Atmospheric
Measurement
Techniques

Predicting ambient aerosol thermal-optical reflectance (TOR) measurements from infrared spectra: organic carbon

A. M. Dillner[1] and S. Takahama[2]

[1]University of California, Davis, Davis, California, USA
[2]Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland

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

Abstract. Organic carbon (OC) can constitute 50 % or more of the mass of atmospheric particulate matter. Typically, organic carbon is measured from a quartz fiber filter that has been exposed to a volume of ambient air and analyzed using thermal methods such as thermal-optical reflectance (TOR). Here, methods are presented that show the feasibility of using Fourier transform infrared (FT-IR) absorbance spectra from polytetrafluoroethylene (PTFE or Teflon) filters to accurately predict TOR OC. This work marks an initial step in proposing a method that can reduce the operating costs of large air quality monitoring networks with an inexpensive, non-destructive analysis technique using routinely collected PTFE filter samples which, in addition to OC concentrations, can concurrently provide information regarding the composition of organic aerosol. This feasibility study suggests that the minimum detection limit and errors (or uncertainty) of FT-IR predictions are on par with TOR OC such that evaluation of long-term trends and epidemiological studies would not be significantly impacted. To develop and test the method, FT-IR absorbance spectra are obtained from 794 samples from seven Interagency Monitoring of PROtected Visual Environment (IMPROVE) sites collected during 2011. Partial least-squares regression is used to calibrate sample FT-IR absorbance spectra to TOR OC. The FTIR spectra are divided into calibration and test sets by sampling site and date. The calibration produces precise and accurate TOR OC predictions of the test set samples by FT-IR as indicated by high coefficient of variation (R^2; 0.96), low bias (0.02 µg m^{-3}, the nominal IMPROVE sample volume is 32.8 m^3), low error (0.08 µg m^{-3}) and low normalized error (11 %). These performance metrics can be achieved with various degrees of spectral pretreatment (e.g., including or excluding substrate contributions to the absorbances) and are comparable in precision to collocated TOR measurements. FT-IR spectra are also divided into calibration and test sets by OC mass and by OM / OC ratio, which reflects the organic composition of the particulate matter and is obtained from organic functional group composition; these divisions also leads to precise and accurate OC predictions. Low OC concentrations have higher bias and normalized error due to TOR analytical errors and artifact-correction errors, not due to the range of OC mass of the samples in the calibration set. However, samples with low OC mass can be used to predict samples with high OC mass, indicating that the calibration is linear. Using samples in the calibration set that have different OM / OC or ammonium / OC distributions than the test set leads to only a modest increase in bias and normalized error in the predicted samples. We conclude that FT-IR analysis with partial least-squares regression is a robust method for accurately predicting TOR OC in IMPROVE network samples – providing complementary information to the organic functional group composition and organic aerosol mass estimated previously from the same set of sample spectra (Ruthenburg et al., 2014).

1 Introduction

Particulate matter (PM) has been implicated in increased morbidity and mortality (Anderson et al., 2012), climate change (Yu et al., 2006) and reduced visibility (Watson, 2002). As a result, its size-resolved chemical composition is measured during episodic measurement campaigns and over longer periods of time in many networks worldwide, includ-

ing the Interagency Monitoring of PROtected Visual Environment (IMPROVE) network (Hand et al., 2012; Malm et al., 1994) in pristine and rural areas in the US, the Chemical Speciation Network/Speciation Trends Network (CSN/STN; Flanagan et al., 2006) in urban and suburban areas in the US, the Southeastern Aerosol Research and Characterization network (SEARCH; Hansen et al., 2003) in urban and rural areas in the southeastern US, the Canadian National Air Pollution Surveillance network (NAPS; Dabek-Zlotorzynska et al., 2011) in primarily urban sites in Canada and the European Monitoring and Evaluation Programme (EMEP; Tørseth et al., 2012) throughout Europe. Typically, organic carbon (OC) and elemental carbon (EC) concentrations are measured on quartz filters using thermal-optical reflectance (TOR; Chow et al., 2007), NIOSH 5040 (Birch and Cary, 1996), European Supersites for Atmospheric Aerosol Research protocol (EUSAAR-2; Cavalli et al., 2010) or similar methods. PM is collected on a quartz filter, and a portion of the filter is subjected to a temperature gradient with two carrier gas regimes that operationally define the organic and elemental carbon (Chow et al., 2007). Charring of organic material during heating is corrected for by using laser reflectance or transmittance (Cavalli et al., 2010; Chow et al., 2007). The measurement artifact caused by gas phase adsorption of organic material on the quartz filter may be corrected for by using blank or back-up quartz filters (Chow et al., 2010; Maimone et al., 2011; Turpin et al., 1994). Organic matter (OM) is estimated by multiplying the reported OC by an assumed OM / OC factor (Pitchford et al., 2007; Turpin and Lim, 2001).

Fourier transform infrared spectroscopy (FT-IR) has been proposed as an alternative for quantification of organic matter in particles collected on filters (Russell, 2003; Ruthenburg et al., 2014). FT-IR measures abundances of bonds connecting carbon atoms with their heteroatoms, leading to characterization of functional groups including aliphatic and aromatic CH, carbonyl (C = O), alcohol OH (C-OH), carboxylic acid OH (C-OH) and others (Blando et al., 2001; Coury and Dillner, 2008; Maria et al., 2003). This bond abundance allows more direct estimates of OM and OM / OC ratios (Russell, 2003; Ruthenburg et al., 2014) compared to using TOR OC and an assumed OM / OC ratio. Organic functional groups in carbonaceous material absorb IR light in (a) specific region(s) of the mid-IR spectrum (4000 to 400 cm^{-1}). The amount of light absorbed is proportional to the moles of functional group. Based on initial work by Allen and colleagues (Allen et al., 1994), researchers (Coury and Dillner, 2008; Reff et al., 2007; Russell et al., 2009; Ruthenburg et al., 2014; Takahama et al., 2013) have shown that organic functional groups can be quantified even in complex mixtures of ambient or indoor aerosols. These studies use laboratory-generated standards as reference material to develop calibration models for quantifying functional group abundance which can be used to calculate OC and OM.

Researchers in other fields have used FT-IR spectra to quantify properties such as total carbon (TC), organic carbon or fatty acid content using calibrations developed from environmental (e.g., soil or food) reference samples. These environmental samples were analyzed by FT-IR alongside an expensive or time-consuming conventional method to measure the property of interest. Partial least-squares regression (PLSR) has been commonly used to develop calibration models that quantitatively predict these properties from the FT-IR spectra. In one example of this approach in the field of soil science (Madari et al., 2005), calibrations were developed for total carbon and organic carbon in soil samples using near-infrared (NIR) and diffuse reflectance mid-infrared spectroscopy (DRIFTS). Over 1000 samples from the Brazilian National Soil Collection were analyzed by a combustion method to determine TC and by a chromate oxidation method to determine OC. Calibrations of DRIFTS spectra developed through spectral pretreatments, and subsets of samples based on carbon content, soil texture and soil class produced accurate predictions of soil TC and OC with high correlation to observations (R^2 of 0.95 and 0.93, respectively).

Another application of this method in the food science field (Vongsvivut et al., 2012) used attenuated total reflectance FT-IR (ATR-FT-IR) spectra of fish oil supplements and PLSR to quantify the fatty acid content of the oil. Fatty acids are composed of organic functional groups including carbonyl groups, carboxylic acid OH groups and aliphatic CH groups. Because gas chromatography (GC), the common method for measuring fatty acids in oils, is time and labor intensive and uses hazardous chemicals, researchers sought a faster, less expensive and more environmentally friendly method. Sixty-four samples were analyzed by GC and ATR-FT-IR, and two-thirds of these were used to develop a calibration for fatty acids using PLSR. Predictive estimations ($R^2 \geq 0.96$ compared to observed values) of total oil, total fatty acids and two specific fatty acids in fish oil samples were made using this technique.

The work presented here proposes a similar approach, in which FT-IR spectra and PLSR are used to predict TOR OC in ambient aerosol samples. As described above, thermal-optical methods such as TOR provide OC measurements in air monitoring network ambient particle matter samples but are destructive and relatively expensive. FT-IR analysis is fast, relatively inexpensive and non-destructive to the samples and can be performed on PTFE filters. The use of PTFE filters for FT-IR analysis has several benefits. While particles collected on PTFE filters likely have similar organic gas phase adsorption as particles collected on quartz filters, PTFE filters have minimal organic gas phase adsorption compared to quartz (Gilardoni et al., 2007; Turpin et al., 1994) and are commonly used in PM monitoring networks, such as the speciation networks mentioned above, for gravimetric mass and elemental analysis. The Federal Reference Method sampling network used for compliance with National Ambient Air Quality Standards for PM mass concentrations in the

United States is a large network that uses PTFE filters which could be analyzed by FT-IR for prediction of TOR OC in locations where speciation monitors are not available. Importantly, many quantities of interest – including organic functional groups, OM and OM / OC – can be quantified from the same FT-IR spectra (Fig. 1). In this work, methods are developed and tested using TOR OC data and FT-IR spectra from parallel PTFE filters from one year of samples from seven IMPROVE sites. Although methods exist for measuring OC directly from FT-IR spectra (Russell, 2003; Ruthenburg et al., 2014), calibrating to TOR OC provides TOR-equivalent OC data that will enable the continuation of long-term trend analysis of particulate pollution and longitudinal epidemiological studies on the effects of particulate pollution on human health.

The objectives of this work are to demonstrate the feasibility of predicting TOR OC from infrared spectra and establish that this prediction can be accomplished with accuracy on par with TOR measurement precision. This work is the first step in proposing a non-destructive method for reducing sampling and analysis costs for large particulate speciation monitoring networks. The method also provides a means of obtaining information about the carbonaceous aerosol at sampling sites that have only Teflon filter samples, provided that new samples have similar aerosol composition to the samples in the calibration set. We will mechanistically explain important differences in sample composition between calibration and test sets that can lead to increased prediction errors; for this we use additional IMPROVE and FT-IR measurements to aid in our interpretation. And, finally, we will demonstrate how sensitivity to sample composition is manifested in predictions for sites which are not included in the calibration set.

2 Methods

2.1 IMPROVE network samples

The IMPROVE filters used in this work were collected at seven sites during 2011. The seven sites are shown in Fig. S1 in the Supplement. The Phoenix, AZ site has two IMPROVE samplers, and filters from both samplers are used in this study. In the IMPROVE network, filters are collected every third day from midnight to midnight local time at a nominal flow rate of 22.8 L min^{-1}, which yields a nominal volume of 32.8 m^3 and produces filter samples of particles smaller than 2.5 µm in diameter (PM$_{2.5}$).

The FT-IR analysis is applied to 25 mm PTFE filters (Teflo, Pall Gelman) that are analyzed for gravimetric mass, elements and light absorption in the IMPROVE network. The sample area is 3.53 cm^2. Quartz filters collected in parallel to the PTFE filters are analyzed by TOR using the IMPROVE_A protocol to obtain OC and EC mass in the IMPROVE network (Chow et al., 2007). Prior to data publi-

cation, the OC values are adjusted to account for charring of organic material during heating (Chow et al., 2007). Organic carbon values are also adjusted to account for the gas phase adsorption artifact by subtracting the monthly median OC value from field blanks collected at a few sites in the network (http://vista.cira.colostate.edu/IMPROVE/Data/QA_ QC/Advisory/da0031/da0031_OC_Artifact.pdf); during 2011 the monthly median OC artifact values ranged from 4.1 to 6.7 µg OC. For this work, the reported TOR OC values are adjusted to account for measured flow differences between the quartz and PTFE filters. IMPROVE data were obtained from the Federal Land Manager Environmental Database (FED, http://views.cira.colostate.edu/fed/Default.aspx) on 1 May 2014. IMPROVE samples lacking either flow records for PTFE filters or TOR measurements are excluded, leaving 794 samples for this analysis.

In order to provide reference performance metrics for the evaluation of the FT-IR to TOR comparisons (see Sect. 2.4 for a description of the metrics), measurements from seven IMPROVE sites with collocated TOR measurements (Everglades, Florida; Hercules Glade, Missouri; Hoover, California; Medicine Lake, Montana; Phoenix, Arizona; Saguaro West, Arizona; Seney, Michigan) are used.

2.2 FT-IR analysis

2.2.1 Spectra acquisition

A total of 794 PTFE ambient samples and 54 PTFE laboratory blank filters are analyzed using a Tensor 27 Fourier transform infrared (FT-IR) spectrometer (Bruker Optics, Billerica, MA) equipped with a liquid-nitrogen-cooled wideband mercury cadmium telluride detector. The samples are analyzed using transmission FT-IR over the mid-infrared wavenumber region of 4000 to 420 cm^{-1} (see Ruthenburg et al., 2014, for more details). Absorbance spectra are calculated using a recent spectrum of the empty sample compartment as a zero reference. Each spectrum is zero-filled (smoothed) with a factor of 8 in the OPUS software. Air free of water vapor and carbon dioxide (delivered by purge-gas generator; PureGas LLC, Broomfield, CO) is used to continuously purge the optical compartments of the instrument and to purge the sample compartment for 4 min before each sample or reference spectrum is acquired. Each sample or reference spectrum takes about 1 min to collect such that the total analytical time per filter is about 5 min. No sample pretreatment is performed.

2.2.2 Spectra preparation

Three different versions of the absorption spectra are used in our analysis (Fig. S2 in the Supplement), corresponding to different pretreatments and wavelength selection: (1) "raw" spectra are unmodified spectra except that values interpolated during the zero-filling process are removed. These spec-
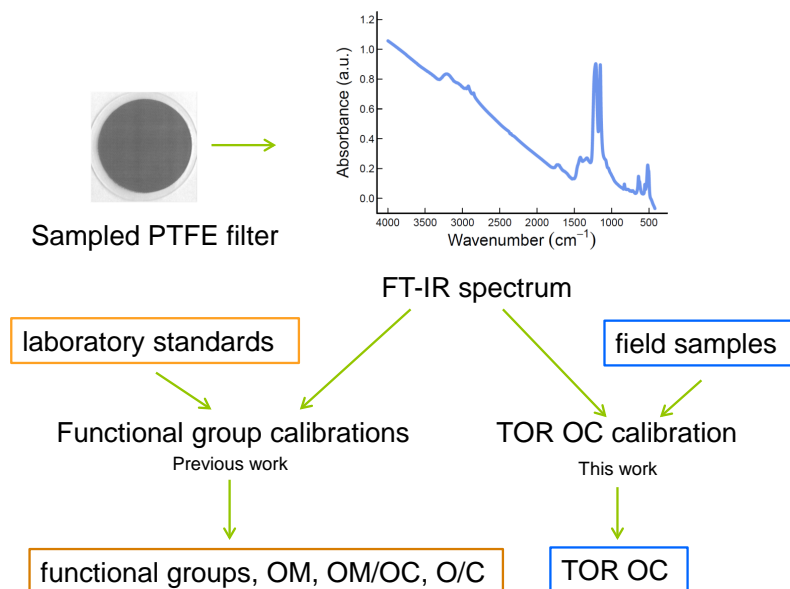
Figure 1. FTIR absorbance spectra from particulate matter collected on PTFE filters can be used for measuring OM; OM / OC; organic functional groups; and, from the work presented here, TOR OC. Previous work on functional group calibrations includes CH groups from saturated (Ruthenburg et al., 2014), unsaturated and ring structures (Gilardoni et al., 2007; Maria et al., 2003; Russell et al., 2009); amine CNH$_2$ (Liu et al., 2009; Maria et al., 2003); alcohol (Ruthenburg et al., 2014) and phenol COH (Bahadur et al., 2010; Russell et al., 2010; Takahama et al., 2013); organosulfate COSO$_3$ by solvent rinsing (Hawkins et al., 2010; Maria et al., 2003); organonitrate CONO$_2$ groups (Day et al., 2010); carboxylic COH (Liu et al., 2009; Ruthenburg et al., 2014; Takahama et al., 2013); and carbonyl CO (Gilardoni et al., 2007; Maria et al., 2003; Ruthenburg et al., 2014; Takahama et al., 2013).

tra contain all 2784 wavenumbers. (2) "Baseline-corrected" spectra include absorbances above 1500 cm^{-1}, and the substrate contribution is removed by subtracting an average blank filter spectrum and then using linear or polynomial baselines by spectral region as described by Takahama et al. (2013). These spectra are standardized to a 2 cm^{-1} resolution and contain 1563 wavenumbers. (3) "Truncated" spectra are the raw spectra interpolated to match the wavenumbers in the baseline-corrected spectra, which excludes the PTFE peaks (the region below 1500 cm^{-1}), and so also contain 1563 wavenumbers.

2.3 Calibration

The FT-IR spectra are calibrated to TOR OC measurements using PLSR (also called projection onto latent structures regression) using the kernel partial least-squared (PLS) algorithm, implemented by the PLS library (Mevik and Wehrens, 2007) for the R statistical package (R Core Team, 2014). In PLSR, the matrix of spectra is decomposed into a product of orthogonal factors (loadings) and their respective contributions (scores); observed variations in the OC mass are reconstructed through a combination of these factors and a set of weights simultaneously developed to relate features in the dependent and independent variables. Candidate models for calibration are generated by varying the number of factors used to represent the matrix of spectra. A common approach

for model selection and assessment is to divide the set of available samples into three groups: a training set for determining model parameters, a validation set for selecting the best model and a test set for evaluating its performance or prediction errors (Hastie et al., 2009; Bishop, 2011; Witten et al., 2011). The first two sets are combined into what is called the calibration set; training and validation is handled by an approach known as K fold cross validation (CV) (Arlot and Celisse, 2010; Hastie et al., 2009). In this approach, the calibration set is partitioned into K segments, and each of the K segments is used for validation while the remaining K-1 segments are used to train the model.

The minimum root mean square error of prediction (RMSEP; Mevik and Cederkvist, 2004) is used to select the model with least prediction error. A value of K between 5 and 10 has often been chosen empirically for CV (Hastie et al., 2009); evaluation of FTIR OC estimates for $K = 5$, 8 and 10 showed very little difference in prediction error (Supplement, Sect. S3), so a value of $K = 10$ is fixed for our protocol. This CV procedure permits development and selection of PLSR models using only the samples in the calibration set, and it guards against overfitting to a single set of samples. Blind evaluation is then carried out on the test set, which imposes no influence on the model development or selection.

We follow the common approach of using two-thirds of the total filters in the calibration set (Arlot and Celisse, 2010; Hastie et al., 2009) for the "Base case" (described in the fol-

lowing paragraph) and other cases used to evaluate which parameters impact prediction quality. Included in this set are spectra from ambient samples and blank laboratory filters, and the corresponding OC mass (which is assumed to be 0 for the blank laboratory filters). Samples with TOR OC values below its method minimum detection limit (MDL) are excluded from the calibration set so as to not train the model to values with low signal-to-noise ratios. The total number of samples in the test set is one-third of the ambient and blank samples. The test set is used to assess the prediction quality and is not used in calibration development. Predicted FT-IR OC values for the laboratory blank samples in the test set are used to calculate the MDL. Performance metrics used to assess the quality and MDL determination are described in Sect. 2.4.

Multiple calibrations are developed by varying the spectral type used and by selecting filters for the calibration and test sets using different ordering regimes. We define a Base case reference scenario, where the samples are chronologically stratified per site (i.e., ordered by date for each site), prior to selecting every third sample for inclusion in the test set. The remaining samples are placed in the calibration set. The Base case is also defined to use the raw spectra. Other calibration models are described in the results section.

2.4 Methods for evaluating the quality of calibration

The quality of each calibration is evaluated by calculating four performance metrics: bias, error, normalized error and the coefficient of variation (R^2) of the linear regression fit of the predicted FT-IR OC to measured TOR OC. FT-IR OC is the OC predicted from the FT-IR spectra and the PLSR calibration model. TOR OC is the artifact-corrected OC reported from TOR and available on the FED website. The bias is the median difference between measured (TOR) and predicted (FT-IR) OC for the test set. Error is the median absolute bias. The normalized error for a single prediction is the error divided by the TOR OC value. The median normalized error is reported. The performance metrics are also calculated for the collocated TOR observations and compared to those of the FT-IR OC to TOR OC regression. The MDL and precision of the FT-IR and TOR methods are calculated and compared. The MDL of the FT-IR method is 3 times the standard deviation of the laboratory blanks in the test set (18 blank filters). The MDL for the TOR method is 3 times the standard deviation of 514 blanks (Desert Research Intitute, 2012). Precision for both FTIR and TOR is calculated using the 14 parallel samples in the test set at the Phoenix, AZ site.

3 Results

3.1 Predicting TOR OC from infrared spectra

Figure 2 compares predicted FT-IR OC to measured TOR OC for the calibration and test set for the Base case. The

performance metrics for the calibration and test sets show good agreement between measured and predicted OC values. Prediction of the calibration set is expected to be better than the test set as the model is trained on these values. An ANOVA analysis between the calibration set predictions and the test set predictions indicates that the predictions are not statistically different, although the bias ($p = 0.08$) and error ($p < 0.001$) are. The performance metrics for the collocated TOR samples show good agreement between TOR samples collected at the same site and time. The precision between TOR samples is expected to be better than that between FT-IR OC and TOR OC because the TOR samples are collected on the same filter type and analyzed by the same method. However, since the collocated observations are from different sites than the FT-IR OC and TOR OC comparison (except Phoenix), a direct comparison (and ANOVA analysis) is not possible. The distribution of normalized errors for the calibration and test set and the collocated precision for the TOR samples is quite similar (Fig. S4 in the Supplement). Additional calibrations are created using fewer samples in the calibration set, and the error in the test set is independent of the number of samples in the calibration set as long as there are at least one-third of the total samples (~ 250 samples) in the calibration set (see Sect. S5 in the Supplement), indicating that the calibration is robust with respect to the number of samples used to calibrate between one-third and two-thirds of the sample set. The number of samples is not, however, an absolute number but is dependent on the specific set of samples in the calibration and test sets. The analysis shows that the accuracy of FT-IR OC predictions with respect to TOR OC values is comparable to the precision of collocated TOR measurements.

Table 1 compares the MDL and precision of the FT-IR OC predictions and TOR OC measurements. The MDL for the FT-IR OC method using raw spectra (Base case, Fig. 2) is higher than TOR, but both methods have fewer than 3 % of the samples below MDL. For the FT-IR OC method with raw spectra, seven of the 268 ambient samples in the test set are below MDL, and four for TOR. The MDL is calculated from 18 blank filters in the test set with 36 blank filters in the calibration set. However, the MDL is independent of the number (from 0 to 36) of blanks in the calibration set and the number of samples (513 to ~ 100) in the calibration set (see Sect. S5 in the Supplement). The absolute precision for FT-IR OC is on par with TOR OC. The mean predicted value for the blanks filters (last row of Table 1) is an order of magnitude lower than the 1st percentile of predicted OC values in this data set.

3.2 Predicting TOR OC using different spectral types

The analysis shown in Fig. 2 is performed on the raw spectra. Figure 3 shows the same prediction capability of the method using baseline-corrected spectra and truncated spectra. All other inputs, including the samples used for the calibration
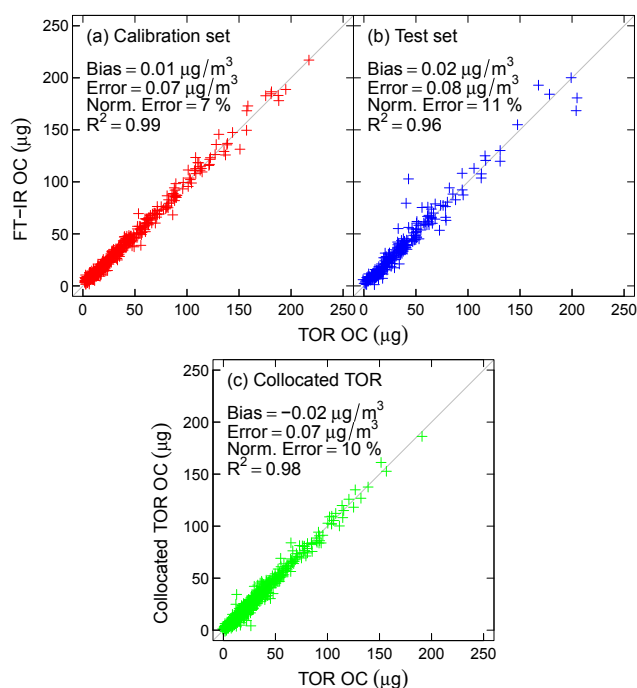
Figure 2. Predicted OC for calibration set **(a)** and test set **(b)**. The collocated TOR samples **(c)** are from sites with parallel quartz filters that are both analyzed by TOR. Only the Phoenix site has samples in the calibration, test and collocated data sets. There are 521 samples in the calibration set **(a)**, 265 samples in the test set **(b)** and 431 samples in the collocated TOR data set **(c)**. Concentration units of $\mu g\,m^{-3}$ for bias and error are based on the IMPROVE nominal volume of $32.8\,m^3$.

Table 1. MDL and precision for FT-IR OC and TOR.

	TOR OC	FT-IR OC raw spectra	FT-IR OC baseline-corrected spectra	FT-IR OC truncated spectra
MDL $(\mu g\,m^{-3})^a$	0.05	0.14	0.11	0.08
% below MDL	1.5	2.6	0.7	0.7
Precision $(\mu g\,m^{-3})^a$	0.14	0.12	0.21	0.12
Mean blank (μg)	NR[b]	0.1 ± 1.5	1.9 ± 1.2	2.8 ± 0.9

[a] Concentration units of $\mu g\,m^{-3}$ for MDL and precision are based on the IMPROVE volume of $32.8\,m^3$. [b] Not reported.

and test sets, are not changed. The performance metrics (test set panel in Fig. 2 for raw spectra) are of the same order for all three cases. An ANOVA analysis of these three predictions produces p values of 0.99 (R^2), 0.53 (bias) and 0.61 (error), indicating that the quality of predictions are not statistically different for these three spectra pretreatments. The distribution of normalized errors for the calibration and test set for both spectral pretreatments are quite similar to the distribution of normalized errors when using the raw spectra and the collocated precision for TOR samples (Fig. S4 in the Supplement).
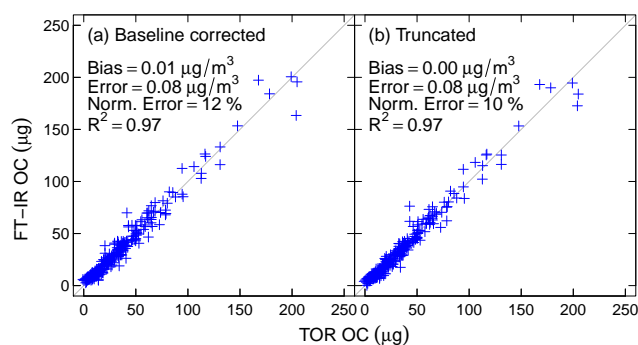

Figure 3. Predicted FT-IR OC versus measured TOR OC for the Base case test set with **(a)** baseline-corrected and **(b)** truncated spectra. Concentration units of $\mu g\,m^{-3}$ for bias and error are based on the IMPROVE nominal volume of $32.8\,m^3$.

Table 1 shows the MDL and precision values for these two cases. When compared to the raw spectra calibration, the MDLs for these two cases are lower than the raw spectra; both have only two samples below MDL. The mean blank values for the baseline-corrected and truncated spectra cases are higher and not centered around 0 as is the raw spectra calibration. For baseline-corrected cases, the mean blank is less than half of the 1st percentile of predicted OC values; for the truncated spectra, the mean blank is of the same order as the 1st percentile of predicted values (3.7 μg). The precision is poorest using baseline-corrected spectra. ANOVA of the blank values indicates that the blank predictions are significantly different ($p < 0.001$ for prediction, bias and error).

3.3 Evaluating causes of bias and error by selecting the calibration and test sets based on measured parameters

In this section, we consider the role of the distribution of TOR OC, OM / OC and ammonium / OC on FT-IR OC predictions. The magnitude of TOR OC is considered since this is the property to be quantified. OM / OC is considered since it is indicative of the mix of primary and secondary organic aerosol composition. OM / OC is obtained from FT-IR analysis calibrated with laboratory standards (Ruthenburg et al., 2014). Ammonium can be an interferant in FT-IR analysis; the absorption band of the N-H stretching vibrations overlaps with several vibrational modes of organic functional groups. We use the ratio of ammonium to OC mass loadings to isolate the effect of ammonium because the magnitude of its interference is dependent on its mass with respect to the organic material mass collected on the filter. Because ammonium is not measured in the IMPROVE network, the ammonium mass is estimated assuming full neutralization solely by ammonium of reported sulfate and nitrate concentrations reported in the IMPROVE network data. The assumption may be an over- or underestimation of ammonium depending on the amount of neutralization and other species present; how-
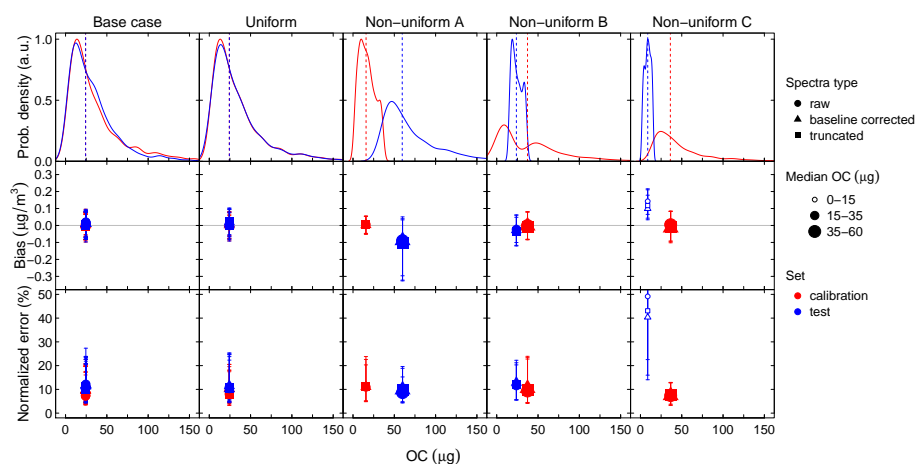
Figure 4. The probability density distribution of OC and bias and normalized error (with the interquartile range shown by error bars) in the calibration (red) and test (blue) sets for five calibration cases: the Base case, the Uniform OC case and three Non-uniform OC cases. Vertical lines are the median of the OC mass distributions color-coded for calibration and test sets.

ever we expect that for the purpose of our study, the errors in this assumption will not significantly alter our evaluation. Separate calibrations are developed for each parameter: OC, OM / OC and ammonium / OC.

To investigate the role of the distribution of each parameter, samples are arranged in ascending order by the parameter of interest prior to selection of filters for the calibration and test sets. Every third sample in the ordered list is put into the test set, and the remaining samples are put into the calibration set. These cases are called the Uniform OC case, Uniform OM / OC case and Uniform ammonium / OC case. Three Non-uniform cases are also considered for TOR OC: samples with TOR OC in the lowest two-thirds of the TOR OC range are used to predict samples with TOR OC in the highest one-third of the TOR OC range (Non-uniform A), samples with the highest and lowest one-third TOR OC are used to predict samples in the middle one-third TOR OC (Non-uniform B) and samples with the highest two-thirds TOR OC are used to predict samples with the lowest one-third TOR OC mass (Non-uniform C). Similarly, three Non-uniform cases are modeled for OM / OC and ammonium / OC.

The top row of subplots in Fig. 4 shows the distribution of OC in the test and calibration sets for the Base case (for reference), the Uniform OC case and the three Non-uniform cases. For the Base case and the Uniform OC case, the distribution of OC is quite similar in the test and calibration set, but for the Non-uniform cases the distributions are different and reflect the algorithm used to select the filters for each case. The median and 25th to 75th percentiles (interquartile range) of the bias and normalized error are shown in the lower two rows of Fig. 4 for each of the three spectral types. Small, open symbols are used for sets with low median OC mass. Larger, closed symbols represent sets that have higher median OC mass. For the Base and Uniform cases, the median bias is close to 0 and the interquartile range is similar

and small for the test and calibration sets. The median normalized error and the interquartile range for these two cases are also small and similar for the test and calibration sets. The bias and error indicate that the test set is well predicted for both the Base and Uniform cases. Similarly, for the case where the lowest and highest thirds of the values are used to predict the middle third (Non-uniform B), the bias and normalized error median and interquartile range are similar and small, indicating good prediction of the test set. For the case when low-OC-mass samples are used to predict high-OC-mass samples (Non-uniform A), there is a small negative bias ($-0.10\,\mu g\,m^{-3}$) and a larger range in bias for the test set. However, the normalized error is small and similar for the two sets, highlighting the linearity of the calibration. For all of these cases, median OC masses for both sets are greater than 15 µg. For the case when high-OC-mass samples are used to predict low-OC-mass samples (Non-uniform C), the median OC mass is less than 15 µg in the test set. For this case the median bias is 0.10 to $0.14\,\mu g\,m^{-3}$ and the normalized error is between 40 and 50 % depending on the spectral types used. The range of errors (the higher errors are outside the bounds of the plot) is also considerably larger. The positive bias and normalized errors for low-OC-mass samples is expected due to some combination of higher analytical TOR and FT-IR errors, including TOR blank correction and PLSR fitting errors at low concentrations. For the samples below 15 µg, the actual measurement artifact may be considerably less than the monthly median value used (Sect. 2.1), leading to an underestimate of TOR OC which contributes to the positive bias in the FT-IR OC. The large sample-to-sample variability in measurement artifact in TOR may contribute to the higher variability in the error.

The top row of subplots in Fig. 5 shows the distribution of OM / OC in the test and calibration sets for the Base case, the Uniform OM / OC case and the three Non-uniform OM / OC
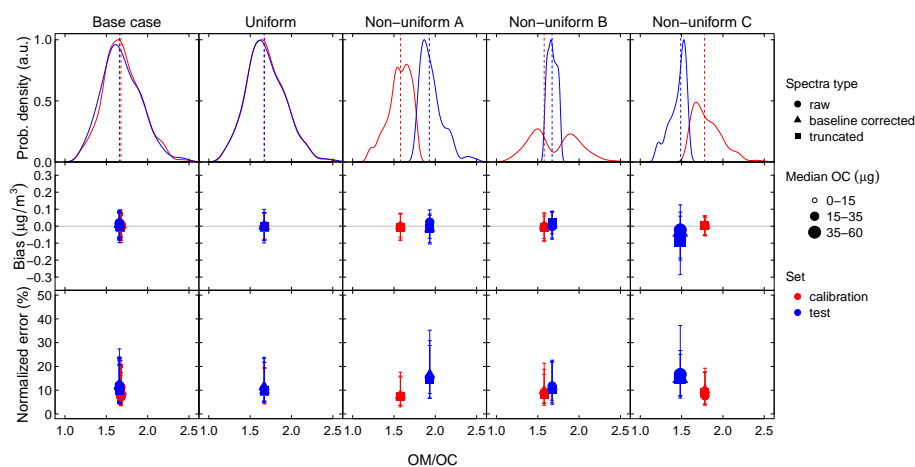
Figure 5. The probability distribution of OM / OC and bias and normalized error (with the interquartile range shown by error bars) in the calibration (red) and test (blue) sets for five calibration cases: the Base case, the Uniform OM / OC case and three Non-uniform OM / OC cases. Vertical lines on the probability distributions are the color-coded median of the OM / OC distributions.

cases. The Base and Uniform cases have similar OM / OC distributions, a median bias of 0 and low normalized error in the test and calibration sets, indicating good prediction of the test set. When the highest and lowest one-third of the samples is used to predict the middle third (Non-uniform B), the median OM / OC is somewhat different between the calibration and test set, but the test set has low bias and error, indicating good prediction. However when there is a larger difference in OM / OC between the test and calibration sets (Non-uniform A and Non-uniform C), the bias is still near 0 (no more than $0.03 \, \mu g \, m^{-3}$) – except for the Non-uniform C, truncated case ($0.09 \, \mu g \, m^{-3}$) – but the normalized error and its range are higher for the test set (14–17 %) than for the calibration set (7–9 %). The higher error is due to difference in the chemical composition of the aerosol in the test and calibration sets. High OM / OC indicates that the carbonaceous aerosol is oxidized and has considerable functionality as would be expected of secondary organic aerosol. Primary organic aerosol has a low OM / OC because there is less oxygen and functionality in the molecules. The difference in composition leads to an increase in the median normalized error in the test set and increases the likelihood of larger errors for some samples as indicated by the larger error bars. This analysis is carried out for OC / EC and is shown in Sect. S6 in the Supplement. OC / EC has been used as an indicator of organic composition (Turpin and Huntzicker, 1995) and follows a similar pattern to OM / OC.

The impact of ammonium is evaluated using Uniform and Non-uniform calibrations of ammonium / OC (Fig. 6). Similar to OC and OM / OC, the Base case, Uniform case and Non-uniform B case have near-zero bias and low normalized error. When low ammonium / OC samples are used to predict samples with high ammonium / OC (Non-uniform A), the bias increases to $0.1 \, \mu g \, m^{-3}$ and the normalized error increases from 8 % in the calibration set to 24 % in the test set.

In this case, the calibration set is not trained to disregard ammonium in the prediction of OC, so some of the ammonium is likely reported to be OC. In the Non-uniform C case, the calibration set is trained to disregard ammonium, the prediction of low ammonium / OC samples is slightly biased low (0.04 to $0.06 \, \mu g \, m^{-3}$), the range of the bias increases and the error increases by 3 or 4 % from the calibration set to the test set, but the range is similar for the two sets. This suggests that a small amount of OC may be incorrectly assigned to ammonium, so the predictions are biased slightly low and the error increases slightly. The distribution of OC, OM / OC, ammonium / OC and EC / OC for the test and calibration sets for the Base, Uniform and Non-uniform cases are shown in Sect. S7 in the Supplement.

3.4 Understanding error in samples with low OC mass

As least-squares algorithms minimize the squared magnitude of residuals, normalized errors for low-mass samples may be large when high mass samples are included in the calibration set. A calibration model localized to the lowest one-third of the OC masses (OC $\leq 15 \, \mu g$) is developed to evaluate our capability to predict OC in samples with these low masses. This calibration model is called the Low Uniform OC calibration model. The test set contains 89 ambient samples that are in the lowest one-third of the OC mass distribution. The lowest one-third mass OC calibration set is made up of 168 ranked OC samples which are in the lowest one-third of the OC mass range plus blanks. The prediction of the test set by the Low Uniform OC calibration is compared to the prediction of the same test set by Uniform OC calibration (Sect. 3.3) which includes the full range of OC. The distribution of OM / OC and ammonium / OC in the test and calibration sets for these cases are similar (Sect. S7 in the Supplement), indicating that the error in the low-OC sam-
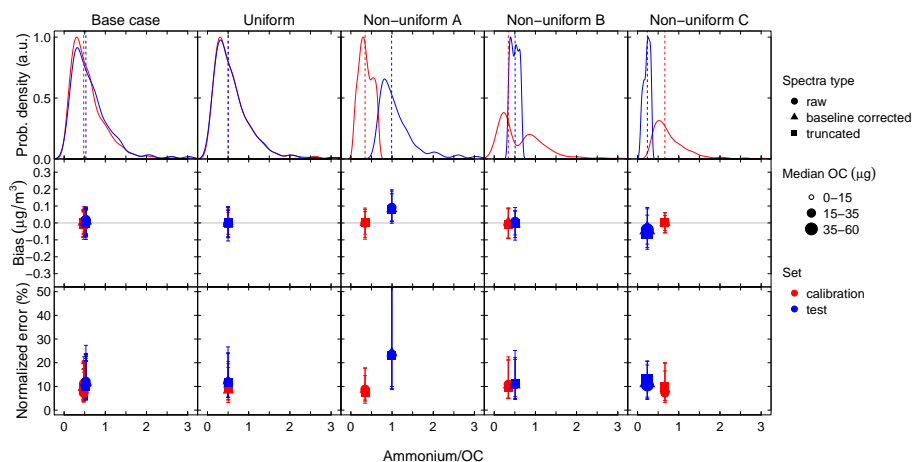
Figure 6. The probability distribution of ammonium / OC and bias and normalized error (with the interquartile range shown by error bars) in the calibration (red) and test (blue) sets for five calibration cases: the Base case, the Uniform ammonium / OC case and three Non-uniform ammonium / OC cases. Vertical lines on the probability distributions are the median of the ammonium / OC distributions.

ples is not due to differences in chemical composition or ammonium in the test and calibration sets. Figure 7 shows the mean error and MDL for the Uniform OC calibration and the Low Uniform OC calibration for each of the three spectral types. Collocated TOR precision for samples in the same mass range as the Low Uniform OC calibration (OC ≤ 15 µg) is shown for comparison in Fig. 7. The mean error does not significantly decrease when using samples with low OC mass in the calibration, and it is comparable to the collocated TOR precision. Improvement in the reported detection limits for the raw and truncated spectra model is observed when using samples with low OC mass, suggesting that samples with masses near MDL may benefit from this alternative calibration model. However, because the average prediction error for these low-mass samples is not significantly improved according to any of these calibrations over the Uniform OC case model, the Uniform OC case calibration is suitable for most samples (further discussion on the distribution of errors is provided in Sect. S8 of the Supplement). Since we are fitting the FT-IR spectra to TOR OC measurements, the error in FT-IR OC cannot be lower than the error in TOR OC itself. However, this analysis suggests that the FT-IR analytical and PLS fitting errors do not impose a significant addition to the TOR analytical and artifact-correction errors already present in the OC measurements.

3.5 Using differences in OC mass and aerosol composition in the test and calibration sets to explain the quality of TOR OC predictions at specific sites

Calibrations are developed using all ambient samples in the calibration set except samples from one site which is predicted. For five sites, the distributions of OC in the test and calibration set, and the median and interquartile range of bias
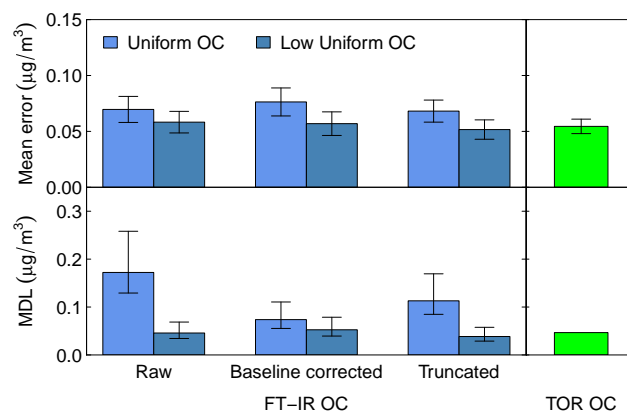

Figure 7. Mean error and MDL for the Uniform OC (which includes the full range of OC) and the Low Uniform OC cases for all three spectral types. Collocated TOR OC precision and MDL for TOR samples in the same mass range as the Low Uniform OC case (OC < 15 µg) are shown for reference. The error bars indicate the 95 % confidence intervals on these point estimates. Absolute errors are compared directly because the same test set is used for both cases (FT-IR OC), or for samples in the same concentration range (TOR OC).

and normalized error are shown in Fig. 8. Three sites – Mesa Verde, Olympic and Trapper Creek – have median OC mass below 15 µg (shown with open symbols) and have the highest median and range of normalized error. As shown with the low-OC calibration and comparison to collocated TOR samples (Sect. 3.4), these errors are due primarily to TOR analytical and artifact-correction errors. All other sites have higher OC mass and are expected to be predicted well, based on OC mass alone. St. Marks and Proctor Maple Research Facility are both well predicted (Fig. 8). Distributions of OC, OM / OC, OC / EC and ammonium / OC for the test and cal-
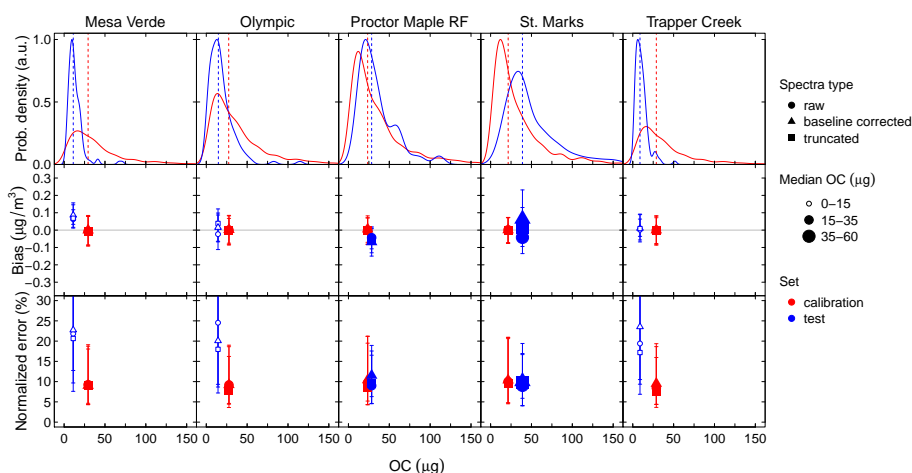
Figure 8. The distribution of OC and the bias and normalized error (with the interquartile range shown by error bars) in the calibration (red) and test (blue) sets for calibrations developed for each of five sites. Each calibration has all samples in the calibration set except for the site to be predicted. Vertical lines are the color-coded median of the OC distributions.
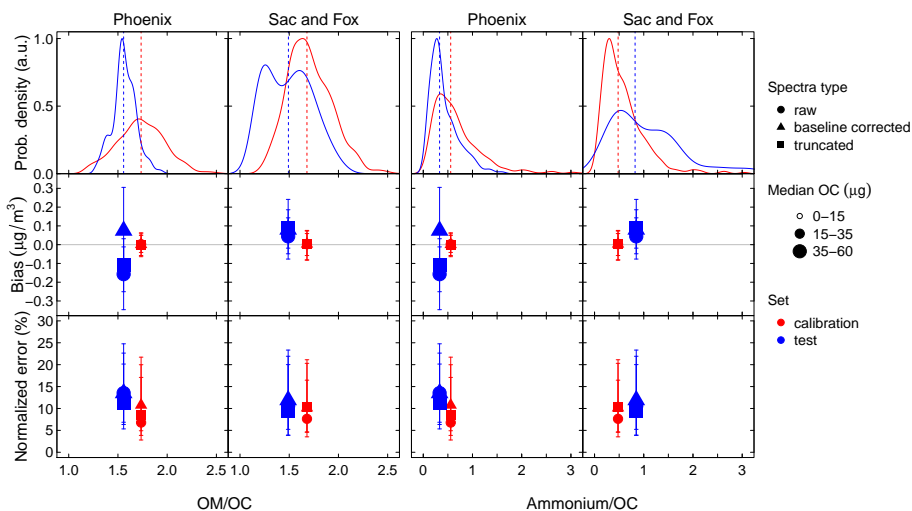

Figure 9. The OM / OC and ammonium / OC distributions and the bias and normalized error (with the interquartile range shown by error bars) in the calibration (red) and test (blue) sets for calibrations developed for Phoenix and Sac and Fox. Each calibration has all samples in the calibration set except for the site to be predicted. Vertical lines are the median of the OM / OC or ammonium / OC distributions.

ibration sets for all sites are shown in Sect. S7 of the Supplement.

Figure 9 shows the OM / OC and ammonium / OC distributions for the two remaining sites, Phoenix and Sac and Fox. Phoenix, an urban site, and Sac and Fox have lower OM / OC than the rest of the sites, which indicates that there is more primary OM at these sites than in the rest of the sites. For Sac and Fox, the median OM / OC is lower than the rest of the sites (calibration set), but the distribution is bimodal such that many of the Sac and Fox samples are in the same range of OM / OC as the other sites, minimizing the impact of the difference in median OM / OC. The median and range of the bias is higher for Sac and Fox than for the other sites, but the error is very similar to the other sites, indicating only

a slightly poorer prediction than for the calibration set. For Phoenix, the difference in composition produces predictions that are more biased (the direction of the bias depends on the type of spectra used) and the range of bias is large, which means that more samples have larger biases than in the calibration set. However, the median OC for Phoenix is nearly 50 µg, so the bias is small relative to the OC mass. The normalized error is also slightly higher for the Phoenix samples than the rest of the samples although the distribution of errors is similar for the calibration and test set, indicating only a small effect on error. Phoenix has the largest difference in composition between it and the rest of the sites, yet the impact on the calibration metrics is small. This analysis is car-

ried out for OC / EC and shows similar trends (Sect. S6 in the Supplement).

Only the Phoenix and Sac and Fox sites show differences in ammonium / OC between the test and calibration set; these are the same two sites impacted by OM / OC differences (Fig. 9). The calibration set for predicting Phoenix has higher ammonium / OC than Phoenix, the same pattern as Non-uniform C for ammonium / OC, which was shown to have only a small impact on predicted values. This suggests that the increased bias and error in Phoenix is due primarily to differences in organic composition, not to ammonium interference. The calibration set for Sac and Fox has lower ammonium / OC than Sac and Fox. This is similar to Non-uniform A for ammonium / OC, in which the calibration is not trained to disregard ammonium when determining OC, so a positive bias is observed and a larger normalized error and range of errors. Sac and Fox has only a small positive bias and increase in error and no increase in the range of error, so the impact of ammonium, if present is small. However, the impact of the difference in OM / OC produces similar changes in bias and error to ammonium / OC, so for Sac and Fox the small increases in bias and error compared to the calibration set may be due to OM / OC, ammonium / OC or some combination of both.

We can therefore predict how well a site not included in the calibration will be predicted, based on the OC, OM / OC and ammonium / OC for the site. However, even for the most poorly predicted sites the median normalized errors are still fairly low; 17–25 % for sites with low OC mass; 11–14 % for Phoenix, which has low OM / OC; and 9–12 % for Sac and Fox due to some combination of low OM / OC and high ammonium / OC.

4 Conclusions

PTFE filters routinely collected in the IMPROVE network are non-destructively analyzed by FT-IR. The FT-IR spectra and parallel TOR OC measurements are used in partial least-squares regression to develop calibrations to predict TOR OC. All three spectral types produce high-quality predictions. Blank filters in the test set are used to calculate MDL. The calibration sets developed from samples ordered by site date, OC, OM / OC or ammonium / OC produce nearly bias-free predictions with low error. Samples with low OC mass predict OC in samples with high OC mass with low error because the calibration is linear. Errors for samples with low OC mass (less than 15 µg or 0.45 µg m^{-3}) are high primarily due to TOR OC analytical errors and artifact-correction errors. The higher errors in the low-OC-mass samples suggest that the use of a single value to artifact-correct all samples collected in a month induces additional error in low-OC samples. The low error in most samples suggests that the charring correction is consistently applied such that it can be accounted for with the statistics used to develop the calibra-

tion models. Using the lowest one-third of OC samples in the calibration set may improve the prediction for some samples near the MDL, but this modification to the calibration does not improve the overall performance of the calibration. Errors and bias are kept to a minimum by including samples in the calibration set that have a similar range of organic composition, as indicated by OM / OC or OC / EC, and a similar range of ammonium / OC to the samples to be predicted. Using a calibration set in which samples do not span the full range of OM / OC or ammonium / OC in the test set leads to higher bias and errors, but the bias and errors are still small. Therefore, we conclude that FT-IR spectra calibrated to TOR OC using partial least-squares regression is a robust method for predicting TOR organic carbon from particulate matter samples. Future work includes establishing that the calibration developed using samples from one year can be used to predict TOR OC during other years and developing a calibration that includes samples with a broader range of aerosol composition.

References

Allen, D. T., Palen, E. J., Haimov, M. I., Hering, S. V., and Young, J. R.: Fourier-transform infrared-spectroscopy of aerosol collected in a low-pressure impactor (LPI/FTIR) – method development and field calibration, Aerosol Sci. Technol., 21, 325–342, doi:10.1080/02786829408959719, 1994.

Anderson, J. O., Thundiyil, J. G., and Stolbach, A.: Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health, Journal of Medical Toxicology, 8, 166–175, doi:10.1007/s13181-011-0203-1, 2012.

Arlot, S. and Celisse, A.: A survey of cross-validation procedures for model selection, Statistics Surveys, 4, 40–79, 2010.

Bahadur, R., Uplinger, T., Russell, L. M., Sive, B. C., Cliff, S. S., Millet, D. B., Goldstein, A., and Bates, T. S.: Phenol Groups in Northeastern US Submicrometer Aerosol Particles Produced from Seawater Sources, Environ. Sci. Technol., 44, 2542–2548, doi:10.1021/es9032277, 2010.

Birch, M. E. and Cary, R. A.: Elemental carbon-based method for occupational monitoring of particulate diesel exhaust:

Methodology and exposure issues, Analyst, 121, 1183–1190, doi:10.1039/an9962101183, 1996.

Bishop, C. M.: Pattern recognition and machine learning, Springer, 2011.

Blando, J. D., Porcja, R. J., and Turpin, B. J.: Issues in the quantitation of functional groups by FTIR spectroscopic analysis of impactor-collected aerosol samples, Aerosol Sci. Technol., 35, 899–908, doi:10.1080/02786820126852, 2001.

Cavalli, F., Viana, M., Yttri, K. E., Genberg, J., and Putaud, J.-P.: Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol, Atmos. Meas. Tech., 3, 79–89, doi:10.5194/amt-3-79-2010, 2010.

Chow, J. C., Watson, J. G., Chen, L. W. A., Chang, M. C. O., Robinson, N. F., Trimble, D., and Kohl, S.: The IMPROVE-A temperature protocol for thermal/optical carbon analysis: maintaining consistency with a long-term database, J. Air Waste Manage., 57, 1014–1023, doi:10.3155/1047-3289.57.9.1014, 2007.

Chow, J. C., Watson, J. G., Chen, L.-W. A., Rice, J., and Frank, N. H.: Quantification of $PM_{2.5}$ organic carbon sampling artifacts in US networks, Atmos. Chem. Phys., 10, 5223–5239, doi:10.5194/acp-10-5223-2010, 2010.

Coury, C. and Dillner, A. M.: 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., 42, 5923–5932, doi:10.1016/j.atmosenv.2008.03.026, 2008.

Dabek-Zlotorzynska, E., Dann, T. F., Martinelango, P. K., Celo, V., Brook, J. R., Mathieu, D., Ding, L. Y., and Austin, C. C.: Canadian National Air Pollution Surveillance (NAPS) $PM_{2.5}$ speciation program: Methodology and $PM_{2.5}$ chemical composition for the years 2003–2008, Atmos. Environ., 45, 673–686, doi:10.1016/j.atmosenv.2010.10.024, 2011.

Day, D. A., Liu, S., Russell, L. M., and Ziemann, P. J.: Organonitrate group concentrations in submicron particles with high nitrate and organic fractions in coastal southern California, Atmos. Environ., 44, 1970–1979, doi:10.1016/j.atmosenv.2010.02.045, 2010.

Desert Research Intitute: DRI Model 2001 Thermal/Optical Carbon Analysis (TOR/TOT) of Aerosol Filter Samples – Method IMPROVE_A, Desert Research Institute, Standard Operating Procedure, 2012.

Flanagan, J. B., Jayanty, R. K. M., Rickman, J. E. E., and Peterson, M. R.: $PM_{2.5}$ Speciation Trends Network: Evaluation of Whole-System Uncertainties Using Data from Sites with Collocated Samplers, J. Air Waste Manage., 56, 492–499, doi:10.1080/10473289.2006.10464516, 2006.

Gilardoni, S., Russell, L. M., Sorooshian, A., Flagan, R. C., Seinfeld, J. H., Bates, T. S., Quinn, P. K., Allan, J. D., Williams, B., Goldstein, A. H., Onasch, T. B., and Worsnop, D. R.: Regional variation of organic functional groups in aerosol particles on four US east coast platforms during the International Consortium for Atmospheric Research on Transport and Transformation 2004 campaign, J. Geophys. Res.-Atmos., 112, D10S27, doi:10.1029/2006jd007737, 2007.

Hand, J. L., Schichtel, B. A., Pitchford, M., Malm, W. C., and Frank, N. H.: Seasonal composition of remote and urban fine particulate matter in the United States, J. Geophys. Res.-Atmos., 117, D05209, doi:10.1029/2011jd017122, 2012.

Hansen, D. A., Edgerton, E. S., Hartsell, B. E., Jansen, J. J., Kandasamy, N., Hidy, G. M., and Blanchard, C. L.: The Southeastern Aerosol Research and Characterization Study: Part 1 – Overview, J. Air Waste Manage., 53, 1460–1471, doi:10.1080/10473289.2003.10466318, 2003.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning – Data Mining, Inference, and Prediction, 2nd Edn., Springer Series in Statistics, Springer, 745 pp., 2009.

Hawkins, L. N., Russell, L. M., Covert, D. S., Quinn, P. K., and Bates, T. S.: Carboxylic acids, sulfates, and organosulfates in processed continental organic aerosol over the southeast Pacific Ocean during VOCALS-REx 2008, J. Geophys. Res.-Atmos., 115, D13201, doi:10.1029/2009jd013276, 2010.

Liu, S., Takahama, S., Russell, L. M., Gilardoni, S., and Baumgardner, D.: Oxygenated organic functional groups and their sources in single and submicron organic particles in MILAGRO 2006 campaign, Atmos. Chem. Phys., 9, 6849–6863, doi:10.5194/acp-9-6849-2009, 2009.

Madari, B. E., Reeves, J. B., Coelho, M. R., Machado, P., and De-Polli, H.: Mid- and near-infrared spectroscopic determination of carbon in a diverse set of soils from the Brazilian National Soil Collection, Spectrosc. Lett., 38, 721–740, doi:10.1080/00387010500315876, 2005.

Maimone, F., Turpin, B. J., Solomon, P., Meng, Q. Y., Robinson, A. L., Subramanian, R., and Polidori, A.: Correction Methods for Organic Carbon Artifacts When Using Quartz-Fiber Filters in Large Particulate Matter Monitoring Networks: The Regression Method and Other Options, J. Air Waste Manage., 61, 696–710, doi:10.3155/1047-3289.61.6.696, 2011.

Malm, W. C., Sisler, J. F., Huffman, D., Eldred, R. A., and Cahill, T. A.: Spatial and seasonal trends in particle concentration and optical extinction in the United-States, J. Geophys. Res.-Atmos., 99, 1347–1370, doi:10.1029/93jd02916, 1994.

Maria, S. F., Russell, L. M., Turpin, B. J., Porcja, R. J., Campos, T. L., Weber, R. J., and Huebert, B. J.: Source signatures of carbon monoxide and organic functional groups in Asian Pacific Regional Aerosol Characterization Experiment (ACE-Asia) submicron aerosol types, J. Geophys. Res.-Atmos., 108, doi:10.1029/2003jd003703, 2003.

Mevik, B.-H. and Cederkvist, H. R.: Mean squared error of prediction (MSEP) estimates for principal component regression (PCR) and partial least squares regression (PLSR), J. Chemometr., 18, 422–429, doi:10.1002/cem.887, 2004.

Mevik, B. H. and Wehrens, R.: The pls package: Principal component and partial least squares regression in R, J. Stat. Softw., 18, 1–24, 2007.

Pitchford, M., Malm, W., Schichtel, B., Kumar, N., Lowenthal, D., and Hand, J.: Revised algorithm for estimating light extinction from IMPROVE particle speciation data, J. Air Waste Manage., 57, 1326–1336, doi:10.3155/1047-3289.57.11.1326, 2007.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2014.

Reff, A., Turpin, B. J., Offenberg, J. H., Weisel, C. P., Zhang, J., Morandi, M., Stock, T., Colome, S., and Winer, A.: A functional group characterization of organic $PM_{2.5}$ exposure: Results from the RIOPA study, Atmos. Environ., 41, 4585–4598, doi:10.1016/j.atmosenv.2007.03.054, 2007.

Russell, L. M.: Aerosol organic-mass-to-organic-carbon ratio measurements, Environ. Sci. Technol., 37, 2982–2987, doi:10.1021/es026123w, 2003.

Russell, L. M., Bahadur, R., Hawkins, L. N., Allan, J., Baumgardner, D., Quinn, P. K., and Bates, T. S.: Organic aerosol characterization by complementary measurements of chemical bonds and molecular fragments, Atmos. Environ., 43, 6100–6105, doi:10.1016/j.atmosenv.2009.09.036, 2009.

Russell, L. M., Hawkins, L. N., Frossard, A. A., Quinn, P. K., and Bates, T. S.: Carbohydrate-like composition of submicron atmospheric particles and their production from ocean bubble bursting, P. Natl. Acad. Sci. USA, 107, 6652–6657, doi:10.1073/pnas.0908905107, 2010.

Ruthenburg, T. C., Perlin, P. C., Liu, V., McDade, C. E., and Dillner, A. M.: 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., 86, 47–57, doi:10.1016/j.atmosenv.2013.12.034, 2014.

Takahama, S., Johnson, A., and Russell, L. M.: Quantification of Carboxylic and Carbonyl Functional Groups in Organic Aerosol Infrared Absorbance Spectra, Aerosol Sci. Technol., 47, 310–325, doi:10.1080/02786826.2012.752065, 2013.

Tørseth, K., Aas, W., Breivik, K., Fjæraa, A. M., Fiebig, M., Hjellbrekke, A. G., Lund Myhre, C., Solberg, S., and Yttri, K. E.: Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972–2009, Atmos. Chem. Phys., 12, 5447–5481, doi:10.5194/acp-12-5447-2012, 2012.

Turpin, B. J. and Huntzicker, J. J.: Identification of secondary organic aerosol episodes and quantitation of primary and secondary organic aerosol concentrations during SCAQS, Atmos. Environ., 29, 3527–3544, doi:10.1016/1352-2310(94)00276-q, 1995.

Turpin, B. J. and Lim, H. J.: Species contributions to $PM_{2.5}$ mass concentrations: Revisiting common assumptions for estimating organic mass, Aerosol Sci. Technol., 35, 602–610, doi:10.1080/02786820152051454, 2001.

Turpin, B. J., Huntzicker, J. J., and Hering, S. V.: Investigation of organic aerosol sampling artifacts in the Los-Angeles basin, Atmos. Environ., 28, 3061–3071, doi:10.1016/1352-2310(94)00133-6, 1994.

Vongsvivut, J., Heraud, P., Zhang, W., Kralovec, J. A., McNaughton, D., and Barrow, C. J.: Quantitative determination of fatty acid compositions in micro-encapsulated fish-oil supplements using Fourier transform infrared (FTIR) spectroscopy, Food Chem., 135, 603–609, doi:10.1016/j.foodchem.2012.05.012, 2012.

Watson, J. G.: Visibility: Science and regulation, J. Air Waste Manage., 52, 628–713, 2002.

Witten, I. H., Frank, E., and Hall, M. A.: Data Mining: Practical Machine Learning Tools and Techniques, 3rd Edn., Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2011.

Yu, H., Kaufman, Y. J., Chin, M., Feingold, G., Remer, L. A., Anderson, T. L., Balkanski, Y., Bellouin, N., Boucher, O., Christopher, S., DeCola, P., Kahn, R., Koch, D., Loeb, N., Reddy, M. S., Schulz, M., Takemura, T., and Zhou, M.: A review of measurement-based assessments of the aerosol direct radiative effect and forcing, Atmos. Chem. Phys., 6, 613–666, doi:10.5194/acp-6-613-2006, 2006.