# 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 Debus[1] , Satoshi Takahama[2] , Andrew T. Weakley[1],
Kelsey Seibert[1], and Ann M. Dillner[1]

## Abstract

Matching the spectral response between multiple spectrometers is a mandatory procedure when developing robust calibrations whose prediction is independent of instrument-related signal variations. A viable alternative to complex calibration transfer methods consists of matching the instrument spectral response by controlling a set of key instrumental and environmental parameters. This paper discusses the applicability of such an approach to three Fourier transform infrared (FT-IR) spectrometers used for the routine assessment of carbonaceous particulate matter concentrations in the Interagency Monitoring of PROtected Visual Environments (IMPROVE) speciation network. The effectiveness of the proposed matching procedure is evaluated by comparing the spectral response for each individual instrument in order to characterize the extent, and nature, of the remaining inter-instrument spectral dissimilarities. Instrument-related contributions to the signal were determined to be small compared with the spectral variability induced by the filter type used for sample collection. The impact of spectral differences on prediction was addressed through the comparison of model performance derived from multiple calibration scenarios. A hybrid model yielding accurate and homogeneous prediction regardless of the instrument was proposed for organic carbon (OC) and elemental carbon (EC), two major constituents of atmospheric particulate matter. Coefficients of determination of 0.98 (OC) and 0.90 (EC) with median biases not exceeding 0.20 μg (OC) and 0.07 μg (EC) are reported. The long-term stability, assessed from weekly measurements of reference samples, shows a deviation in predicted concentrations of less than ±5% over a 2.5-year period for most of the data collected. Extending OC and EC hybrid models to the prediction of ambient samples collected during the two subsequent years provides satisfactory performance. The proposed instrument matching procedure coupled with the relative simplicity of the hybrid model is an alternative to computationally advanced calibration transfer methodologies for the characterization of carbonaceous particulate matter using multiple FT-IR instruments.

## Keywords

## Introduction

Fourier transform infrared (FT-IR) spectroscopy, which captures structural information associated with molecular composition in a short amount of time, is an attractive analytical technique for the routine monitoring of process controlled operations in food,[1,2] fuel,[3] and pharmaceutical industries.[4] Besides speed, FT-IR is nondestructive, requires

[1]Air Quality Research Center, University of California Davis, Davis, CA, USA
[2]ENAC/IIE, Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland

Corresponding author:
Ann M. Dillner, 1 Shields Avenue, UC Davis, Davis, CA 95616, USA.
Email: amdillner@ucdavis.edu

little sample preparation, does not involve hazardous solvents, and is relatively robust, making it applicable to indoor and outdoor applications.[5,6] Due to its numerous benefits, including a high sensitivity to organic functional groups, FT-IR spectroscopy has garnered interest in recent years for the characterization of carbonaceous and inorganic particulate matter (PM) deposited on polytetrafluoroethylene (PTFE) filters.[7] A detailed description of the spectral fingerprint of functional groups commonly encountered in PM and absorbing at specific wavenumbers throughout the mid-IR range is given in the Supplemental Material (Section 1). Due to the complex nature of PM, quantifying function groups such as carbonyl (1650–1800 $cm^{-1}$) or the aliphatic C–H ($\approx 2900\ cm^{-1}$) is usually challenging due to the presence of spectral interferences caused by either inorganic compounds or filter-induced light scattering. To address this issue, multivariate regression techniques, such as partial least squares (PLS),[8] are often employed to develop quantitative models able to predict specific properties of ambient samples including organic carbon (OC),[9] elemental carbon (EC),[10] major inorganic ions,[11] organic functional groups and organic matter (OM).[12,13]

Assessing the composition of speciated fine PM emitted or formed in the atmosphere is crucial for understanding its substantial impact on human health,[14,15] climate change,[16] and visibility degradation.[17,18] As PM is known to contain organic compounds, black carbon, soil, sea salt, trace metals, and inorganic ions whose proportion varies temporally and spatially,[19–21] identifying pollution sources and observing local and regional trends in concentrations can best be addressed with consistent measurements across a large region.[22,23] In the United States, routine sampling of atmospheric particles is performed at rural and pristine locations by the Interagency Monitoring of PROtected Visual Environments (IMPROVE, http://vista.cira.colostate.edu/Improve) network while urban and suburban locations are monitored by the Chemical Speciation Network (CSN).[24,25] The CSN comprises about 140 urban sites mostly impacted by anthropogenic emissions related to combustion (fossil fuel burning) and industry while IMPROVE monitors regional impacts and visibility at about 160 rural sites (Supplemental Material, Section 2). Ambient PM samples are collected at every IMPROVE site once every third day and at CSN sites every third or every sixth day generating about 1500 and 1300 samples monthly for IMPROVE and CSN, respectively.

For the sake of maintaining high sample throughput applications, operating multiples spectrometers either in one or several locations is often a prerequisite.[26–28] The main challenge in the development of such measurement platforms lies in the control of the extent of dissimilarities between the response associated with each spectrometer. In practice, any combination of intrinsic factors such as differences in light source aging, optical path, detector response, temperature, or humidity may cause spectral deviations even between nearly identical instruments.[29] From a multivariate calibration standpoint, this implies that a calibration model developed on a specific instrument can rarely be extended to the spectra collected on a different instrument or under different experimental conditions without introducing erroneous or biased predictions.[30] A straightforward, but time-intensive, approach is to calibrate each spectrometer individually using the same set of standard samples. Conventionally, the need for independent calibration per instrument is often circumvented by either correcting or minimizing instrument-dependent signal contributions so that differences in spectral response between instruments have a negligible impact on model predictions. This way, a specific calibration developed on a target (primary) spectrometer can be universally applied to a network of satellite (secondary) instruments. For this purpose, good practices in spectroscopy involve either spectral preprocessing or advanced standardization procedures.[31] The latter, also known as calibration transfer, are a series of mathematical algorithms intended to match the response between two instruments by either manipulating their raw spectra or their predicted values.[32–35]

The aim of this work is to address the challenges associated with the development of a single instrument-agnostic calibration model for three FT-IR spectrometers used for the routine analysis of IMPROVE and CSN ambient samples. The need for calibration transfer is negated by a careful control of key instrumental parameters and environmental conditions.[36] The effectiveness of the instrument matching procedure is investigated by comparing the spectral response of each spectrometer to identify the remaining sources of dissimilarities. The impact of spectral dissimilarities on OC and EC is evaluated through the development of a hybrid model based on IMPROVE samples collected in 2015. The long-term stability of the implemented multivariate calibration model is tested using reference and ambient samples collected in 2016 and 2017 in an attempt to identify instrument drift and foresee situations where instrument recalibration become necessary.

## Materials and Methods

### Fourier Transform Infrared Spectroscopy

Three FT-IR spectrometers, a Tensor 27 (FTIR1) and two Tensor II (FTIR2, FTIR3) instruments (Bruker Optics), are considered in this work. Each system is equipped with a pre-aligned mid-IR source and a liquid nitrogen-cooled, wideband mercury–cadmium–telluride (MCT) detector with the same nominal specifications. Spectra are acquired in transmission mode over the range 4000–420 $cm^{-1}$ with a 6 mm source aperture and a resolution of 4 $cm^{-1}$, yielding 11 138 (zero-filled) data points. A total of 512 scans per sample are collected and averaged to improve the signal-to-noise ratio. The resulting signal is converted into an

absorption spectrum using the most recent empty chamber spectrum, acquired hourly.

All three FT-IR spectrometers are located in the same laboratory and were installed in 2011 (FTIR1), 2014 (FTIR2), and 2017 (FTIR3). In May 2016, the two oldest instruments were transferred to a new laboratory situated in the same building, which now houses all three instruments. Mid-infrared spectra of ambient PM samples are collected simultaneously on all three instruments following a random assignment procedure.

## Instrument Matching

In an effort to minimize the dissimilarities in spectral response between the three spectrometers, an instrument matching procedure was implemented. The procedure includes a careful control of key environmental and instrumental parameters including sample position, gas phase interferences, detector temperature, and acquisition parameters. Fourier transform IR systems are located less than a few meters apart in a dedicated room to ensure that temperature and humidity variations are consistent across instruments. To prevent external factors from affecting the instruments output, every FT-IR spectrometer is equipped with the systems detailed hereafter. The temperature of the detector is kept below $-165\,°C$ using an embedded Dewar that is re-filled automatically every 8 h by a liquid nitrogen micro dosing system (NORHOF, The Netherlands). Signal contributions associated with atmospheric water vapor and carbon dioxide are minimized by continuously purging the instruments with a volume control damper (VCD) series $CO_2$ adsorber/dryer system (PureGas LLC). For an optimal repeatability in sample position relative to the IR beam, a $4.0 \times 5.1 \times 4.5\,cm$ house-built chamber was specifically designed. Mounted directly inside the default instruments sample compartment, this chamber is connected to the VCD series system to enhance purge efficiency and reduce waiting time before acquisition. A full description of the house-built chamber design and characteristics is available in Sections 3 and 4 in the Supplemental Material, respectively. The total measurement procedure per sample includes a 4 min purge period to minimize atmospheric water vapor and carbon dioxide content in the sample compartment followed by a nearly 1 min spectrum acquisition.

In addition to using instruments from the same manufacturer, with the same hardware and under the same controlled environmental conditions, a careful matching of acquisition settings was also established to reduce the impact of instrument-related intrinsic factors on the spectral response. The matching was achieved by first exporting all acquisition parameters defined on FTIR1 using the OPUS software (Bruker Optics). The subsequent file is then loaded onto the two other FT-IR systems by an automatic procedure before each measurement session. This way, neither the length nor the alignment of the wavenumber channels is compromised across instruments thus preventing the need for spectra interpolation.

## Data Sets

In order to address the performance of each spectrometer at different time scales and compare their spectral response, a series of ambient samples, polystyrene standard, field blanks, and reference samples is considered. Ambient samples are routinely collected in the IMPROVE monitoring network by pulling a nominal air volume of $32.8\,m^3$ (24 h sampling time) through a 25 mm polytetrafluoroethylene (PTFE) filter. Samples are collected at nearly 160 pristine sites across the U.S. every third day. Unlike ambient samples, field blanks do not contain PM deposition as the filters are passively exposed during sample collection but no air is pulled through the sample. Finally, reference samples are a fixed set of ambient and blank samples that are analyzed weekly for quality control. For the sake of covering different aspects of the inter-instrument comparison, all available samples are structured into three specific data sets.

The first data set uses spectra from a polystyrene standard and multiple blank filters to identify potential differences in spectral responses between instruments. Polystyrene spectra are acquired daily, once 50 filters have been analyzed, to capture variability in spectral intensity and identify trends over time. A total of 96 (FTIR1), 77 (FTIR2), and 82 (FTIR3) standard polystyrene spectra obtained from February through August 2017 are considered. Additionally, 107 (FTIR1), 113 (FTIR2), and 99 (FTIR3) field blank PTFE spectra measured over the same time period are used to assess the distribution in PTFE signal. To provide further inter-instrument comparison, an ambient PM sample featuring nearly median OC (25.1 μg) and EC (4.2 μg) concentrations is introduced. The latter, sampled on 7 March 2017 at the Wichita Mountains (Oklahoma) IMPROVE site, was analyzed on all three spectrometers.

The second data set is introduced to evaluate the impact of the response of each spectrometer on OC and EC predictions. This set contains spectra from ambient samples analyzed on FTIR1 and FTIR2 as well as duplicate ambient spectra measured on the same instrument. First, a total of 743 ambient samples analyzed on FTIR1 and FTIR2 over a one-year period (2015) were employed to develop and validate a multivariate regression model for OC and EC prediction. Second, duplicate spectra from 265 (FTIR1) and 228 (FTIR2) ambient samples from the same year are used to evaluate instrument precision. Duplicate spectra are acquired within a 1 h time period, once or twice per day per instrument, to assess measurement repeatability. As the number of samples collected on FTIR3 is not yet statistically relevant, the above study is restricted to the two older instruments.

The third data set, designed to monitor the long-term stability of the three spectrometers, is based on a series of four references samples measured weekly as part of a quality control procedure. The former includes two ambient samples collected in Davis, California, USA in 2009, whose composition is stable over time, and two reference blank filters. This study focuses on 1327 spectra of the reference samples measured between February 2015 and September 2017 (FTIR1–FTIR2) and from February 2017 to September 2017 (FTIR3) to identify potential drift in instrument response over time. The ability to predict ambient samples collected in a different year is evaluated by introducing two additional sets of 855 and 391 ambient samples measured in the year 2016 and 2017, respectively. Spectra from 2016 were collected on both FTIR1 and FTIR2 while 2017 spectra were acquired on all three instruments.

## Data Analysis

The impact of the instrumental response on OC and EC predictions was addressed by developing a set of multivariate calibration models from the 743 ambient samples measured on both FTIR1 and FTIR2. Since this data set does not include blank filters, model predictions at low concentrations are not satisfactory. For this purpose, a homogeneously distributed set of 50 samples from the blanks data set, selected with the Kennard–Stone algorithm,[37] was included in the model to improve the minimum detection limit (MDL). Data partitioning into calibration and test sets was achieved following a base case scenario in which ambient and blanks filters are chronologically stratified by site.[9] Two-thirds of the ambient spectra along with a series of 34 blank spectra were used to build the multivariate model, yielding a total of 530 calibration samples. The remaining third of the ambient spectra were merged with 16 blank spectra to form a 263 samples test set employed to evaluate model performance. Before developing the model, spectra were centered by subtracting the mean absorbance value at each wavenumber. Fourier transform IR spectra were calibrated against mean-centered OC and EC concentrations provided by the thermal-optical reflectance (TOR) analysis of collocated quartz filters in agreement with the IMPROVE_A protocol (http://views.cira.colostate.edu/fed).[38]

Although zero-filled spectra (11 138 wavenumbers) are used for visual interpretation, interpolated data points are removed so that only the original data points are used in model development. The processed data yields 2725 channels per spectrum sampled every $1.286\,\text{cm}^{-1}$. Multivariate models were developed using partial least squares (PLS) regression.[8] Selection of the optimal number of factors, yielding a model with the least prediction error, is achieved via a K-fold cross-validation procedure using $K = 5$.[39] This procedure is repeated for a series of models computed by varying the number of factors. For each model, the root

mean square error of cross-validation (RMSECV) is calculated and the number of factors minimizing the RMSECV is selected as optimum.

Data handling and analysis were performed in Matlab R2015a (The MathWorks, Inc.) using the statistics and signal processing toolboxes. Partial least squares calculations were run using the libPLS Matlab package (v.1.9, Changsha Nice City, China).[40]

## Figures of Merit

To compare the performance of the models developed in this work, several figures of merit are derived based on the linear regression of the predicted (FT-IR) to measured (TOR) concentrations of test samples. Calculated metrics include the coefficient of determination ($R^2$), the root mean square error of prediction (RMSEP), the bias, the error, and the normalized error. The bias is calculated as the difference between predicted (FT-IR) and measured (TOR) concentrations while the error is defined as the absolute bias. The normalized error for a single prediction is the error divided by the corresponding TOR value. In this work, median bias, error, and relative error values are reported. Two additional TOR independent figures of merit addressing the minimum detection limit (MDL) and the instrument precision are also introduced. MDL is computed as three times the standard deviation of field blank samples while instrument precision is estimated from duplicate spectra following Eq. 1:

$$Precision = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{(\hat{y}_{i1} - \hat{y}_{i2})^2}{2}} \qquad (1)$$

where $\hat{y}_{i1}$ and $\hat{y}_{i2}$ represents the predicted concentrations of the $i$th spectra pair and $N$ is the total number of spectra measured in duplicate.

# Results and Discussion

## Assessment of Spectral Dissimilarities Between Spectrometers

The presence of spectral dissimilarities across instruments is first assessed by evaluating the distribution of absorbance values in a polystyrene standard. Mean absorbance and standard deviation, derived from the peak at $2849\,\text{cm}^{-1}$, are compared in Fig. 1 for each spectrometer. While nearly identical distributions are observed between FTIR1 ($0.743 \pm 0.005$ a.u) and FTIR2 ($0.744 \pm 0.006$ a.u), slightly different values are reported for FTIR3 ($0.752 \pm 0.003$ a.u). Significance testing of absorbance means ($t$-test) and variances ($F$-test) at the 95% significance level confirms that the two older instruments show similar accuracy and precision for the measurement of polystyrene absorbance values at
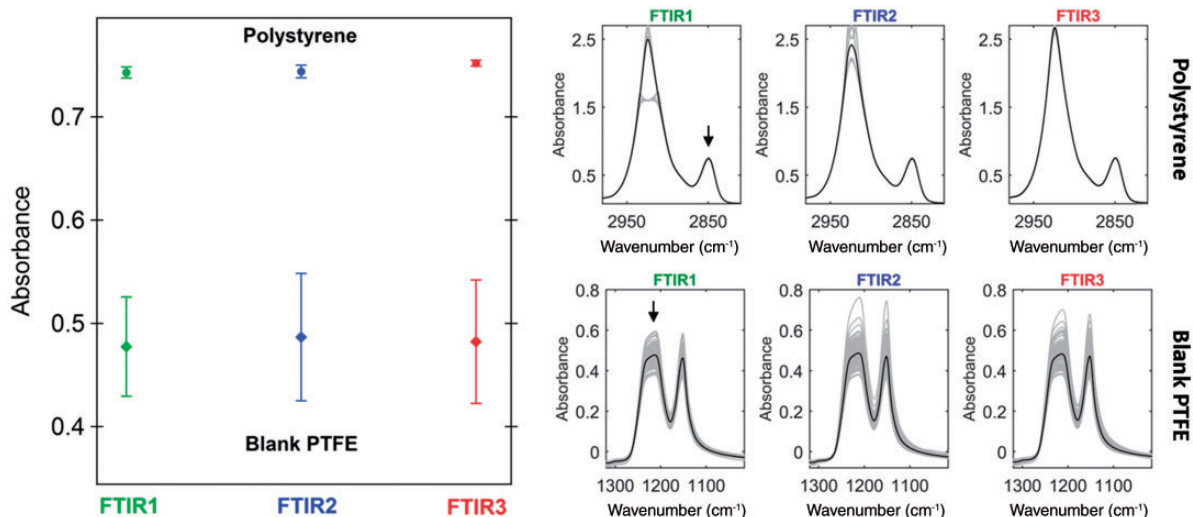
**Figure 1.** Comparison of accuracy and precision across instruments achieved for polystyrene standard and blank PTFE filters absorbances monitored at 2849 cm$^{-1}$ and 1210 cm$^{-1}$, respectively. Dark lines correspond to the mean spectrum per data set and per instrument while peak locations considered to compute distribution parameters are indicated by arrows. The highly absorbing polystyrene peak at 2925 cm$^{-1}$ is associated with a nonlinear response of the MCT detector. Variability in blank PTFE signal is not representative of instrument precision but of the variability in PTFE material.

2849 cm$^{-1}$. However, both older instruments show statistically significant differences with FTIR3 which suggests that (minor) discrepancies may exist even though all three spectrometers share the same instrumentation, environmental conditions, and experimental settings.

To better evaluate the extent of spectral dissimilarities in the case of filter samples, a similar analysis was conducted on the 1210 cm$^{-1}$ peak from blank PTFE filters. The PTFE signal in the range 1320–1000 cm$^{-1}$ is shown in Fig. 1 for all three spectrometers. Compared with polystyrene, the distribution of blank PTFE absorbance values for FTIR1 (0.477 ± 0.048 a.u), FTIR2 (0.487 ± 0.062 a.u), and FTIR3 (0.482 ± 0.060 a.u) is characterized by a larger standard deviation. Significance testing indicates no statistical differences between FTIR2 and FTIR3 blank PTFE absorbance values, suggesting that variability in the observed polystyrene signal is solely attributed to random error. On the other hand, comparing blank PTFE distributions between FTIR1 and FTIR2 or between FTIR1 and FTIR3 reveals statistically significant differences indicating that differences are not just due to random error but also to instrumental or experimental error. The low signal repeatability among filter blanks is mainly attributed to the combination of path length differences and non-uniformity in PTFE fiber morphology,[41] both resulting in a broad distribution of unique optical properties within the same filter and between filters. The substantially larger variation in blank PTFE signal compared to polystyrene leads to the conclusion that spectral dissimilarities across spectrometers are primarily driven by inter-filter variability rather than instrumental factors.

The low repeatability of PTFE signal and its impact on the spectral response of ambient samples is highlighted in Fig. 2a by comparing the FT-IR spectrum of an IMPROVE sample (Wichita Mountains, Oklahoma) collected on all three instruments. The spectral region below 1300 cm$^{-1}$ is characterized by a collection of sharp and highly absorbing peaks ascribed to the various C–F stretching modes of PTFE.[42,43] Beyond this range, the spectrum features a light-scattering induced sloping baseline as well as a series of low absorbing IR bands at 1720 cm$^{-1}$, 2844 cm$^{-1}$, 2911 cm$^{-1}$, and 3050–3340 cm$^{-1}$ associated with the IR fingerprints of organic functional groups (FG). The relatively weak absorbance of FG compared with the magnitude of a blank filter (Fig. 2, inset) indicates that PTFE account for the vast majority of the measured signal. Beside minor variations in relative intensity, more pronounced at the higher end of the spectra where scattering prevails, visual inspection of the raw spectra does not reveal any major dissimilarity. Similar comparison based on second derivative spectra confirms the absence of either spectral or intensity shifts, outside of spectral regions featuring water vapor and PTFE contributions (Section 5, Supplemental Material). A more detailed picture of spectral dissimilarities across instruments is obtained by subtracting the spectrum acquired on FTIR1 from the signal measured on FTIR2 and FTIR3. The resulting difference spectra, reported in Fig. 2b, capture information about the nature of spectral differences involved in the signal of ambient samples. Based on the shape and magnitude of the difference spectra, two conclusions can be drawn. First, even though spectral dissimilarities exist
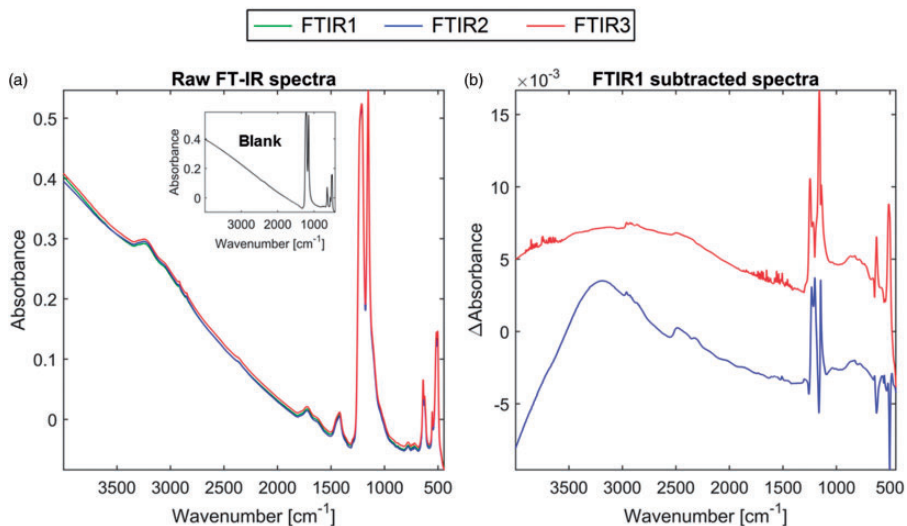
**Figure 2.** (a) Raw and (b) difference FT-IR spectra of the Wichita Mountains, Oklahoma, ambient sample measured on all three spectrometers. Differential spectra are obtained by subtracting both FTIR2 and FTIR3 spectra with the signal collected on FTIR1. Insert representing a blank filter is displayed for the sake of comparison.

across instruments, their respective intensity remain small compared to the raw signal (note the 1000-fold decrease in scale between Fig. 2a and 2b). Second, most of the spectral differences such as the intensity offset and the peak structure at 1200 cm$^{-1}$ are attributed to PTFE-related factors. Notably, the change in PTFE signal observed across spectrometers for the same ambient sample is induced by the non-uniform density of the filter (Section 6, Supplemental Material). From Fig. 2b, it is worth noting that dissimilarities of lower magnitude can also be spotted such as the signature of carbon dioxide (near 2360 cm$^{-1}$), water vapor (3500–4000 cm$^{-1}$, 1540–1800 cm$^{-1}$), and a trace of volatile organic carbon (VOC, 2900 cm$^{-1}$).[44] In single-beam FT-IR studies, minimizing gas-phase absorption is commonly achieved by collecting an empty chamber background before each measurement. In the context of large monitoring networks such as IMPROVE, however, backgrounds are collected hourly to reduce analysis time and accommodate high sample throughput which might cause small fluctuations in water vapor, $CO_2$, and VOC regions on the measured spectrum. Improving the purge efficiency by using the small volume house-built chamber helps to minimize signal contributions of gas-phase interferents relative to other sample-specific variations (i.e., PTFE).

In summary, the strict control of key environmental and instrumental parameters minimizes inter-instrument deviations in spectral response between the three spectrometers. The sources of spectral discrepancies observed in PM filter samples are primarily attributed to PTFE contributions and to a lesser extend to the presence of gas-phase interferents inside the sample compartment.

## Impact of Spectral Dissimilarities on Organic Carbon and Elemental Carbon Prediction

To address the performance of OC and EC predictions given the presence of minor inter-instruments spectral dissimilarities, a series of calibration models are developed and validated based on the set of 793 ambient samples introduced in the Data Analysis section. The latter samples, analyzed on both FTIR1 and FTIR2, are stratified into calibration and test sets using a two-thirds ($n = 530$) and one-third ($n = 263$) ratio, respectively. First, separate calibrations are developed on each FT-IR system from the set of 530 samples collected within the same instrument. Then, model evaluation is performed following two specific scenarios: (1) validation using the 263 ambient spectra collected on the same instrument; or (2) validation using the 263 ambient spectra acquired on the other instrument. Performance metrics associated with every instrument–scenario combination are reported in Table 1.

To prevent subjectivity in the assessment of significant differences between models developed on FTIR1 and FTIR2, the ratio of the squared RMSEP values obtained for each model was computed and compared to the critical value of a two-sided $F$-test at the 95% confidence level.[45,46] Since all calculated ratios fall below the critical value $F_{(n = 262, \ n = 262, 0.95)} = 1.275$, OC and EC models developed separately on each instrument are not considered statistically different. Furthermore, swapping models between spectrometers (2) does not introduce significant differences in prediction metrics with the exception of a slight increase in normalized error. Although model performances are considered equivalent, the impact of spectral dissimilarities

**Table I.** Comparison of OC and EC model performance between FTIR1 and FTIR2 for scenarios (1) and (2), figures of merit are computed from test set FT-IR spectra ($n = 263$).

| | Calibration | Test | PLS components | $R^2$ | RMSEP (µg) | Bias (µg) | Error (µg) | Norm error (%) | MDL (µg) |
|---|---|---|---|---|---|---|---|---|---|
| OC (1) | FTIR1 | FTIR1 | 18 | 0.97 | 5.9 | 0.12 | 2.98 | 11.4 | 3.3 |
| | FTIR2 | FTIR2 | 17 | 0.97 | 6.0 | 0.41 | 3.03 | 12.9 | 6.7 |
| OC (2) | FTIR1 | FTIR2 | 18 | 0.97 | 5.9 | 0.01 | 3.01 | 13.3 | 4.8 |
| | FTIR2 | FTIR1 | 17 | 0.97 | 6.3 | 0.46 | 3.20 | 13.5 | 5.4 |
| EC (1) | FTIR1 | FTIR1 | 30 | 0.92 | 1.7 | 0.01 | 0.86 | 34.4 | 2.5 |
| | FTIR2 | FTIR2 | 32 | 0.91 | 1.8 | 0.12 | 0.81 | 34.5 | 1.8 |
| EC (2) | FTIR1 | FTIR2 | 30 | 0.91 | 1.8 | 0.05 | 0.89 | 33.6 | 1.9 |
| | FTIR2 | FTIR1 | 32 | 0.92 | 1.7 | 0.12 | 0.87 | 36.4 | 2.7 |

between instruments can still be observed in the form of a higher bias and MDL on FTIR2 as well as a slightly different number of PLS components. Despite the apparent variability in MDL, values remain consistent with the ones reported in previous works.[9,47]

Considering that inter-instrument spectral dissimilarities are predominantly influenced by filter anisotropic properties, maintaining high-quality predictions supposes the selection of an appropriate subset of calibration samples which is representative of both PM composition and optical properties of the substrate experienced in test samples. In Table 1, the statistical equivalence between the performance of the models suggests that the size of the calibration set is adequate. However, it is likely that a different number of calibration samples or a different samples combination could achieve equivalent prediction performance. To estimate the optimal (i.e., minimal) size of the calibration set required to maintain high-quality predictions, the number of calibration samples $N_{cal}$ was gradually decreased from 530 to 57 while keeping the number of blanks constant ($n = 34$). For each calibration set, a PLS model was generated and validated using the same test set as in Table 1 ($n = 263$). In Fig. 3, variations in model performance as function of $N_{cal}$ are shown for the case of OC prediction metrics developed from scenario (1). Complementary studies associated with scenario (2) and EC predictions are given in the Supplemental Material (Section 7). No significant degradation in prediction is observed until $N_{cal}$ drops by more than half of its initial size ($n = 244$). Between 530 and 244 calibration samples, nearly identical metrics are reported for both OC scenarios with the exception of a consistently higher bias on FTIR2 ($\approx 0.5$ µg) suggesting slightly overestimated OC concentrations compared with FTIR1. However, the magnitude of such a bias is lower than the 95% confidence interval (CI) of test set blank predicted concentrations ($\pm 0.9$ µg) and could thus be considered negligible. As $N_{cal}$ decreases from 244 to 140, a moderate deterioration in prediction metrics might suggest that calibration samples are no longer representative of the whole distribution in PTFE and PM

composition of test set samples. Further reduction in the number of samples yields increased error with prediction metrics no longer matching between the two spectrometers. The $R^2$ and RMSEP plots in Fig. S7-2 (Supplemental Material) suggest similar conclusions for EC prediction.

To assess the extent of PTFE variability between test and calibration samples, distribution parameters (e.g., mean absorbance) of the PTFE peak at 1210 cm$^{-1}$ are derived from the raw spectra (Fig. S7-3, Supplemental Material). For $N_{cal}$ values $> 250$, PTFE distribution parameters derived from calibration and test spectra are nearly identical and therefore do not influence the prediction quality other than a negligible bias. However, lower $N_{cal}$ values are characterized by a divergence in PTFE distribution parameters between calibration and test spectra which likely contributes to the drop in prediction quality seen in Fig. 3. In addition to the lack of representativeness of PTFE distribution in test samples, the progressive degradation of OC prediction as $N_{cal}$ decreases is also attributed to unmatched PM composition between calibration and test samples. According to Fig. S7-3, at least 300 samples are required in the calibration in order to efficiently diminish the impact of PTFE signal variability on model performance.

Based on the results from the two first scenarios, the variability in PTFE signal appears to have little to no influence on OC and EC prediction as long as the calibration set is representative of both PTFE and PM composition in test samples. The fact that calibration models built on either instrument predict equivalent concentrations without resorting to sophisticated calibration transfer strategies is attributed to two factors. One is the strict control of water vapor and carbon dioxide concentrations, detector temperature, sample position relative to the IR beam, and room conditions which enable the instruments to provide similar responses. The second is the capability of PLS to reduce the remaining sources of variability interfering with OC or EC prediction (i.e., PTFE).

Because maintaining multiple calibration models over time is generally not suitable from a long-term perspective,
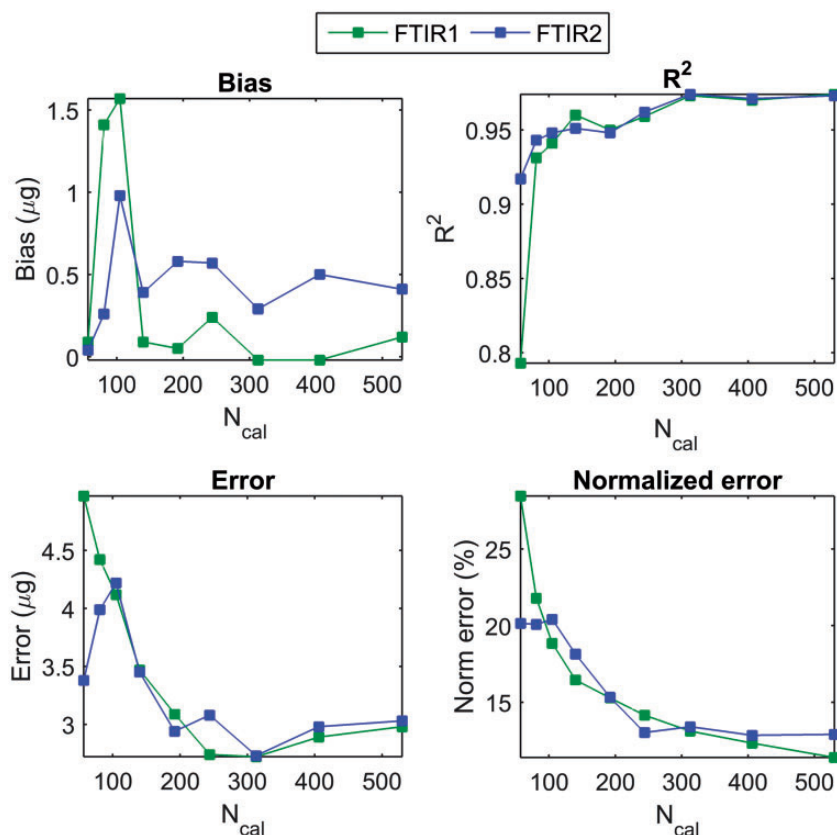
**Figure 3.** Inter-instrument comparison of OC prediction metrics based on scenario (1) as function of the number of samples $N_{cal}$ included in the calibration. Reported performance are derived from test spectra ($n = 263$) collected on the same instrument used to analyze calibration samples. Models are developed and evaluated with samples from the year 2015.

and considering that models developed separately on each instrument are statistically equivalent, it appears that either calibration can be selected as part of a routine measurement procedure. However, neither model accounts for the specific spectral response of both instruments as indicated by the difference in bias reported in Table 1 and in Fig. 3 for FTIR1 and FTIR2. To ensure that instrument specific differences, while small, are accounted for, an instrument-agnostic calibration model, often referred to as a hybrid model,[48] is introduced. The hybrid model is presented in a third scenario (3) by using spectra acquired on both spectrometers within a single calibration set. To ensure that the number of samples, the distribution of sites, and the sample composition remain the same as in Table 1, only samples whose array index features odd numbers on FTIR1 and even numbers on FTIR2 were included in the calibration. Performance associated with the third scenario, reported in Table II, was evaluated from the same set of test spectra used previously on scenarios (1) and (2). As in Table 1, the significance testing of the RMSEP values do not exceed the critical value $F_{(n = 262, n = 262, 0.95)} = 1.275$, indicating that the hybrid model predicts OC and EC concentrations equally well on both spectrometers. In general, similar or slightly improved OC prediction metrics are reported

compared with models built on a single instrument while similar or slightly worse prediction metrics can be claimed for EC. Because EC has low IR activity,[49] the corresponding hybrid model is more prone to overfitting due to the presence of low-magnitude spectral interferents. In particular, difference in water vapor and carbon dioxide content from one instrument to another is assumed to be the main factor explaining the decrease in EC prediction. Nonetheless, the overall performance by the hybrid model is considered satisfactory for providing robust predictions independent of the instrument. Additional benefit of the hybrid model compared with the other two scenarios can be inferred from the absence of significant bias between the two spectrometers. The larger bias previously reported for FTIR2 is corrected for by simply developing a common set of regression coefficients considering the characteristics of both instruments. The third FT-IR instrument will be evaluated for inclusion into a hybrid calibration model using spectra from all three instruments in future studies.

## Repeatability and Long-Term Stability Assessment

Instrument stability within a short time frame, typically on the order of a few hours, is first addressed by considering

**Table II.** Comparison of OC and EC hybrid model performance between FTIR1 and FTIR2 for scenario (3), figures of merit are computed from test set FT-IR spectra ($n = 263$).

| | Calibration | Test | PLS components | $R^2$ | RMSEP (µg) | Bias (µg) | Error (µg) | Norm error (%) | MDL (µg) |
|---|---|---|---|---|---|---|---|---|---|
| OC (3) | Hybrid | FTIR1 | 18 | 0.98 | 5.9 | 0.19 | 3.09 | 12.7 | 3.1 |
| | | FTIR2 | | 0.98 | 5.8 | 0.15 | 2.88 | 12.4 | 4.8 |
| EC (3) | Hybrid | FTIR1 | 28 | 0.91 | 1.8 | 0.03 | 0.93 | 40.1 | 2.1 |
| | | FTIR2 | | 0.90 | 1.9 | 0.07 | 0.93 | 33.9 | 2.5 |

daily duplicate measurements collected once or twice per day on each instrument. After subjecting all replicate spectra to the hybrid models, the subsequent predicted concentrations for each sample pair is used to compute instrument precision, a reliable estimate of measurement repeatability, according to Eq. 1. While equivalent precision values are achieved on FTIR1 (0.79 µg) and FTIR2 (0.75 µg) for OC, EC results suggest that FTIR2 provides more robust measurement repeatability than FTIR1 with precision values of 0.26 µg (FTIR2) and 0.73 µg (FTIR1). Although slightly elevated, calculated precisions are of the same order of magnitude as the 95% CI of test set blank predicted OC ($\pm 0.70$ µg) and EC ($\pm 0.42$ µg) concentrations. Given the narrow gap between precision and the 95% CI values, the stability of both spectrometers over a short period of time is considered satisfactory and justifies the choice to collect backgrounds hourly.

Through variations introduced by environmental conditions or characteristics of the instrument, a change in spectral response over time is a common occurrence in spectroscopic systems. Therefore, monitoring instrument stability is required to ensure the validity of the hybrid model over time, highlight any new potential sources of spectral variations affecting model performance, and foresee situations where a full recalibration may be required. In an attempt to identify long-term drifts in instrument response, we refer to the set of four reference samples selected for this evaluation. Unlike IMPROVE ambient samples, reference filters are always measured at a fixed orientation to minimize the impact of filter anisotropic properties and emphasize spectral variations induced by the instruments. Predicted concentrations, calculated weekly for each reference sample based on the OC hybrid model, are used to develop a group of time series plots displayed in Fig. 4. The series is characterized by (local) non-periodic oscillations in predicted concentrations attributed to minor changes in background properties over time induced by either environmental- or instrument-related perturbations. To assess the extent of those variations, the median OC concentration is computed based on the first month of data collection on all three spectrometers. Intervals corresponding to $\pm 5\%$ (dotted lines) and $\pm 10\%$ (straight lines) variations around the median predicted concentration are then considered for reference

ambient samples while only the median predicted concentration is displayed for blank reference filters (dotted lines). Almost 90% of the predicted concentrations of reference ambient samples measured over a 2.5-year period for FTIR1 and FTIR2 and over an eight-month period for FTIR3 are within the $\pm 5\%$ of the initial concentration while about 10% lie in the $\pm 5$–10% interval. The absence of clear drift in predicted standard OC concentrations over an extended period of time highlights the relative stability of the spectral response on all three FT-IR systems. Similar conclusions can be drawn in the case of EC predictions (data not shown).

Although reference samples provide a reliable screening tool to monitor instrument stability over time, their composition is not necessarily representative of the PM distribution experienced in network ambient samples. To address this issue, a similar study was conducted using 2078 FT-IR spectra of ambient PM samples measured on both FTIR1 and FTIR2 over the same three-year time period. For each year of spectra collection, predicted concentrations achieved on FTIR2 are regressed against their FTIR1 counterpart and a simple ordinary least squares regression procedure is employed to derive slope and intercept parameters. Percentile-based bootstrapped CIs are then developed to test whether or not the calculated slope and intercept are significantly different from one and zero, respectively.[50] In this study, we rely on 95% CIs computed from a series of 1000 bootstrap replicates for each metric. Yearly slope, intercept, and their respective 95% CIs derived for OC prediction on the FTIR2–FTIR1 instrument pair are illustrated in Fig. 5.

Figure 5 shows that the hybrid model still provides acceptable predictions for both 2016 and 2017 IMPROVE samples. This corroborates the findings of Reggente et al. who showed that a single-instrument calibration model can be extended to samples collected and analyzed in a subsequent year.[47] Notably, slope-associated CIs reveal no significant differences in cross-instrument predictions between consecutive years. The high quality of the regression is suggested by the fact that none of the calculated slopes are statistically different than one. The presence of non-zero intercepts, however, discloses a low magnitude bias between the predictions achieved on FTIR1 and FTIR2. While the years 2015 and 2016 are both
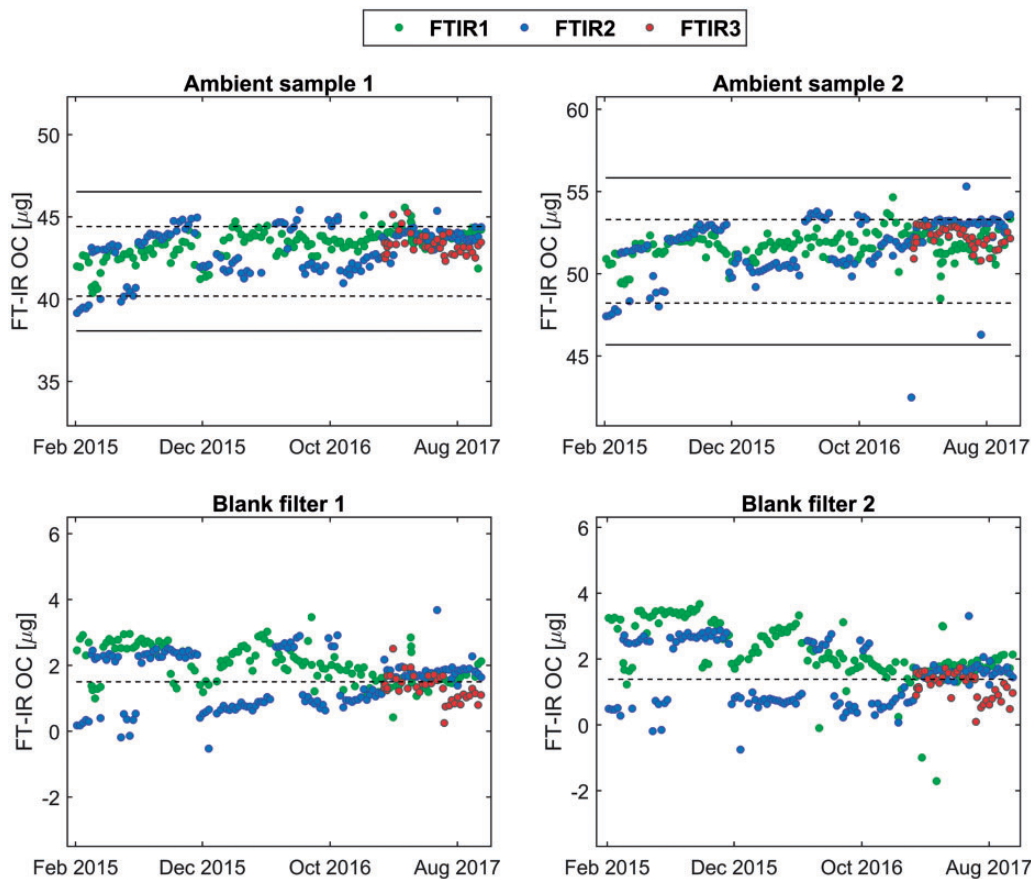
**Figure 4.** Concentrations time series of four reference samples predicted by the hybrid OC model initially developed for the year 2015. The x-axis corresponds to the date when the analysis was performed. Dashed and solid lines for reference ambient samples indicate the ±5% and ±10% error intervals based on the median OC concentration calculated by averaging the first month of data collection for all three instruments. Dashed lines for reference blank filters represent the median concentration.
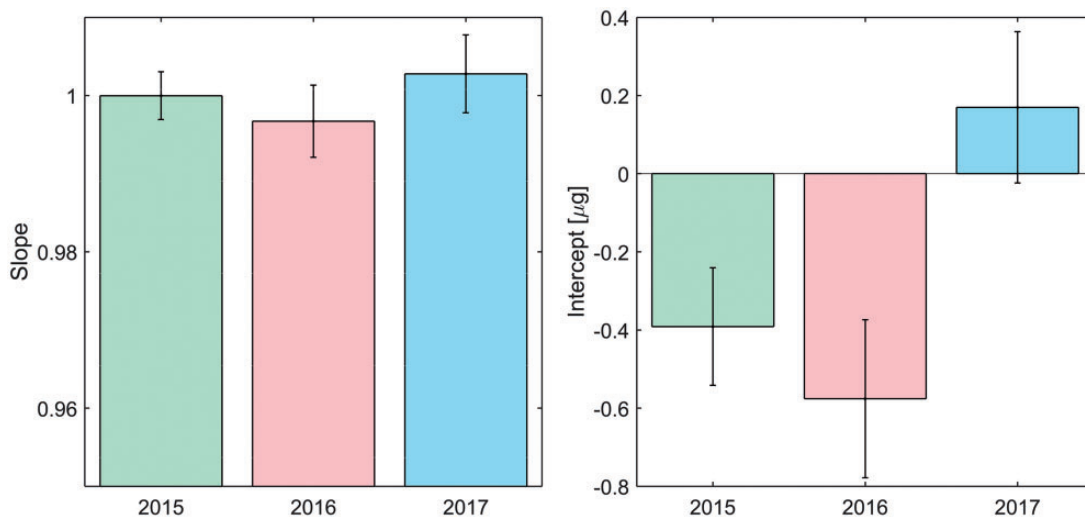


**Figure 5.** Time-dependent slopes (left) and intercepts (right) derived from test set OC predicted concentrations on the FTIR2–FTIR1 instrument pair, uncertainty of each metric is evaluated by developing 95% CIs from 1000 bootstrap replicates.

characterized by negative intercepts, significantly different from zero, the year 2017 features a positive intercept not statistically different from zero. Relative to the year 2015, the moderate increase in the 2016 intercept is tentatively attributed to the transfer of the instruments to a new laboratory room. However, the change in intercept for 2017 data is not yet clearly associated with any known sources of interference but the increase in predicted reference concentrations on FTIR2 (Fig. 4) in early 2017 suggests an instrumental drift. As the intercepts remain below the MDL (Table II), the performance of the OC hybrid model extended to the years 2016 and 2017 are nonetheless considered satisfactory. Similar results are observed for the EC hybrid model (Fig. S8-1, Supplemental Material) for which a significant gap between 2016 (0.14 µg) and 2017 (−0.94 µg) intercepts is also observed. Such a difference, larger than the precision values reported above, pinpoints the limitation and sensitivity of the hybrid EC model to variations in experimental conditions.

Finally, instrument pairs involving the new FTIR3 system were investigated for the restricted 2017 time period. While FTIR3–FTIR1 and FTIR3–FTIR2 pairs provide equivalent OC predictions (Fig. S7-2), different results are reported for EC with the absence of overlap in the 95% confidence intercept interval (Section 8, Supplemental Material). Despite the elevated intercept values, instrument or model recalibration is not suggested as the intercept between instrument-pairs does not exceed EC MDL values in Table II.

## Conclusion

This paper describes the strategy behind the development of a robust, low maintenance, calibration model applicable across multiple FT-IR instruments and multiple years. The need for advanced calibration transfer procedures was circumvented by carefully controlling instrumental and environmental parameters in an attempt to minimize dissimilarities in instrument response. After ensuring that all spectrometers are operated under fixed experimental conditions, the remaining sources of spectral variability were attributed to the heterogeneous structural and optical properties of PTFE filters. However, the effects of PTFE interference have little to no influence on OC and EC predicted concentrations as long as calibration samples are representative of the PTFE variability and PM composition in test samples. Instead of building a specific calibration per instrument, hybrid OC and EC calibration models were developed based on ambient IMPROVE spectra collected on each spectrometer. With the exception of a small decrease in EC prediction capability attributed to variations in water vapor and carbon dioxide content inside the sample compartments, hybrid models provide robust and accurate predictions independent of the spectrometer used for data acquisition. The weekly spectral collection of

duplicate and standard PM samples indicates both high measurement repeatability and long-term stability of the response of the instruments. Extending hybrid OC and EC models, developed for the year 2015, to the prediction of ambient PM samples collected in the two subsequent years does not produce significant changes in prediction metrics further indicating that the proposed hybrid calibration models are suitable for OC and EC prediction in network operations.

## Supplemental Material

All supplemental material mentioned in the text, consisting of Figs. S1–S8, is available in the online version of the journal.

## ORCID iD

Bruno Debus ⬥ http://orcid.org/0000-0003-3635-4217
Satoshi Takahama ⬥ http://orcid.org/0000-0002-3335-8741

## References

1. R. Bauer, H. Nieuwoudt, F.F. Bauer, J. Kossmann, et al. "FTIR Spectroscopy for Grape and Wine Analysis". Anal. Chem. 2008. 80(5): 1371–1379.
2. L. Cassani, M. Santos, E. Gerbino, M. Del Rosario Moreira, et al. "A Combined Approach of Infrared Spectroscopy and Multivariate Analysis for the Simultaneous Determination of Sugars and Fructans in Strawberry Juices During Storage". J. Food Sci. 2017. 83(3): 631–638.
3. S. Yousefinejad, L. Aalizadeh, F. Honarasa. "Application of ATR-FTIR Spectroscopy and Chemometrics for the Discrimination of Furnace Oil, Gas Oil and Mazut Oil". Anal. Methods. 2016. 8(23): 4640–4647.
4. A.A. Bunaciu, H.Y. Aboul-Enein, S. Fleschin. "Application of Fourier Transform Infrared Spectrophotometry in Pharmaceutical Drugs Analysis". Appl. Spectrosc. Rev. 2010. 45(3): 206–219.
5. C. Koch, A.E. Posch, H.C. Goicoechea, C. Herwig, et al. "Multi-Analyte Quantification in Bioprocesses by Fourier-Transform-Infrared Spectroscopy by Partial Least Squares Regression and Multivariate Curve Resolution". Anal. Chim. Acta. 2014. 807: 103–110.
6. P.E.P. Chang, J.C.R. Yang, W. Den, C.F. Wu. "Characterizing and Locating Air Pollution Sources in A Complex Industrial District Using Optical Remote Sensing Technology and Multivariate Statistical Modeling". Environ. Sci. Pollut. Res. 2014. 21(18): 10852–10866.
7. L.M. Russell. "Aerosol Organic-Mass to Organic-Carbon Ratio Measurements". Environ. Sci. Technol. 2003. 37(13): 2982–2987.
8. S. Wold, M. Sjöström, L. Eriksson. "PLS-Regression: A Basic Tool of Chemometrics". Chemom. Intell. Lab. Syst. 2001. 58(2): 109–130.

9. A.M. Dillner, S. Takahama. "Predicting Ambient Aerosol Thermal-Optical Reflectance (TOR) Measurements from Infrared Spectra: Organic Carbon". Atmos. Meas. Tech. 2015. 8(3): 1097–1109.

10. A.M. Dillner, S. Takahama. "Predicting Ambient Aerosol Thermal-Optical Reflectance Measurements from Infrared Spectra: Elemental Carbon". Atmos. Meas. Tech. 2015. 8(10): 4013–4023.

11. C. Coury, A.M. Dillner. "A Method to Quantify Organic Functional Groups and Inorganic Compounds in Ambient Aerosols Using Attenuated Total Reflectance FTIR Spectroscopy and Multivariate Chemometric Techniques". Atmos. Environ. 2008. 42(23): 5923–5932.

12. T.C. Ruthenburg, P.C. Perlin, V. Liu, C.E. McDade, et al. "Determination of Organic Matter and Organic Matter to Organic Carbon Ratios by Infrared Spectroscopy with Application to Selected Sites in the IMPROVE Network". Atmos. Environ. 2014. 86(Suppl. C): 47–57.

13. K.M. George, T.C. Ruthenburg, J. Smith, L. Yu, et al. "FT-IR Quantification of the Carbonyl Functional Group in Aqueous-Phase Secondary Organic Aerosol from Phenols". Atmos. Environ. 2015. 100(Suppl. C): 230–237.

14. P.A. Solomon, M. Costantini, T.J. Grahame, M.E. Gerlofs-Nijland, et al. "Air Pollution and Health: Bridging the Gap from Sources to Health Outcomes: Conference Summary". Air Qual. Atmos. Health. 2012. 5(1): 9–62.

15. H. Zhao, X. Li, Q. Zhang, X. Jiang, et al. "Effects of Atmospheric Transport and Trade on Air Pollution Mortality in China". Atmos. Chem. Phys. 2017. 17(17): 10367–10381.

16. T. Glotfelty, Y. Zhang, P. Karamchandani, D.G. Streets. "Will the Role of Intercontinental Transport Change in A Changing Climate?" Atmos. Chem. Phys. 2014. 14(17): 9379–9402.

17. Y. Liu, Z. Liu, R. Maghirang, E. Santos, et al. "Apportionment of Visibility Degradation to Sources of PM2.5 in Rural and Urban Kansas". Paper presented at 2017 ASABE Annual International Meeting, Spokane, Washington; July 16–19, 2017.

18. J.G. Watson. "Visibility: Science and Regulation". J. Air Waste Manag. Assoc. 2002. 52(6): 628–713.

19. J.L. Jimenez, M.R. Canagaratna, N.M. Donahue, A.S.H. Prevot, et al. "Evolution of Organic Aerosols in the Atmosphere". Science. 2009. 326(5959): 1525–1529.

20. J.F. Hamilton, P.J. Webb, A.C. Lewis, J.R. Hopkins, et al. "Partially Oxidised Organic Components in Urban Aerosol Using GCXGC-TOF/MS". Atmos. Chem. Phys. 2004. 4(5): 1279–1290.

21. Z. Tao, H. Yu, M. Chin. "Impact of Transpacific Aerosol on Air Quality Over the United States: A Perspective from Aerosol–Cloud–Radiation Interactions". Atmos. Environ. 2016. 125(Part A): 48–60.

22. G. Snider, C.L. Weagle, R.V. Martin, A. Van Donkelaar, et al. "SPARTAN: A Global Network to Evaluate and Enhance Satellite-Based Estimates of Ground-Level Particulate Matter for Global Health Applications". Atmos. Meas. Tech. 2015. 8(1): 505–521.

23. K. Tørseth, W. Aas, K. Breivik, A.M. Fjæraa, et al. "Introduction to the European Monitoring and Evaluation Programme (EMEP) and Observed Atmospheric Composition Change During 1972–2009". Atmos. Chem. Phys. 2012. 12(12): 5447–5481.

24. W.C. Malm, B.A. Schichtel, M.L. Pitchford, L.L. Ashbaugh, et al. "Spatial and Monthly Trends in Speciated Fine Particle Concentration in the United States". J. Geophys. Res. 2004. 109(D3): D03306.

25. P.A. Solomon, D. Crumpler, J.B. Flanagan, R.K.M. Jayanty, et al. "U.S. National PM2.5 Chemical Speciation Monitoring Networks—CSN and IMPROVE: Description of Networks". J. Air Waste Manag. Assoc. 2014. 64(12): 1410–1438.

26. R.G. Dambergs, M.D. Mercurio, S. Kassara, D. Cozzolino, et al. "Rapid Measurement of Methyl Cellulose Precipitable Tannins Using Ultraviolet Spectroscopy with Chemometrics: Application to Red Wine and Inter-Laboratory Calibration Transfer". Appl. Spectrosc. 2012. 66(6): 656–664.

27. Y. Ge, C.L.S. Morgan, S. Grunwald, D.J. Brown, et al. "Comparison of Soil Reflectance Spectra and Calibration Models Obtained Using Multiple Spectrometers". Geoderma. 2011. 161(3–4): 202–211.

28. C. Grelet, J.A. Fernández Pierna, P. Dardenne, V. Baeten, et al. "Standardization of Milk Mid-Infrared Spectra from A European Dairy Network". J. Dairy Sci. 2015. 98(4): 2150–2160.

29. M. Sohn, F.E. Barton Ii, D.S. Himmelsbach. "Transfer of Near-Infrared Calibration Model for Determining Fiber Content in Flax: Effects of Transfer Samples and Standardization Procedure". Appl. Spectrosc. 2007. 61(4): 414–418.

30. F. Sales, M.P. Callao, F.X. Rius. "Multivariate Standardization Techniques Using UV–Vis Data". Chemom. Intell. Lab. Syst. 1997. 38(1): 63–73.

31. J.J. Workman Jr. "A Review of Calibration Transfer Practices and Instrument Differences in Spectroscopy". Appl. Spectrosc. 2018. 72(3): 340–365.

32. R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, et al. "Transfer of Multivariate Calibration Models: A Review". Chemom. Intell. Lab. Syst. 2002. 64(2): 181–192.

33. H. Swierenga, W.G. Haanstra, A.P. De Weijer, L.M.C. Buydens. "Comparison of Two Different Approaches Toward Model Transferability in NIR Spectroscopy". Appl. Spectrosc. 1998. 52(1): 7–16.

34. E. Bouveresse, C. Hartmann, D.L. Massart, I.R. Last, et al. "Standardization of Near-Infrared Spectrometric Instruments". Anal. Chem. 1996. 68(6): 982–990.

35. O.E. De Noord. "Multivariate Calibration Standardization". Chemom. Intell. Lab. Syst. 1994. 25(2): 85–97.

36. I.S. Adhihetty, J.A. McGuire, B. Wangmaneerat, T.M. Niemczyk, et al. "Achieving Transferable Multivariate Spectral Calibration Models: Demonstration with Infrared Spectra of Thin-Film Dielectrics on Silicon". Anal. Chem. 1991. 63(20): 2329–2338.

37. R.W. Kennard, L.A. Stone. "Computer Aided Design of Experiments". Technometrics. 1969. 11(1): 137–148.

38. J.C. Chow, J.G. Watson, L.C. Pritchett, W.R. Pierson, et al. "The Dri Thermal/Optical Reflectance Carbon Analysis System: Description, Evaluation and Applications in U.S. Air Quality Studies". Atmos. Environ. Gen. Top. 1993. 27(8): 1185–1201.

39. S. Arlot, A. Celisse. "A Survey of Cross-Validation Procedures for Model Selection". Stat. Surv. 2010. 4: 40–79.

40. H.-D. Li, Q.-S. Xu, Y.-Z. Liang. "LibPLS: An Integrated Library for Partial Least Squares Regression and Linear Discriminant Analysis". Chemom. Intell. Lab. Syst. 2018. 176: 34–43.

41. C. Quarti, A. Milani, C. Castiglioni. "Ab Initio Calculation of the IR Spectrum of PTFE: Helical Symmetry and Defects". J. Phys. Chem. B. 2013. 117(2): 706–718.

42. C.Y. Liang, S. Krimm. "Infrared Spectra of High Polymers. III. Polytetrafluoroethylene and Polychlorotrifluoroethylene". J. Chem. Phys. 1956. 25(3): 563–571.

43. R.E. Moynihan. "The Molecular Structure of Perfluorocarbon Polymers. Infrared Studies on Polytetrafluoroethylene". J. Am. Chem. Soc. 1959. 81(5): 1045–1050.

44. U.J. Kim, C.A. Furtado, X. Liu, G. Chen, et al. "Raman and IR Spectroscopy of Chemically Processed Single-Walled Carbon Nanotubes". J. Am. Chem. Soc. 2005. 127(44): 15437–15445.

45. C.F. Pereira, M.F. Pimentel, R.K.H. Galvão, F.A. Honorato, et al. "A Comparative Study of Calibration Transfer Methods for

Determination of Gasoline Quality Parameters in Three Different Near Infrared Spectrometers". Anal. Chim. Acta. 2008. 611(1): 41–47.

46. C.M. Gryniewicz-Ruzicka, S. Arzhantsev, L.N. Pelster, B.J. Westenberger, et al. "Multivariate Calibration and Instrument Standardization for the Rapid Detection of Diethylene Glycol in Glycerin by Raman Spectroscopy". Appl. Spectrosc. 2011. 65(3): 334–341.

47. M. Reggente, A.M. Dillner, S. Takahama. "Predicting Ambient Aerosol Thermal-Optical Reflectance (TOR) Measurements from Infrared Spectra: Extending the Predictions to Different Years and Different Sites". Atmos. Meas. Tech. 2016. 9(2): 441–454.

48. D. Ozdemir, M. Mosley, R. Williams. "Hybrid Calibration Models: An Alternative to Calibration Transfer". Appl. Spectrosc. 1998. 52(4): 599–603.

49. J.L. Figueiredo, M.F.R. Pereira, M.M.A. Freitas, J.J.M. Órfão. "Modification of the Surface Chemistry of Activated Carbons". Carbon. 1999. 37(9): 1379–1389.

50. B. Efron, R.J. Tibshirani. An Introduction to the Bootstrap. New York; London: Chapman and Hall, 1993.

# 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 Debus[1], Satoshi Takahama[2], Andrew T. Weakley[1], Kelsey Seibert[1], Ann M. Dillner[1]*

[1] Air 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**



*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**
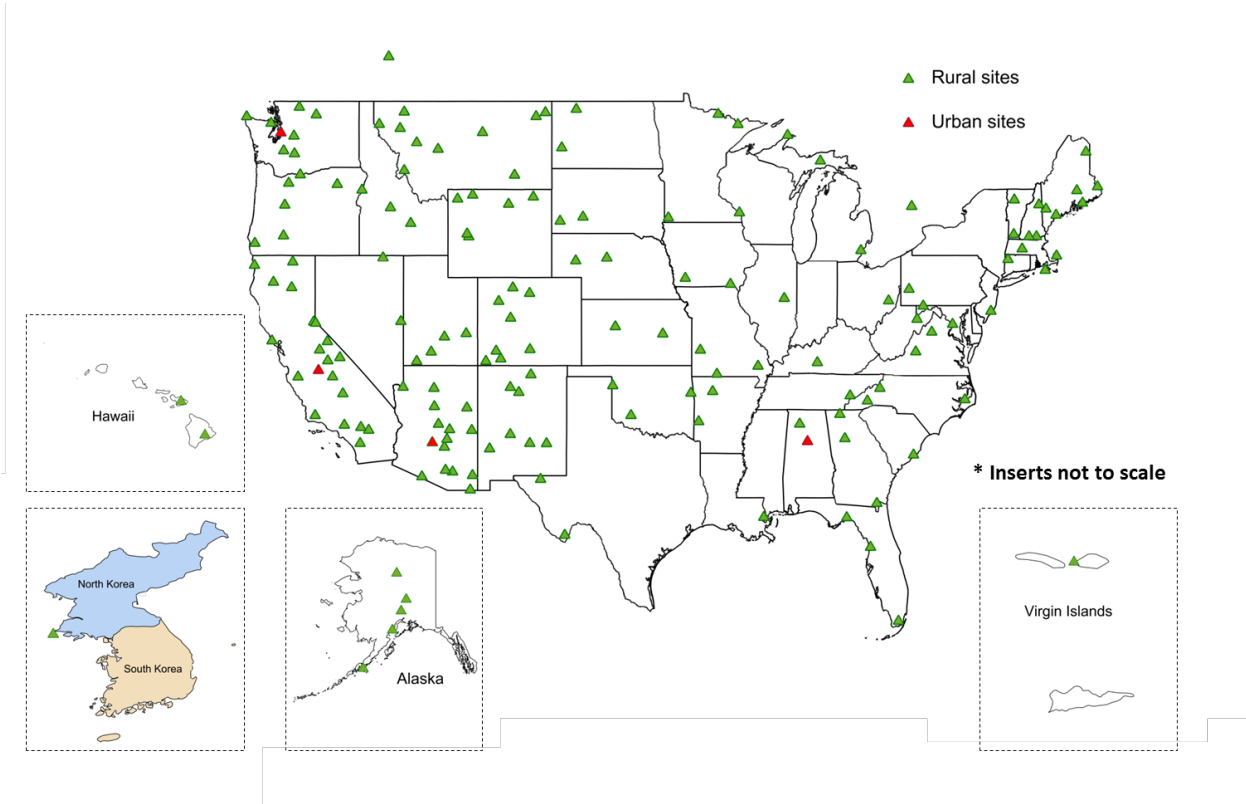


*Figure S2. PM sampling locations from the IMPROVE network including 159 rural sites (green markers) and 4 urban sites (red markers). With the exception of the heavily polluted site in South Korea, ambient samples collected across the whole network were considered in the present study.*

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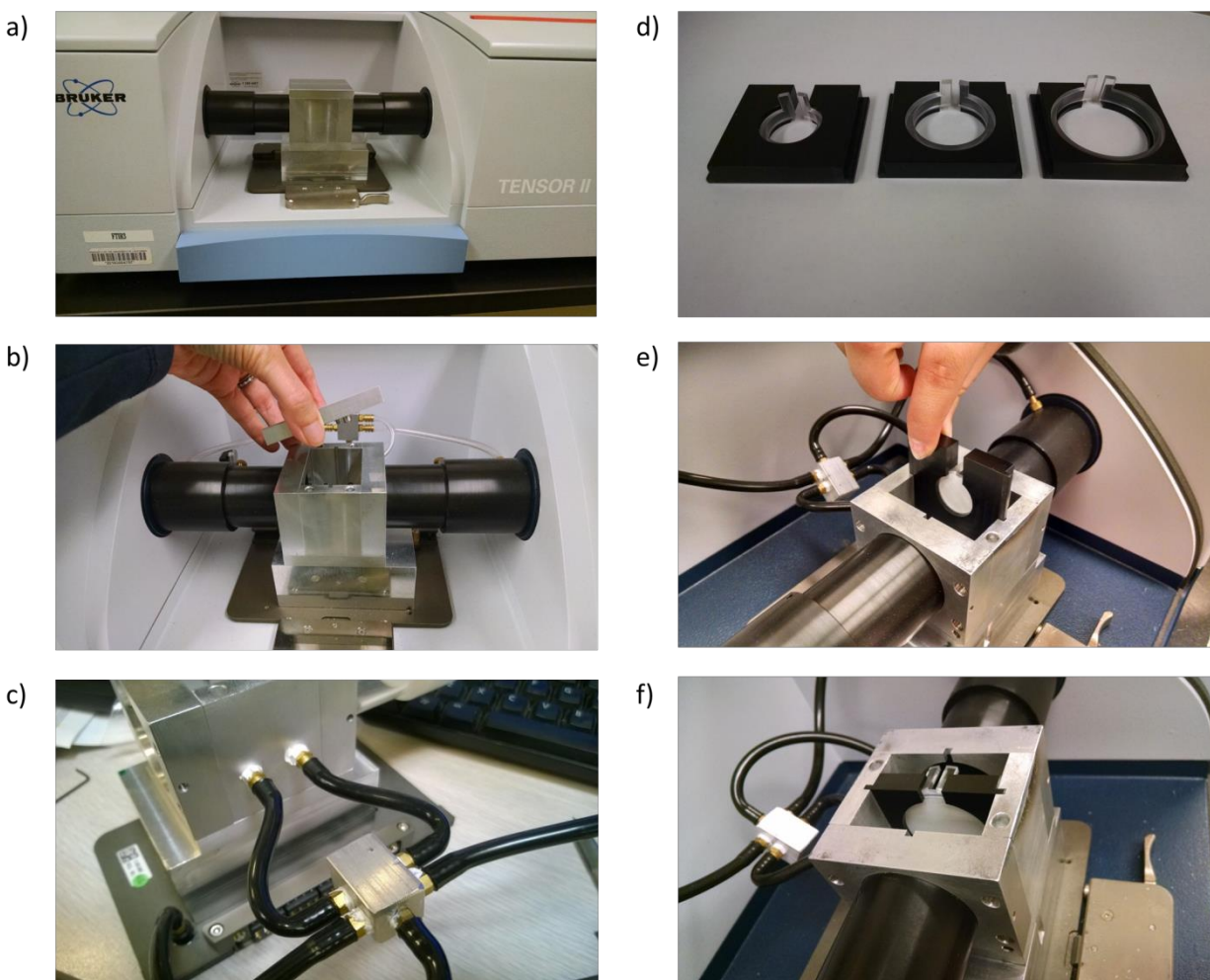


*Figure S3-1. a) Home-built UC Davis sample chamber installed in a Bruker Tensor II instrument, b) chamber lid and support block are made of aluminum while arms are made of black delrin, c) back of the sample chamber with purge lines shown, d) 25 mm, 37 mm and 47 mm filter holders made from black delrin (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 amdillner@ucdavis.edu.

FTIR4831

FTIR4836

FTIR4835

FTIR4833

FTIR4832

FTIR4833

FTIR4832

FTIR4830

FTIR4833
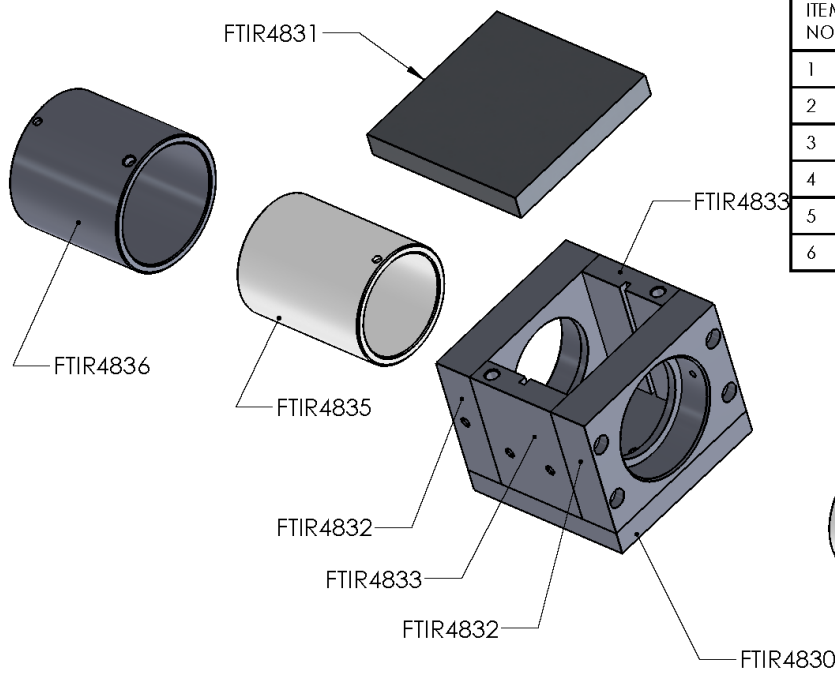
FTIR4835

FTIR4836

| ITEM NO. | PART NUMBER | DESCRIPTION | QTY. |
|---|---|---|---|
| 1 | FTIR4830 | FTIR CHAMBER BASE | 1 |
| 2 | FTIR4832 | FTIR CHAMBER- SIDE PORT | 2 |
| 3 | FTIR4831 | LID- FTIR CHAMBER | 1 |
| 4 | FTIR4833 | FTIR CHAMBER- SLIDE SUPPORT SIDES | 2 |
| 5 | FTIR4836 | FTIR CHAMBER- OUTER SLEEVE | 2 |
| 6 | FTIR4835 | FTIR CHAMBER- INNER SLEEVE | 2 |

| | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER |
|---|---|---|---|---|---|---|
| | | DIMENSIONS ARE IN INCHES TOLERANCES: FRACTIONAL ± ANGULAR: MACH ± BEND ± TWO PLACE DECIMAL ± THREE PLACE DECIMAL ± | DRAWN | | | UNIVERSITY OF CALIFORNIA DAVIS |
| | | | CHECKED | | | TITLE: |
| | | | ENG APPR. | | | ASSEMBLY DRAWING |
| | | | MFG APPR. | | | |
| | | INTERPRET GEOMETRIC TOLERANCING PER: | Q.A. | | | |
| | | | COMMENTS: | | | |
| | | MATERIAL | | | | SIZE DWG. NO. REV |
| NEXT ASSY | USED ON | FINISH | | | | A FTIR4835a |
| APPLICATION | | DO NOT SCALE DRAWING | | | | SCALE: 1:8   WEIGHT:   SHEET 1 OF 1 |

5

2.680

1.680

2X ∅ .125 ▽ .300
∨ ∅ .140 X 90°, NEAR SIDE

#10-32 Tapped Hole

1.340

2.180

∅ .159 ▽ .540
10-32 UNF - 2B ▽ .380
∨ ∅ .240 X 90°, NEAR SIDE

#4-40 Tapped Hole

1.625

2.680

2.000

.265

2.088

2X ∅ .089 ▽ .350
4-40 UNC - 2B ▽ .220
∨ ∅ .125 X 90°, NEAR SIDE

2.680

2.353

3.930

.327

| | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER |
|---|---|---|---|---|---|---|
| | | DIMENSIONS ARE IN INCHES TOLERANCES: FRACTIONAL± 1/64 ANGULAR: MACH± 0.5 BEND ±2 TWO PLACE DECIMAL ±.01 THREE PLACE DECIMAL ±.005 | DRAWN | BMD | 9/26/2014 | UNIVERSITY OF CALIFORNIA DAVIS |
| | | | CHECKED | | | TITLE: |
| | | | ENG APPR. | | | |
| | | | MFG APPR. | | | FTIR CHAMBER-RISER BLOCK |
| | | INTERPRET GEOMETRIC TOLERANCING PER: | Q.A. | | | |
| | | | COMMENTS: | | | |
| | | MATERIAL ALUMINUM 6061 | | | | SIZE **A** | DWG. NO. FTIR4829 | REV |
| NEXT ASSY | USED ON | FINISH √64 | | | | |
| APPLICATION | | DO NOT SCALE DRAWING | | | | SCALE: 1:2 | SHEET 1 OF 1 |

2.750

8X ⌀ .150 THRU ALL
⌴ ⌀ .281 ⤓ .158

⌀ .201 THRU ALL
⌴ ⌀ .375 ⤓ .190

.188

.500

1.375

2.375

1.750

2.750

1.375

.188

.875

.375

1.875

2.563

2X ⌀ .125 THRU ALL
⌵ ⌀ .140 X 90°, NEAR SIDE
⌵ ⌀ .140 X 90°, FAR SIDE

| | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER UNIVERSITY OF CALIFORNIA DAVIS |
|---|---|---|---|---|---|---|
| | | DIMENSIONS ARE IN INCHES TOLERANCES: FRACTIONAL± 1/64 ANGULAR: MACH± 0.5 BEND ±2 TWO PLACE DECIMAL ±.01 THREE PLACE DECIMAL ±.005 | DRAWN | BMD | 9/15/2014 | TITLE: FTIR-CHAMBER BASE |
| | | | CHECKED | | | |
| | | | ENG APPR. | | | |
| | | | MFG APPR. | | | |
| | | INTERPRET GEOMETRIC TOLERANCING PER: | Q.A. | | | |
| | | | COMMENTS: | | | |
| | | MATERIAL ALUMINUM ALLOY 6061 | | | | SIZE **A** | DWG. NO. FTIR4830 | REV |
| NEXT ASSY | USED ON | FINISH √ 32 | | | | |
| APPLICATION | | DO NOT SCALE DRAWING | | | | SCALE: 1:1 | SHEET 1 OF 1 |

7

2X ⌀ .188 ⤓ .25
⌵ ⌀ .22 X 90°, NEAR SIDE

.188

.875

2.563

1.875

2.750

2.750

.375



| | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER |
|---|---|---|---|---|---|---|

UNIVERSITY OF CALIFORNIA DAVIS

DIMENSIONS ARE IN INCHES
TOLERANCES:
FRACTIONAL±
ANGULAR: MACH±    BEND ±
TWO PLACE DECIMAL   ±
THREE PLACE DECIMAL  ±

DRAWN

CHECKED

ENG APPR.

MFG APPR.

Q.A.

COMMENTS:

TITLE:

LID-FTIR
CHAMBER

INTERPRET GEOMETRIC
TOLERANCING PER:

MATERIAL
ALUMINUM 6061

FINISH
√32

NEXT ASSY

USED ON

APPLICATION

DO NOT SCALE DRAWING

SIZE
A

DWG. NO.
FTIR4831

REV

SCALE: 1:1   WEIGHT:    SHEET 1 OF 1

8

2.750

.175

.500

2.250

4X Ø .150 THRU ALL
⌴ Ø .281 ▼ .158

2X Ø .107 ▼ .440
6-32 UNC - 2B ▼ .280
∨ Ø .150 X 90°, NEAR SIDE

.625

1.000

.250

2.000

.188

A

A

Ø 1.500

.500

1.500

2.563

Ø 1.750

.107 THRU
6-32 UNC - 2B THRU
∨ Ø .150 X 90°, NEAR SIDE

.500

.030 X 45.00°

SECTION A-A

UNLESS OTHERWISE SPECIFIED:

DIMENSIONS ARE IN INCHES
TOLERANCES:
FRACTIONAL± 1/64
ANGULAR: MACH± 0.5 BEND ±2
TWO PLACE DECIMAL ± .01
THREE PLACE DECIMAL ± .005

INTERPRET GEOMETRIC
TOLERANCING PER:

MATERIAL
ALUMINUM ALLOY 6061

FINISH
√32

NEXT ASSY          USED ON

APPLICATION

DO NOT SCALE DRAWING

| | NAME | DATE |
|---|---|---|
| DRAWN | BMD | 9/16/2014 |
| CHECKED | | |
| ENG APPR. | | |
| MFG APPR. | | |
| Q.A. | | |
| COMMENTS: | | |

AIR QUALITY RESEARCH CENTER
UNIVERSITY OF CALIFORNIA DAVIS

TITLE:

FTIR CHAMBER-SIDE
PORT

SIZE **A**

DWG. NO.
FTIR4832

REV

SCALE: 1:1          SHEET 1 OF 1

2X ⌀ .107 ▽ .340
6-32 UNC - 2B ▽ .280
∨ ⌀ .150 X 90°, NEAR SIDE

.188

.250

1.250

1.500

.188

.500

.375

.750

.500

1.500

1.125

2X ⌀ .107 ▽ .440
6-32 UNC - 2B ▽ .280
∨ ⌀ .150 X 90°, NEAR SIDE

.100

.130

.250

⌀ .188 ▽ .250
∨ ⌀ .215 X 90°, NEAR SIDE

.188

.188

1.500

2.000

2X ⌀ .107 ▽ .440
6-32 UNC - 2B ▽ .280
∨ ⌀ .150 X 90°, NEAR SIDE

.375

2X ⌀ .107 THRU ALL
6-32 UNC - 2B THRU ALL
∨ ⌀ .150 X 90°, NEAR SIDE

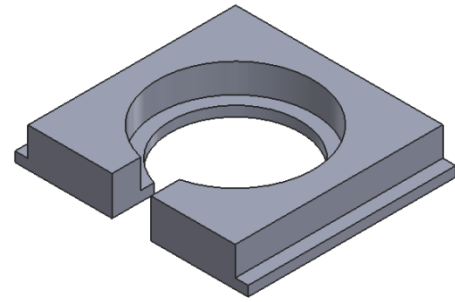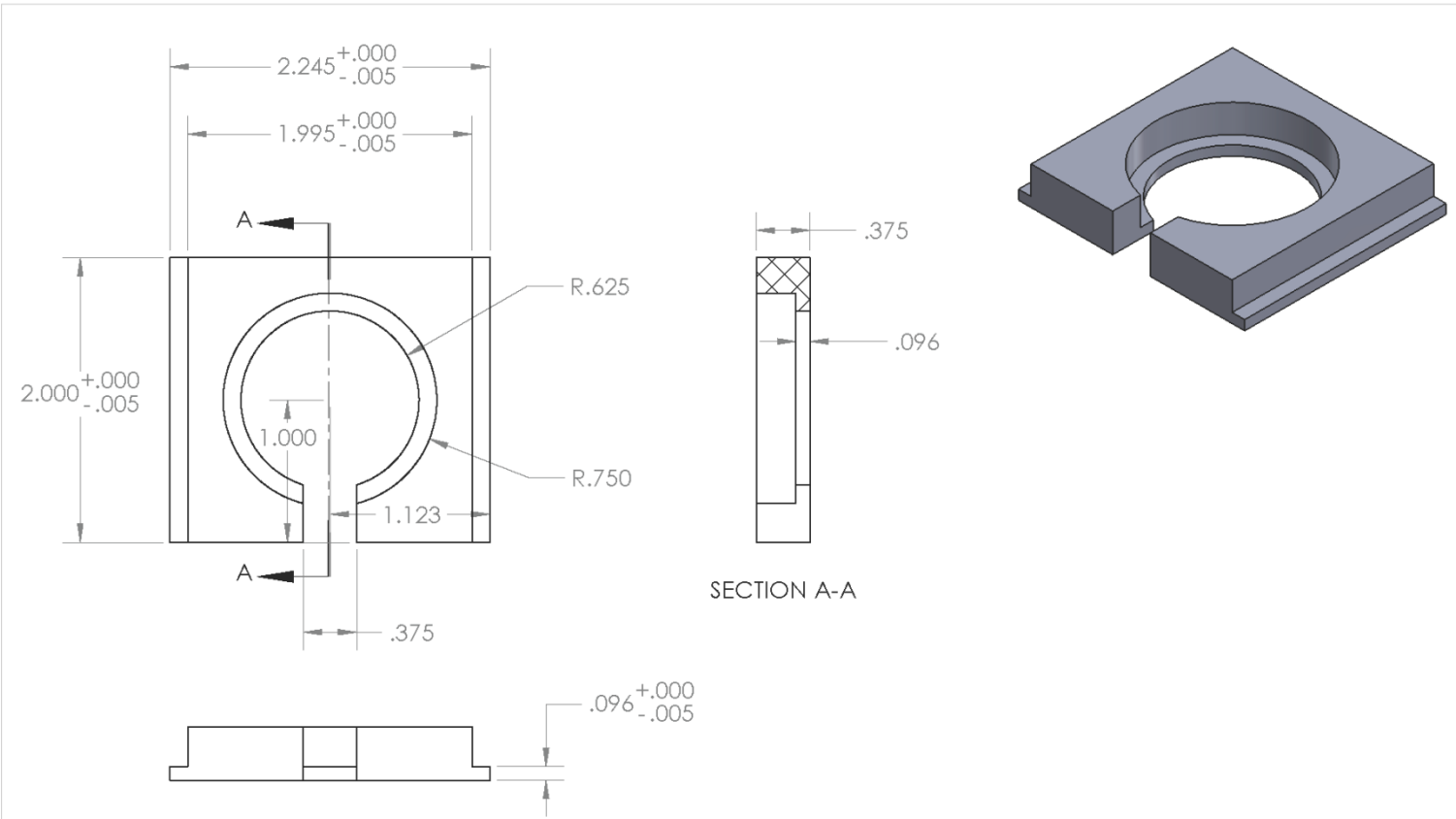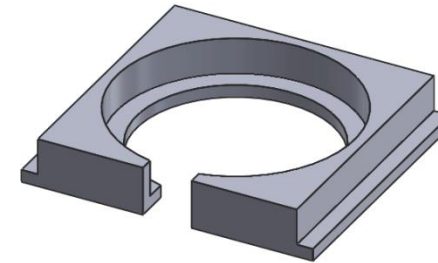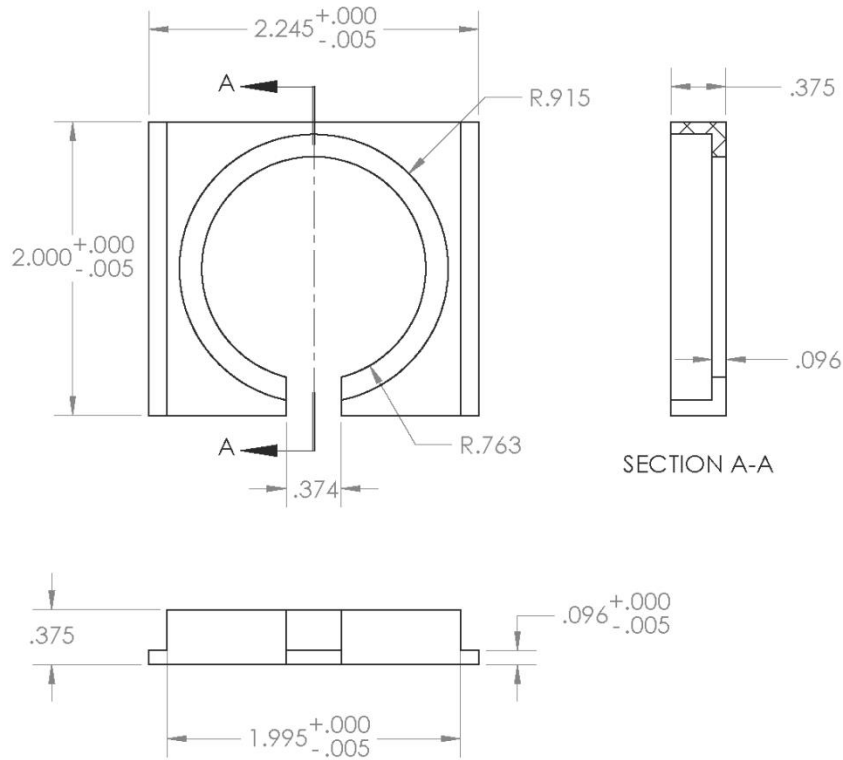| | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER |
| | | | | | | UNIVERSITY OF CALIFORNIA DAVIS |
| | | DIMENSIONS ARE IN INCHES | DRAWN | BMD | 9/16/2014 | TITLE: |
| | | TOLERANCES: | CHECKED | | | |
| | | FRACTIONAL± 1/64 | | | | FTIR CHAMBER- |
| | | ANGULAR: MACH± 0.5 BEND ±2 | ENG APPR. | | | SLIDE SUPPORT |
| | | TWO PLACE DECIMAL ±.01 | | | | SIDES |
| | | THREE PLACE DECIMAL ±.005 | MFG APPR. | | | |
| | | | Q.A. | | | |
| | | INTERPRET GEOMETRIC | | | | |
| | | TOLERANCING PER: | COMMENTS: | | | |
| **PROPRIETARY AND CONFIDENTIAL** | | MATERIAL | | | | |
| The information contained in this drawing is the sole property of the University of California, Davis. ANY REPRODUCTION IN PART OR AS A WHOLE WITHOUT THE WRITTEN PERMISSION OF the University of California, Davis IS PROHIBITED. Contact A. Dillner at amdillner@ucdavis.edu. | | ALUMINUM ALLOY 6061 | | | SIZE | DWG. NO. | REV |
| | | FINISH | | | | |
| | NEXT ASSY | USED ON | √32 | | | A | FTIR4833 | |
| | APPLICATION | | DO NOT SCALE DRAWING | | | SCALE: 1:1 | SHEET 1 OF 1 |

.020 X 45.00°

2.250

.125

.020 X 45.00°

$\emptyset$ .125 $\overline{\underline{\vee}}$ .062

.020 X 45.00°

.020 X 45.00°

SECTION A-A

A
A

$\emptyset\ 1.500\ ^{+.005}_{-.005}$

$\emptyset\ 1.750\ ^{+.000}_{-.005}$

| | | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER UNIVERSITY OF CALIFORNIA DAVIS |
|---|---|---|---|---|---|---|---|
| | | | DIMENSIONS ARE IN INCHES TOLERANCES: FRACTIONAL± 1/64 ANGULAR: MACH± 0.5 BEND ±2 TWO PLACE DECIMAL ±.01 THREE PLACE DECIMAL ±.005 | DRAWN | BMD | 9/29/2014 | TITLE: |
| | | | | CHECKED | | | FTIR CHAMBER- INNER SLEEVE |
| | | | | ENG APPR. | | | |
| | | | | MFG APPR. | | | |
| | | | INTERPRET GEOMETRIC TOLERANCING PER: | Q.A. | | | |
| | | | | COMMENTS: | | | |
| | | | MATERIAL BLACK DELRIN | | | | SIZE **A** | DWG. NO. FTIR4835 | REV |
| NEXT ASSY | USED ON | | FINISH $\sqrt{32}$ | | | | |
| | APPLICATION | | DO NOT SCALE DRAWING | | | | SCALE: 1:1 | SHEET 1 OF 1 |

11

$\varnothing$ .125 $\overline{\vee}$ .060

.020 X 45.00°

$\varnothing$ .159 THRU
10-32 UNF - 2B THRU
$\vee$ $\varnothing$ .190 X 90°, NEAR SIDE

2.000

1.875

.175

.020 X 45.00°

.020 X 45.00°

.020 X 45.00°

SECTION A-A

A

A

$\varnothing$ 2.000 $^{+.005}_{-.005}$

$\varnothing$ 1.750 $^{+.005}_{-.000}$

| | | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER |
|---|---|---|---|---|---|---|---|
| | | | DIMENSIONS ARE IN INCHES | DRAWN | BMD | 9/29/2014 | UNIVERSITY OF CALIFORNIA DAVIS |
| | | | TOLERANCES: FRACTIONAL± 1/64 | CHECKED | | | TITLE: |
| | | | ANGULAR: MACH± 0.5 BEND ±2 TWO PLACE DECIMAL ±.01 | ENG APPR. | | | |
| | | | THREE PLACE DECIMAL ±.005 | MFG APPR. | | | FTIR CHAMBER- OUTER SLEEVE |
| | | | INTERPRET GEOMETRIC TOLERANCING PER: | Q.A. | | | |
| | | | | COMMENTS: | | | |
| | | | MATERIAL BLACK DELRIN | | | | SIZE **A** DWG. NO. FTIR4836 REV |
| NEXT ASSY | USED ON | | FINISH $\sqrt{32}$ | | | | |
| APPLICATION | | | DO NOT SCALE DRAWING | | | | SCALE: 1:1    SHEET 1 OF 1 |

2.245 +.000 / -.005

1.995 +.000 / -.005

A

2.000 +.000 / -.005

R.438

R.503

A

.375

.375

.096

SECTION A-A

.096 +.000 / -.005

.375

UNLESS OTHERWISE SPECIFIED:

DIMENSIONS ARE IN INCHES
TOLERANCES:
FRACTIONAL± 1/64
ANGULAR: MACH± 0.5 BEND ±2
TWO PLACE DECIMAL ±.01
THREE PLACE DECIMAL ±.005

INTERPRET GEOMETRIC
TOLERANCING PER:

MATERIAL
BLACK DELRIN

FINISH
√32

NEXT ASSY | USED ON

APPLICATION

DO NOT SCALE DRAWING

| | NAME | DATE |
|---|---|---|
| DRAWN | BMD | 11-5-2014 |
| CHECKED | | |
| ENG APPR. | | |
| MFG APPR. | | |
| Q.A. | | |
| COMMENTS: | | |

AIR QUALITY RESEARCH CENTER
UNIVERSITY OF CALIFORNIA DAVIS

TITLE:
FTIR-
25MM FILTER SUPPORT

SIZE **A**

DWG. NO.
FTIR4837

REV

SCALE: 1:1

SHEET 1 OF 1

2.245 +.000 / -.005

1.995 +.000 / -.005

A

R.625

2.000 +.000 / -.005

1.000

1.123

A

R.750

.375

.375

.096

SECTION A-A

.096 +.000 / -.005

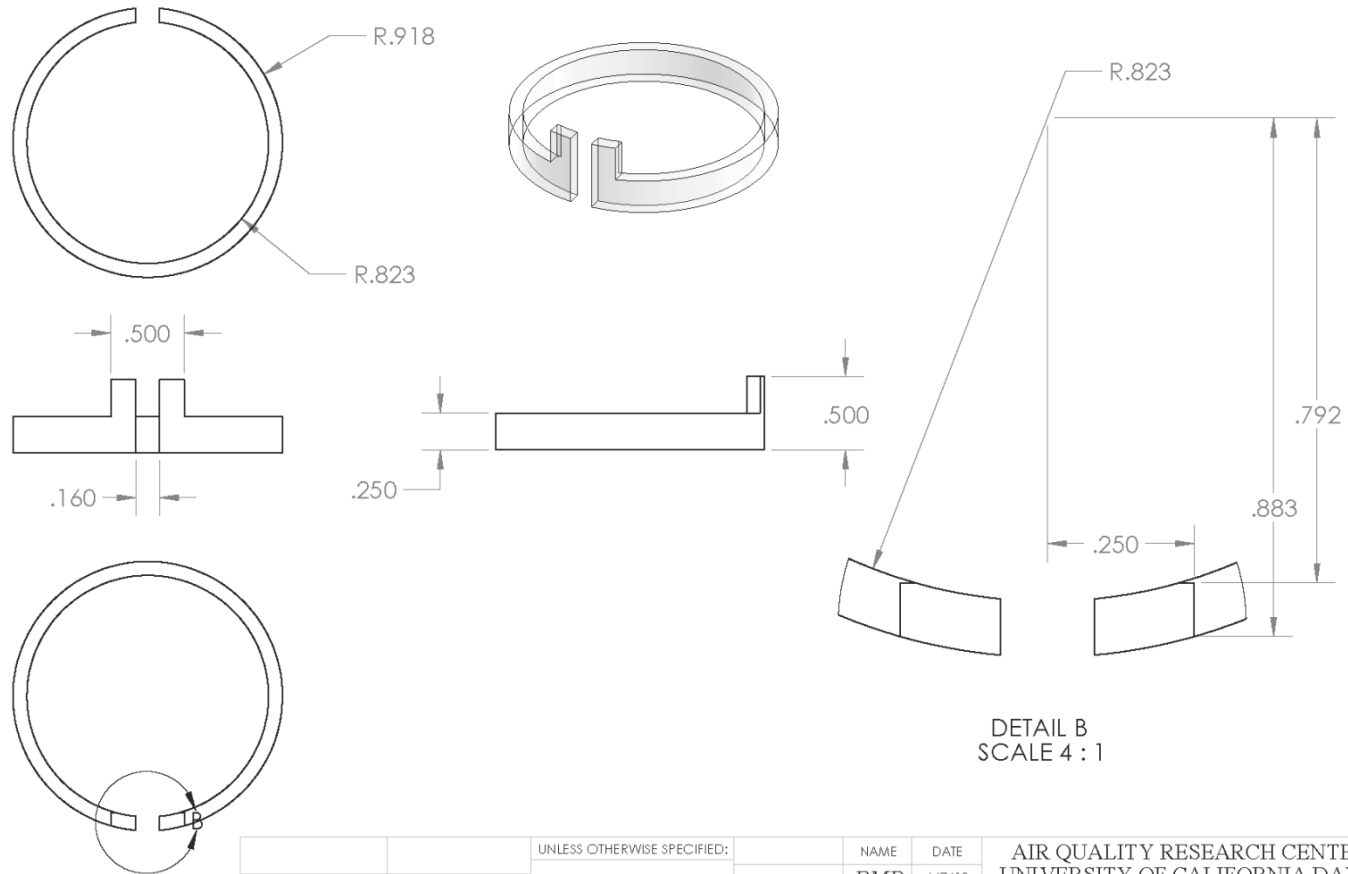| | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER UNIVERSITY OF CALIFORNIA DAVIS |
|---|---|---|---|---|---|---|
| | | DIMENSIONS ARE IN INCHES TOLERANCES: FRACTIONAL± 1/64 ANGULAR: MACH± 0.5 BEND ±2 TWO PLACE DECIMAL ±.01 THREE PLACE DECIMAL ±.005 | DRAWN | BMD | 12/18/2014 | TITLE: |
| | | | CHECKED | | | FTIR- 37MM FILTER SUPPORT |
| | | | ENG APPR. | | | |
| | | | MFG APPR. | | | |
| | | INTERPRET GEOMETRIC TOLERANCING PER: | Q.A. | | | |
| | | | COMMENTS: | | | |
| | | MATERIAL DELRIN-BLACK | | | | SIZE A | DWG. NO. FTIR4838 | REV |
| NEXT ASSY | USED ON | FINISH √32 | | | | |
| APPLICATION | | DO NOT SCALE DRAWING | | | | SCALE: 1:1 | WEIGHT: | SHEET 1 OF 1 |

2.245 +.000 / -.005

R.915

2.000 +.000 / -.005

.374

R.763

A — A

.375

.096

SECTION A-A

.375

.096 +.000 / -.005

1.995 +.000 / -.005

| | | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER |
|---|---|---|---|---|---|---|---|
| | | | DIMENSIONS ARE IN INCHES TOLERANCES: | DRAWN | BMD | | UNIVERSITY OF CALIFORNIA DAVIS |
| | | | FRACTIONAL± 1/64 | CHECKED | | | TITLE: |
| | | | ANGULAR: MACH± 0.5 BEND ±2 | ENG APPR. | | | |
| | | | TWO PLACE DECIMAL ± .01 | | | | FTIR |
| | | | THREE PLACE DECIMAL ± .005 | MFG APPR. | | | 47 MM FILTER SUPPORT |
| | | | INTERPRET GEOMETRIC TOLERANCING PER: | Q.A. | | | |
| | | | | COMMENTS: | | | |
| | | | MATERIAL Black Delrin | | | | SIZE A DWG. NO. FTIR4827 REV |
| NEXT ASSY | USED ON | | FINISH | | | | |
| | APPLICATION | | DO NOT SCALE DRAWING | | | | SCALE: 1:1 WEIGHT: SHEET 1 OF 1 |

Ø .838    Ø 1.014

.364

.495

.180

1/2"

.200

| UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER |
|---|---|---|---|---|
| DIMENSIONS ARE IN INCHES | DRAWN | MV | 8/21/2017 | UNIVERSITY OF CALIFORNIA DAVIS |
| TOLERANCES: | CHECKED | | | TITLE: |
| FRACTIONAL± 1/64 | ENG APPR. | | | |
| ANGLE: MACH± 0.5° BEND ± 2° | MFG APPR. | | | FTIR-FILTER HOLDER |
| TWO PLACE DECIMAL  ± .01 | | | | 25mm CIRCLIP |
| THREE PLACE DECIMAL  ± .005 | Q.A. | | | |
| MATERIAL  ACRYLIC | COMMENTS: | | | |

THIRD ANGLE
PROJECTION

DO NOT SCALE DRAWING

| SIZE | DWG. NO. | | REV. |
|---|---|---|---|
| A | FTIR4840 | | |
| SCALE:2:1 | WEIGHT: 0.00 | SHEET 1 OF 1 | |

R.625

R.753

R.625

.590

.710

.500

.160

DETAIL A
SCALE 4 : 1

.500

.250

A

UNLESS OTHERWISE SPECIFIED:

DIMENSIONS ARE IN INCHES
TOLERANCES:
FRACTIONAL± 1/64
ANGULAR: MACH± 0.5 BEND ±2
TWO PLACE DECIMAL ±.01
THREE PLACE DECIMAL ±.005

INTERPRET GEOMETRIC
TOLERANCING PER:

MATERIAL
CLEAR ACRYLIC

FINISH
√32

NEXT ASSY | USED ON

APPLICATION

DO NOT SCALE DRAWING

| | NAME | DATE |
|---|---|---|
| DRAWN | BMD | 12/18/2014 |
| CHECKED | | |
| ENG APPR. | | |
| MFG APPR. | | |
| Q.A. | | |
| COMMENTS: | | |

AIR QUALITY RESEARCH CENTER
UNIVERSITY OF CALIFORNIA DAVIS

TITLE:

FTIR- 37MM FILTER
SUPPORT RETAINING
RING

SIZE **A** | DWG. NO. FTIR4839 | REV

SCALE: 2:1 | WEIGHT: | SHEET 1 OF 1

R.918

R.823

.500

.160

.250

.500

R.823

.792

.250

.883

DETAIL B
SCALE 4 : 1

B

| | | UNLESS OTHERWISE SPECIFIED: | | NAME | DATE | AIR QUALITY RESEARCH CENTER |
|---|---|---|---|---|---|---|
| | | DIMENSIONS ARE IN INCHES TOLERANCES: FRACTIONAL± 1/64 ANGULAR: MACH± 0.5 BEND ±2 TWO PLACE DECIMAL ±.01 THREE PLACE DECIMAL ±.005 | DRAWN | BMD | 6/7/13 | UNIVERSITY OF CALIFORNIA DAVIS |
| | | | CHECKED | | | TITLE: |
| | | | ENG APPR. | | | FTIR |
| | | | MFG APPR. | | | 47mm FILTER SUPPORT |
| | | INTERPRET GEOMETRIC TOLERANCING PER: | Q.A. | | | RETAINING RING |
| | | | COMMENTS: | | | |
| | | MATERIAL ACRYLIC | | | | SIZE A / DWG. NO. FTIR4828 / REV |
| NEXT ASSY | USED ON | FINISH 32 | | | | |
| | APPLICATION | DO NOT SCALE DRAWING | | | | SCALE: 1:1 — SHEET 1 OF 1 |

18

**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.
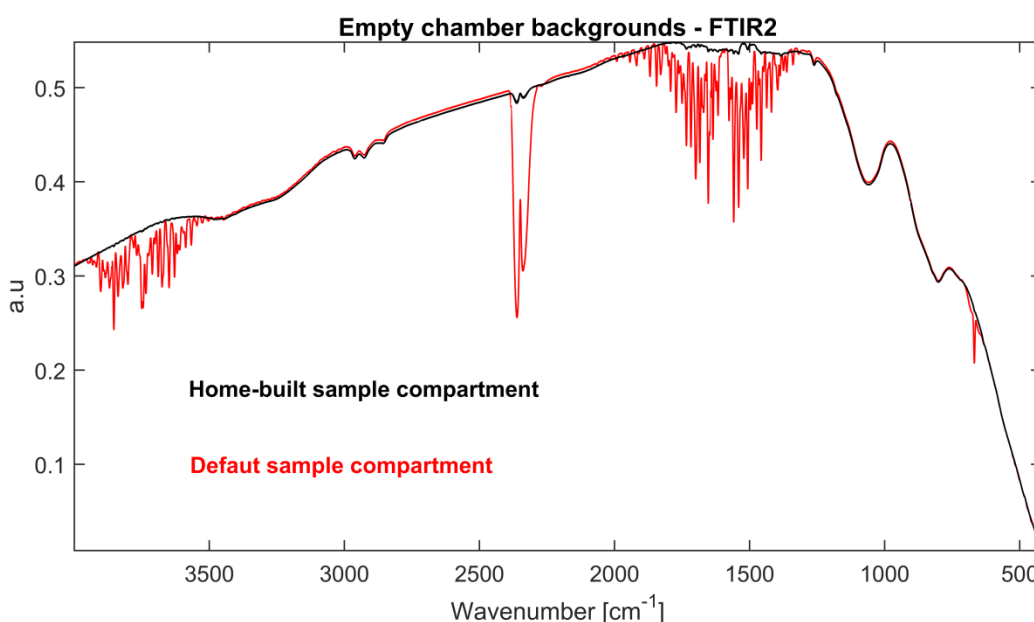


*Figure S4-1. Sets of 48 empty chamber backgrounds collected on FTIR2 with and without the home-built sample compartment during the repeated measurement procedure. Backgrounds are acquired by a liquid nitrogen-cooled wide-band mercury cadmium telluride detector, identical for all three spectrometers, characterized by a detectivity D\*(500, 1000, 1) of 1.10¹⁰ cm Hz¹ᐟ² W⁻¹ 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):

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

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.
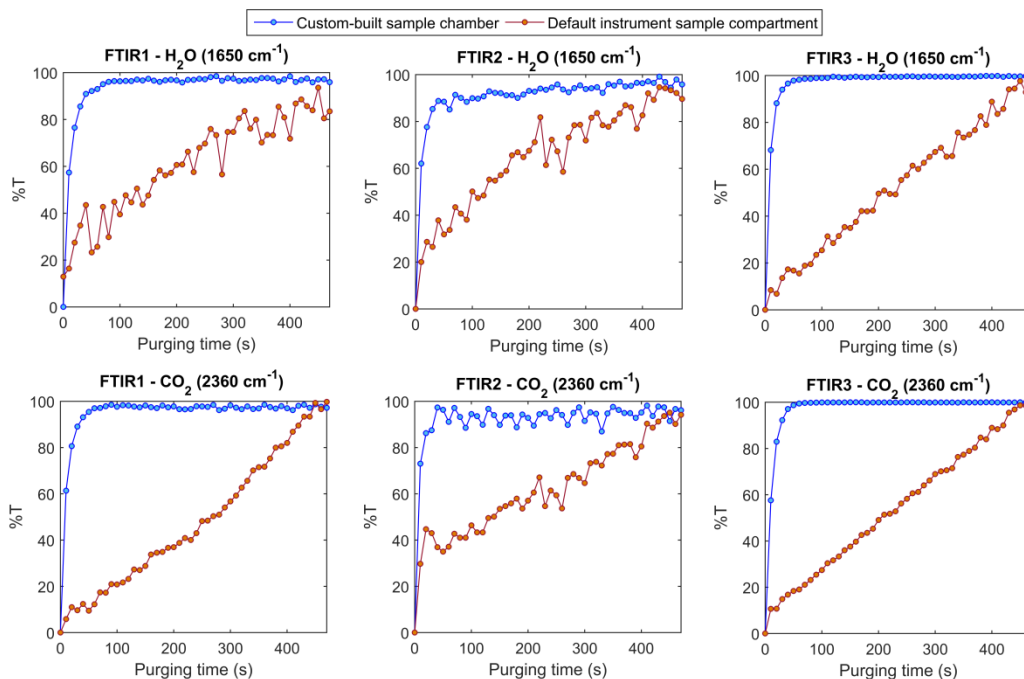


***Figure S4-2. Comparison of time-dependent transmittance for water vapor (1650 cm$^{-1}$) and carbon dioxide (2360 cm$^{-1}$) peaks as function of purge time on all three spectrometers.***

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 2<sup>nd</sup> derivative spectra of the Wichita Mountains (OK) ambient sample**
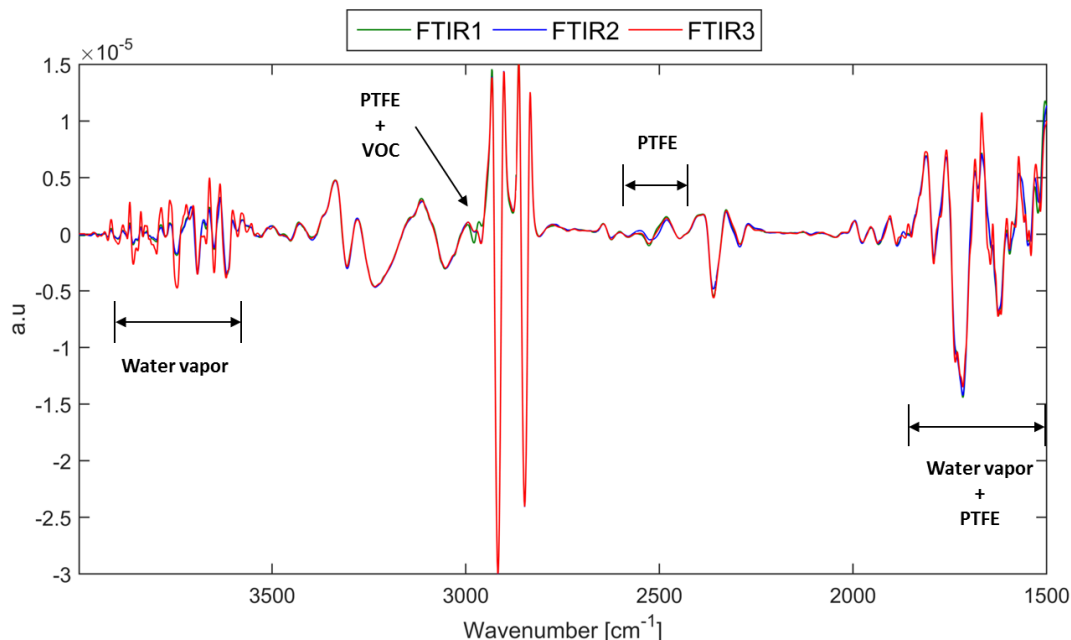


*Figure S5. Spectra of the Wichita Mountains (OK) ambient sample on all three instruments obtained after computing their second derivative. Arrows indicate spectral regions featuring significant spectral differences between instruments. The region below 1500 cm$^{-1}$, moslty comprised of PTFE absorbance, 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.

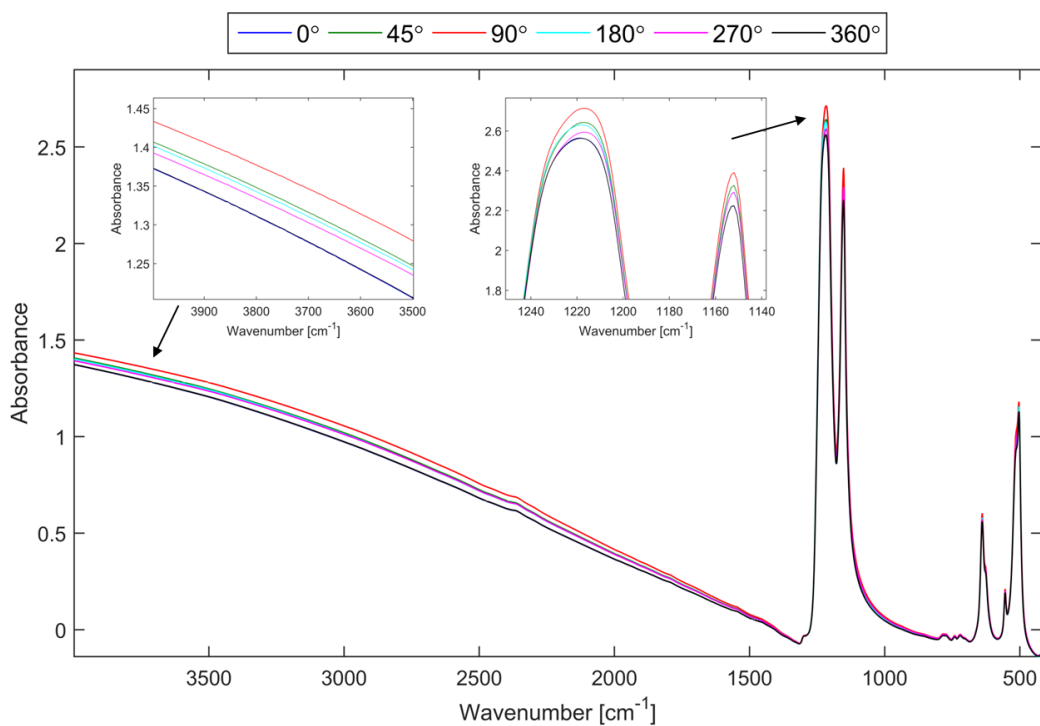**Section 6: Anisotropic optical properties of PTFE filters**



*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.

**Section 7: OC and EC models performance as function of the number of calibration samples**
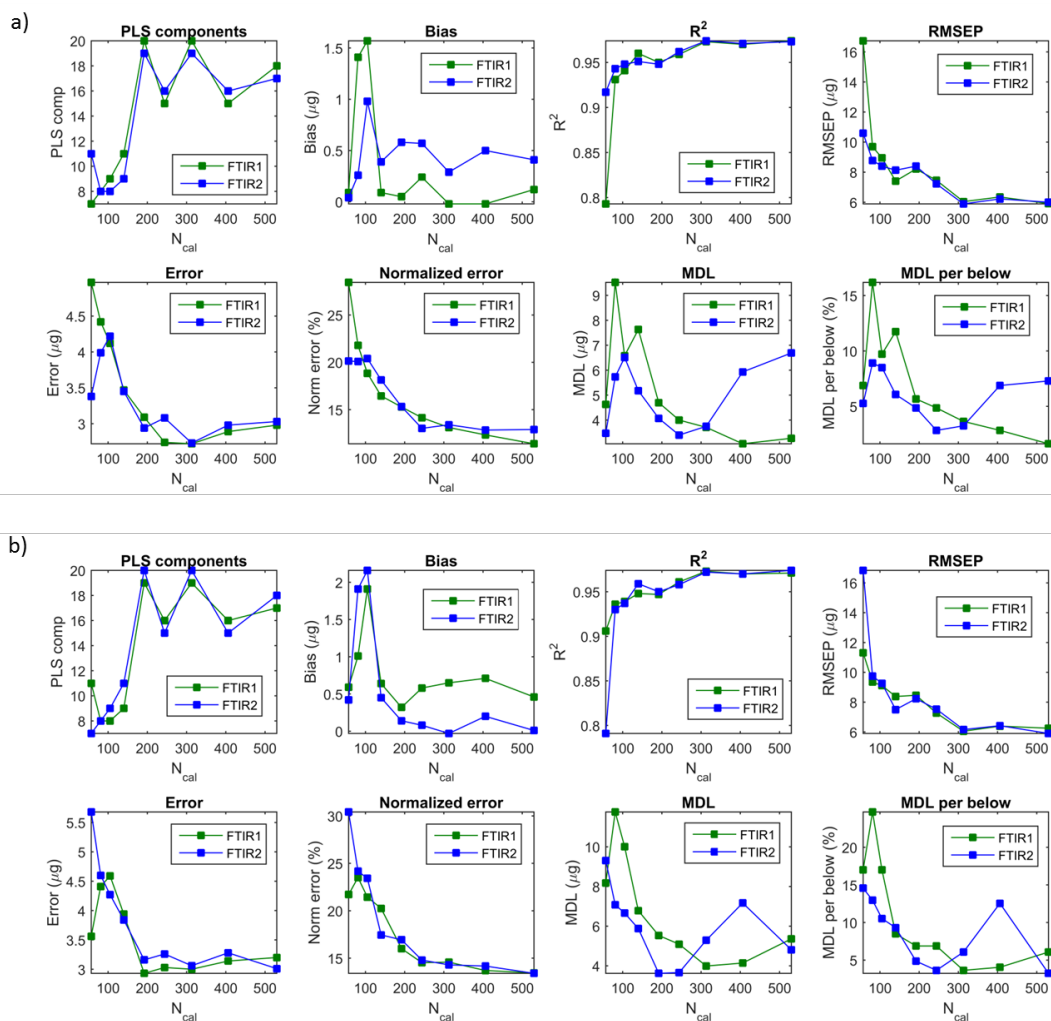


*Figure S7-1. Variation in OC prediction metrics for the first (a) and second (b) scenarios as function of the number of calibration samples (N$_{cal}$) included in the model. In scenario (i), a calibration developed with spectra from one instrument are used to predict test samples from the same instrument 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 (≈ 0.5 μg) 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.*
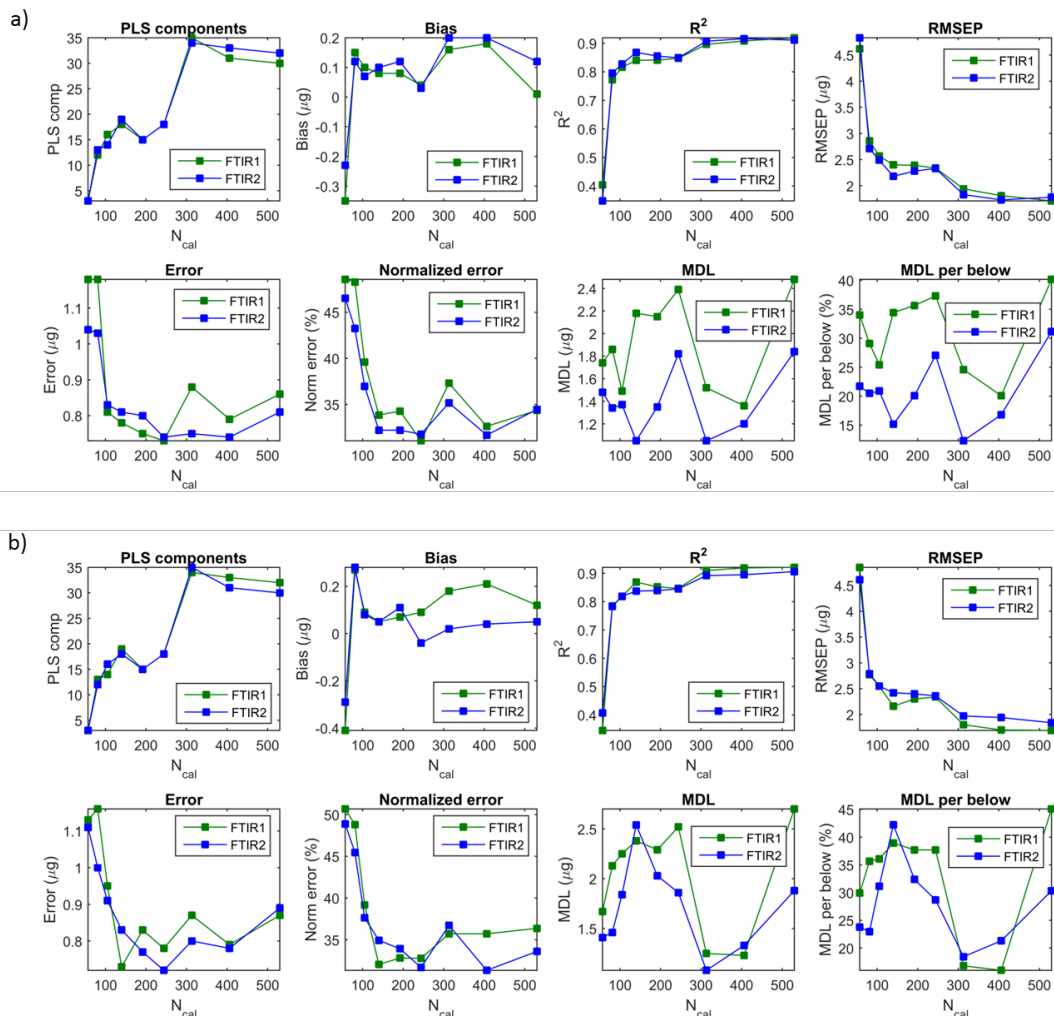
*Figure S7-2. Variation in EC prediction metrics for the first (a) and second (b) scenarios as function of the number of calibration samples ($N_{cal}$) included in the model. Even though increasing the number of calibration samples yield better model performance, no significant changes are observed 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.*
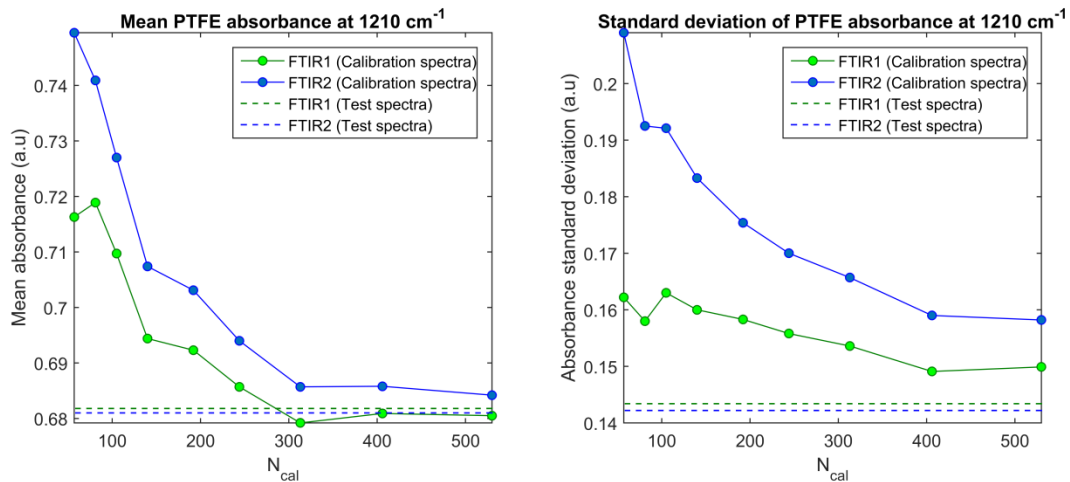
*Figure S7-3. Evolution in PTFE mean absorbance and standard deviation between calibration and test spectra as function of $N_{cal}$. Distribution parameters are derived from the 1210 $cm^{-1}$ PTFE peak for FTIR1 and FTIR2, respectively. Dotted lines correspond to the PTFE distribution parameters in 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.*

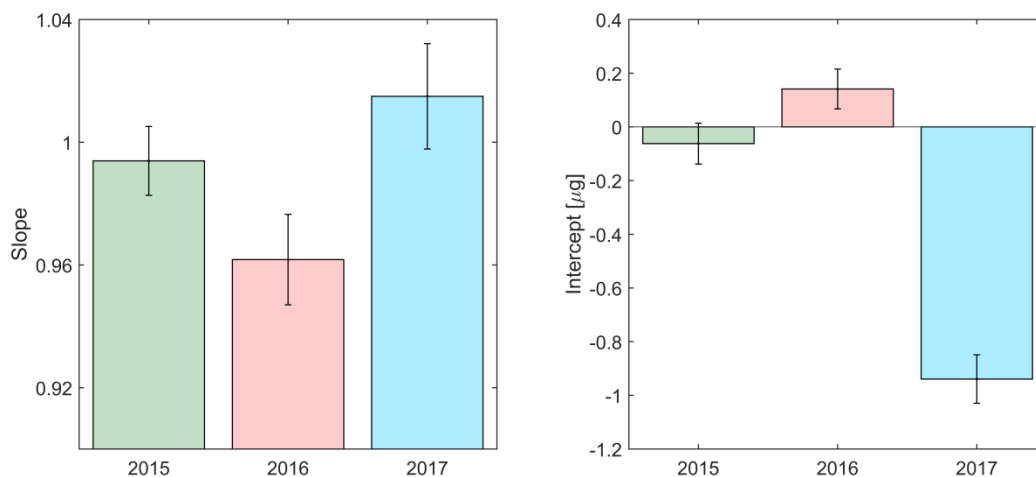**Section 8: Performance of the hybrid models extended to the 2016 and 2017 ambient samples**



*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.*
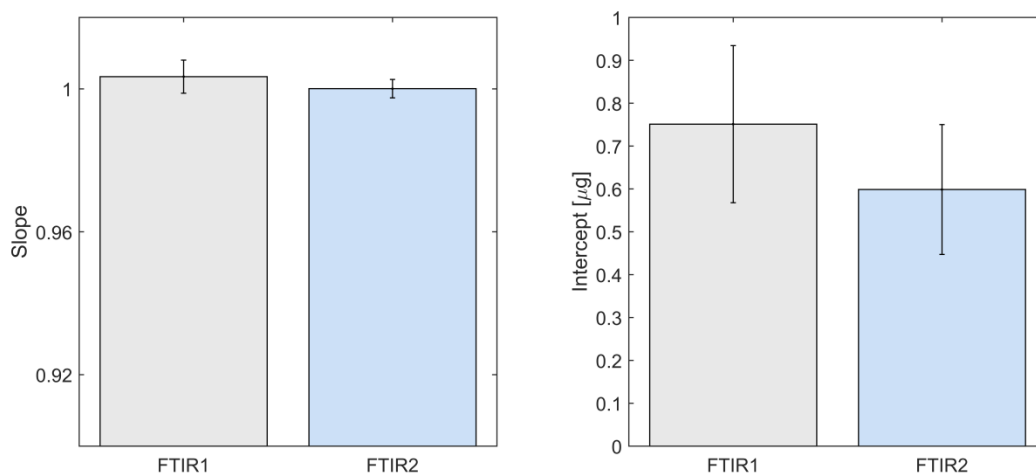


*Figure S8-2. Time-dependent slopes (left) and intercepts (right) derived from test set OC predicted concentrations on the FTIR3 – FTIR1 and FTIR3 – FTIR2 instrument-pairs during 2017, uncertainty of each metric is evaluated by developing 95 % confidence intervals from 1000 bootstrap replicates. The 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.*
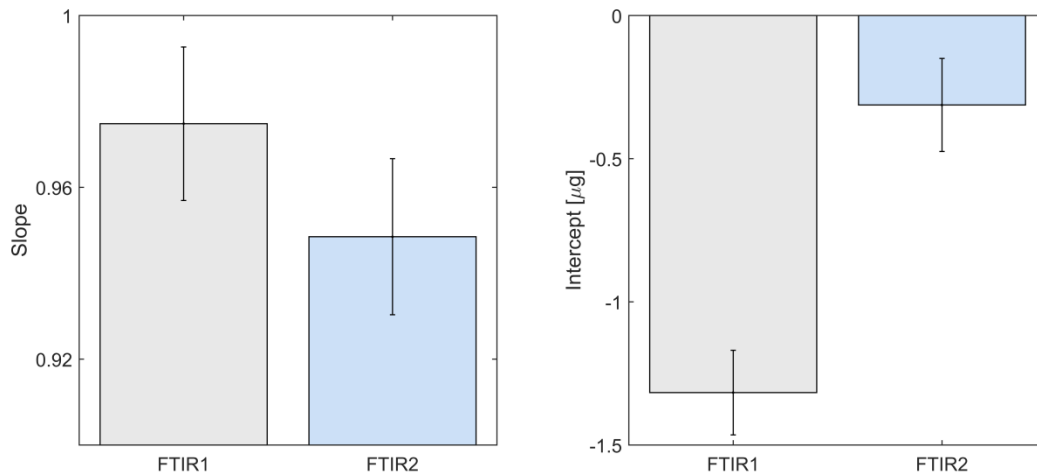
*Figure S8-3. Time-dependent slopes (left) and intercepts (right) derived from test set EC predicted concentrations on the FTIR3 – FTIR1 and FTIR3 – FTIR2 instrument-pairs, uncertainty of each metric is evaluated by developing 95 % confidence intervals from 1000 bootstrap replicates. The larger intercept 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.*