Editor's review

Sources of error in open-path FTIR measurements of N2O and CO2 emitted from agricultural fields

Cheng-Hsien Lin[1], Richard H. Grant[1], Albert J. Heber[2], and Cliff T. Johnston[1,3]

This paper is presented as an analysis of systematic errors in quantitative measurements of N_2O and CO_2 in air using OP-FTIR spectroscopy. As such it can only be useful if the detailed methods used in the analyses are clearly documented in such a way that future practitioners could repeat the analyses on their own data when using the same CLS and PLS techniques. Unfortunately the paper falls short of this requirement, the descriptions in their current form are not sufficiently detailed to be of wider use, and the results pertain only to this particular dataset and analysis. They do provide a useful, qualitative guide to the magnitude of errors that may be encountered.

The authors have adequately responded to some of the referees' comments, but unfortunately not all. In addition and as indicated in my earlier access review I have some editor comments of my own which in many cases overlap those of the referees. I combine these comments below and ask that the authors address these comments before the paper is acceptable for publication in AMT.

Page and line references below are to the "amt-2019-263-author_response-version2.pdf" document, which includes responses to the two referees' comments.

Abstract, P8L18-19

This paper provides a useful qualitative guide to practitioners of OP-FTIR spectroscopy for atmospheric analysis of the systematic errors and biases that arise with commercial low resolution spectrometers using commercial analysis packages based on CLS and PLS chemometric methods. However the authors provide no evidence that the quantitative errors observed and documented in detail here will translate to other instruments and field setups with different resolutions, instrument lineshapes, pathlengths and other conditions. I therefore do not agree with the last sentence of the abstract, which should be removed – this studies serves as a qualitative guide, but not a reference for other users (see also RC2 page C2).

P10L13 (RC2 page C2 ref P3L17) This comment has not been addressed in the authors' response. It is essentially the same comment as made in my initial access review. This work demonstrates significant and complex biases in quantitation using CLS and PLS, including non-linearity and cross-dependency between variables – in this sense it is a useful contribution to the field for the users of commercial chemometric software packages. These sources of error are already well known and recognised from earlier studies. They are less prevalent in a least squares fitting approach to spectrum analysis as now referenced in the papers on L14. RC2's question "What advantage does CLS and PLS offer over NLLS as implemented in the works cited here?" should be addressed, perhaps by a short review paragraph which points out the pros and cons of each approach.

P10L22: "The influences…" The meaning of this sentence is unclear. HITRAN provides a database of absorption line parameters from which an absorption spectrum for any combination of temperature, pressure and gas composition (including mixtures) can be calculated. This is the approach used in NLLS analysis in which the spectrum is iteratively calculated until a best fit to the measured spectrum is obtained. Please clarify.

P11 section 2.1.1 and P12 section 2.2 – sample and atmospheric pressure:

I cannot find any reference to pressure measurement or control, or pressures used in the CLS/PLS generation of calibration models or analysis of unknowns. Pressure has an important effect on the spectra, as does temperature (for example linewidths are proportional to pressure). In the lab measurements I presume the pressure has been measured and controlled to be the same for calibration and analysis measurements. If so, please state so. But how was pressure included in the open path analysis? Atmospheric pressure will change from hour to hour and day to day, yet the calibration models are presumably built at a single pressure and temperature. Pressure and temperature have two separate effects on retrieved mole fractions in air:
(1) the spectroscopic analysis fundamentally determines a concentration*pathlength product, from which the concentration (in mol/m^3 or similar) is determined. To convert to a mole fraction (eg ppm or ppb) requires the density of air, P/RT. How is this done with the software used?
(2) molecular lineshapes are both pressure and temperature dependent, leading to errors of the calibration spectra and unknown spectra are measured at different pressures and temperatures. This spectroscopic effect is independent of the density effect (1).

The treatment of pressure and pressure variability should therefore be addressed.

P11L17: The synthetic reference mixtures were prepared using N2 as a buffer gas rather than air. Line-broadening coefficients for N2 and air are different, and this will add a systematic bias to the reference measurements of N2O relative to air. This error should be addressed and corrected, or included in the error analysis

P12 L21. Zero filling (RC1): Different FTIR manufacturers and users define "ZFF" differently and it is probably better to avoid this term. From Figure 3 it appears that the point spacing is approx. 0.25 cm-1 or half the 1/maxOPD resolution of 0.5 cm-1. This means the interferogram is minimally sampled and corresponds to NO zero filling, ie the interferogram is extended with zeros only up to the next power of 2 points beyond the highest OPD point (a filling factor of 1, but for example the Bruker and RC1 definition would call this ZFF=2). The best way to resolve this ambiguity is to say "no zero filling" rather than quoting a ZFF value. With no zero filling the point spacing is 0.5/maxOPD.

P11 2.1.2 (RC2) I agree with RC2's comment that the CLS and PLS calibration details are lacking and should be expanded in the paper along the lines in the response – it is not sufficient to respond only to RC2, but to make the methods used clear to all readers in the paper, with sufficient detail that they could repeat the analysis given the same set of data, or their own data. The respective software packages for quantitative analysis have only been presented as a "black box". The claim in the abstract that this paper can be a reference for future work could only stand up if all details of the analysis are presented.

P12 L28: How were the single gas reference spectra calculated from HITRAN data? This is a non-trivial calculation (see eg Griffith 1996), and it is very important to know details of both the molecular spectra and the convolution with the FTIR instrument function if these spectra are to be quantitatively compared with measured spectra (via the CLS/PLS models). It is also not clear exactly what the IMACC software has been used for, or how. In the manual provided in the supplement there is only a description of the user interface, not the underlying calculations or physics behind them. Was IMACC used to calculated SB spectra from Hitran data, or only to generate ABS spectra from measured spectra and background? If not, how were they calculated? None of this is clear. As above, the description should be in principle sufficient for another practitioner to follow your method and achieve the same results.

P13 L16: What is the accuracy (not precision) of the IRIS and Licor analysers used to measure the S_OPS samples? How were they calibrated. Any systematic bias or error in these calibrations feeds directly into the biases calculated for OP measurements via Eq. 1. (This was answered in a response to RC2, please include in the manuscript)

On the same topic, gas concentrations along the open path will vary with time and location. If the S-OPS measurements are not coincident in time (ie simultaneous) or space (eg along the whole path rather than 150m OP vs 50m S-OPS) an estimate of the potential error needs to be made. At present S-OPS seems to be used as absolute truth without full justification. RC2 points out that this is potentially a very valuable feature of this study, but it cannot be fully exploited without this part of the error analysis.

P14 L26: Increasing resolution will decrease the non-linearity effects of Beer's Law breakdown, but at the cost of signal:noise and hence detection limits and precision. The optimal resolution is a trade-off – this should be included in the discussion here.

P16L18: "Detector saturation at short distances was avoided…" The meaning is unclear. If higher signal levels at short open path distances resulted in saturation (as seen in signal below the detector cut-off), how was this addressed or corrected? Were these spectra rejected? Was the intensity reduced so the detector operated in the linear regime?

Further comments to RC2

Section 2.2.2 – please address this point in the manuscript, not only in the response to the referee. The same applies to other comments from RC2 – they need to be addressed in the manuscript.