I band is inevitably affected by the water vapor bands and its complex structure is prone to the vapor correction . All methods of the resolution enhancement routinely used in the protein FTIR spectroscopy (e.g. second derivative, Fourier self-deconvolution, etc.), increase also signals originating in the improperly subtracted vapor-related bands [3, 4, 5, 6]. Thus, the vapor correction is such an important and not trivial task which has to be performed in the best possible way.

The water vapor spectrum is composed of many sharp peaks which are significantly sensitive to even slight changes of temperature in the measurement chamber. Even the smallest difference between temperatures corresponding to the vapor spectrum and other spectra can result in the occurrence of narrow differential bands, really difficult to compensate. Purging is a good choice for a short series of measurements, but the moisture content and the temperature can vary significantly during a long experiment. It should be used routinely during such demanding studies, however, it does not provide a complete vapor correction on its own. The use of vacuum is expensive and may be difficult with liquid samples. Software algorithms use a database of vapor spectra which can be, and usually are inadequate to the conditions prevailing in the laboratory and may significantly affect measured spectra in an uncontrollable way and introduce artifacts. The smoothing is never a way to deal with interference by water vapor as some spectral information and resolution can be lost.

A direct registration and manual subtraction of the vapor spectrum is usually the best choice if the time distance between the spectral series of interest and the measured vapor spectrum is short. However, if the the vapor spectrum is measured after the series (i.e. at the end of experiment) it rarely can be perfectly subtracted from the first measured spectrum in the series. The same applies if the vapor spectrum is measured before the series – it is difficult to subtract it properly from the last spectrum of interest. Another problem arises if the number of measured spectra to be corrected is very high. An experienced researcher can subtract the vapor spectrum from each of the spectra manually with an arbitrarily chosen subtraction coefficient, however, because only one vapor spectrum can be subtracted at a time, the compensation of temperature-related changes in the vapor bands is virtually impossible. Such a procedure is also time consuming.

The method described in this paper originates in the last described case. However, the researcher's participation is limited and despite this it gives a surprisingly good results without any assumption of subtraction coefficients.

2. Materials and methods

2.1. Method description

The presented method has two significant differences in comparison to a standard way of spectral series collection and treatment: 1) additional steps during measurements are introduced, and 2) the vapor spectra fitting to the raw spectra uses a special residual function.

Experimentally, the method differs only slightly from the standard way of spectra series collection. Instead of a single water vapor spectra measurement, either at the beginning or at the end of the experiment, several spectra of an empty measurement chamber (or a clean internal reflection element if ATR spectra are measured) are collected between samples. The number of these spectra depends on the experimental conditions (e.g.the quality of inert gas used to purge the spectrometer, humidity, temperature, etc.), However, usually only two to three vapor spectra have a considerable contribution to the overall spectral series.

The computational algorithm is fairly simple and employs the least squares curve fitting procedure. In contrast to the band fitting procedure, routinely used in FTIR spectroscopy, the residual function used for minimization is not a simple difference between original spectrum and a sum of vapor spectra multiplied by appropriate coefficients. It should be a difference between original spectrum (Y_ν) and a sum of n vapor spectra ($V_{\nu,n}$ with appropriate subtraction coefficients a_n) and an ideal vapor-corrected spectrum $\bar{Y}_\nu$:

$$r_\nu = Y_\nu - \Big(\sum_n a_n V_{n,\nu} + \bar{Y}_\nu\Big) \tag{1}$$

The equation can be represented in an equivalent form:

$$r_\nu = \Big(Y_\nu - \sum_n a_n V_{\nu,n}\Big) - \bar{Y}_\nu \tag{2}$$

where $(Y_\nu - \sum_n a_n V_{\nu.n})$ is the original spectrum from which n vapor spectra were subtracted, and $\bar{Y}_\nu$ is the spectrum perfectly devoid of the vapor contribution. However, $\bar{Y}_\nu$ is not available before subtraction and its shape has to be estimated. In each optimization step the vapor-corrected spectrum is approximated by smoothing of the difference between the original spectrum and the sum of vapor spectra multiplied by initial or optimized coefficients, i.e. $(Y_\nu - \sum_n a_n V_{\nu.n})$. If these subtraction coefficients are close to the ideal ones, or become so during the optimization procedure, vapor spectra are subtracted almost perfectly and the smoothing procedure diminishes only the experimental error. This way, the smoothed spectrum becomes as close to the ideal vapor-corrected spectrum as possible, and r_ν can be minimized. The least squares minimization procedure with such a residual function results with a set of optimized vapor spectra subtraction coefficients. Finally, these coefficients can be used to remove the vapor contribution from the original spectrum. All these steps are additionally presented in Figure 1.

2.2. General remarks

In the Python script[1] available on the web page of the article the Savitzky–Golay smoothing algorithm is applied [7], however, any other efficient smoothing procedures can be employed to this purpose. Additionally, a quadratic baseline is implemented in the script during the optimization procedure to compensate any baseline drifts in measured spectra. However, only vapor components are finally subtracted from the raw spectra, not the fitted baseline.

It seems that the presented method is not very sensitive to the choice of the smoothing parameters[2] making it easier to apply by an inexperienced researcher. It must be stressed, however, that the smoothing step is employed only for estimation of the ideal spectrum shape and is not used to artificially improve the shape of original spectrum or the resultant vapor-corrected spectrum.

The method gives also the best results if applied to a selected part of the FTIR spectrum where the vapor noise affects the analyzed bands, and not to the full range mid-FTIR spectrum. Thus, in the following example the region of 1950–1350 cm^{-1} was selected.

---

[1] vaporfit.py and sample data files can be found online at Supplementary Material web page of the article.

[2] During the development of the method, I tried various combinations of polynomial degrees and number of points used in the Savitzky–Golay smoothing algorithm. The results were equally good if the polynomial degrees were higher than 2 and the number of points was less than 50. The improper subtraction was apparent if those parameters were very high or very low. However, the useful range of those parameters was still very broad and the choice did not affect the quality of results.

2.3. Experimental setup of the example application

An example series of 36 aqueous solutions of *N*-methylacetamide (NMA, 0.00 to 1.01 $mol \cdot dm^{-3}$) were measured by means of the Nicolet 8700 FTIR spectrometer (Thermo) equipped with a single reflection GoldenGate ATR accessory (Specac). The accessory had a diamond single internal reflection element opened to the outside of the FTIR spectrometer. The interior of the accessory was also purged with dry nitrogen. Such a setup allowed to change samples without the need to open the measuring chamber and minimized a possible interruption in the flow of dry gas. The resolution of all spectra was set to 2 cm^{-1} and the number of independent scans was set to 256. The spectrometer was purged with dry nitrogen two hours before and during the experiment. One background spectrum was measured prior to the series. During the measurement, nine additional vapor spectra were collected after each four spectra of the series.

The series of spectra of aqueous solutions of NMA is a part of a larger experimental setup which will provide information on interactions of this small peptide-mimicking molecule with other co-solutes. The scope of the article is, however, focused on the method of vapor spectra subtraction, thus the meaning of changes visible in these spectra is not given in this paper. This particular experiment will serve to determine changes caused in the shape of the amide I band caused only by the change in NMA concentration. In this case, the band shape is composed of two main contributions: carbonyl and amide bands of NMA, and OH bending bands of water. Other examples of vapor correction are available in Supplementary Material.

3. Results and discussion

Results of the correction of NMA spectral series with nine vapor spectra are presented in Supplementary Material file Figure (the number of visible spectra shown in this figure has been reduced for better clarity). The selected wavenumber range corresponds to the amide I band of the NMA and bending vibrations of water molecules. Other examples are also available in Supplementary Material with two additional data files to test the vaporfit.py script.

A selected single raw spectrum has been overlaid with the result of vapor-correction in Figure 2a. As can be seen in Figure 2b, two first vapor spectra have the highest impact on the correction as their subtraction coefficients are significantly different from 0. However, their signs are opposite, suggesting that none of the corresponding spectra could perfectly compensate vapor bands visible in Figure 2a on its own and the correction is improved by interchanging of these two vapor spectra. In this case, these subtraction coefficients indicate that the vapor contribution is in-between first and second vapor spectrum and a subtraction of single selected vapor spectrum could be insufficient. The most probable reason for such a set of coefficients is that the vapor properties, expressed also in the measured sample spectra, varied mostly between first and second measurements of vapor spectra. The properties and spectral shape of other vapor spectra was apparently fairly constant. However, the fitting procedure involving numerous parameters can result in a completely different set of coefficients yet the most important goal of the method is the most accurate vapor subtraction, not the determination of subtraction coefficients itself.

Yet the method can also be used if only one vapor spectrum is available. In such a situation, it just finds the optimal subtraction coefficient of the spectrum from the spectra of interest. The time consuming step of by hand vapor correction is omitted and the whole spectra analysis can be significantly sped up. An example of such a case is presented in Figure 3. Panels a) and b) show a fragment of the series of aqueous solutions of NMA spectra in the range of 1850–1750 cm^{-1} with corresponding nine vapor spectra. Spectra from a) were corrected in two ways: c) by using only one vapor spectrum and d) with all available vapor spectra. The difference could be barely recognizable on a full-range spectrum, and in most cases such a subtraction would be really satisfying, yet the magnification reveals that the use of all available vapor spectra results in a better correction.

4. Conclusions

Because many vapor spectra, corresponding to variable environmental conditions of the experiment, are subtracted simultaneously, the subtraction provides a much better correction. The method uses vapor spectra measured before, after and between samples which gives a perfect picture of environmental changeability during the experiment. It does not rely on the researcher's experience, no coefficients are arbitrarily chosen or tweaked more or less by hand, thus such results can be more trustworthy and more accurate. The procedure is almost automatic and can save time as the vapor spectra are not subtracted one by one by a researcher. Its implementation in a computational environment (e.g. Python, Matlab) is easy as it involves only an appropriate residual function for the least squares method including the smoothing step. No calibration step is needed and there is no need for a more sophisticated methods of data analysis (e.g. Principal Component Analysis).

5. Acknowledgements

This work was funded by the National Science Centre, Poland (grant 2017/26/D/NZ1/00497).

6. Declaration of interest

Declarations of interest: none

7. References

[1] S. W. Bruun, A. Kohler, G. D. Sockalingum, H. Martens, I. Adt, M. Manfait, Correcting Attenuated Total Reflection–Fourier Transform Infrared Spectra for Water Vapor and Carbon Dioxide, Applied Spectroscopy 60 (2006) 1029–1039.

[2] Y. Chen, H. S. Wang, Y. Morisawa, Y. Ozaki, Concept and properties of an infrared hybrid single-beam spectrum and its application to eliminate solvent bands and other background interferences, Talanta 119 (2014) 105–110.

[3] D. Perez-Guaita, J. Kuligowski, G. Quintás, S. Garrigues, M. De La Guardia, Atmospheric compensation in fourier transform infrared (FT-IR) spectra of clinical samples, Applied Spectroscopy 67 (2013) 1339–1342.

[4] G. Vedantham, H. G. Sparks, S. U. Sane, S. Tzannis, T. M. Przybycien, A holistic approach for protein secondary structure estimation from infrared spectra in H2O solutions, Analytical Biochemistry 285 (2000) 33–49.

[5] S. E. Reid, D. J. Moffatt, J. E. Baenziger, The selective enhancement and subsequent subtraction of atmospheric water vapour contributions from Fourier transform infrared spectra of proteins, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 52 (1996) 1347–1356.

[6] E. Goormaghtigh, J.-M. Ruysschaert, Subtraction of atmospheric water contribution in Fourier transform infrared spectroscopy of biological membranes and proteins, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 50 (1994) 2137–2144.

[7] A. Savitzky, M. J. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Analytical Chemistry 36 (1964) 1627–1639.

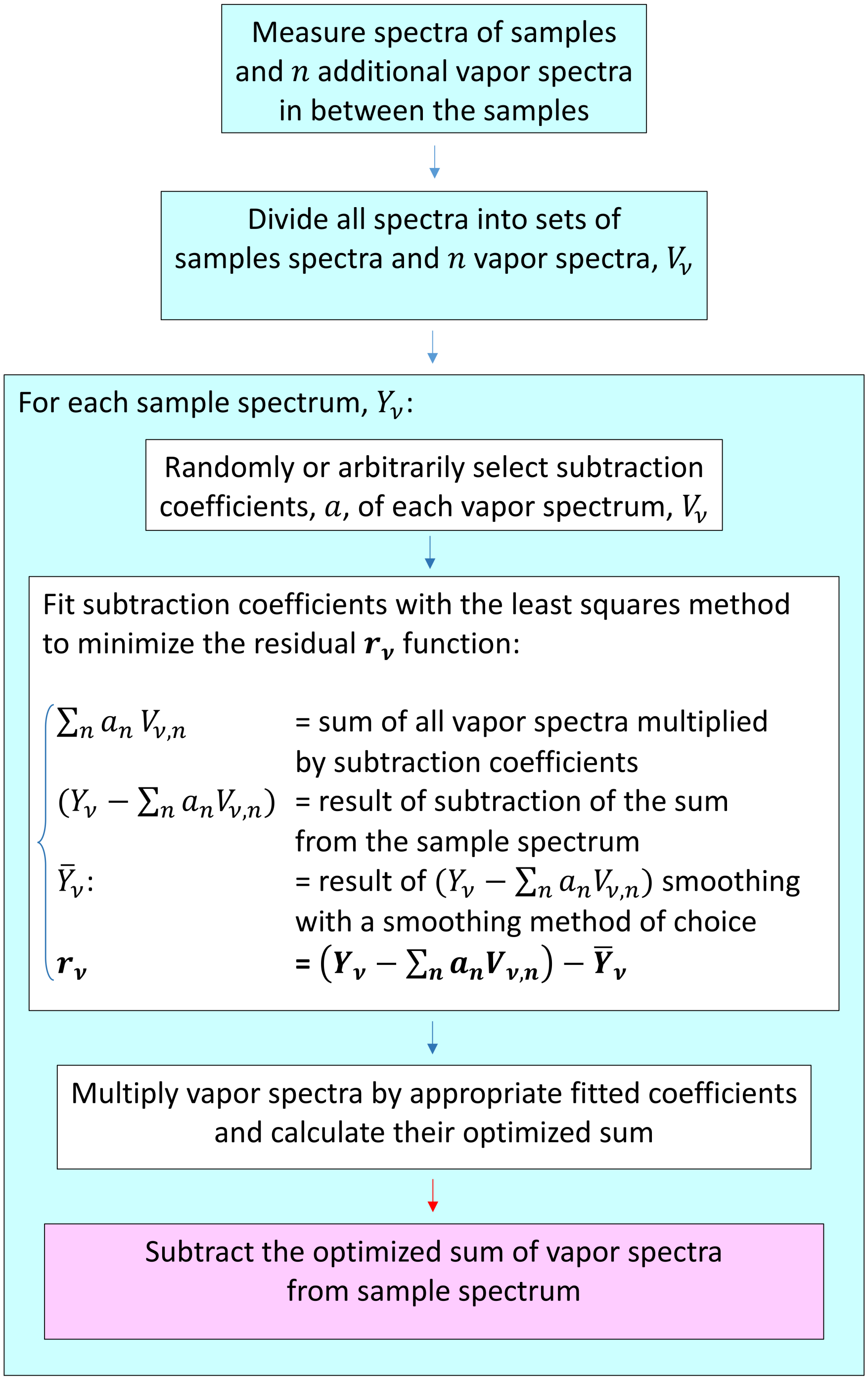

Figure 1: Chart representing main steps of the vapor spectra subtraction method.

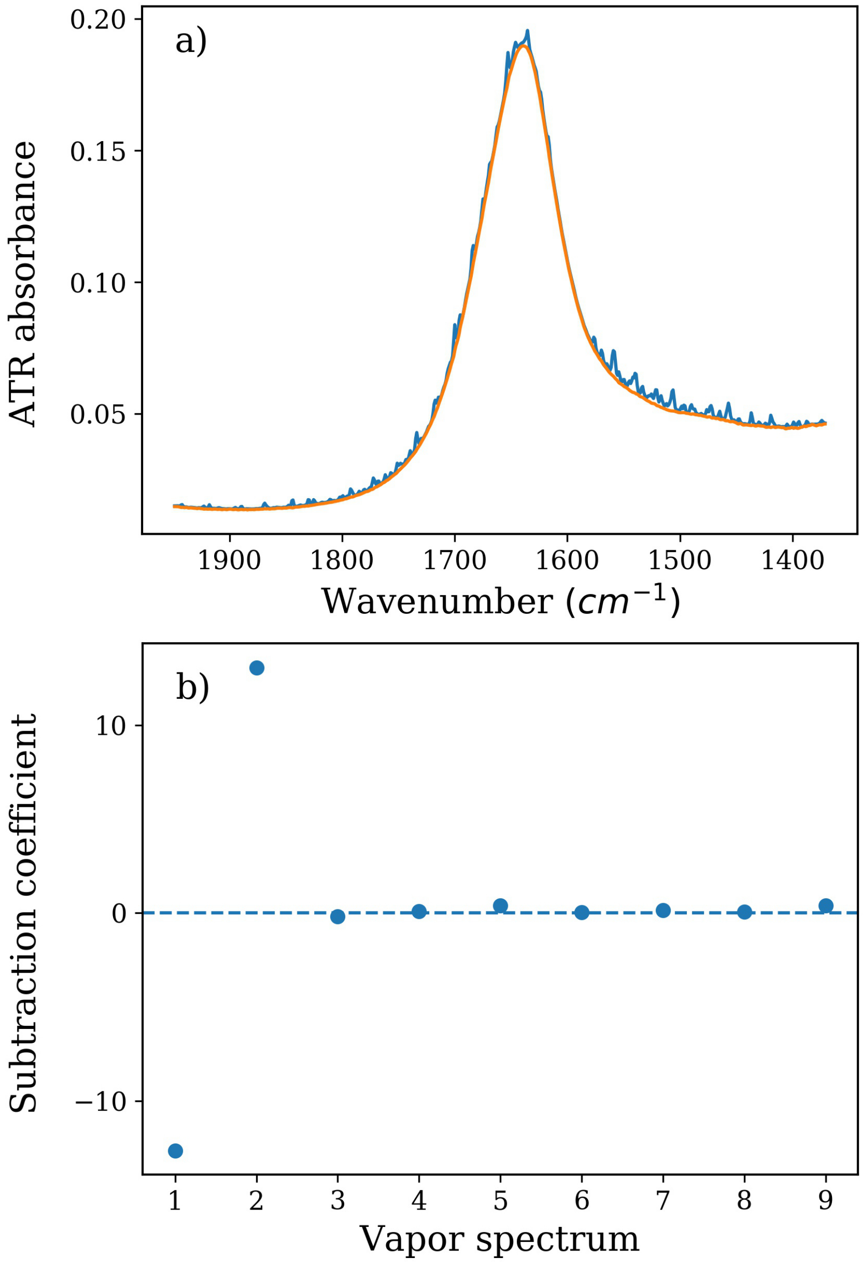

Figure 2: A single spectrum (blue) overlaid with a corresponding vapor-corrected spectrum (orange), and b) subtraction coefficients of consecutive vapor spectra optimized during the subtraction procedure.

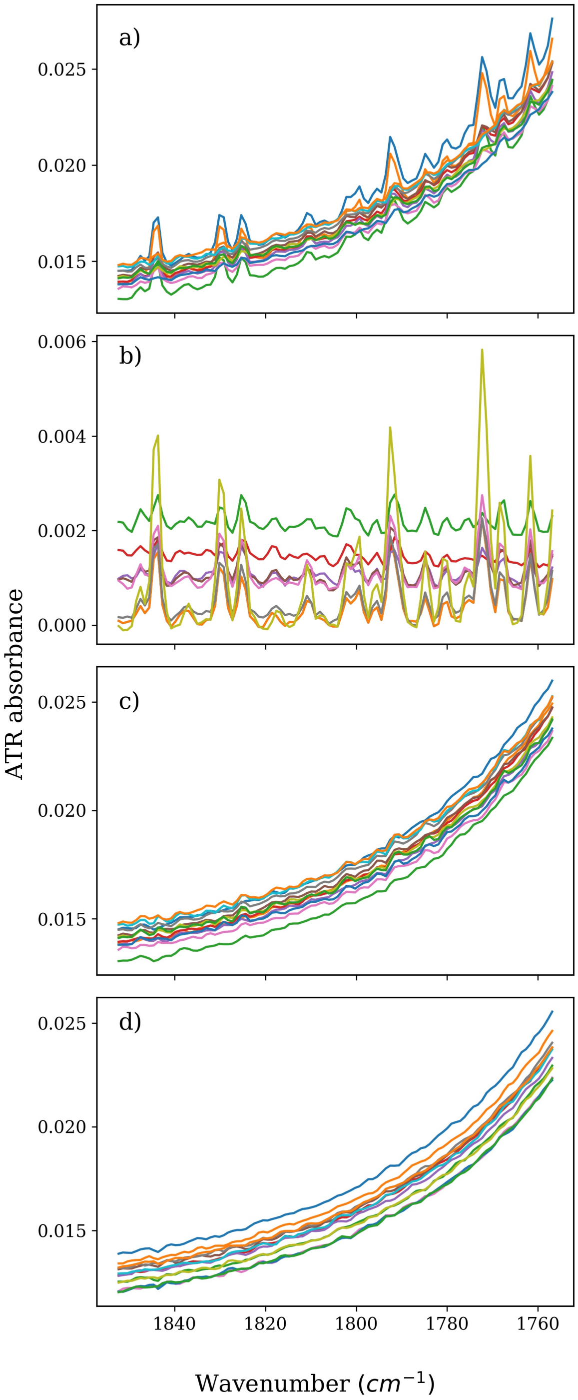

Figure 3: a) A magnified fragment of the spectra series from Figure 1a with b) nine vapor spectra in the same spectral region. c) The result of subtraction of a single vapor spectrum from the series in comparison to d) the result of all nine vapor spectra subtraction.