Supplemental material for Predicting Ambient Aerosol Thermal Optical Reflectance (TOR) Measurements from Infrared Spectra: Elemental Carbon

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

[1]University of California - Davis, Davis, California

[2]Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland

Correspondence to: A. M. Dillner (amdillner@ucdavis.edu) and S. Takahama (satoshi.takahama@epfl.ch)

S1. Location of sampling sites

Figure S1. Aerosol sampling locations: Mesa Verde,CO (37.1984, -108.4907) 119 samples Jan - Dec 2011, Olympic, WA (48.0065, -122.9727) 120 samples Jan - Dec 2011, Phoenix, AZ (33.5038, -112.0958) 100 samples Jan - Dec 2011, Phoenix, AZ second sampler (33.5038, -112.0958) 99 samples Jan - Dec 2011, Proctor Maple Research Facility, VT (44.5284, -72.8688) 106 samples Jan - Dec 2011, Sac and Fox, KS (39.9791, -95.5682) 53 Samples Jan - Jun 2011, St. Marks, FL (30.0926, -84.1614) 108 Samples Jan - Dec 2011, Trapper Creek, AK (63.3153, -150.3156) 110 samples Jan - Dec 2011.

S2. Spectral types

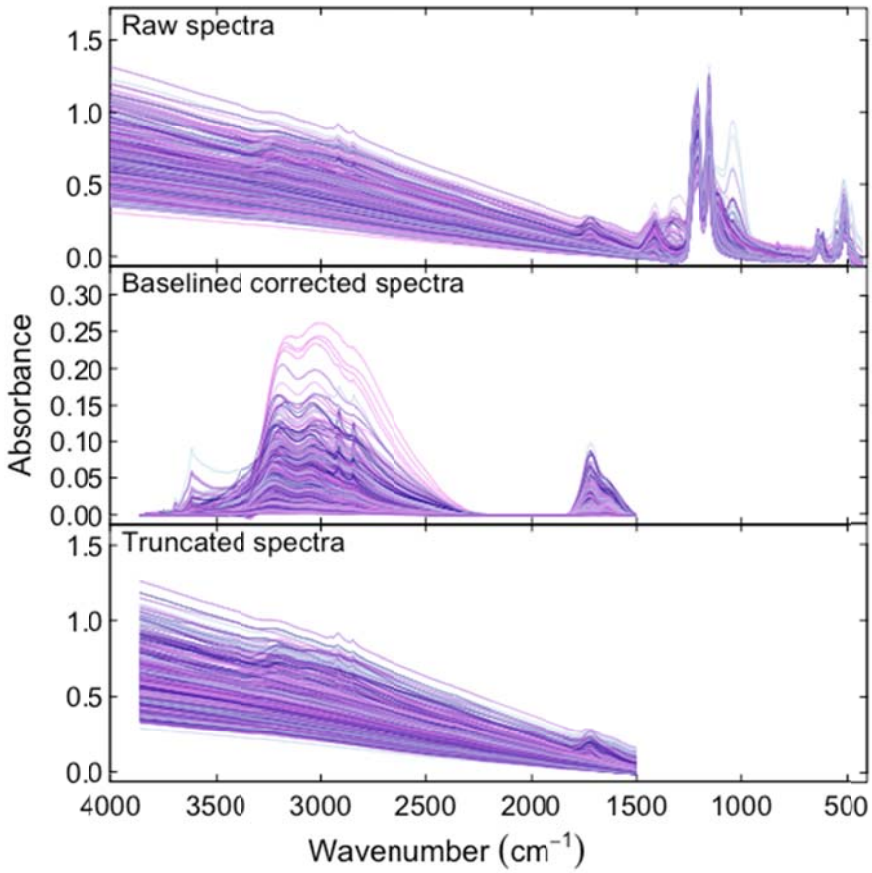

Figure S2. Spectra of 794 IMPROVE PTFE samples, shown as original absorbances ("raw" spectra) and after pretreatment ("baseline corrected" and "truncated" spectra) as described in Section 2.2. Each spectrum is differentiated by color.

S3. Error in samples used in OC Base case but corrected in this paper

In Dillner and Takahama (2015), four samples were inadvertently left out of the base case run only. The four samples are Mesa Verde_20111126, Olympic_20110106, Olympic__20110614 and Trapper Creek_20111024. All four samples are above MDL. The bias, error and normalized error of the predicted values in the test set are the same with and without these four samples but the R^2 changes from 0.96 (reported in Dillner and Takahama (2015)) to 0.95 when the samples are included. The figure below shows the comparison of the predicted values in the test set with and without the excluded filters for the three spectral types. R^2 values are 1 with little to no error or bias. The exclusion of the samples made only very small difference in the fit metrics and does not change the conclusions of the paper.

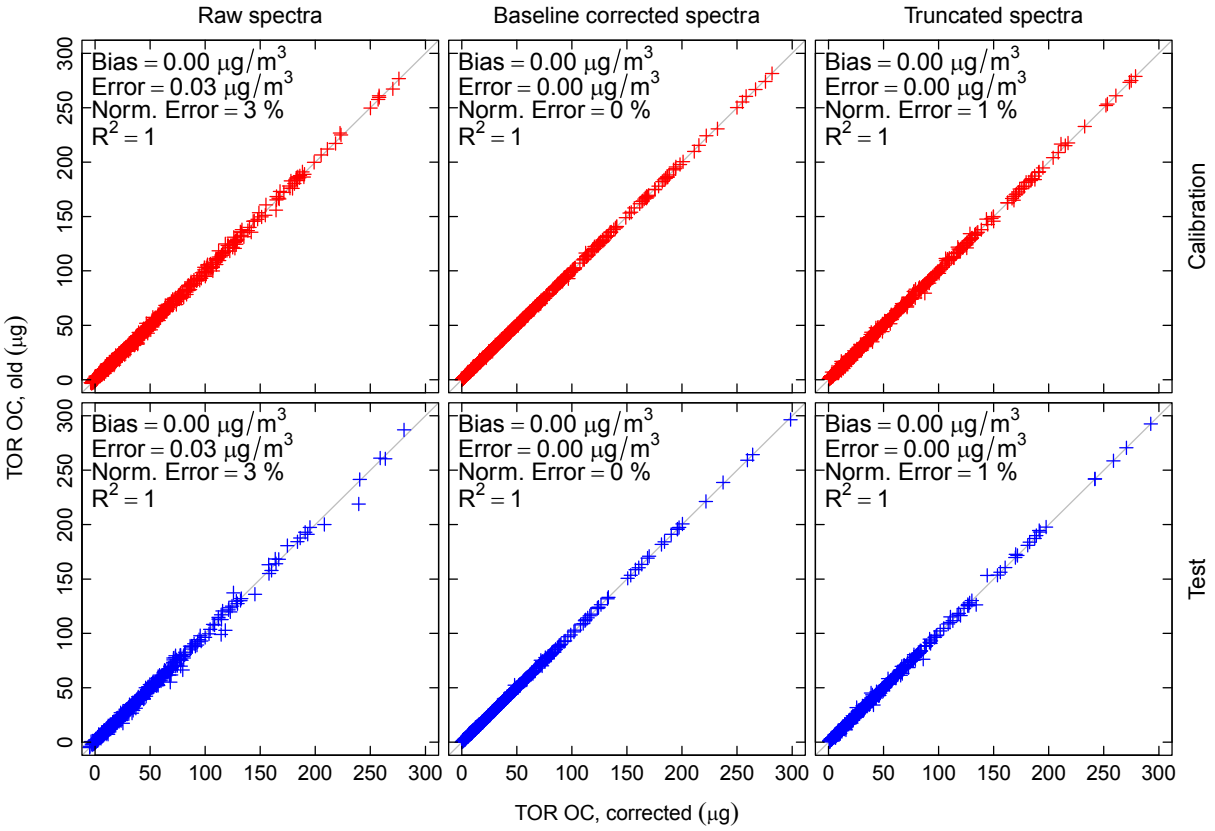

Figure S3a. Comparison of calibration and test set predictions of OC when all 521 samples are included in the calibration (corrected) and when 517 samples are included in the calibration set (old) for all three spectral types.

S4. Hybrid base case calibrations

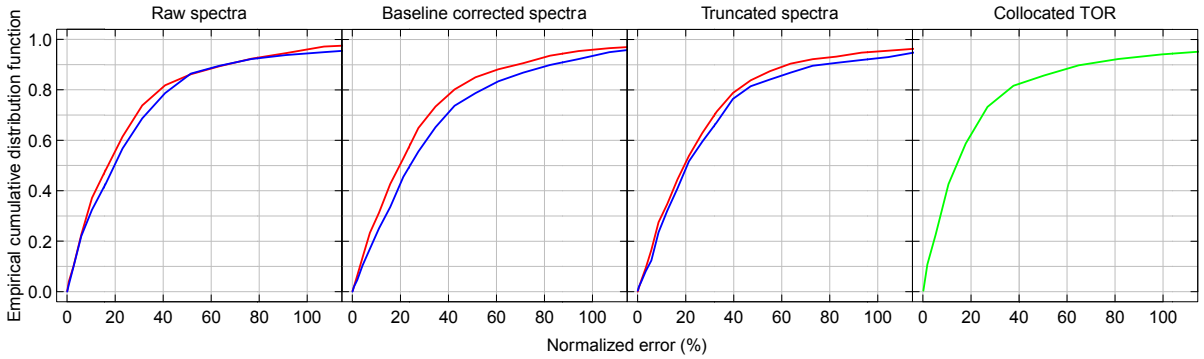

Figure S4a. Distributions of normalized error for the hybrid base case for each spectral type and precision for collocated TOR EC samples. Calibration set is in red and the test set is in blue.

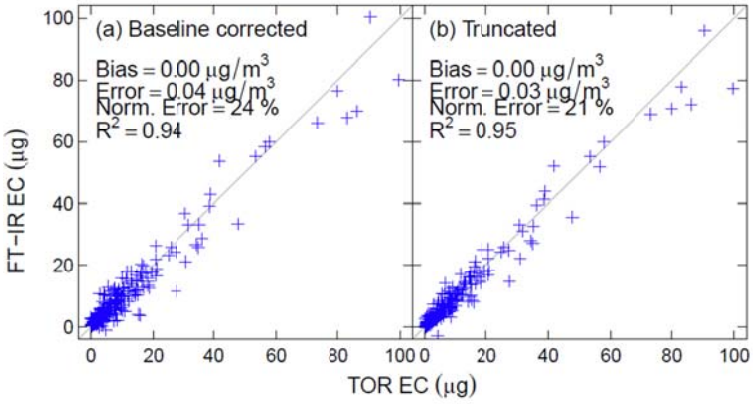

Figure S4b. Predicted FT-IR EC versus measured TOR EC for the hybrid base case test set with (a) baseline corrected and (b) truncated spectra. Concentration units of $\mu g/m^3$ for bias and error are based on the IMPROVE nominal volume of 32.8 m^3.

S5. Samples and blank filters in calibration

To evaluate the number of samples and blanks needed to provide good predictions of the test and MDL respectively, the uniform hybrid EC calibration is used. The number of blanks in the calibration is varied from zero to 36 to evaluate their impact on MDL. The samples and blanks in the test set and the samples in the calibration set remained constant for each case. The MDL does not correlate with the number of blanks in the calibration set as shown in Figure S5a.

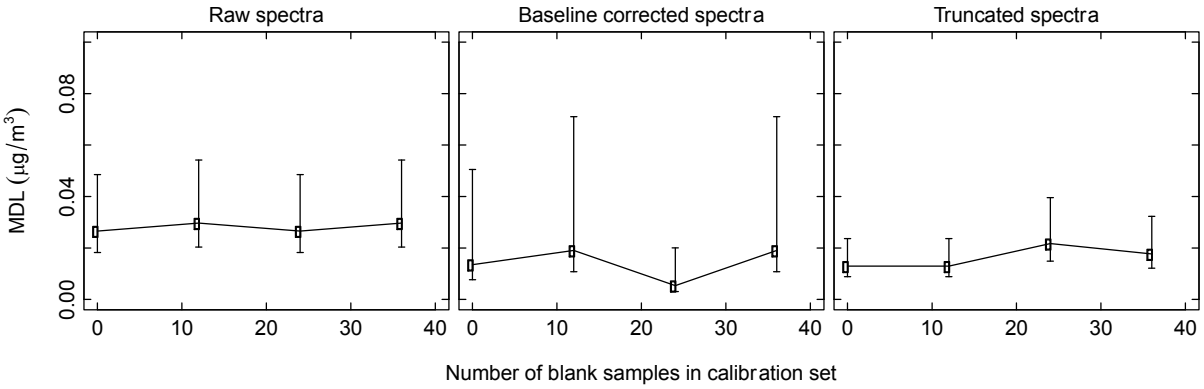

Figure S5a. MDL calculated from the same 16 blank filters in the test set with the number of blanks in the calibration set ranging from zero to 36. Ambient samples in the calibration and test set are the same for each case.

The number of samples in the calibration was varied from 507 samples (2/3 of the total samples) to less than 200 (less than 1/3) to evaluate the quality of calibration with decreasing calibration samples. Figure S5b shows that the mean error and MDL are independent of the number of filters over this range for the test set used in this analysis.

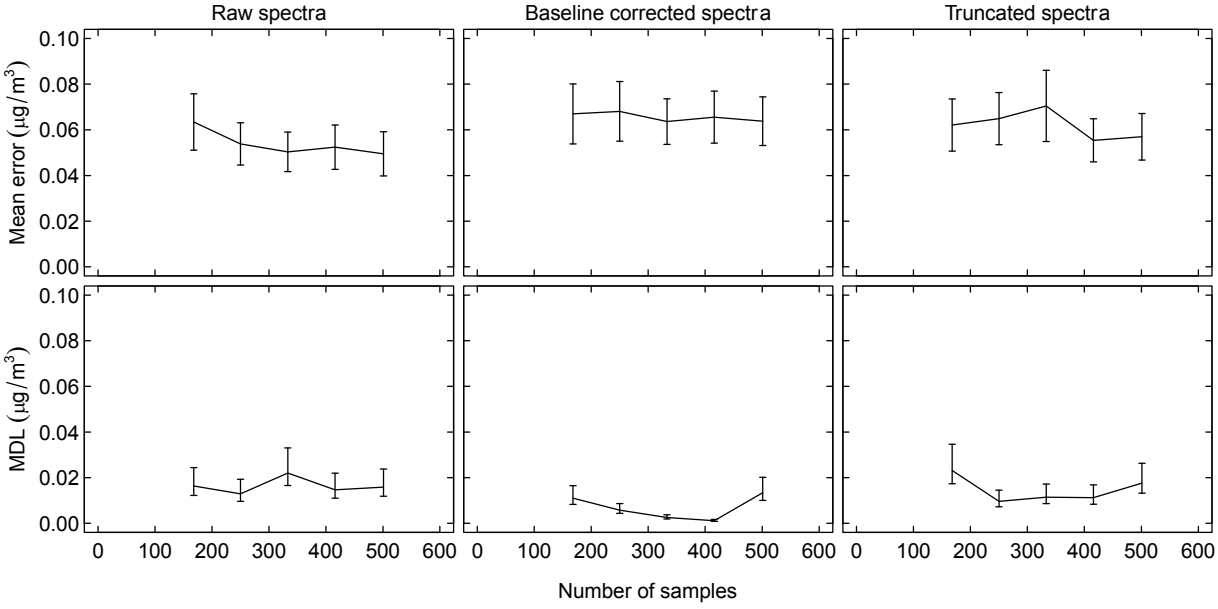

Figure S5b. Mean error and MDL for calibrations using 2/3 of the samples (data point at far right of each subplot) and fewer samples. The samples and blanks in the test set are the same for each calibration. All three spectra types are shown.

S6. Predicting TOR EC using TOR TC minus TOR OC

In evolved gas analysis (Chow et al., 2007) the total carbon (TC) volatilized or vaporized under various temperature and oxidizing conditions is recorded. The respective fractions of OC and EC from TC is determined by the operating conditions under which the gas-phase carbon (C) was measured, with a posteriori correction for an estimated quantity of pyrolyzed or charred carbon (OP) derived from an optical property of the sample (reflectance for TOR analysis). Given the anticipated error in TOR EC, estimation of EC from the subtraction of OC (which has high relative precision) from TC (which does not suffer from OP correction) predicted by PLS calibration has been considered as an alternative to building a calibration model for TOR EC directly.

However, we conclude that the direct fitting approach outlined in this manuscript is the more accurate approach for quantification of EC, as 1) an estimation method developed with TOR OC already includes the estimated OP fraction, and 2) subtraction of OC from TC predicted by PLS will lead to large uncertainties in EC owing to a combination of two sets of prediction errors. Let us define the various forms of carbon reported by TOR:

$$OC = \sum_{i=1}^{4} C_i + OP$$
$$EC = \sum_{i=5}^{7} C_i - OP$$
$$TC = \sum_{i=1}^{7} C_i = OC + EC \quad ,$$

and EC* as the estimate of EC obtained by subtraction of OC from TC:

$$EC^* = TC - OC = \sum_{i=5}^{7} C_i - OP \quad .$$

Next, let us consider estimation errors. For each variable X, we denote X_{TOR} as the reported TOR measurement to which a PLS model is calibrated. Each quantity can be represented by its predicted value according to PLS ($\hat{X}$), and its residual ($\delta X = \hat{X} - X_{\mathrm{TOR}}$), which represents fitting or prediction error:

$$OC_{\mathrm{TOR}} = \hat{OC} + \delta OC$$
$$EC_{\mathrm{TOR}} = \hat{EC} + \delta EC$$
$$TC_{\mathrm{TOR}} = \hat{TC} + \delta TC \quad .$$

Letting $\mathbb{E}\{\cdot\}$ represent the expectation operator, we establish that we can construct calibration models such that $\mathbb{E}\{\hat{OC} - OC_{\mathrm{TOR}}\}$, $\mathbb{E}\{\hat{EC} - EC_{\mathrm{TOR}}\}$, and $\mathbb{E}\{\hat{TC} - TC_{\mathrm{TOR}}\}$ are within TOR measurement precision (Dillner and Takahama, 2015 and main body of this paper); PLS predictions are shown in Figure

S6a, which are on the order of TC precision of 0.12 μg/m^3). Therefore, we expect that, on average $\hat{EC}^* \approx \hat{EC} \approx EC_{TOR}$.

Next, we compare δEC^* and δEC. Linear propagation of uncorrelated errors (Skoog et al., 2006) suggests that the estimated error of EC from subtraction is a combination of the errors from TC and OC such that $\mathbb{E}\{(\delta EC^*)^2\} = \mathbb{E}\{\delta TC^2\} + \mathbb{E}\{\delta OC^2\}$. Letting $\sigma\{X\} = \left(\mathbb{E}\{\delta X^2\}\right)^{1/2}$ (also referred to as the root mean square error, or RMSE), the expected value of the error can be written in a more familiar form:

$$\sigma\{EC^*\} = \left(\sigma^2\{TC\} + \sigma^2\{OC\}\right)^{1/2}$$

Since errors for EC and TC are correlated, the sum on the right hand side is likely to be an overestimate. However, individually, the magnitude of both $\sigma\{TC\}$ and $\sigma\{OC\}$ are larger than $\sigma\{EC\}$ (Figure S6b), so we can anticipate that $\sigma\{EC^*\} \geq \sigma\{EC\}$. Predictions of EC* compared against observations are shown in Figure S6c, and comparison of estimation errors for the two estimation methods for EC in Figure S6d confirm that $\sigma\{EC^*\} \geq \sigma\{EC\}$.

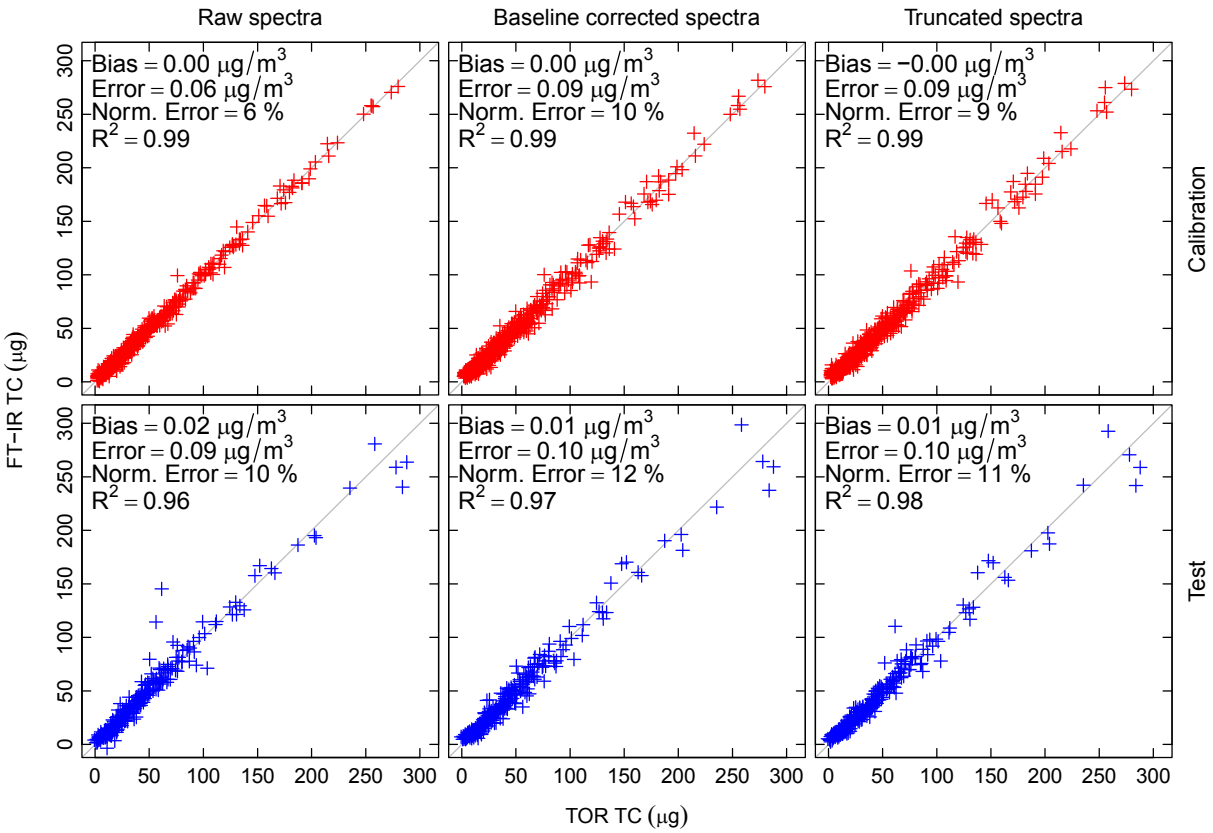

Figure S6a. Comparison of FT-IR and TOR TC for calibration samples (top row) and test set samples (bottom row) for each type of spectra preparation (shown along columns) in the Base case scenario.

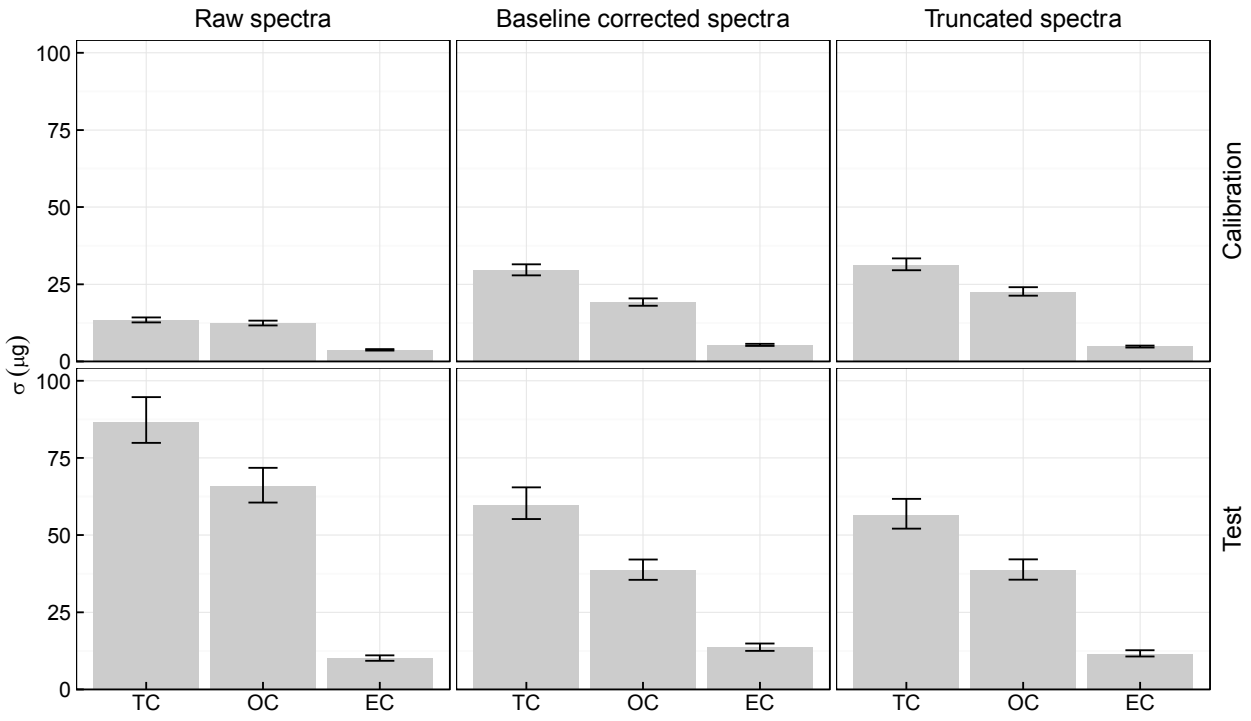

Figure S6b. RMSE values (between predicted and observed concentrations) for each type of variable and spectra preparation (shown across columns) for calibration and test set samples (shown across rows) in the Base case scenario.

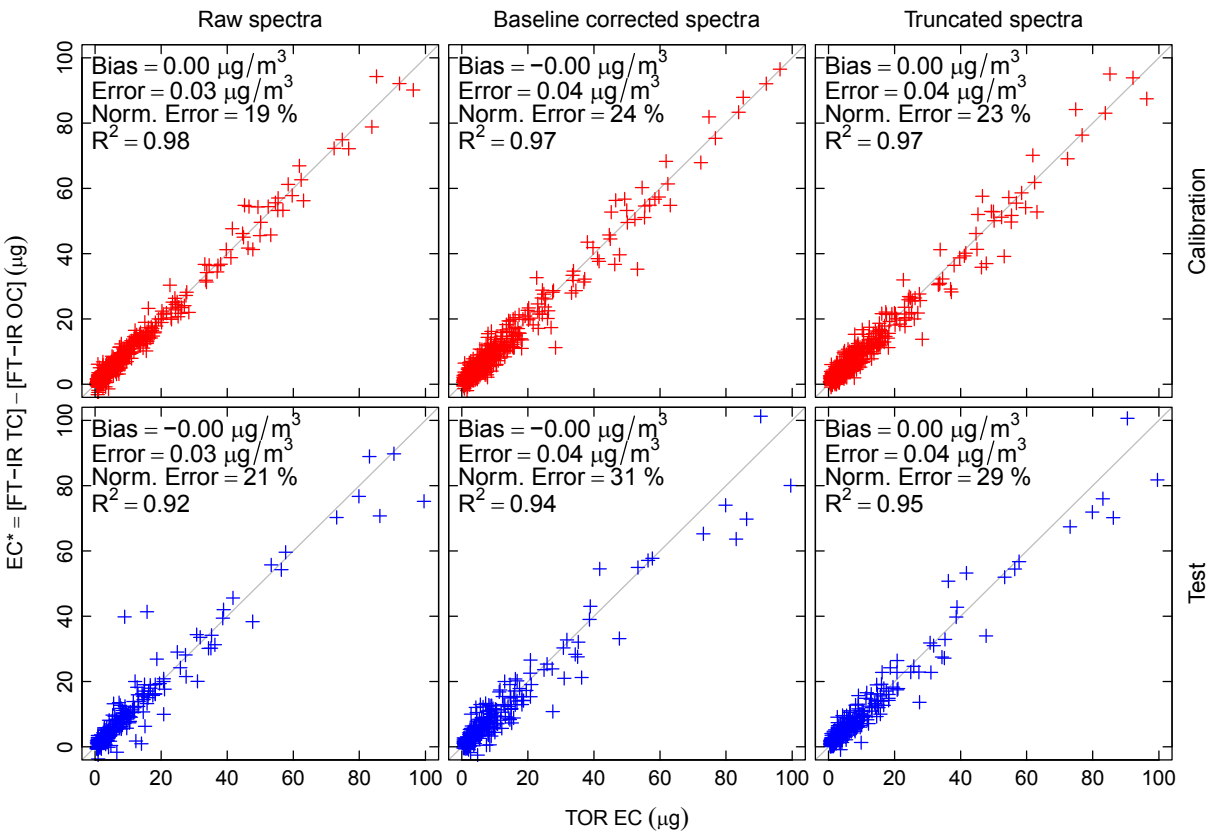

Figure S6c. Comparison of estimated EC* from TC minus OC and TOR EC for calibration samples and test set samples (shown along rows) for each type of spectra preparation (shown along columns) in the Base case scenario.

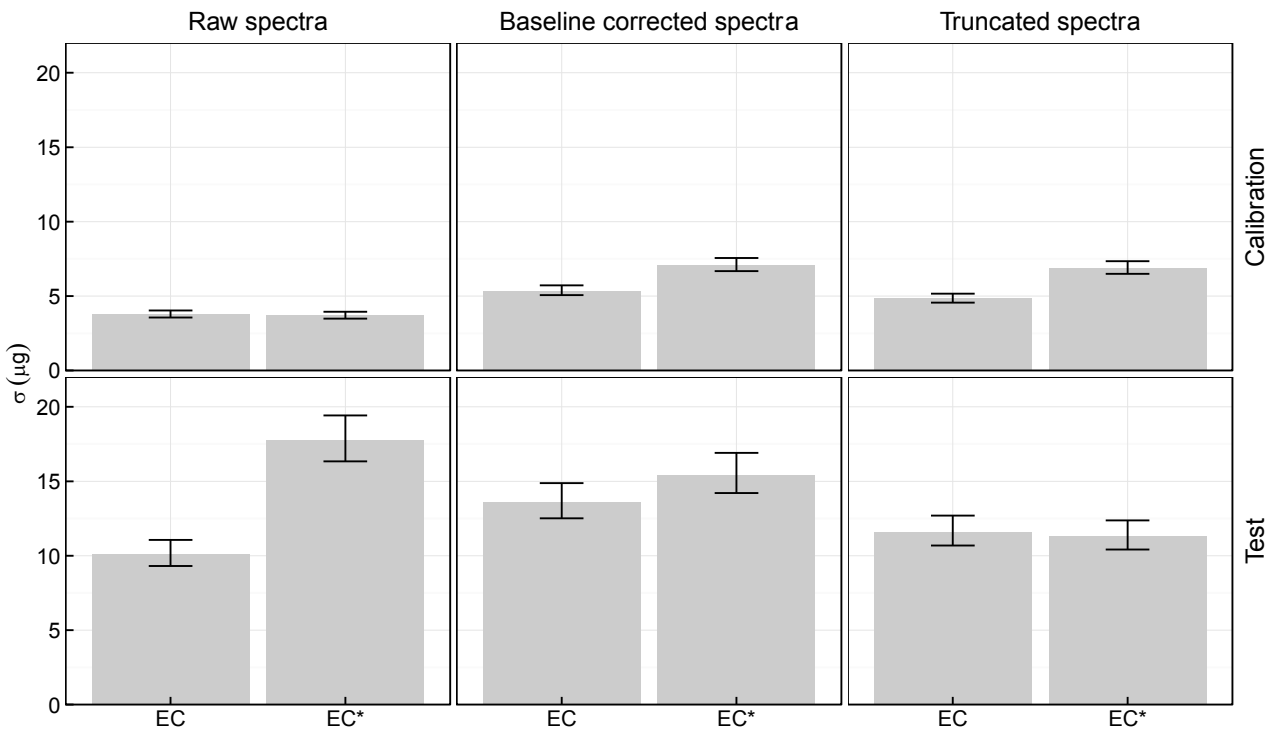

Figure S6d. RMSE (predicted minus observed) of EC by direct calibration, and EC* obtained from subtraction of FT-IR OC from TC in the Base case scenario. Results from calibration and test set samples shown in top and bottom and top rows, respectively, and results from spectra preparation shown across columns.

S7. Probability density functions of EC, OC/EC and Ammonium/OC for the test and calibration sets for all calibrations developed.

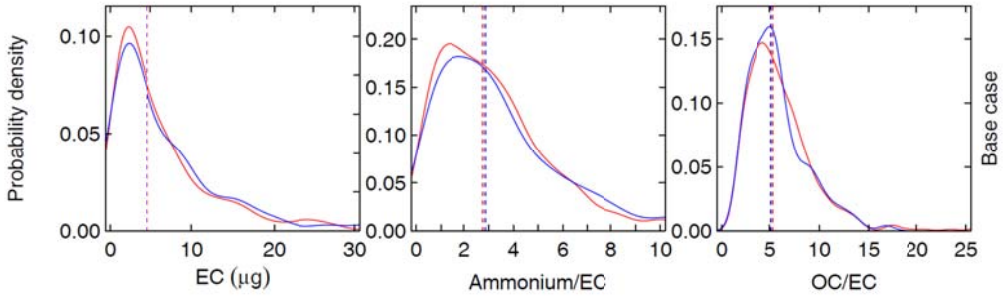

Figure S6a. Distribution of EC, Ammonium/EC and OCEC for the Base case.

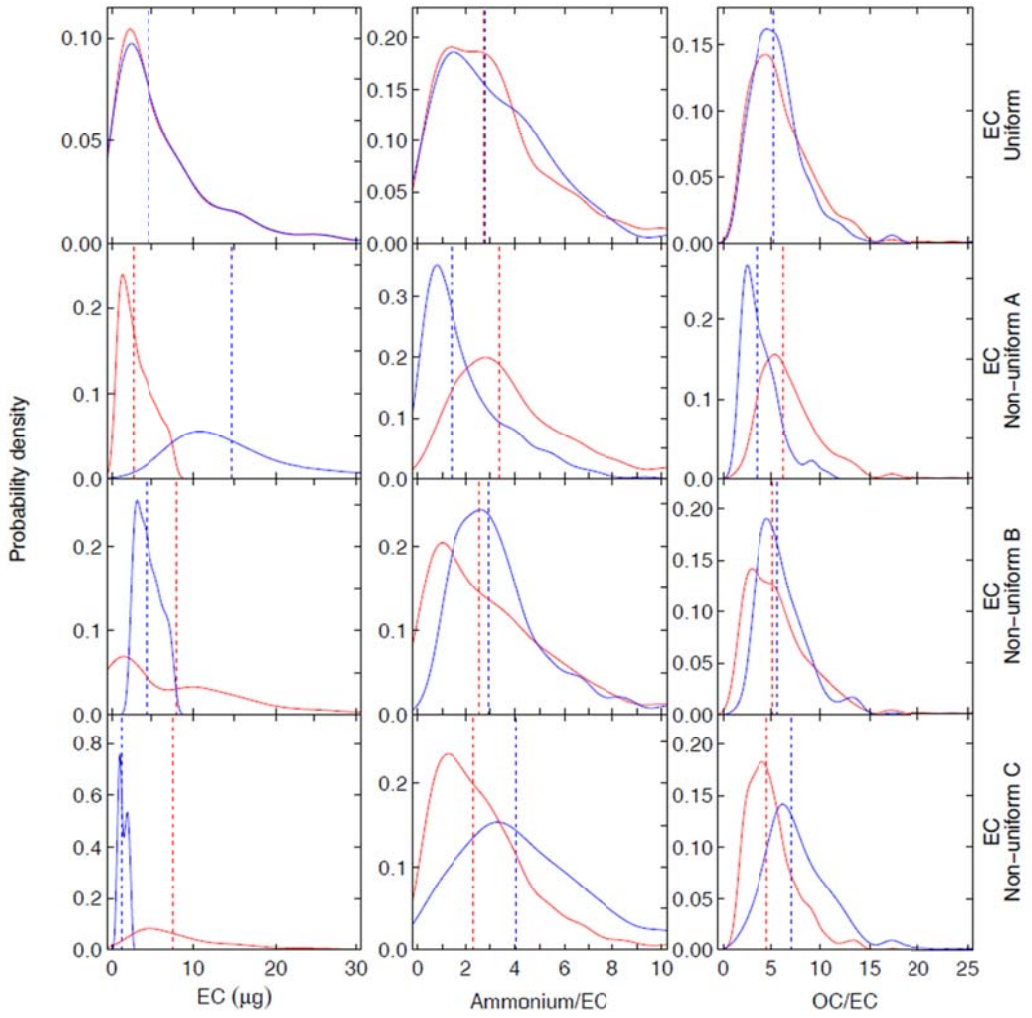

Figure S6b. Distribution of EC, Ammonium/EC and OCEC for the Uniform and Non-Uniform EC cases.

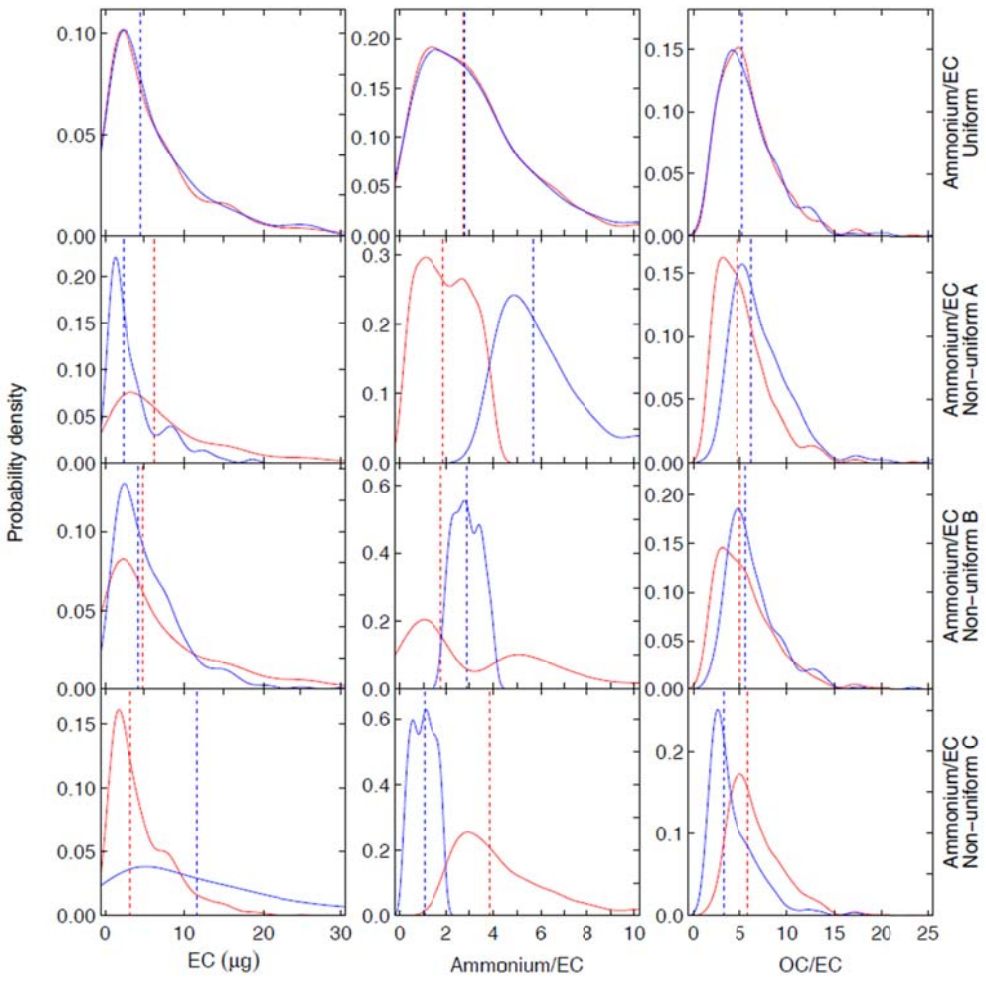

Figure S6c. Distribution of EC, Ammonium/EC and OCEC for the Uniform and Non-Uniform Ammonium/EC cases.

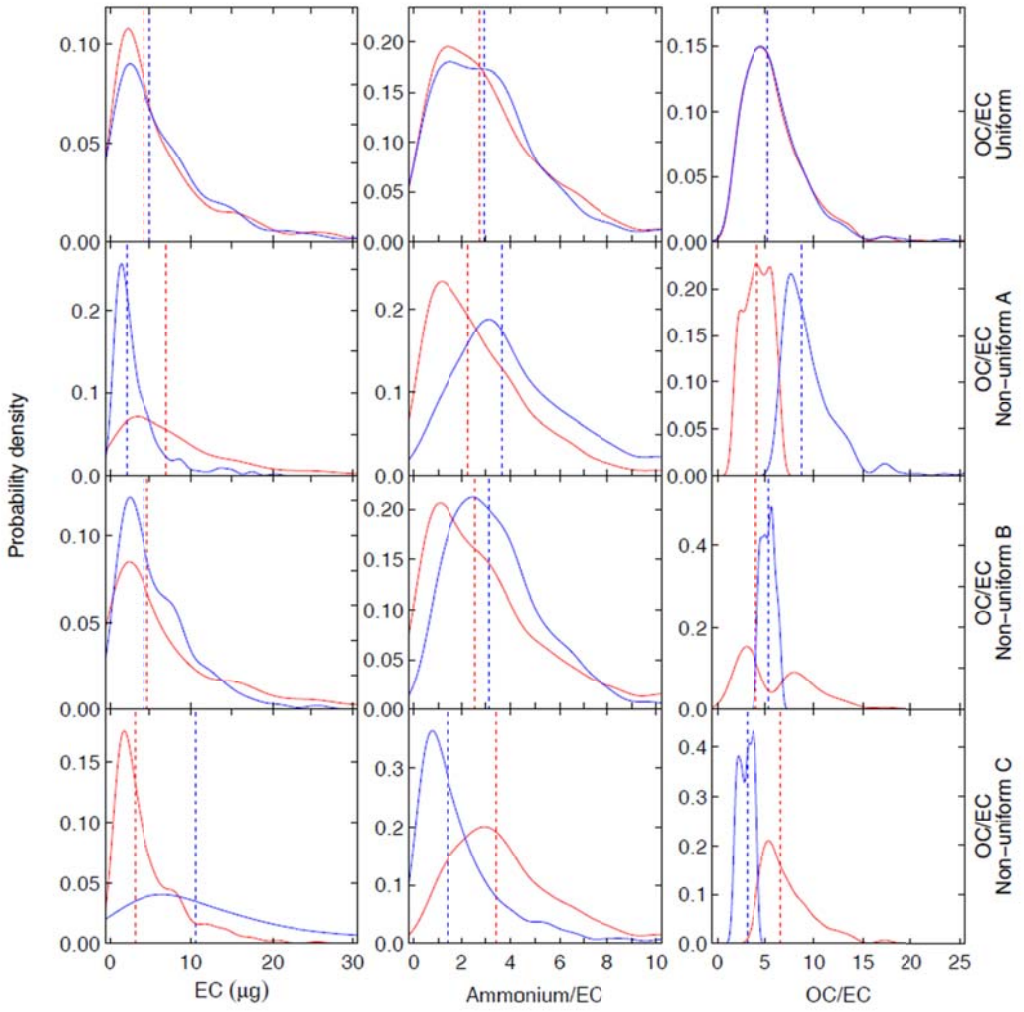

Figure S6d. Distribution of EC, Ammonium/EC and OCEC for the Uniform and Non-Uniform OC/EC cases.

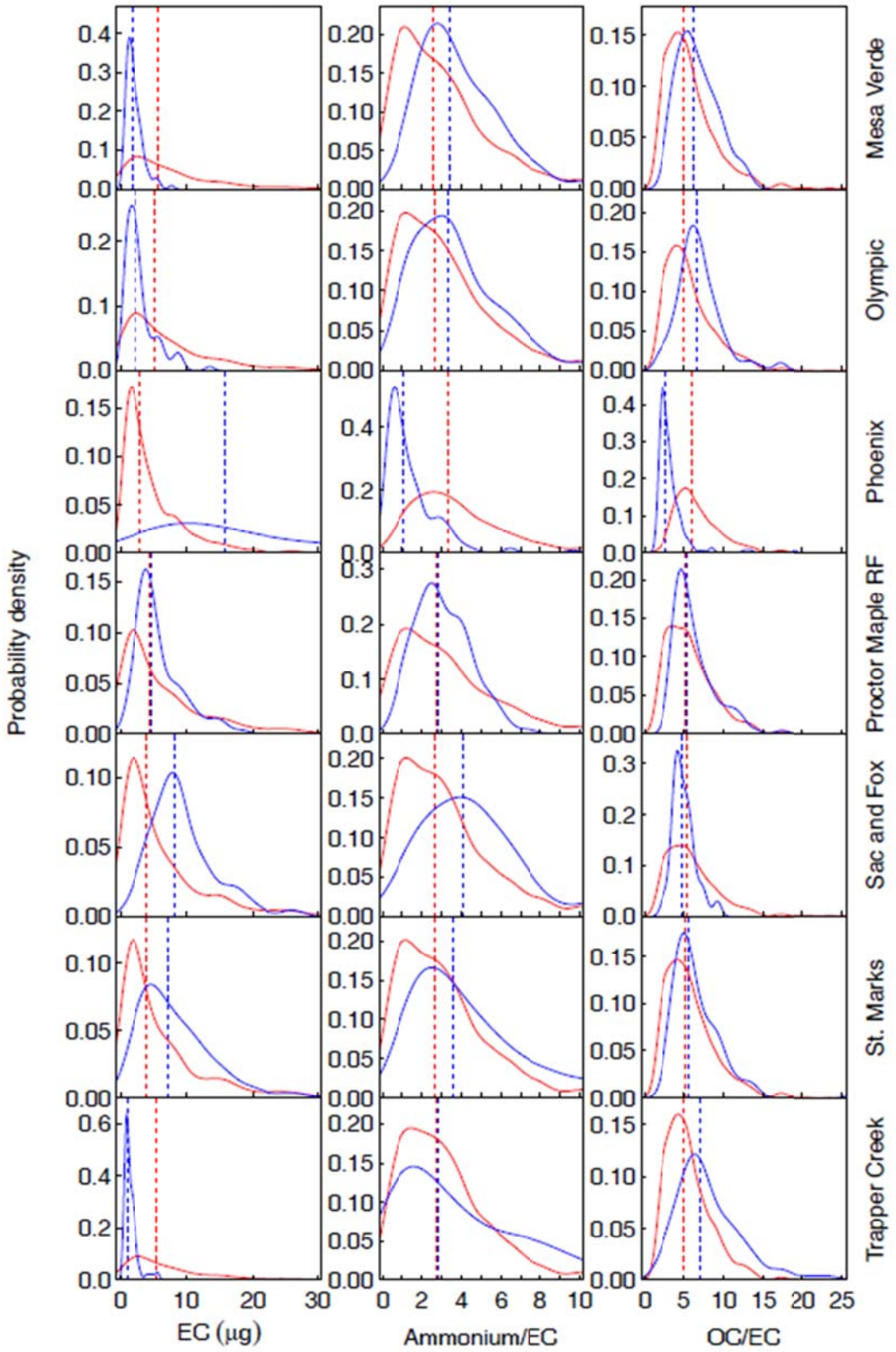

Figure S7e. Distribution of EC, Ammonium/EC and OCEC for the Sites cases.

S8. Residuals

In this section, we examine the structure of residuals that are not reflected in fit metrics, such as mean bias or mean error, discussed in the main text. The reason for examining this structure is to examine the congruence of PLS model assumptions with our specific application, as summarized by Dillner and Takahama (2015) (Section S8). We find that as with the Uniform OC calibration models (Dillner and Takahama, 2015), residuals in the Uniform EC calibration are heteroscedastic (Figure S8a) lead to a long-tailed distribution (Figure S8b) when compared against a normal distribution (*p*-values are greater than 0.05, indicating rejection of the hypothesis that the residuals are normally distributed). Even with this residual distribution, fit metrics indicate a suitable predictive capability overall (Section 3.2 from the main body of this manuscript). However, regression models for EC localized or restricted to a smaller range in concentrations (Low Uniform EC) leads to normality in residual distribution (Figure S8b), and improved fitting metrics for low-concentration samples and MDLs as discussed in Section 2.4. This improvement is in contrast with findings for the development of the OC calibration model (where localized regression did not improve prediction quality), indicating the higher sensitivity of EC predictions to the selection of samples for calibration.

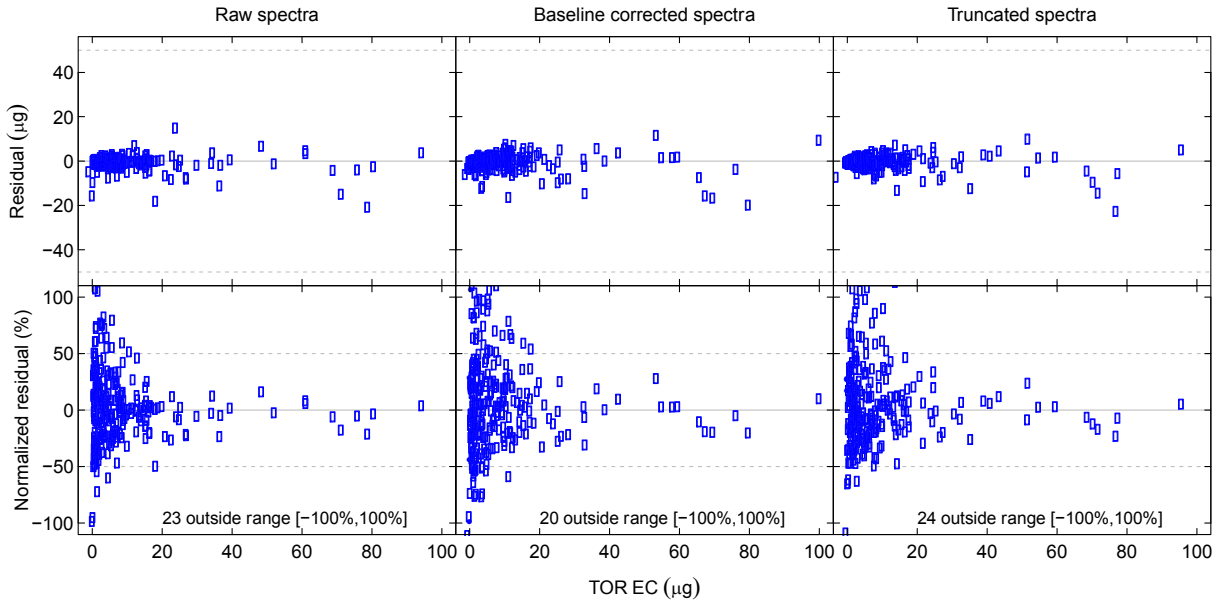

Figure S8a. Absolute and normalized residuals as a function of TOR EC mass concentrations.

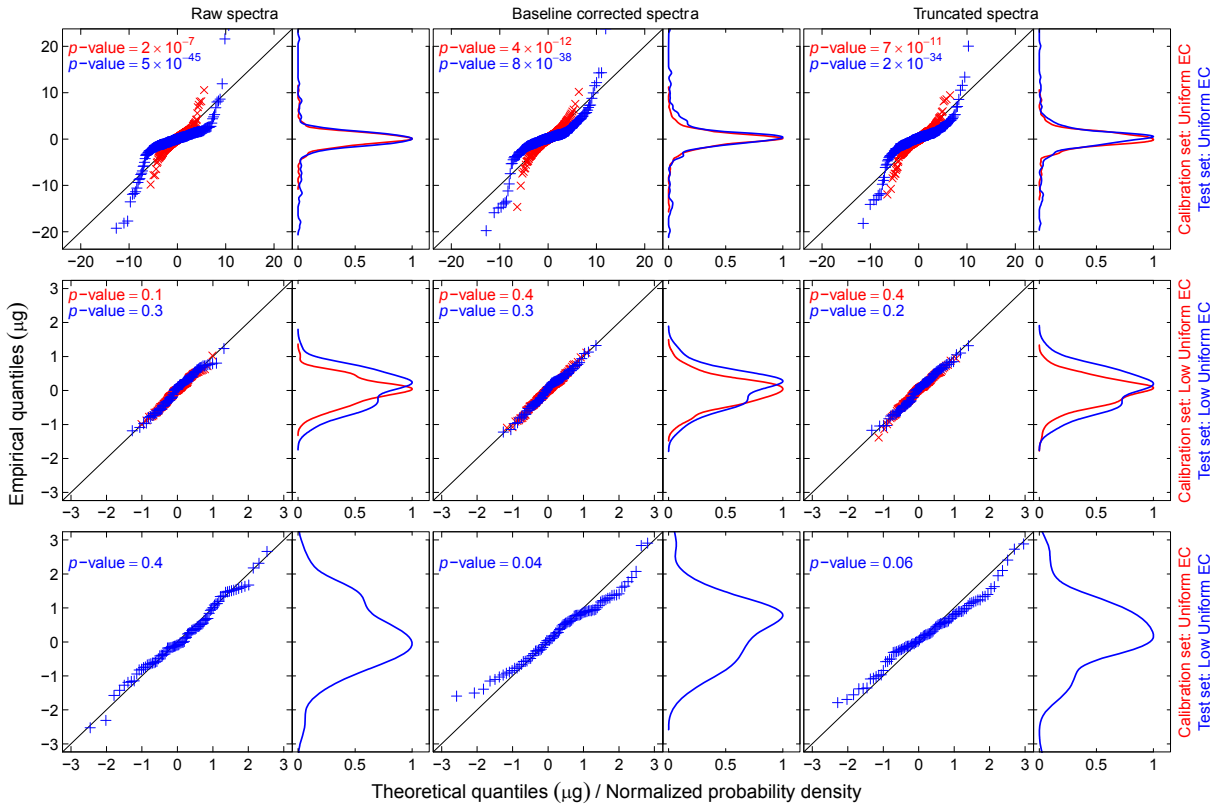

Figure S8b. Distributions of regression residuals, and corresponding comparison of empirical and theoretical quantiles of a normal distribution. Red symbols correspond to calibration set samples and blue symbols correspond to test set samples. *P*-values are calculated from the chi-squared test for normality. Residuals for the calibration set a in the last row are identical to those in the first row and are therefore not shown.

S9. Calibration protocol

In this paper, we show that several methods can be used to calibrate FT-IR spectra for accurate prediction of TOR EC using PLSR, so long as chemical composition of the calibration set spans the range of the chemical composition of the samples to be predicted. We summarize a canonical protocol for calibration and evaluation in this section. It is not necessary to baseline correct or modify the sample spectra produced by the FT-IR, beyond removing absorbance values arising from interpolation during the zero-filling process (which is done to reduce computational cost). We do not include blanks in the calibration model because for this data set, the inclusion of blanks did not impact the MDL calculation. However, blanks can be included as discussed in the methods section of the paper.

1. Select the calibration and test sets by ordering the samples by TOR EC mass and selecting every third sample for the test set and the rest of the sample for the calibration set.
2. Put the calibration set TOR data and raw spectra into a PLS model and select the model with the lowest RMSECV. This is called the Uniform EC calibration.
3. Order the samples in the lowest 1/3 of predicted FTIR EC range and put every third into the “low EC” test set and the rest into the low EC calibration set. Include blanks in the test set.

4. Develop a PLS model with the low EC calibration set and use it to predict the low EC test sets. This is called the Low Uniform EC calibration. Use the predictions from the Low Uniform EC calibration for the low EC test samples. Use the prediction of the blanks to calculate the MDL.
5. Use the predictions from Uniform EC calibration as the predictions for the rest of the samples.

References

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, Journal of the Air & Waste Management Association, 57, 1014-1023, Doi 10.3155/1047-3289.57.9.1014, 2007.

Dillner, A. M., and Takahama, S.: Predicting Ambient Aerosol Thermal Optical Reflectance (TOR) Measurements from Infrared Spectra: Organic Carbon, Atmospheric Measurement Techniques, 8, 1097-1109, 2015.

Skoog, D. A., Holler, F. J., and Crouch, S. R.: Principles of Instrumental Analysis, 6 edition ed., Cengage Learning, Belmont, CA, 1056 pp., 2006.