

Interactive comment on “Analysis of functional groups in atmospheric aerosols by infrared spectroscopy: functional group quantification in US measurement networks” by Matteo Reggente et al.

Anonymous Referee #1

Received and published: 16 December 2018

Summary:

In this work, Reggente et al. compare two methods to derive functional group abundance from FTIR spectra of aerosol collected on filters. The methods compared are peak fitting (PF) and partial least squares regression (PLSR). Total Organic Carbon (OC) was validated against Thermal Optical Reflectivity (TOR-OC)—which separates organic mass from refractory black carbon. Also validated was the NH functional group concentrations against sulfate-plus-nitrate concentrations.

C1

Overall, the general topic of this work is of interest to the Atmospheric Measurement Techniques community. Devolving complex spectra into their components is a powerful technique that is applicable beyond even FTIR. There are, however, several major and minor comments I have about this work, which are summarized below.

Major Comments:

My main major comment about this work is that much of it has been developed and validated in previous papers. The authors do not clearly outline this in their introduction or methods. In fact, readers are required to check at least three separate papers to fully grasp what is novel about this work. The three previous papers mentioned are as follows:

Takahama et al., 2013 developed and validated the peak fitting analysis for COH and CO.

Ruthenberg et al., 2014 analyzed the same dataset as this paper and ran PLS (raw) on the samples. They also validated their data using TOR-OC.

Kuzmiakova et al., 2016 developed background correction protocols for peak fitting analysis. They also tested the background-corrected peak fitting analysis against TOR-OC.

In summary, none of the data, functional group fitting techniques, or baseline corrections presented in this paper are novel. One new product is the updated molar absorption coefficients. Thus, the only real new analysis are the baseline corrected PLSR data. Conceivably, this work could be labelled as an intercomparison; however, that is difficult to justify for only two methods and no additional functional group verification. Ultimately, this paper would benefit for an explicit description of what has been done previously, and a similarly explicit description of how this paper improves upon previous techniques or instructs the field on how to best use these functional group analyses for filter-based FTIR spectra.

C2

My second major comment is related to functional group verification. While the focus of this paper is determining the functional group abundance in US measurement networks—only one functional group validation (NH) was made. Ultimately, the authors conclude that "further progress in parameter selection [i.e., molar absorption coefficients for PF and number of latent variables for PLSR] is required." This belies the entire purpose of this article, which is to give solidarity to these techniques and their ability to estimate functional group abundances from filters collected in US networks. Perhaps the authors could assess the current literature, or find additional data products from the IMPROVE network other than TOR OC and ammonium, to help validate their measurements.

Finally, because this work relies heavily on methods used in previously published papers from this group, the authors often do not fully describe their techniques. While this is understandable for details of the technique, oftentimes entire concepts are missing. This forces the reader to read several other papers to grasp the main concepts of the techniques used in paper. As much as I could, I have tried to outline this in the minor comments section; however, the authors should also pay close attention to this—especially when describing both the specifics of PF algorithms and PLSR in the context of identifying functional groups in FTIR spectra.

Minor Comments:

P3L9-17: There are 32 (!) references in these two sentences. Ultimately, it is more distracting than helpful. I suggest keeping one or two of the most representative references for each functional group/type.

P3L23: This apportionment of detectable vs non-detectable fractions is hard to follow here. It might be more instructive to provide an example. While the Takahama and Ruggeri 2017 paper is referenced, this work should, at least conceptually, stand on its own.

P3L33: Is it not true that the peak shapes for single or simple components are also

C3

Gaussian? Certainly, you must fit Gaussians to your simple components for calibration.

P4L34: The objective of this paper is to evaluate the robustness in estimated abundances. What is your measure of robustness—can you quantify this?

P8L4: It seems slightly confusing to me to use $n_{ik}(a)$ here as the number of moles of bond n , since it was used as areal surface density earlier.

P8L17: This "apportioning" method needs to be described explicitly. As of now, the reader must go to the Takahama et al., 2013b reference to even understand the concept of the method.

P8L31 (Figure 3): I am not certain how you are calculating the error bars on your absorption coefficients. Just by eye, both the errors of the individual components and the assumed standard deviation of all the samples looks to be much larger than what you reported.

P9L29: How are you quantifying "significantly worse?"

P12L2: It might be clearer to remind the readers that the calibration curves are from 158 of the 238 laboratory standards and are used to derive a calibrated absorption coefficient. Then, the remaining laboratory standards are used in the predicted concentrations to validate the calibrated absorption coefficients. This is largely omitted from the text of this paragraph and requires the reader to rely only upon Figure 2 to interpret these crucial results.

P12L13: This paragraph would be clearer if the authors explicitly outline why they decided to re-run the molar absorption, and how it differs from both Russell and coworkers and Takahama et al., 2013. A difference is prescribed by the authors to a different baseline correction, but there are differences between just Russell and coworkers and Takahama et al., 2013.

P13L3: This should be a main result of this paper; however, as mentioned in the main comments, there are no ways to validate the functional group molar densities;

C4

therefore, there is no way to validate the molar absorption coefficients. This calls into serious question the utility of this work.

P13L26: Could it also be that both are generally underpredicting by 40% or more, but the PLSr has some erroneous contribution from the PTFE filter? More discussion should be added here as PLSbc is one of the new techniques introduced to this paper.

P14L3: Is there some sort of tradeoff for not using PLSbc*? The authors mentioned overfitting leading to unrealistic results in their methods section—but it seems like the PLSbc* results shown here are more realistic.

P14L32: Is there any direct evidence that PLS2 outperforms PLS1 for identification of aerosol functional groups in FTIR?

Section 3.5: Again, these are useful comparisons—but without an external validation method, it is difficult to choose which methods are most robust. For example, PLSr seems to be doing well in predicting TOR-OC (Figure 2) and potentially OM/OC, but its OC value is low. The other methods have low TOR OC and possibly too high OM/OC, but realistic O/C. Which method are we to trust as the reader? Which method is most robust? Additionally, the TOR-OC normalized results seem more realistic—but they are hard to evaluate without external validation.

Section 3.6: This entire section, while interesting, seems like it would be better suited to go into the Supporting Information.

Technical Comments:

P1L2: I'm not sure that FTIR needs to be defined, as it is commonly known acronym.

P1L7: The phrase "including their model parameters" is parenthetical and could be enclosed by commas.

P7L9: OC is used here, but has not been previously defined.

P1L11: It would be much more constructive to explain what the "series of possible

C5

input parameters" refers to. As it reads now, it is not instructive and, therefore, is unnecessary.

P1L14: TOR OC is used here, but it has not been defined.

P2L9: OM is used here, but it has not been previously defined.

P2L27: Like the FTIR calibration, readers should be generally familiar with the most general principle of FTIR. This could be cut for brevity.

P3L23: Should FGss be FGs?

P4L16: PLS and PF "have," not "has"

P7L6: There should be a space before "A single parameter ..."

P10L16: The phrase "is depends" should be just "depends"

P14L24: Should the phrase read "no gross overestimate, or that underestimation is unlikely?"

p19L5: There is an extra "a" at the beginning of this line.

p19L30: There should be a space after "Phoenix," and before "is." Figure 10. The models do not need to be defined here as they are defined earlier and used often since.

Figure 2. It might be clearer to use different colors in the bottom row to highlight the fact that you've used the test samples to validate the calibration made with the calibration samples. Figure 7. The urban x-axis value on the right side is overlapping with the rural x-axis on the left side. Also, is the correlation coefficient shown here R or R²?

Figure 10. This figure would be much clearer if you centered the descriptions at the top (e.g., "From FG calibrations") and made them larger or bold.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-331, 2018.

C6