Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2018-70-RC2, 2018

Atmospheric Measurement Techniques Discussions

Interactive comment on "Atmospheric particulate matter characterization by Fourier Transform Infrared spectroscopy: a review of statistical calibration strategies for carbonaceous aerosol quantification in US measurement networks" *by* Satoshi Takahama et al.

Anonymous Referee #3

Received and published: 29 August 2018

General Comments:

Takahama et al (2018), hereafter referred to as T2018 reviews methods used to determine quantitative measurements of OC/EC using FTIR spectroscopy on PTFE filter samples taken from sampling networks such as the Chemical Speciation Network (CSN) and the Interagency Monitoring of Protected Visual Environments (IMPROVE) network. The topic and scope of the article are very appropriate for AMT. There have

been many papers on this topic in the past few years, so a review article summarizing the calibration, processing, and evaluation techniques is timely. However, this paper does not stand alone. It requires knowledge of the authors' previous papers, as well as other information. There are too many gaps in the information provided. Examples include such as why PTFE is used and not quartz fiber filters, the lack of general references for the measurement networks, and the lack of legends on the plots. Additionally, there are some organizational problems, ranging from unclear section titles, to different datasets used in different ways, requiring effort from the reader to keep everything straight. Stated differently, this paper seems to focus on what the authors do and how they do it, which is important, but the authors need to emphasize why this is important, how it fits within the larger body of relevant literature, and make the paper stand on its own, while being clear to read. Thus, while this paper definitely should be published in AMT, it requires major revisions to improve the clarity of writing.

Specific Comments:

I can't list all the areas where the authors could improve the organization, only a few are mentioned here.

Site selection and use: It's not always clear what datasets are used when, the map on Figure 1 suggests that the BYIS and FRES IMPROVE sites are used only as testing samples, but at the end of the paper the BYIS and FRES sites are used for training a new predictor. The 'data' section mentions calibration data from IMPROVE 2011, and CSN 2013, but it isn't clear that each of these datasets will be used separately and at different points within the paper. Thus, at various points in the paper: 2011 IMPROVE data are used to test calibrations using 2011 IMPROVE data, 2013 IMPROVE measurements are used to test a calibrations with 2011 IMPROVE data, 2013 IMPROVE data (For FRES and BYIS) are used to test calibrations with 2013 IMPROVE data, and 2013 CSN data are used to test calibrations with 2013 CSN data. In a paper this long, it can be challenging to remember what data is used where.

Some other site selection questions: A user might wonder why CSN and IMPROVE data are not used together to develop a model: there are some differences in the collection method between CSN and IMPROVE (Weakley et al., 2016), but no discussion of these differences is found in the paper. The Elisabeth (ELLA) required a special calibration. Weakley et al. (2016) noted that this site was located near a refinery but no discussion of the potential influence of the refinery on different spectra is discussed.

Section titles could be improved. Section 4.2 is an example of this, a more descriptive title like "Applicability of calibrations developed under one set of conditions to samples measured under new conditions" would be more descriptive.

One of of the use cases for this technique is in a network where OC/EC measurements are not made using the standard EGA technique (Page 6, Line 30). The discussion in 4.2 directly touches this concept - but the topic mentioned in the introduction isn't really brought up in that section.

The data availability does not mention where to obtain the FTIR Spectra. Also, I could not find the source code at http://airspec.epfl.ch.

The authors should include a list of acronyms as an appendix.

There are no calibration sites in the Midwestern states (e.g. at longitudes between Birmingham, Alabama and Mesa Verde, CO), other than the Sac and Fox site which only has one half years' worth of data. Is this a problem for applicability of the model?

The authors briefly mention meteorological influences on the calibrations – could model error be used to infer something about the variability due to meteorological conditions? Are there better performances across the different months – e.g. both meteorological patterns, as well as well as combustion patterns, are different between summer and winter. How does calibration data during only a short period (such as at the Sac and Fox site) bias the results?

Specific Minor Comments:

Page 3, Line 9 is written as P3L9 P3L9 – needs a citation.

P9L15: The authors should cite Malm and Hand (2007) or Solomon et al. (2014).

P9L21: The authors should cite (Solomon et al., 2014)

P10L1: Are the two sites that were collocated from IMPROVE, or was one IMPROVE site collocated with one CSN site?

P20L1: The Situation with the ELLA predictions is problematic – how many locations will need specific trainings associated with them? Probably should mention that the Elizabeth site is located next to a refinery (Weakley et al., 2016).

P32L31: Samples from these sites – is this for 2011 or 2013?

P38L31 – NOx isn't a surrogate for EC – but it may give an indication or be useful for prediction of EC. Somewhere – should mention that PTFE filters should have lower gas-phase adsorption than quartz fiber filters (Turpin et al 1994, Gliardoni et al 2007). The BYIS site is very interesting as it is on another continent – this is an interesting use case, and basically is an example of using a model trained in one location or with a network to apply to another location – and as noted, this method does not do well under that scenario.

P45L20: Also, Figure 1 indicates that BYIS is only used as a test dataset, not calibration dataset. Is this correct? So then wouldn't we expect some errors in BYIS since that location is on a different continent than the trainings? What is similar about FRES and BYIS? What about South Korea – it's very interesting that this location is not an outlier, given that the models were trained on an entirely different continent.

Table 3: Why are OC combined for FRES and BYIS, but separated out for EC?

Figure 1: It would be helpful to also include the site abbreviation, either in the location labels in this figure, or in a table. The authors identify the abbreviations the first time they are used, which is good, but the paper is long enough I found myself having to

refer back.

Figure 12 - The authors need to clarify in the paper what the different meanings of the two types of triangles - I assume it is just meant to show two different similar types of spectra, associated with wildfires. For the wildfire cases, are there significant differences between the red spectra (those with predicted-observed difference > combined uncertainties) and the black spectra? Finally, the subplots should be labeled a,b,. . ..f, and there should be a legend identifying black, red, and the two different triangle containing spectra. It might be good to mention the importance of aliphatic stretches and carbonyl vibrations.

Figure 14 – add the filled circle to the legend.

Figure 15 – Similar comments to Figure 15 – please restate the descriptions of triangles and red samples, and include a legend in the plot. Also, subplot labels will help with interpretation.

Technical Corrections:

P3L25: smog chamber -> smog chambers

P6L23: measured by EGA EC is an -> measured by EGA, EC is an

P6L25: and therefore less influenced -> and therefore is less influenced

P9L31: “Change to artifact correction method for OC carbon fractions” – not sure what this is.

P20L5: calibrations models -> calibration models

P30L6: for new smaples. -> for new samples

P36L10: for the rediction standard error -> for the prediction standard error

References: There were many errors in the references section, and it would take too much time for me find and fix all of the problems. A non-exhaustive list follows, with

some examples. Please redo the references section.

- Many of the references appear to be missing journals, and then the title ends up being formatted like the journal. Examples include Cunningham et al (1976), Efron and Tibshirani (1996).

- The Debus et al (2018) citation has no information other than author, title, and year, and the fact that it was accepted in some journal. This is a problem because section 2.3 cites Debus et al (2018) heavily. I tried to find it by searching online but could not.

- Book and grey literature references are incorrect, e.g. Mahalanobis, P “On the generalised distance in statistics”, Tibshirani (2014)

References: Malm, W. C., and Hand, J. L.: An examination of the physical and optical properties of aerosols collected in the IMPROVE program, Atmospheric Environment, 41, 3407-3427, https://doi.org/10.1016/j.atmosenv.2006.12.012, 2007.

Solomon, P. A., Crumpler, D., Flanagan, J. B., Jayanty, R. K. M., Rickman, E. E., and McDade, C. E.: U.S. National PM2.5 Chemical Speciation Monitoring Networksâ ĂŤCSN and IMPROVE: Description of networks, Journal of the Air & Waste Management Association, 64, 1410-1438, 10.1080/10962247.2014.956904, 2014.

Weakley, A. T., Takahama, S., and Dillner, A. M.: Ambient aerosol composition by infrared spectroscopy and partial least-squares in the chemical speciation network: Organic carbon with functional group identification, Aerosol Science and Technology, 50, 1096-1114, 10.1080/02786826.2016.1217389, 2016.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-70, 2018.