Advances in Data Processing for Open-Path Fourier Transform Infrared Spectrometry of Greenhouse Gases†

Limin Shao

Department of Chemistry, University of Science and Technology of China, Hefei, Anhui, 230026, PROC

Peter R. Griffiths*

Department of Chemistry, University of Idaho, Moscow, Idaho 83844-2343

April B. Leytem

United States Department of Agriculture, Agricultural Research Service, Northwest Irrigation and Soils Research Laboratory, Kimberly, Idaho 83341

The automated quantification of three greenhouse gases, ammonia, methane, and nitrous oxide, in the vicinity of a large dairy farm by open-path Fourier transform infrared (OP/FT-IR) spectrometry at intervals of 5 min is demonstrated. Spectral pretreatment, including the automated detection and correction of the effect of interrupting the infrared beam, is by a moving object, and the automated correction for the nonlinear detector response is applied to the measured interferograms. Two ways of obtaining quantitative data from OP/FT-IR data are described. The first, which is installed in a recently acquired commercial OP/FT-IR spectrometer, is based on classical least-squares (CLS) regression, and the second is based on partial least-squares (PLS) regression. It is shown that CLS regression only gives accurate results if the absorption features of the analytes are located in very short spectral intervals where lines due to atmospheric water vapor are absent or very weak; of the three analytes examined, only ammonia fell into this category. On the other hand, PLS regression works allowed what appeared to be accurate results to be obtained for all three analytes.

It is well-known that greenhouses function because the glass from which they are constructed transmits short-wavelength solar radiation which is then absorbed by the earth inside the greenhouse and heats it above the outside temperature. Since the vast majority of the radiation emitted by a 300 K blackbody is at wavelengths longer than 4 \( \mu \text{m} \) (i.e., wavenumbers <2500 \( \text{cm}^{-1} \)) and glass is opaque to all wavelengths longer than 4 \( \mu \text{m} \), the thermal radiation emitted by the earth is absorbed by the glass, heating it up so that the air inside is heated by conduction and convection. Molecules present in the atmosphere with infrared absorption bands at wavelengths longer than 4 \( \mu \text{m} \) act in an analogous manner to the glass of a greenhouse, hence the name greenhouse gases.

The greenhouse gases that are present at the highest concentration are water and carbon dioxide. These molecules give rise to the strongest greenhouse effect as they are opaque over wide intervals of the mid-infrared spectrum, as shown in Figure 1. Indeed, were it not for the presence of water and CO\(_2\) the earth would be much cooler. Not as commonly recognized is the positive feedback in the greenhouse effect. As the temperature of atmospheric water and CO\(_2\) increases, the population of the higher rotational states is increased, so that the spectral intervals over which the rotational transitions in their vibration—rotation spectrum absorb strongly are extended away from the center of the absorption bands, thereby increasing the opacity over a wider spectral interval and further contributing to the warming of the Earth. Molecules that absorb in the “atmospheric windows” between the bands of H\(_2\)O and CO\(_2\) also contribute to the greenhouse effect, and this paper will cover the measurement of the concentration of three of these gases (NH\(_3\), CH\(_4\), and N\(_2\)O). To be significant contributors to
the greenhouse effect, these molecules should be present at a concentration of several parts-per-billion by volume (ppbv) or else their absorption, and hence the amount by which they heat the atmosphere, will be insignificant.

Open-path Fourier transform infrared (OP/FT-IR) spectrometry is sensitive to most volatile compounds that have absorption bands in atmospheric windows; the typical limit of detection of OP/FT-IR spectrometry for most greenhouse gases is between 1 and 10 ppbv.1–4 The greater the absorptivity of a band absorbing in an atmospheric window, the more it contributes to the greenhouse effect (for a given concentration) and the lower is the detection limit of OP/FT-IR measurements. OP/FT-IR spectrometry is "blind" in spectral intervals where atmospheric water vapor and CO$_2$ absorb strongly; however, molecules that absorb only in these regions do not contribute significantly to the greenhouse effect, and so, measurement of their concentrations is largely unnecessary. As a result, OP/FT-IR spectrometry is an almost ideal technique for measuring the concentration of greenhouse gases as it is reasonably sensitive, noninvasive, fast, capable of determining multiple compounds simultaneously, and only measures those compounds that contribute to global warming. Furthermore, the instrumentation is rugged and relatively easy to handle in the field.5–9

OP/FT-IR spectrometry has also been used for monitoring hazardous air pollutants10 and airborne toxic industrial chemicals11 from various sources such as industrial plants, agricultural operations, burning of biomass, engines of motor vehicles or aircraft, and volcanoes.12 Wide applications of OP/FT-IR spectrometry are found in occupational and environmental air monitoring and military and homeland security. In the past decade, the application of OP/FT-IR spectrometry for atmospheric monitoring has seen a significant increase in Asian countries such as China and South Korea. However, on a worldwide scale, and especially in the USA, this technique is still undergoing a slow acceptance, largely because of the difficulty of processing the data when commercial instruments are used.

The atmosphere is a complex, multicomponent system, the study of which is complicated by uncontrolled or unpredicted factors such as wind, rain, snow, and dust. Furthermore, for OP/FT-IR spectrometers mounted at close to ground level, the beam can be interrupted by vehicular and pedestrian traffic or even by birds. As a result, although a typical OP/FT-IR spectrum contains the information on the path-integrated concentration of greenhouse gases, the determination of these molecules is hampered by various types of interference in addition to the omnipresent absorption of the beam by water vapor and CO$_2$. To add to the analytical challenge, continuous atmospheric monitoring by OP/FT-IR spectrometry yields a huge amount of data from which the concentrations of the various analytes must be extracted. Unfortunately, most of the software of commercial OP/FT-IR spectrometers requires considerable expertise and time on the part of the operator. Furthermore, the chemometric technique that is usually employed, namely classical least-squares (CLS) regression, is far more accurate when all molecules in the spectral interval being investigated (including the lines in the spectrum of water vapor) obey Beer’s Law, which is seldom the case with atmospheric spectra where interfering water lines are often very intense and are measured with an instrument line shape function that is broader than the spectral lines. As a result, nonlinear Beer’s law behavior is often observed, and algorithms to correct for nonlinearity have been described.13 Nonetheless, these algorithms have not been incorporated in all commercially available OP/FT-IR spectrometers, and it is usually necessary to only use very short intervals of the spectrum between the stronger water lines to perform the analysis of compounds for which the spectrum is located in intervals of strong absorption by water vapor.3

A more general and user-friendly software program for OP/FT-IR spectroscopy that is not based on CLS regression and that handles nonlinear behavior of band intensities would increase the quality and automation of the data processing and enable researchers without spectroscopic expertise to use the OP/FT-IR technology for atmospheric monitoring more readily.14,15 In this paper, we will show the results obtained with such a system. The data processing techniques that we use were either adopted or specifically developed in our laboratories both to improve the accuracy of the result and to decrease the spectroscopic expertise on the part of the operator that is required to obtain a reliable result.

**Distortion of the Interferograms by the Extraneous Factors.** The frequencies at which the infrared wavelengths are modulated by the interferometer in a typical OP/FT-IR spectrometer (the so-called Fourier frequencies) are usually greater than 500 Hz. Most of the variations in the signal at frequencies below 500 Hz are caused by either adverse weather conditions or the blockage of the IR beam by a moving object. Such interference can be removed by the application of a high-pass filter to the interferogram. Three years ago, we reported a simple procedure to reject invalid interferograms that had been subjected to high-pass filtering from further analysis in order to eliminate erroneous
results.16 This procedure operates automatically and effectively but at the cost of losing the information at the times those rejected measurements were measured. It is noteworthy that the effect of fog or smoke is completely removed by this approach17 and regions where the frequency of the interference is less than the lowest Fourier frequency are corrected.18

Reducing the Effect of the Nonlinear Response of MCT Detectors. Another source of error is caused by the nonlinear response of the mercury cadmium telluride (MCT) detector. It is well-known that the response of MCT detectors used for FT-IR spectrometry can be nonlinear; i.e., the response is not linearly proportional to the incident flux. This nonlinearity produces a baseline offset in the calculated single-beam spectrum and, hence, introduces a severe error in the absorbance of strong spectral features including many of the lines in OP/FT-IR spectra of atmospheric water vapor.19–21 Therefore, the effect of the nonlinear response must be corrected in order to improve the quantitative accuracy of OP/FT-IR spectrometry. Several correction methods based on hardware and software implementation are reviewed in ref 22. Most of them require iterative calculations, and none of them works well if the energy that is incident on the detector changes significantly, as it often does with OP/FT-IR spectrometry.

We have developed a rapid numerical method in which the correction is implemented by modifying the signal intensities of just three points on the affected interferogram, namely, the centerburst (i.e., the largest data point) and the adjacent points on either side.22 Since these points have larger values than all of the other points in the interferogram, they are the most adversely affected by a nonlinear detector response. Results demonstrated the effectiveness of this correction method. This approach also features fast and automatic computation since an iterative process that is common to most nonlinearity–correction algorithms is not employed. Unlike other algorithms, the one that we use can be applied in real time.

There is another benefit of the nonlinearity correction for OP/FT-IR monitoring. In unattended measurements, we found that some factors, especially the effect of wind, sometimes significantly increased the optical alignment of the instrument and resulted in signals so intense as to overflow the analog-to-digital converter. In our prior experiments, such interferograms were rejected despite the very high centerburst signal.15 As a result, during the initial optical setup, we had to leave room for an incidental alignment improvement. Nonetheless, ever since the integration of the nonlinearity correction into our data processing software package, the effect of a varying optical alignment under windy conditions is a less serious issue than before.

Calculation of the Sample Spectrum. After the above procedures have been carried out, the remaining interferograms are ready for the calculation of the single-beam spectra. The operator first defines the resolution, apodization function, zero-filling factor, and number of scans to be averaged. The single-beam spectrum (SBS) of the sample is, thus, obtained and must be ratioed against a suitable background SBS to yield the transmittance spectrum of the atmosphere, T(v).

In most previous work, the operators attempt to measure the sample and background SBS over the same optical path length with the background spectrum having the same temperature and humidity for both measurements but with all analytes absent. Were this condition to be met for field measurements, when the ratio of the sample and background spectra is calculated, the baseline would be at exactly 100% T, the spectral features due to atmospheric water and CO2 would be completely removed, and the spectra of other molecules that absorb in the atmospheric windows could be observed with no interference other than the effect of noise. In practice, however, the measurement of sample and background spectra with everything identical except the presence of the analyte is never possible, so that, for example, the protocols specified by the American Society for Testing and Materials,2 the United States Environmental Protection Agency (USEPA),3 and the European Committee for Standardization4 rarely, if ever, yield quantitatively accurate results.

Hart and Griffiths showed that, provided that the data are processed by partial least-squares regression (PLS, vide infra), the use of a background SBS measured over a very short path length (i.e., with the retroreflector only about 1 m from the telescope) yields spectra that are far more easily acquired and results that are more accurate than if a background SBS of the same path length as the sample SBS is used.23,24 Not only does this procedure make setting up measurements in the field far easier, but also, for the measurement of important greenhouse gases that have a global average concentration exceeding about 10 ppb, such as methane, the short-path background SBS must be used; otherwise, the absorption due to methane in the long-path background SBS would partially cancel out the corresponding bands in the sample SBS and result in a lower calculated concentration than the true value.

Quantitative Analysis. The most straightforward algorithm for quantitative analysis of OP/FT-IR spectra is classical least-squares regression (CLS), which relies on Beer’s law being obeyed by all molecules absorbing in the spectral range under investigation. To the best of our knowledge, this algorithm is incorporated in all commercial OP/FT-IR spectrometers. However, Beer’s law is a limiting law, and the linearity of the measured absorbance with respect to the concentration of a particular molecule is affected by several factors of which the most important for OP/FT-IR measurements is the effect of resolution on strongly absorbing lines such as those in the vibration—rotation spectra of atmospheric water vapor and CO2.25–28 Generally speaking, the peak absorbance of bands or lines due to trace molecules in


the atmosphere is sufficiently weak in OP/FT-IR spectra, so that the measured absorbance is proportional to concentration even when the spectrum is measured at low resolution. Thus, provided that the spectral interval being used for the analysis is free of interference by water and CO$_2$, CLS gives accurate results. If this is not the case, the quantitative accuracy can be drastically impaired, as we will show later in this paper.

In practice, the complete elimination of lines in the spectrum of H$_2$O and CO$_2$ is far more easily said than done. Several techniques designed to accomplish this goal are outlined in the USEPA guidance document for OP/FT-IR spectroscopy, however, all of them are essentially impossible to implement in practice. Indeed, to the best of our knowledge, there has never been a report of an OP/FT-IR spectral measurement where all the lines in the vibration–rotation spectrum of water vapor have been compensated. Hong and Cho reported a technique to compensate for the water absorption in the measured spectrum with the reference of water that was impaired, as we will show later in this paper.

To implement this approach, single-beam background spectra were measured through a pristine atmosphere over a 9 month period using a wide range of path lengths. During this time, the temperature varied from below $-5$ to $35 \, ^\circ C$; however, since these measurements were all taken in Idaho, the relative humidity rarely exceeded 50% on warm days. Each of these spectra was ratioed against a background SBS measured with the retroreflector mounted very close to the telescope. All interferograms were corrected for the response nonlinearity of the MCT detector. The spectra were then converted to the absorbance format. To each of these spectra, one or more reference spectra of the analytes being investigated were added. To ensure that there are no resolution errors, we took advantage of the fact that at high resolution ($\Delta \tilde{\nu} < \text{fwhm}$) Beer’s law is obeyed over a wide concentration range. After first measuring a reference absorbance spectrum of each analyte at high resolution (0.125 cm$^{-1}$ in our research), these spectra were then scaled by different factors to represent the absorbance spectrum at different concentrations in the range of interest. These spectra were then converted to transmittance, and their Fourier transform was calculated. The Fourier domain array was then truncated to an optical path difference equal to that at which the OP/FT-IR spectra were measured in the field and multiplied by the appropriate apodization function, so that reference spectra of the analytes at various concentrations were generated at the resolution at which the OP/FT-IR spectra were acquired (1 or 4 cm$^{-1}$ in our work).

Finally, the spectra calculated in this way are added to open-path background spectra measured through pristine atmospheres in order to generate the calibration set for PLS regression. For molecules that are known to be present at a detectable concentration in pristine air, the global average concentration was added to the concentration of the added analyte.

The fact that all the single-beam OP/FT-IR spectra that are measured over long paths in our investigations are ratioed against a short path-length background SBS obviates the need to measure a background spectrum over exactly the same path length and with the same temperature and relative humidity as the sample spectrum but with all analytes absent. Because the vibration–rotation lines in the spectrum of water vapor are seen throughout the

---


spectrum, CLS regression cannot be easily applied to obtain quantitative data from such spectra without fitting the water lines in the way that has been described by Hong and Cho29 or Griffith et al.32,33 Consequently, CLS is usually carried out over very short spectral ranges in which absorption by water lines is minimal. To circumvent such problems, we use partial least-squares regression (PLS) as the quantitative algorithm by first building calibration data sets as described in the previous paragraph using OP/FT-IR spectra that were measured through a pristine atmosphere under as wide a range of path length, temperature, and humidity as possible. The calibration spectra contain all the absorption features due to water vapor and CO2 acquired over a wide range of relative humidity, temperature, and path length as well as the spectra of the analytes under investigation. We have used the same set of background spectra for several years without observing any degradation of the results, and concentrations calculated from spectra measured with the two instruments showed only small differences, even though they were manufactured more than 20 years apart.

**EXPERIMENTAL SECTION**

To assess the differences in measuring concentrations of methane, ammonia, and nitrous oxide utilizing either the PLS technique described above or a CLS method, two OP/FT-IR spectrometers were used. Each contained a Globar source, a cube-corner interferometer of the type incorporated in the ABB-Bomem MB-100 FT-IR spectrometer, and a beamsplitter and optics that diverted about 50% of the energy leaving the interferometer to a 25 cm beam-expanding telescope. The expanded beam was passed over an approximately 100 m path to a cube-corner array retroreflector that returned the beam into the telescope and beamsplitter, from which about 50% of the energy was passed to an MCT detector. One instrument (AirSentry, Cerex Monitoring Solutions, Atlanta, GA) had been recently purchased and the other (donated to the University of Idaho after over 10 years of operation by du Pont) was over 25 years old. Over the path lengths studied in this project, the signal-to-noise ratio (SNR) of spectra measured by both instruments was very similar, although the SNR of spectra measured on the older instrument over total path lengths greater than 400 m was poorer.

Spectra were collected over a 200 m total path length (100 m between the telescope and retroreflector) in an area that was located adjacent (to the east) of a large (~13,000 cows) open-freestall dairy located in southern Idaho, as shown in Figure 2. The barns run east and west and are each almost 1 km in length. At the time that these measurements were taken (January, 2010), the winds were generally from an easterly direction. In this case, the two OP/FT-IR spectrometers were located upwind of the cattle but several times the wind changed direction so that it blew from a more westerly direction. OP/FT-IR spectra were acquired continuously at intervals of 5 min over a period of 44 h. Over the measurement period, the average temperature was $-1.3\,^\circ C$, with a daytime high of 5.8 $^\circ C$ and a nighttime low of $-4.4\,^\circ C$. The average relative humidity was 87%, with a high of about 94% (at nighttime, when the temperature was lowest) and a low of 58%. The average wind speed was 1.2 m s$^{-1}$ and never exceeded 2.8 m s$^{-1}$ while these measurements were being taken. Because both instruments were located to the east of the dairy, the analyte concentrations were expected to be near their global mean when the wind was from the east but should show the influence of the dairy emissions when the wind was predominantly from the southwest and west.

The data acquired by the older of the two instruments were processed with the PLS-based software described above. Data from the other instrument were processed with the CLS-based software that was installed in the spectrometer by the vendor.

**RESULTS AND DISCUSSION**

The results obtained by PLS are shown in Figure 3a, which shows that ammonia was present at or below its detection limit over much of the duration of these measurements (when the wind was predominantly from the east) but increased to well above the limit of detection during periods when the wind shifted to the southwest and west and blew the emissions from the dairy into the infrared beam, as illustrated in Figure 4. The concentration of methane calculated by PLS was at about the global average for much of the measurement period but, as would be expected, rose.
during the same time periods as the ammonia concentration rose, when the winds blew from the west. The concentration of nitrous oxide remained approximately constant at the global average over the 44 h measurement period. This behavior is typical of what would be found adjacent to a dairy production facility.34,35

Whereas the concentration of ammonia calculated by the CLS-based software supplied by the vendor with the instrument was similar to the values calculated by PLS at all times, the results calculated for methane and nitrous oxide were dramatically different, as shown in Figure 3b. Methane was frequently calculated to have a negative concentration although the calculated concentration was in approximate agreement with the concentration calculated by PLS when methane was present at a concentration that was appreciably greater than the global average. The calculated concentration of N\textsubscript{2}O varied so wildly that only a few data points fit on the scale of this figure and were never close to their expected values.

The reason for these results is speculated to be as follows. The spectral intervals used in the two programs are shown in Table 1. Ammonia absorbs in a spectral interval where water only absorbs very weakly (see Figure 5), and so, it is possible to find short spectral intervals (such as the one chosen for the CLS analysis) where the presence of atmospheric water vapor does not interfere at all. (Thus, estimating the concentration of ammonia by CLS is essentially a univariate analysis, even though a multivariate algorithm is being applied.) Conversely, water lines appear quite strongly in the spectral intervals that are used to calculate methane and nitrous oxide by the CLS-based program (1300–1307 and 1235–1283 cm\textsuperscript{-1}, respectively). The effect of these atmospheric interferences is clearly not handled adequately when the data are processed by CLS. The spectral intervals that are used in the PLS calculations are far wider and were selected in part because the strongest lines in the spectra of the analytes were found in these intervals. Even though there is significant overlap by rotational lines due to both the water and CO\textsubscript{2} spectra, their effect is taken care of much better using the PLS approach and the result of these calculations appears to be very accurate.

**CONCLUSION**

In order to facilitate the analysis of OP/FT-IR measurements, good software is a necessity. As noted by Heise et al., although commercially available software for OP/FT-IR is user-friendly, it generally lags behind the state of the art of chemometrics and, hence, is less effective than the programs developed in research.

**Table 1. Spectral Intervals in the CLS and PLS Calculations**

<table>
<thead>
<tr>
<th>Analyte</th>
<th>CLS (cm\textsuperscript{-1})</th>
<th>PLS (cm\textsuperscript{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>methane</td>
<td>1300–1307</td>
<td>2850–3200</td>
</tr>
<tr>
<td>ammonia</td>
<td>955–969</td>
<td>750–1250</td>
</tr>
<tr>
<td>nitrous oxide</td>
<td>1235–1283</td>
<td>2160–2260</td>
</tr>
</tbody>
</table>

Figure 3. Concentrations of methane (solid line), ammonia (dashed line), and nitrous oxide (dotted line) calculated with (a) PLS and (b) CLS regression.

Figure 4. (○) Concentration of ammonia calculated by PLS regression. (●) Wind direction at each time the spectrum was measured; when the direction is 90°, the wind is coming from the east.

Figure 5. (Above) Absorbance spectrum of water vapor measured in pristine air over a path length of 408 m, showing the vendor-recommended spectral intervals for CLS analysis of each analyte. (Below) Reference spectra of (A) N\textsubscript{2}O, (B) CH\textsubscript{4}, and (C) NH\textsubscript{3}.
laboratories, while the latter are often difficult to use by inexperienced users. In our opinion, the software should be user-friendly, but a totally “black-box” program is not as helpful as much of the intermediate data that helps experienced users to interpret the results is inaccessible. We kept these concepts in mind while developing a software package with the numerical techniques described above incorporated. The result is a software package (available on request) with a straightforward graphical user interface, which provides both the analytical results and all the intermediate data including the processed interferograms, single-beam spectra, and absorbance spectra in a popular format for easy access.

ACKNOWLEDGMENT

This work was funded by the National Natural Science Foundation in China (Grant No. 20705032) and partly by the Chinese Universities Scientific Fund. The development of the software at the University of Idaho was funded in part by Cooperative Agreement 58-5368-3-269 with the USDA/NWISL, Agricultural Research Service, Northwest Irrigation and Soils Research Laboratory, Kimberly, Idaho, and in part by Contract W91ZLK08P0739 from the Edgewood Chemical Biological Center (ECBC), Edgewood Arsenal, U.S. Army.

Received for review June 29, 2010. Accepted September 14, 2010.