Supplemental Material

Long Term Strategy for Assessing Carbonaceous Particulate Matter Concentrations from Multiple Fourier Transform Infrared (FT-IR) Instruments: Influence of Spectral Dissimilarities on Multivariate Calibration Performance

Bruno Debus¹, Satoshi Takahama², Andrew T. Weakley¹, Kelsey Seibert¹, Ann M. Dillner^{1*}

¹Air Quality Research Center, University of California Davis, Davis, CA, USA

² ENAC/IIE, Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland

Correspondence to: Ann. M. Dillner (amdillner@ucdavis.edu)

Table of contents:

- Section 1: Common functional groups resolved in a single FT-IR spectrum of ambient PM samples
- Section 2: The IMPROVE network locations of sampling sites
- Section 3: House-built sample chamber for the FT-IR analysis of ambient PM samples
- Section 4: Impact of the house-built chamber on purge efficiency
- Section 5: Inter-instrument comparison of 2nd derivative spectra of the Wichita Mountains (OK) ambient sample
- Section 6: Anisotropic optical properties of PTFE filters
- Section 7: OC and EC models performance as function of the number of calibration samples
- Section 8: Performance of the hybrid models extended to the 2016 and 2017 ambient samples





Figure S1. FT-IR spectrum of a blank PTFE sample a) and a heavy loaded ambient sample b) collected in Fresno (CA) on February 2nd 2015. Infrared bands associated with functional groups commonly found in atmospheric samples are labelled.

Section 2: The IMPROVE network – locations of sampling sites



Figure S2. PM sampling locations from the IMPROVE network including 159 rural sites (green markers) and 4 urban sites (red markers). With the exception of the heavily polluted site in South Korea, ambient samples collected across the whole network were considered in the present study.

Section 3: House-built sample chamber for the FT-IR analysis of ambient PM samples



Figure S3-1. a) Home-built UC Davis sample chamber installed in a Bruker Tensor II instrument, b) chamber lid and support block are made of aluminum while arms are made of black delrin, c) back of the sample chamber with purge lines shown, d) 25 mm, 37 mm and 47 mm filter holders made from black delrin (support rings are made from acrylic), e) 25 mm filter holder (with PTFE filter included) being inserted in the sample chamber, f) sample chamber with filter holder in place (example with a 47 mm filter).

Engineering drawings and parts list of the house-made sample chamber are provided at the end of this section. All engineering drawings included in this document are the sole property of the Air Quality Research Center, University of California, Davis, CA. Any reproduction in part or as a whole without the written permission of the University of California, Davis is prohibited. Please contact the corresponding author at <u>amdillner@ucdavis.edu</u>.





























Section 4: Impact of the house-built chamber on purge efficiency

The advantages of the home-built sample chamber are twofold. The system ensures that the filter position relative to the IR beam is fixed for each analysis. This is achieved by using filter holders, tightly inserted into a solid support block attached to the instrument. In addition, the small volume of the chamber (\approx 92 cm³) enables shorter and more stable purge prior to collecting an FT-IR spectrum. To evaluate the benefit of the chamber for purge duration and stability, repeated measurements of empty chamber backgrounds are performed with and without the home-built chamber. In Figure S4-1, visual inspection of backgrounds spectra collected with and without the house-built sample compartment indicates a significant decrease of water vapor and carbon dioxide contributions when the built-in system is applied.



Figure S4-1. Sets of 48 empty chamber backgrounds collected on FTIR2 with and without the homebuilt sample compartment during the repeated measurement procedure. Backgrounds are acquired by a liquid nitrogen-cooled wide-band mercury cadmium telluride detector, identical for all three spectrometers, characterized by a detectivity D*_(500, 1000, 1) of 1.10¹⁰ cm Hz^{1/2} W¹ and a cut-off wavelength of nearly 24 μm.

To quantify the impact of the house-built chamber over time, time-dependent variations in background intensity are monitored at 1650 cm⁻¹ and 2360 cm⁻¹ associated with water vapor and carbon dioxide, respectively. The backgrounds are collected every 10 seconds during 470 seconds, yielding 48 data points. Each background is computed by averaging 10 consecutive scans acquired with a resolution of 4 cm⁻¹. The set of *x* intensity values obtained at both wavenumbers is then converted into transmittance following Eq. (2):

$$\% T_i = 100 \times \left(\frac{x_i - \min(x)}{\max(x) - \min(x)}\right)$$
⁽²⁾

with x_i the intensity of the *i*th recorded empty chamber background and $\%T_i$ the corresponding transmittance. The time-profile comparison between the home-built and the default instrument sample compartments are shown in Figure S4-2 for all three spectrometers. Each profile is computed by averaging three different runs performed over a one-week period.



Figure S4-2. Comparison of time-dependent transmittance for water vapor (1650 cm⁻¹) and carbon dioxide (2360 cm⁻¹) peaks as function of purge time on all three spectrometers.

At t = 0 s, the lid of chamber is briefly opened and the background intensity reaches a minimum corresponding to a large absorbance of water vapor and carbon dioxide. As purge time increases, the transmittance inside the house-built sample compartment rapidly increases to a nearly hundred percent value after only 100 s and remains stable afterward. In contrast, a slow and gradual increase in transmittance is reported for the default sample compartment due to the larger chamber volume. The results confirm that a purge time of approximately 240 s (4 minutes) is sufficient enough to have low and stable water vapor and carbon dioxide concentrations within the sample chamber.

It is worth noting that FTIR3 achieves a more stable signal than the two other instruments. Differences in stability can be tentatively attributed to the fact that FTIR3 is on a separate purge line while FTIR1 and FTIR2 are both connected to the same purge-gas generator.

Section 5: Inter-instrument comparison of 2nd derivative spectra of the Wichita Mountains (OK) ambient sample



Figure S5. Spectra of the Wichita Mountains (OK) ambient sample on all three instruments obtained after computing their second derivative. Arrows indicate spectral regions featuring significant spectral differences between instruments. The region below 1500 cm⁻¹, moslty comprised of PTFE absorbance, is not represented here for the sake of clarity.

From Figure S5, no significant shift in either intensity or wavenumber dimension could be observed which indicate that the strict control of acquisition parameters efficiently reduces inter-instruments differences in spectral response. Although a few (minor) deviations attributed to a change in volatile organic carbon (VOC), water vapor and PTFE signal are observed, a reasonable match between derivative spectra is achieved in regions associated with the spectral signature of PM such as 3000 – 3450 cm⁻¹ and 2735 – 2962 cm⁻¹. In particular, the absence of band shifts in those particular regions confirms that all spectrometers register absorbance values at the exact same set of wavenumbers.

Section 6: Anisotropic optical properties of PTFE filters



Figure S6. Variability in optical properties of a single PTFE blank filter for different orientation angles. Rotation is performed along the dimension that is orthogonal to the IR beam.

As the filter orientation around the incident beam is not perfectly reproducible between instruments, rotating the filter by a few degrees produces a slightly different spectral response on account for the non-uniform density of the filter.





Figure S7-1. Variation in OC prediction metrics for the first (a) and second (b) scenarios as function of the number of calibration samples (N_{cal}) included in the model. In scenario (i), a calibration developed with spectra from one instrument are used to predict test samples from the same instrument and in scenario

(ii) calibratons developed on one instrument are used to predict test samples collected on the other instrument. Even though increasing the number of calibration samples yield better model performance, no significant changes are observed when N_{cal} exceed 300 samples. With the exception of a small bias (\approx

 $0.5 \mu g$) between the two spectrometers, similar figures of merit are obtained independently of which instrument or scenario is used. Overall, the second scenario is characterized by only a slight increase in prediction metrics indicating that swaping models between instruments does not significantly impact prediction. The mistmatch between MDL metrics for the two instruments suggest different instrument sensitivity at low concentration levels.



Figure S7-2. Variation in EC prediction metrics for the first (a) and second (b) scenarios as function of the number of calibration samples (N_{cal}) included in the model. Even though increasing the number of calibration samples yield better model performance, no significant changes are observed when Ncal exceed 300 samples. Unlike OC prediction in Figure S6-1, no significant bias between FTIR1 and FTIR2 is observed. Overall, the first and second scenarios are characterized by similar prediction metrics indicating that swaping models between instruments do not significantly affect prediction. The mistmatch between MDL metrics suggest different instrument sensitivity at low concentration levels.



Figure S7-3. Evolution in PTFE mean absorbance and standard deviation between calibration and test spectra as function of N_{cal}. Distribution parameters are derived from the 1210 cm⁻¹ PTFE peak for FTIR1 and FTIR2, respectively. Dotted lines correspond to the PTFE distribution parameters in test spectra. Despite measuring the same set of samples across instruments at a fixed sample position (house-built chamber), uneven PTFE distributions are observed across instruments due to the lack of repeatability in filter orientation (anisotropic properties). For N_{cal} values above 300, similar distribution parameters are experienced in both calibration and test spectra so that PTFE interference has little to no impact on model prediction. However, as N_{cal} values decrease, calibration and test PTFE distribution no longer overlap which is likey one of the reasons why model performance decrease when using a low number of calibration samples.





Figure S8-1. Time-dependent slopes (left) and intercepts (right) derived from test set EC predicted concentrations on the FTIR1 – FTIR2 instrument-pair, uncertainty of each metric is evaluated by developing 95 % confidence intervals from 1000 bootstrap replicates.



Figure S8-2. Time-dependent slopes (left) and intercepts (right) derived from test set OC predicted concentrations on the FTIR3 – FTIR1 and FTIR3 – FTIR2 instrument-pairs during 2017, uncertainty of each metric is evaluated by developing 95 % confidence intervals from 1000 bootstrap replicates. The good agreement within uncertainty of all three instruments throughout 2017 indicates that all spectrometers have very consistent OC predictions. Linear regression parameters were computed by regressing FTIR3 OC predicted concentrations against either FTIR1 or FTIR2 predictions.



Figure S8-3. Time-dependent slopes (left) and intercepts (right) derived from test set EC predicted concentrations on the FTIR3 – FTIR1 and FTIR3 – FTIR2 instrument-pairs, uncertainty of each metric is evaluated by developing 95 % confidence intervals from 1000 bootstrap replicates. The larger intercept observed for the FTIR3 – FTIR1 instrument-pair suggests that FTIR1 and FTIR3 are the most dissimilar spectrometers throughout the year 2017. Linear regression parameters were computed by regressing FTIR3 EC predicted concentrations against either FTIR1 or FTIR2 predictions.