Atmospheric
Measurement
Techniques

Open Access

EGU

Supplement of

Estimating mean molecular weight, carbon number, and OM/OC with mid-infrared spectroscopy in organic particulate matter samples from a monitoring network

Amir Yazdani et al.

Correspondence to: Satoshi Takahama (satoshi.takahama@epfl.ch)

S1 Mixture effect on absorption peak width

The mid-infrared spectrum of a mixture has particular features compared to that of a pure compound. For instance, a comparison between the spectrum of n-hexane and mineral oil (Nujol) indicated that Nujol, which is a complex mixture of high-molecular-weight alkanes has a simpler mid-infrared spectrum than hexane (Mayo et al., 2004). This is while the number of normal vibrational modes of a molecule increases with the number of atoms existing in the molecule, thus more vibrational modes are expected for Nujol molecules. The simplification of the Nujol spectrum over that of n-hexane, however, can be explained by considering the fact that in the mid-infrared spectrum of a large mixture of compounds, only those absorption bands which have highly consistent wavenumber values absorb above the background level (Mayo et al., 2004). In addition to some absorption bands disappearing in mixture spectrum, the absorbance profile, especially absorption line width, can vary between mixture and pure compound. This effect, which is similar to inhomogeneous broadening in pure compounds, is mainly caused by the slightly different energy levels of the same functional group in different molecules of a mixture.

In this section, the mixture effect on absorption peak width is investigated through a simple statistical simulation. For this purpose, we have assumed that there is a mixture composed of many compounds with hypothetical Lorentzian ($y = \frac{1}{1+((x-\mu)/\sigma)^2}$) absorbance profiles (Kelly, 2013) with varying mean (μ) and standard deviation (σ) (equivalent to peak frequency and width, respectively). The mid-infrared spectrum of this mixture is the superposition of Lorentzian profiles for which a certain distribution of mean and standard deviation is considered ($\mu = 0 + N(0,m)$ and $\sigma = 1 + N(0,n)$, where N is a normal distribution). After superposing and normalizing, the resulting peak is compared to the reference Lorentzian peak ($y = \frac{1}{1+(x)^2}$; Fig. S1). We observe that if the variation of mean (m) is small compared to variation in standard deviation (n), the resulting normalized peak appears to be sharper than the reference profile close to the peak (Fig. S1, right). However, if the variation of mean is very larger compared to peak width ($m >> n$, left) then the resulting peak will be broader (Fig. S1, right). In other words, peak width in mid-infrared spectrum of a mixture may vary form that of the pure compound due to superposition of slightly different peaks.
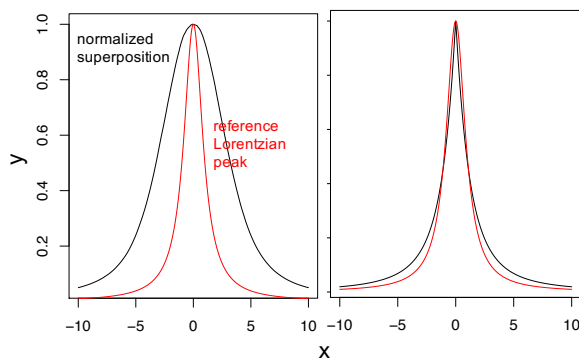

Figure S1. Normalized profile resulted from superposition of 10000 individual Lorentzian profiles ($y = \frac{1}{1+((x-\mu)/\sigma)^2}$ where $\mu = 0 + N(0,m)$ and $\sigma = 1 + N(0,n)$; black profile) and reference Lorentzian profile ($y = \frac{1}{1+(x)^2}$; red profile) when $m << n$ (right), and when $m >> n$ (left).

Previously it was discussed that absorption peak width is also affected by the OM/OC ratio and the phase state. An interesting case of peak width is rural samples which have generally narrower absorption, especially close to the peak, compared to urban ones although they are more oxidized (Ruthenburg et al., 2014; Bürki et al., 2020; Zhang et al., 2007). This suggests that either the mixture effect or the physical phase has dominant narrowing effect on the peak width.

S2 Estimating atmospheric organic aerosol phase state

In this study, the measured melting point of the laboratory standards was taken as reference and standards with melting temperature below the laboratory temperature were considered liquid and vice versa. In order to develop statistical models to estimate phase state of atmospheric aerosols, first, PCA was applied on the normalized aliphatic $C-H$ profiles of the laboratory standards. Thereafter, a logistic regression with a step-wise parameter selection algorithm was used to model their phase state using the PC scores. The resulting logistic regression only uses PC3 scores as can be seen from Fig. S2 (liquid and solid laboratory standards are completely separated based on the PC3 scores).
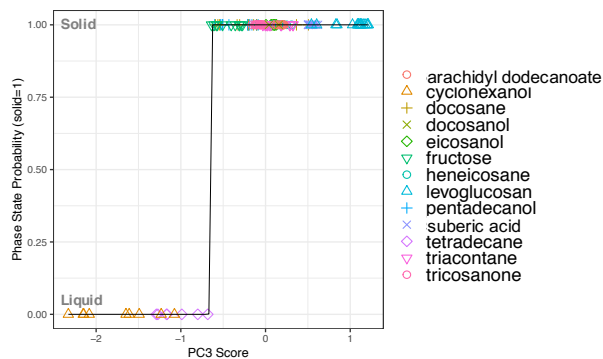
Figure S2. Phase state probability of the laboratory standards (liquid $= 0$, solid $= 1$) against their PC3 scores. Black curve indicates the phase state probability estimated by the logistic regression.

In order to estimate the phase state of the atmospheric samples, their normalized aliphatic $C-H$ profiles were projected onto the PC space to calculate their PC3 scores. The logistic regression predicts that urban samples with lower OM/OC show more liquid-like characteristics (Fig. S3). The glass transition temperature (T_g) (parametrized by Shiraiwa et al., 2017) was also calculated for the atmospheric samples using the estimated mean molecular weight and the O/C ratio of this study. As can be seen from Fig. S3, PC3 scores and calculated T_g are roughly correlated, suggesting a physical connection between T_g and PC3 scores (i.e., spectral features related to phase state). Although the logistic regression appears to capture qualitative phase state behavior, it should be noted that a calibration set with a limited number of samples was used in this study and a more diverse calibration set is suggested for future phase state studies.
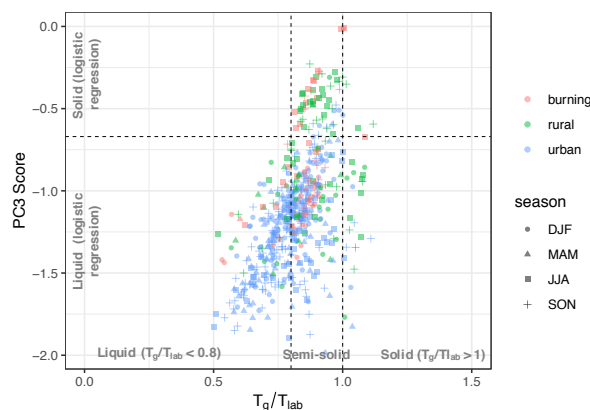
Figure S3. PC3 score (used in logistic regression for modeling phase state) regressed against glass transition temperature (T_g) calculated using the estimated mean molecular weight and the O/C ratio of this study.

S3   Individual samples

In order to have a better understanding of the way the PLSR models function, the spectra of some individual atmospheric samples with different estimated properties were compared in this section. For this purpose, we took the estimates of the PLSR model as reference and compared the spectra of two samples for which one of the carbon number or the OM/OC ratio estimates was almost identical and the other was different. Since only two parameters among molecular weight (MW), carbon number (n_C) and the OM/OC ratio are independent and the third one can be derived from the other two, we essentially observe the effect of moving horizontally and diagonally in the estimated MW-n_C space (Fig. S4) by this choice of samples. This kind of analysis is useful as molecular weight and carbon number are highly correlated in the calibration set and differentiating the factors affecting each is not straightforward.
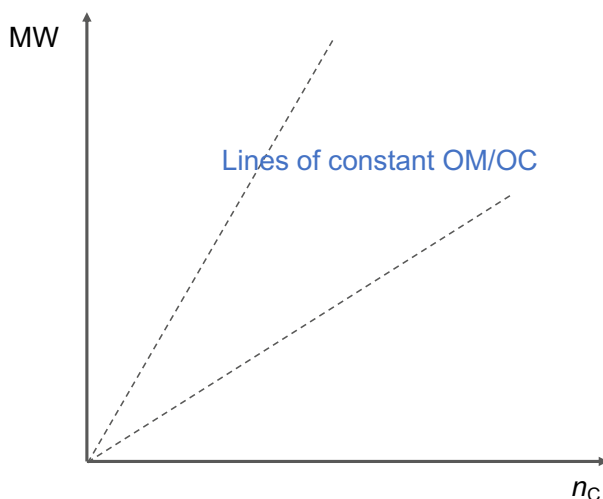
Figure S4. MW-n_C space considering the molecular weight (MW) and carbon number (n_C) of organic aerosols as independent variables. Dashed lines show constant OM/OC ratio lines.

S3.1   Similar OM/OC

Here, two samples having similar estimated OM/OC ratio (≈ 1.4) and slightly different molecular weight and carbon number are considered (Table S1). Both samples have similar peak width and frequency. However, the sample with higher carbon number and molecular weight has a shorter first (A_1) and third peak (A_3) (Fig. S5), which is consistent with what observed in the calibration set. This example shows that A_1 is the main factor used by the models for carbon number estimates.
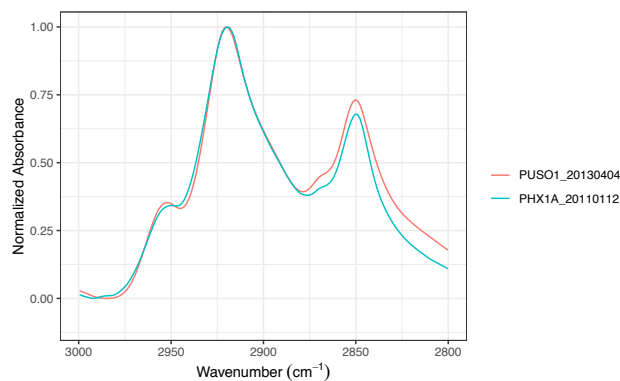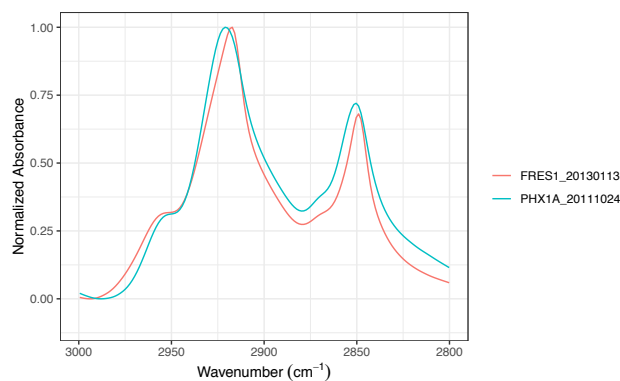
Figure S5. Normalized baseline-corrected spectra of two ambient samples which have similar estimated OM/OC ratios, but different estimated molecular weight and carbon number.

Table S1. Predicted molecular weight, carbon number and OM/OC ratio for individual samples having similar OM/OC ratio.

Site name/date	$\mathbf{MW}$ (g mol^{-1})	n_C	**OM/OC**
Phoenix 2011.01.12	288	17	1.40
Puget Sound 2013.04.04	268	16	1.40

S3.2 Similar carbon number (n_C):

Two samples with similar estimated carbon number and fairly different molecular weight and OM/OC ratio are considered (Table S2). The sample with higher molecular weight has slightly lower peak frequency (Fig. S6) although having the same carbon number. However, this sample has lower frequency which is consistent with observations in the calibration set as well. This shows that molecular weight estimation is probably based on peak ratio (due to its correlation with carbon number) as well as peak frequency besides some unidentified fine features.

Figure S6. Normalized baseline-corrected spectra of two ambient samples which have similar carbon number but different molecular weight and OM/OC ratio.

Table S2. Predicted molecular weight, carbon number and OM/OC ratio for individual samples having similar carbon number.

Site name/date	MW (g mol^{-1})	n_C	OM/OC
Fresno 2013.01.13	280	16	1.5
Phoenix 2011.10.24	257	16	1.4

To conclude, studying the individual spectra shows that the predictions made by the PLS models in atmospheric samples are generally consistent with the trend of the basic features in the calibration set. Nonetheless, there are also instances that the predictions cannot be justified suggesting that the PLS models rely on a combination of basic feature (peak ratio, frequency and width) as well as some fine structures in the spectra that are not captured by basic features.

References

Bürki, C., Reggente, M., Dillner, A. M., Hand, J. L., Shaw, S. L., and Takahama, S.: Analysis of Functional Groups in Atmospheric Aerosols by Infrared Spectroscopy: Method Development for Probabilistic Modeling of Organic Carbon and Organic Matter Concentrations, Atmos. Meas. Tech., 13, 1517–1538, https://doi.org/10.5194/amt-13-1517-2020, 2020.

5   Kelly, A. M.: Condensed-Phase Molecular Spectroscopy and Photophysics, John Wiley & Sons, Inc., Hoboken, NJ, 1st edn., 2013.

Mayo, D. W., Miller, F. A., and Hannah, R. W.: Course Notes on the Interpretation of Infrared and Raman Spectra, John Wiley & Sons, Hoboken, NJ, 2004.

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, https://doi.org/10.1016/j.atmosenv.2013.12.034, 2014.

10

Shiraiwa, M., Li, Y., Tsimpidi, A. P., Karydis, V. A., Berkemeier, T., Pandis, S. N., Lelieveld, J., Koop, T., and Pöschl, U.: Global Distribution of Particle Phase State in Atmospheric Secondary Organic Aerosols, Nat. Commun., 8, 15 002, https://doi.org/10.1038/ncomms15002, 2017.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M., Sun, Y. L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P. F., Salcedo, D., Onasch, T., Jayne, J. T., Miyoshi, T., Shimono, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R. J., Rautiainen, J., Sun, J. Y., Zhang, Y. M., and Worsnop, D. R.: Ubiquity and Dominance of Oxygenated Species in Organic Aerosols in Anthropogenically-Influenced Northern Hemisphere Midlatitudes, Geophys. Res. Lett., 34, L13 801, https://doi.org/10.1029/2007GL029979, 2007.

15