

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

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

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

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

29 The water vapor spectrum is composed of many sharp peaks which are sig-



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

44 A direct registration and manual subtraction of the vapor spectrum is
45 usually the best choice if the time distance between the spectral series of
46 interest and the measured vapor spectrum is short. However, if the the vapor
47 spectrum is measured after the series (i.e. at the end of experiment) it rarely
48 can be perfectly subtracted from the first measured spectrum in the series.
49 The same applies if the vapor spectrum is measured before the series – it is
50 difficult to subtract it properly from the last spectrum of interest. Another
51 problem arises if the number of measured spectra to be corrected is very
52 high. An experienced researcher can subtract the vapor spectrum from each
53 of the spectra manually with an arbitrarily chosen subtraction coefficient,
54 however, because only one vapor spectrum can be subtracted at a time, the



55 compensation of temperature-related changes in the vapor bands is virtually
56 impossible. Such a procedure is also time consuming.

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

60 **2. Materials and methods**

61 *2.1. Method description*

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

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

75 The computational algorithm is fairly simple and employs the least squares
76 curve fitting procedure. In contrast to the band fitting procedure, routinely
77 used in FTIR spectroscopy, the residual function used for minimization is
78 not a simple difference between original spectrum and a sum of vapor spec-

79 tra multiplied by appropriate coefficients. It should be a difference between
80 original spectrum (Y_ν) and a sum of n vapor spectra ($V_{\nu,n}$ with appropriate
81 subtraction coefficients a_n) and an ideal vapor-corrected spectrum \bar{Y}_ν :

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

82 The equation can be represented in an equivalent form:

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

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



98 *2.2. General remarks*

99 In the Python script¹ available on the web page of the article the Savitzky–
100 Golay smoothing algorithm is applied [7], however, any other efficient smooth-
101 ing procedures can be employed to this purpose. Additionally, a quadratic
102 baseline is implemented in the script during the optimization procedure to
103 compensate any baseline drifts in measured spectra. However, only vapor
104 components are finally subtracted from the raw spectra, not the fitted base-
105 line.

106 It seems that the presented method is not very sensitive to the choice of
107 the smoothing parameters² making it easier to apply by an inexperienced re-
108 searcher. It must be stressed, however, that the smoothing step is employed
109 only for estimation of the ideal spectrum shape and is not used to artificially
110 improve the shape of original spectrum or the resultant vapor-corrected spec-
111 trum.

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

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

²During 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.



116 *2.3. Experimental setup of the example application*

117 An example series of 36 aqueous solutions of *N*-methylacetamide (NMA,
118 0.00 to 1.01 mol · dm⁻³) were measured by means of the Nicolet 8700 FTIR
119 spectrometer (Thermo) equipped with a single reflection GoldenGate ATR
120 accessory (Specac). The accessory had a diamond single internal reflection
121 element opened to the outside of the FTIR spectrometer. The interior of the
122 accessory was also purged with dry nitrogen. Such a setup allowed to change
123 samples without the need to open the measuring chamber and minimized a
124 possible interruption in the flow of dry gas. The resolution of all spectra was
125 set to 2 cm⁻¹ and the number of independent scans was set to 256. The
126 spectrometer was purged with dry nitrogen two hours before and during the
127 experiment. One background spectrum was measured prior to the series.
128 During the measurement, nine additional vapor spectra were collected after
129 each four spectra of the series.

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



140 3. Results and discussion

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

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



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

177 4. Conclusions

178 Because many vapor spectra, corresponding to variable environmental
179 conditions of the experiment, are subtracted simultaneously, the subtraction
180 provides a much better correction. The method uses vapor spectra mea-
181 sured before, after and between samples which gives a perfect picture of
182 environmental changeability during the experiment. It does not rely on the
183 researcher's experience, no coefficients are arbitrarily chosen or tweaked more
184 or less by hand, thus such results can be more trustworthy and more accu-
185 rate. The procedure is almost automatic and can save time as the vapor
186 spectra are not subtracted one by one by a researcher. Its implementation
187 in a computational environment (e.g. Python, Matlab) is easy as it involves
188 only an appropriate residual function for the least squares method including



189 the smoothing step. No calibration step is needed and there is no need for
190 a more sophisticated methods of data analysis (e.g. Principal Component
191 Analysis).

192 **5. Acknowledgements**

193 This work was funded by the National Science Centre, Poland (grant
194 2017/26/D/NZ1/00497).

195 **6. Declaration of interest**

196 Declarations of interest: none

197 **7. References**

- 198 [1] S. W. Bruun, A. Kohler, G. D. Sockalingum, H. Martens, I. Adt, M. Man-
199 fait, Correcting Attenuated Total Reflection–Fourier Transform Infrared
200 Spectra for Water Vapor and Carbon Dioxide, *Applied Spectroscopy* 60
201 (2006) 1029–1039.
- 202 [2] Y. Chen, H. S. Wang, Y. Morisawa, Y. Ozaki, Concept and properties of
203 an infrared hybrid single-beam spectrum and its application to eliminate
204 solvent bands and other background interferences, *Talanta* 119 (2014)
205 105–110.
- 206 [3] D. Perez-Guaita, J. Kuligowski, G. Quintás, S. Garrigues, M. De La
207 Guardia, Atmospheric compensation in fourier transform infrared (FT-
208 IR) spectra of clinical samples, *Applied Spectroscopy* 67 (2013) 1339–
209 1342.



- 210 [4] G. Vedantham, H. G. Sparks, S. U. Sane, S. Tzannis, T. M. Przybycien, A
211 holistic approach for protein secondary structure estimation from infrared
212 spectra in H₂O solutions, *Analytical Biochemistry* 285 (2000) 33–49.
- 213 [5] S. E. Reid, D. J. Moffatt, J. E. Baenziger, The selective enhancement
214 and subsequent subtraction of atmospheric water vapour contributions
215 from Fourier transform infrared spectra of proteins, *Spectrochimica Acta*
216 *Part A: Molecular and Biomolecular Spectroscopy* 52 (1996) 1347–1356.
- 217 [6] E. Goormaghtigh, J.-M. Ruysschaert, Subtraction of atmospheric wa-
218 ter contribution in Fourier transform infrared spectroscopy of biological
219 membranes and proteins, *Spectrochimica Acta Part A: Molecular and*
220 *Biomolecular Spectroscopy* 50 (1994) 2137–2144.
- 221 [7] A. Savitzky, M. J. Golay, Smoothing and Differentiation of Data by
222 Simplified Least Squares Procedures, *Analytical Chemistry* 36 (1964)
223 1627–1639.



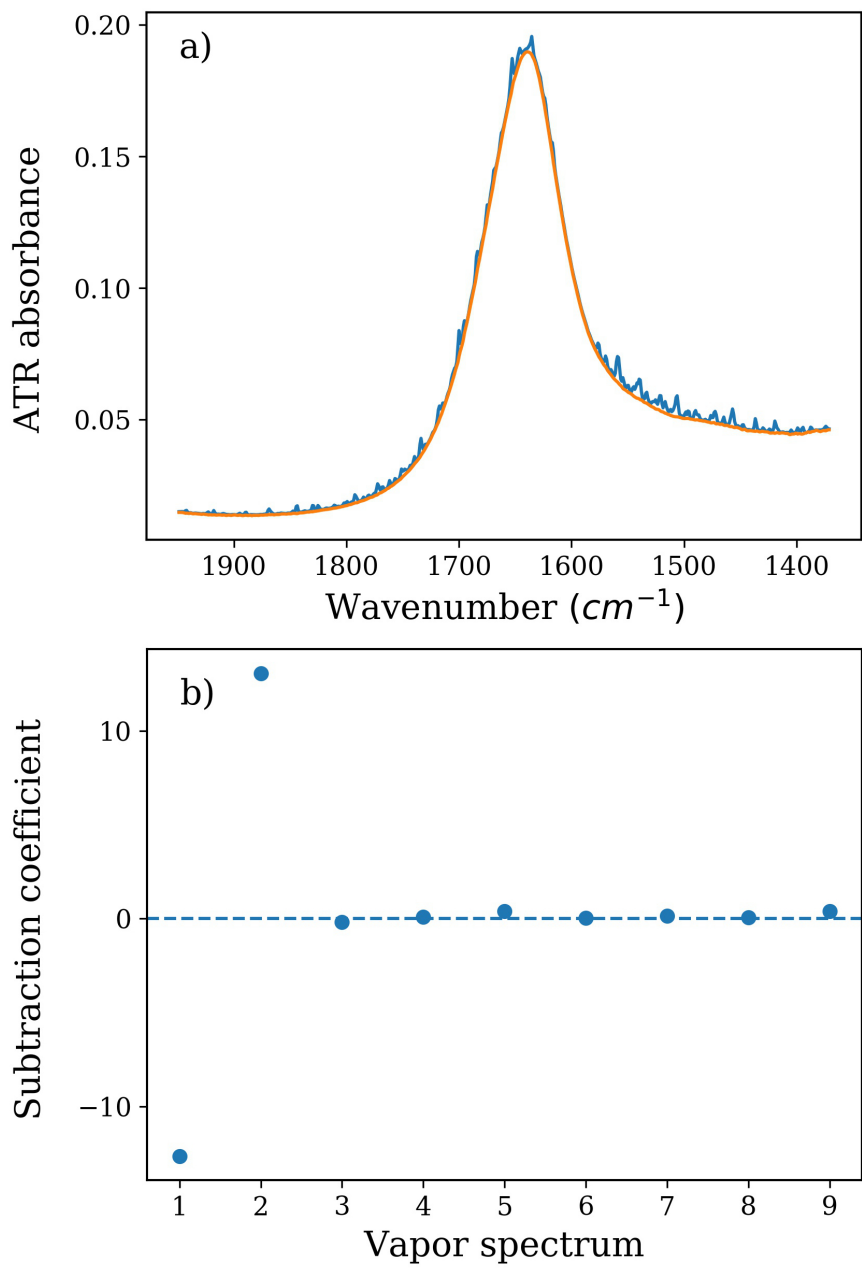


Figure 2: A single spectrum (blue) overlaid with a corresponding vapor-corrected spectrum (orange), and b) subtraction coefficients of consecutive vapor spectra optimized during the subtraction procedure.

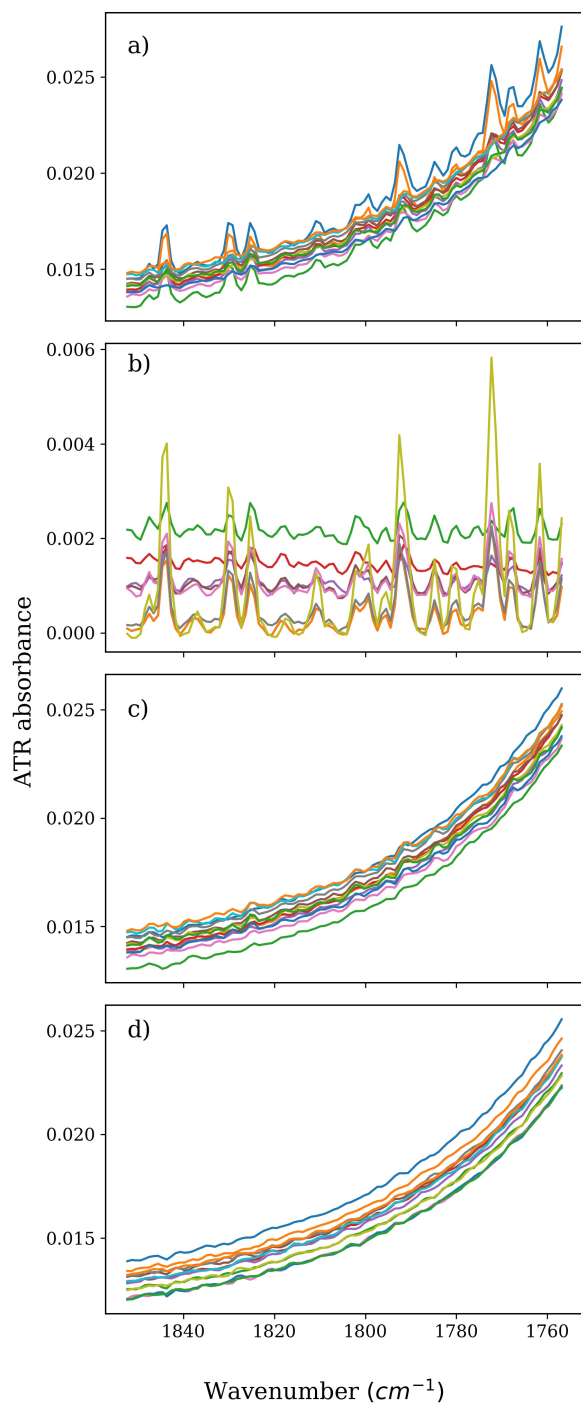


Figure 3: a) A magnified fragment of the spectra series from Figure 1a with b) nine vapor spectra in the same spectral region. c) The result of subtraction of a single vapor spectrum from the series in comparison to d) the result of all nine vapor spectra subtraction.