Thèse n° 9382

Chemical characterization of organic aerosols with a focus on biomass burning and mid-infrared spectroscopy

Présentée le 24 janvier 2022

Faculté de l'environnement naturel, architectural et construit
Laboratoire de recherche sur les particules atmosphériques
Programme doctoral en génie civil et environnement

pour l'obtention du grade de Docteur ès Sciences

par

Amir YAZDANI

Acceptée sur proposition du jury

Prof. J. Schmale, présidente du jury
Dr S. Takahama, directeur de thèse
Prof. M. Kalberer, rapporteur
Dr S. Gilardoni, rapporteuse
Dr C. Hüglin, rapporteur

École
polytechnique
fédérale
de Lausanne

2022

Any particle that is in the air or on the land,
look at it carefully, it is charmed like us
— Rumi

To my family and Claire-Lise…

Acknowledgements

First, I would like to thank my Ph.D. supervisor Prof. Satoshi Takahama that gave me the chance to continue my studies in his group. He is indeed one of the most patient and knowledgeable persons I have worked with. He enthusiastically explained different aspects of aerosol science and infrared spectroscopy during our (sometimes very long but interesting) meetings. He taught me a lot regarding the aerosols, spectroscopy, and also the general qualities required as a scientist such as critical thinking. Although Ph.D. studies are stressful in nature, I enjoyed doing research in the past four years predominantly thanks to the atmosphere he created. I also wish to thank Prof. Nenes who gave me the chance to work in his group during the last two years of my studies. He always inspired me to be self confident and motivated. Being a member of his group was very enriching and gave me the opportunity to work on interesting projects and collaborate with some of the most knowledgeable people in the community.
My colleagues and friends Nikunj Dudani, Maria Lbadaoui-Darvas, Pabitro Ray, Andrea Arangio, Matteo Reggente, Rudolf Höhn, Charlotte Bürki, and Yaling Zeng were a great help during my studies. I would like to thank them for our scientific discussions and also the fun moments they created, especially, during coffee pauses. During the last one and a half years, I have been working mainly from home due to the Covid-19 pandemic. For that, I should thank Claire-Lise who created a calm environment at home that helped me continue my research efficiently.

Neuchâtel, December 15, 2021 Amir Yazdani

Abstract

Aerosols play a significant role in the atmosphere through affecting the radiative budget, cloud condensation nuclei activity, and visibility. They also cause adverse health effects leading to premature deaths. A major fraction of aerosols is organic matter (OM), which has a complex composition, is produced by various sources, and undergoes chemical transformation in the atmosphere, making its full characterization difficult. Biomass burning has become a major source of OM with an increasing effect due to more frequent large wildfires around the world resulting partly from climate change.

In this thesis, Fourier transform infrared spectroscopy (FTIR) has been employed in six different projects as the main tool to characterize OM from various sources especially biomass burning in order to understand its formation and evolution in the atmosphere. This technique has been used to analyze the functional group (FG) composition of OM, specific spectral profiles from different sources, and biomass burning marker signatures in a nondestructive and cost-efficient manner. FTIR has also been combined with other analytical techniques. First, we used the spectral profiles in the aliphatic CH region of mid-infrared spectra to extract information about the molecular structure of atmospheric organic aerosols in terms of their mean carbon number and molecular weight. We found that urban, rural, and biomass burning aerosols have distinct mean molecular weights and carbon numbers. FTIR and aerosol mass spectrometry (AMS) were used in the second project to characterize burning aerosols in an environmental simulation chamber. We showed the agreement of these two instruments in terms of the OM mass concentration and elemental ratios (H:C, and O:C). We found that AMS spectra contained functional group information that agreed with that of FTIR even for moderately aged aerosols. In the third project, we used univariate and multivariate statistics to combine FTIR and AMS measurements to better understand and interpret the complex AMS mass spectra in terms of the FG composition and to estimate the high-time-resolution FG composition of combustion aerosols during the course of aging. In another environmental chamber study (fourth project), we developed a procedure to estimate the often-neglected evolution of primary biomass burning aerosols during daytime and nighttime chemical agings. We found that at least 15 % of the primary aerosol mass undergoes chemical transformation at relatively short time scales in the order of a day in the atmosphere. Organic biomass burning markers were among the fastest decaying species, making the identification of aged biomass burning aerosols in the atmosphere challenging. In the fifth project, we used the biomass burning marker signatures in the FTIR spectra for the first time to identify and quantify atmospheric samples affected by wood smoke. The FTIR-based identification method agreed

well with those using ion chromatography and satellite observations. This method, which is one of the few scalable to large air pollution monitoring networks, was applied to around 20,000 filters collected across the US in 2015 to estimate the impact of biomass burning. In the sixth project, we used FTIR to characterize OM emitted from different cookstoves and fuels. We found similarities between the spectra of unburned fuels and OM emissions, and measured high abundances of aromatics and polycyclic aromatic hydrocarbons in the particulate emissions.
Key words: aerosol mass spectrometry, organic aerosol, biomass burning, cookstove, Fourier transform infrared spectroscopy, functional group, environmental simulation chamber.

Résumé

Les aérosols jouent un rôle important dans l'atmosphère en affectant le bilan radiatif, l'activité des noyaux de condensation des nuages, et la visibilité. Ils provoquent également des effets néfastes sur la santé qui causent décès prématurés. Une fraction majeure des aérosols est la matière organique (MO), qui a une composition complexe. MO est produite par diverses sources et subit une transformation chimique dans l'atmosphère. La combustion de la biomasse est devenue une source majeure de MO avec un effet croissant en raison des incendies de forêt plus fréquents dans le monde résultant du changement climatique.

Dans cette thèse, la spectroscopie infrarouge à transformée de Fourier (FTIR) a été utilisée dans six projets différents en tant qu'outil principal pour caractériser la MO à partir de diverses sources, en particulier la combustion de la biomasse. Cette technique a été utilisée pour analyser la composition du groupe fonctionnel (GF) de la MO, les profils spectraux spécifiques de différentes sources et les signatures des marqueurs de combustion de la biomasse de manière non destructive et rentable. Premièrement, nous avons utilisé les profils spectraux dans la région aliphatique CH des spectres IR pour extraire des informations sur la structure moléculaire des aérosols organiques atmosphériques. Nous avons constaté que les aérosols urbains, ruraux, et de combustion de biomasse ont un poids moléculaire moyen et un nombre de carbone distincts. La FTIR et la spectrométrie de masse d'aérosols (AMS) ont été utilisés dans le deuxième projet pour caractériser les aérosols de la combustion dans une chambre de simulation environnementale. Nous avons montré la concordance de ces deux instruments et nous avons trouvé que les spectres AMS contenaient des informations de groupe qui concordaient avec celles du FTIR même pour les aérosols d'âges modérés. Dans le troisième projet, nous avons utilisé des modèles statistiques pour combiner les mesures FTIR et AMS afin d'estimer la composition GF à haute résolution temporelle des aérosols pendant vieillissement. Dans le quatrième projet, nous avons développé une procédure pour estimer l'évolution souvent négligée des aérosols de combustion de biomasse primaire pendant les vieillissements chimiques diurnes et nocturnes, en utilisant AMS et FTIR. Nous avons constaté qu'au moins 15 % de la masse d'aérosols primaires subit une transformation chimique même à des échelles de temps relativement courtes. Les marqueurs de combustion de la biomasse étaient parmi les espèces à la décomposition la plus rapide. Ce dernier rend difficile l'identification de la bbOA âgée dans l'atmosphère. Dans le cinquième projet, nous avons utilisé pour la première fois les signatures des marqueurs de combustion de biomasse dans les spectres FTIR pour identifier et quantifier des échantillons atmosphériques affectés par la fumée de bois. La méthode d'identification basée sur le FTIR concordait bien avec celles utilisant la chromatographie

ionique et les observations satellitaires. Cette méthode, qui est évolutives aux grands réseaux de surveillance, a été appliquée à environ 20 000 filtres collectés aux États-Unis pour estimer la contribution de bbOA. Dans le sixième projet, nous avons utilisé le FTIR pour caractériser la MO émise par différents fourneaux (cookstoves) et nous avons constaté des abondances élevées d'aromatiques et d'hydrocarbures aromatiques polycycliques.
Mots clés : spectrométrie de masse d'aérosols, aérosol organique, combustion de biomasse, cuisinière, spectroscopie infrarouge à transformée de Fourier, groupe fonctionnel, chambre de simulation environnementale.

Contents

Acknowledgements **i**

Abstract (English/Français) **iii**

List of Figures **xiii**

List of Tables **xxi**

1 Introduction **1**

Introduction **1**

1.1 Atmospheric aerosols 1

1.1.1 Atmospheric organic aerosols 1

1.1.2 Biomass burning 2

1.2 Measurement techniques for organic matter 3

1.2.1 Comparison of analytical techniques 4

1.2.2 FTIR spectroscopy 5

1.3 Research objectives and dissertation structure 8

2 Estimating mean molecular weight, carbon number, and OM/OC with mid-infrared spectroscopy in organic particulate matter samples from a monitoring network **11**

2.1 introduction 12

2.1.1 Organic aerosols and measurement methods 12

2.1.2 Aliphatic C–H absorption and the molecular structure 14

2.2 Methods 17

2.2.1 IMPROVE network monitoring sites (sampling and analysis) 17

2.2.2 Laboratory standards (sampling and analysis) 18

2.2.3 Baseline correction and normalization 20

2.2.4 Building the calibration models 20

2.2.5 Interpreting the calibration models using the basic spectral features . . . 21

2.3 Results and discussions 21

2.3.1 Basic features 22

2.3.2 Developing and evaluating the calibration models 27

2.3.3 Applying the calibration models to atmospheric samples 28

2.3.4 Calibration model interpretation 33

2.4 Concluding remarks 35

3 Characterization of primary and aged wood burning and coal combustion organic aerosols in an environmental chamber and its implications for atmospheric aerosols 37

3.1 Introduction 38
3.2 Methods 40
3.2.1 Laboratory experimental set-up and procedure 41
3.2.2 On-line PM measurement 42
3.2.3 Off-line PM sampling and measurement 43
3.2.4 Atmospheric samples (IMPROVE network) 43
3.2.5 Post-processing of MIR spectra to identify and quantify functional groups in laboratory and IMPROVE samples 44
3.2.6 Dimensionality reduction of AMS mass spectra 45
3.3 Results and discussions 46
3.3.1 Wood burning – functional group composition 46
3.3.2 Wood burning – primary signatures 49
3.3.3 Coal combustion – functional group composition 50
3.3.4 Coal combustion – primary signatures 51
3.3.5 Comparison between AMS mass spectra and FG composition of OA . . . 51
3.3.6 Atmospheric biomass burning-influenced aerosols 54
3.4 Concluding remarks 60

4 Fragment ion-functional group relationships in organic aerosols using aerosol mass spectrometry and mid-infrared spectroscopy 63

4.1 Introduction 64
4.2 Methods 66
4.2.1 Laboratory experimental set-up and procedure 66
4.2.2 Online AMS PM measurements 66
4.2.3 Offline MIR PM measurements 67
4.2.4 Identifying influential MIR absorbances for AMS OM 67
4.2.5 Identifying FG-ion fragment relationships 69
4.2.6 Interpolating FG abundances using AMS mass spectra 69
4.3 Results and discussions 70
4.3.1 Combination and comparison of OM measurements 70
4.3.2 AMS fragment ion-MIR FG relationships 75
4.3.3 MIR FG interpolation using AMS mass spectra 82
4.4 Concluding remarks 84

5 Chemical evolution of primary and formation of secondary biomass burning aerosols during daytime and nighttime 87

5.1 Introduction 88
5.2 Methods 90
5.2.1 Experimental set-up and procedure 90

5.2.2 On- and off-line PM measurements . . . 90
5.2.3 Data analysis . . . 91
5.3 Results . . . 93
5.3.1 Evolution of primary biomass burning aerosols . . . 93
5.3.2 Characterization of oxidized biomass burning aerosols . . . 98
5.3.3 Implications for atmospheric aerosols . . . 101
5.4 Conclusion . . . 103

6 Identification of smoke-impacted $PM_{2.5}$ samples with mid-infrared spectroscopy in a monitoring network 105
6.1 Introduction . . . 106
6.2 Methods . . . 108
6.2.1 Sampling $PM_{2.5}$ aerosols in the IMPROVE network . . . 108
6.2.2 Fire and smoke plume detection using remote sensing . . . 109
6.2.3 Smoke detection using HPAEC-PAD and WSOC . . . 109
6.2.4 FTIR analysis, post-processing, and smoke detection . . . 111
6.2.5 Development of a smoke classifier . . . 111
6.3 Results and discussions . . . 112
6.3.1 Selectivity and LOD of FTIR signatures of bbOA marker . . . 112
6.3.2 Comparison of FTIR and satellite smoke detection . . . 115
6.3.3 Comparison of FTIR and HPAEC-PAD smoke marker measurements . . . 117
6.3.4 Analysis of IMPROVE samples from 2015 using the FTIR smoke detection method . . . 121
6.3.5 Indications of aged bbOA . . . 123
6.4 Concluding remarks . . . 125

7 Quantified Functional Group Compositions in Household Fuel Burn Emissions using FTIR 127
7.1 Introduction . . . 128
7.2 Materials & Methods . . . 129
7.2.1 Cookstove emissions testing & sampling . . . 130
7.2.2 Measurements . . . 131
7.2.3 Post-processing . . . 132
7.3 Results & Discussion . . . 133
7.3.1 OC, EC, and $PM_{2.5}$ emission factors . . . 134
7.3.2 FTIR spectra . . . 135
7.3.3 Influential group frequencies for OC . . . 138
7.3.4 Quantifying functional groups from FTIR spectra . . . 139

8 Conclusions and outlook 143
8.1 Summary and conclusion . . . 143
8.2 Outlook and improvements . . . 144

A Appendix **147**
A.1 Appendix to chapter 2 . . . 148
A.1.1 Correlation matrix heat map . . . 148
A.1.2 Relating mixture property to pure compound property . . . 148
A.1.3 Elemental carbon and carbon number . . . 150
A.1.4 Mixture effect on absorption peak width . . . 150
A.1.5 Estimating atmospheric organic aerosol phase state . . . 152
A.1.6 Individual samples . . . 154
A.2 Appendix to chapter 3 . . . 157
A.2.1 Experimental set-up . . . 157
A.2.2 Blank subtraction . . . 157
A.2.3 Dimensionality reduction . . . 158
A.2.4 Fourier self-deconvolution . . . 159
A.2.5 The MIR and AMS spectra and chemical composition of the chamber experiments . . . 160
A.2.6 Aromatics and Polycyclic aromatic hydrocarbons (PAHs) . . . 161
A.2.7 Levoglucosan detection and quantification using MIR spectroscopy . . . 165
A.2.8 Lignin signatures . . . 168
A.2.9 Atmospheric smoke impacted samples . . . 169
A.3 Appendix to chapter 4 . . . 171
A.3.1 Methodology diagrams . . . 171
A.3.2 AMS fragment ion-FG connections . . . 173
A.3.3 PLS models . . . 175
A.4 Appendix to chapter 5 . . . 179
A.4.1 Residual spectrum calculation . . . 179
A.4.2 AMS-FTIR comparison . . . 179
A.4.3 FTIR spectra of primary biomass burning aerosols . . . 181
A.4.4 Investigating lignin-related fragments . . . 181
A.4.5 POA transformation . . . 182
A.4.6 Investigating the volatility of lignin-like compounds . . . 183
A.4.7 Ammonium subtracted spectra . . . 185
A.4.8 f_{44}:f_{43} ratios in residual spectra . . . 186
A.4.9 Dimensionality reduction of AMS mass spectra . . . 187
A.4.10 Atmospheric smoke impacted FTIR spectrum . . . 187
A.5 Appendix to chapter 6 . . . 189
A.5.1 Contribution of fresh bbOA to OM . . . 189
A.5.2 The carbonyl peak as proxy to WSPC . . . 189
A.5.3 Levoglucosan fingerprint in FTIR spectra of laboratory standards . . . 190
A.5.4 Selectivity of levoglucosan fingerprint signatures. . . . 191
A.5.5 Satellite vs. FTIR smoke detection . . . 192
A.5.6 HPAEC-PAD vs. FTIR smoke detection . . . 194
A.5.7 Alternative smoke classifiers . . . 196

A.5.8 Smoke analysis of IMPROVE 2015 data 200
A.6 Appendix to chapter 7 203
A.6.1 WBT protocol 203
A.6.2 EC, OC, and $PM_{2.5}$ emissions 203
A.6.3 Mid-infrared spectra of $PM_{2.5}$ separated by stove type 206
A.6.4 Fuels Analysis, Original Fuel Compositions 207
A.6.5 Using out-of-plane vibrations for quantifying aromatic CH 209
A.6.6 EC information in the FTIR spectra 209
A.7 Appendix to conclusion 213

Bibliography **215**

Curriculum Vitae **243**

List of Figures

1.1 The mechanisms of VOC degradation in the atmosphere. 2
1.2 Absorption wavelengths of different covalent bonds in the infrared spectrum. . 6

2.1 Normalized aliphatic C−H spectra 16
2.2 Diagram showing the relation between spectral features and molecular/physical properties. The way previous approaches (e.g. Ruthenburg et al., 2014; Takahama et al., 2013) and the current approach use mid-infrared spectrum to estimate different parameters is shown in blue and red boxes, respectively. Highlighted molecular properties can only be estimated using the current approach. 17
2.3 The location of IMPROVE sites used for this work (the USA and Alaska); the year at which samples are taken is differentiated by color and the type of the site by point shape. 18
2.4 Percentage of the samples which were recovered from each category (sample type and season) after baseline correction. The number of samples in each category is shown in red. 18
2.5 A sample C−H spectrum showing the convention of peak parameters used in this study 22
2.6 Scatter plot showing the variation of the second peak wavenumber ($\tilde{\nu}_2$) with molecular weight (MW) in the calibration set, affected by the OM/OC ratio and phase state. 23
2.7 Kernel density estimate of second peak wavenumber ($\tilde{\nu}_{\mathbf{2}}$), the ratio of peak heights of symmetric CH_2 to asymmetric CH_2 stretching (A_1/A_2), the ratio of peak heights of asymmetric CH_3 to asymmetric CH_2 stretching (A_3/A_2), and the second peak width (w_2) of the aliphatic C−H band in mid-infrared spectra of the atmospheric samples segregated based on sample type and season. 24
2.8 Scatter plots showing the relation between carbon number (n_C) and the ratio of peak heights of symmetric CH_2 to asymmetric CH_2 stretching (A_1/A_2, upper panel), and the ratio of peak heights of asymmetric CH_3 stretching to asymmetric CH_2 stretching (A_3/A_2, lower panel), averaged for each substance in laboratory standards. 26
2.9 The average value of second peak width (w_2) measured for each compound in the calibration set versus the OM/OC ratio, colored based on compound phase state at laboratory condition (25 °C). 27

2.10 Bi-plots showing the scores of normalized spectra of laboratory standards (color) and normalized spectra of atmospheric samples (filled circles) projected onto the first six principal components. 28
2.11 Scatter plot of fitted (predicted) indirect OM/OC ratio, molecular weight (MW), and carbon number (n_C) against the values from chemical formula of the calibration set (test set). 29
2.12 Bar chart showing median OM/OC ratio calculated for each season based on samples collected in the Phoenix, AZ, monitoring site using our method and the one used by Bürki et al. (2020). 31
2.13 Comparison between carbon-oxygen grid simulated by Jathar et al. (2015) for Atlanta and Los Angeles with sample points estimated for IMPROVE network (2011 and 2013) from the molecular weight and carbon number estimates of this study. The dashed lines show the range of simulated carbon and oxygen and the triangles indicate the location of the highest $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ concentrations for the simulations of Jathar et al. (2015). 33
2.14 Kernel density estimates of indirect OM/OC ratio, molecular weight (MW) and carbon number (n_C) estimated from normalized aliphatic C−H mid-infrared absorbances by PLSR models (segregated by sample type and season). 34
2.15 Regression tree of molecular weight (MW) and carbon number (n_C) estimates in atmospheric samples based on the basic spectral features: second peak frequency ($\tilde{\nu}_2$), the ratio of peak heights of symmetric CH_2 stretching to asymmetric CH_2 stretching (A_1/A_2), the ratio of peak heights of asymmetric CH_3 to asymmetric CH_2 stretching (A_3/A_2) and second peak width (w_2) of aliphatic C−H band. 35

3.1 Averaged OM concentration and OM:OC estimates from MIR and AMS. 47
3.2 The MIR spectra of primary and aged (with OH and NO_3) wood burning (WB) and coal combustion (CC) OAs and their parent compounds. 48
3.3 Loadings of the first three principal components and the normalized mean AMS mass spectrum. 52
3.4 Biplots of PC2–PC1 (a), and PC3–PC1 (b) scores. 55
3.5 Scatter plots of f_{44} against f_{43}, f_{57}, f_{60}, and f_{29}. 56
3.6 Normalized (Euclidean norm) mean spectra of ambient aerosols (rural, urban, residential wood burning, and wildfire) and chamber WB aerosols. 58
3.7 Scatter plot comparing total $PM_{2.5}$ and OM mass in atmospheric samples. . . . 61

4.1 Statistical relations and strategy for comparison of MIR and AMS measurements. 70
4.2 VIP scores of the MIR absorbances regressed against the AMS OM concentration. 72
4.3 Scatter plot comparing OM concentration and OM:OC estimates by AMS and MIR for primary and aged aerosols of wood burning and coal combustion in the environmental chamber experiments. 73

4.4 Comparison of AMS van Krevelen (H:C vs O:C) aging trajectories for wood burning (WB) and coal combustion (CC) experiments with MIR estimates for aerosols collected on PTFE filters before and after aging. 75
4.5 Bar plots showing covariances of normalized AMS fragment ion concentrations and normalized FG abundances. 79
4.6 Bar plots showing the Pearson correlation coefficients of normalized individual fragment ion concentrations and normalized FG abundances. 80
4.7 VIP scores of MIR absorbances regressed against AMS fragment ion concentrations. 81
4.8 Time series of normalized concentration of functional groups interpolated using AMS mass spectra and time series of O:C ratios calculated from the interpolated functional groups. 84

5.1 Wall-loss-corrected and normalized concentrations of individual fragments ions in different aging scenarios. 95
5.2 FTIR spectra of aged aerosols, their apparent primary fraction, and the residual OA for different aging experiments. 98
5.3 Normalized AMS residual spectra of for biomass burning emissions aged with UV lights and the nitrate radical in dry and humid conditions. 99
5.4 Oxidation trajectories of residual spectra of wood burning (WB) and pellet burning (PB) emissions using UV lights and nitrate radical in dry and humid conditions. 101
5.5 PC1-PC2 and PC1-PC3 biplots of primary and aged chamber OA, residual OA, and ambient PMF factors. 102

6.1 The IMPROVE network sites where $PM_{2.5}$ samples of this study were collected. 109
6.2 The calibration curve used to derive the absorption coefficient of levoglucosan (a) and the box and whisker plot showing the variability of absorbances (converted to concentrations) at the location of the three levoglucosan peaks in blank filters after baseline correction and blank subtraction. 114
6.3 Mean spectra of samples classified as smoke-impacted by both satellite observations and FTIR for the prescribed burn and wildfire datasets. The gray band shows the PTFE absorption region.Vertical lines show the signatures of levoglucosan (LG) and lignin-like compounds. 116
6.4 Scatter plot of FTIR equivalent levoglucosan concentrations (from $P_{\mathrm{LG,\,ave}}$) versus the sum of anhydrosaccharides measured by HPAEC-PAD. 118
6.5 Scatter plot of FTIR absorbances related to levoglucosan ($P_{\mathrm{LG,\,ave}}$) and lignin-like compounds (P_{Lig}). 119
6.6 Scatter plot comparing the smoke classifier that uses FTIR smoke signatures and the carbonyl peak to that using WSOC and HPAEC levoglucosan measurements. 120
6.7 Maximum monthly concentrations of primary bbOA in the US in 2015 estimated from FTIR levoglucosan signatures. 122

6.8 Fraction of smoke impacted-samples in each month in 2015 determined by the FTIR smoke classifier. 123
6.9 The averaged normalized spectrum of six clusters of smoke-impacted samples. Gray band show the mean ± one standard deviation. 125

7.1 Flow chart from sample collection to measurements, post-processing, and results. 129
7.2 Photo showing 15 stoves used in this work. 130
7.3 Artifact-corrected emission factors (mg/MJ) of $PM_{2.5}$ (gravimetric), TOT OC, and EC separated by fuel type and test phase. 134
7.4 Average mid-infrared spectra of unburned fuels and their particulate emissions separated by source and phase. Emission spectra are shown in terms of emission factors. 136
7.5 VIP scores of TOT OC regressed against baseline-corrected mid-infrared absorbances for each fuel type. 138
7.6 OM emission factors and OM:OC ratios calculated from mid-infrared spectroscopy, separated by functional group contribution, and averaged over each phase. 140
7.7 Scatter plot comparing OC (areal density on filters) estimated from TOT and FTIR. 141

A.1 Correlation matrix heat map (absolute values) of mid-infrared spectra of the laboratory standards in aliphatic C−H region. 148
A.2 Scatter plot showing the relationship between collocated measurements of EC concentration and carbon number estimates by PLSR models in the IMPROVE network in 2011 and 2013. 150
A.3 Normalized profile resulted from superposition of 10000 individual Lorentzian profiles. 151
A.4 Phase state probability of the laboratory standards (liquid = 0, solid = 1) against their PC3 scores. 152
A.5 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. 153
A.6 MW-n_C space considering the molecular weight (MW) and carbon number (n_C) of organic aerosols as independent variables. 154
A.7 Normalized baseline-corrected spectra of two ambient samples which have similar estimated OM/OC ratios, but different estimated molecular weight and carbon number. 155
A.8 Normalized baseline-corrected spectra of two ambient samples which have similar carbon number but different molecular weight and OM/OC ratio. 155
A.9 Schematic of the experimental set-up used for photo-oxidation and dark oxidation of biomass burning and coal combustion emissions. 157
A.10 Comparison between baseline-corrected (red) and blank-subtracted baseline-corrected spectra (blue) of WB aerosols. 157

A.11 Heat map of correlation matrix of mass fragments. 158
A.12 Example of Fourier self-deconvolution applied to the 1500–1800 cm^{-1} region of FTIR spectra of primary (a) and aged (b) wood burning aerosols using UV lights. 159
A.13 Baseline-corrected spectra of wood burning (WB) and coal combustion (CC) emissions separated by emission source and aerosol type. 160
A.14 Bar plots comparing AMS and MIR OM concentration and OM:OC estimates separated by functional group. 161
A.15 Loadings of the first three principal components and the normalized mean AMS mass spectrum. 162
A.16 AMS normalized mass spectra (shown up to m/z 80) averaged during filter sampling for 20 minutes (left) and normalized mid-infrared spectra (Euclidean norm) of WB (a) and CC (b) on filter. 162
A.17 Biplots of PC2–PC1 (a), PC3–PC1 (b). 163
A.18 Scatter plot comparing MIR absorbance at 754 cm^{-1} (attributed the aromatic CH out-of-plane vibration) and sum of AMS fragment ions attributed to aromatics over the filter sampling periods for each experiment. 164
A.19 Baseline corrected spectra of six different organic compounds containing the aCOH group (e.g. sugars, sugar derivatives, cyclic and straight-chain alcohols). 166
A.20 Scatter plot comparing the MIR absorbances attributed to levoglucosan and AMS $C_2H_4O_2{}^+$ fragment ion concentration averaged over the filter sampling periods. 167
A.21 MIR spectra of hardwood lignin (KBr Pellet) from Boeriu et al. (2004), fresh beech wood burning aerosols (this work), coniferyl alcohol (film) from Boeriu et al. (2004), and some other substituted aromatic compounds (neat) from SpectraBase database (https://spectrabase.com/). 168
A.22 Scatter plot comparing the MIR 1520 cm^{-1} peak absorbance, attributed to lignin, and the AMS $C_8H_9O_2{}^+$ fragment ion concentration averaged over the filter sampling periods ($R^2 = 0.68$). 169
A.23 Scatter plot comparing total $PM_{2.5}$ and OM mass concentration in atmospheric samples. 169
A.24 Box plot comparing potassium concentration of smoke-impacted and non-smoke-impacted samples in the IMPROVE network from X-ray fluorescence measurements. 170
A.25 Diagram showing the MIR and AMS measurements of bulk OM and the initial postprocessing. 171
A.26 Diagram showing how the MIR spectra and AMS OM estimates are combined to determine the influential FGs. 171
A.27 Diagrams showing how (a) MIR FG abundances or (b) MIR spectra are combined with the concentration of AMS fragment ions to determine which FGs produce certain fragment. 172
A.28 Diagram showing how high-temporal-resolution FG compositions are calculated by combining AMS and MIR measurements. 172

A.29 Scatter plot highlighting the correlation between major AMS mass fragments before and after normalization by total AMS OM. 173
A.30 Barplot showing the Pearson correlation coefficients of lignin, levoglucosan and out-of-plane aromatic peaks in the MIR spectra and the concentration of AMS fragment ions. 173
A.31 Barplot showing the Pearson correlation coefficient of $RONO_2$ peak in MIR spectra and the concentration of AMS organic fragment ions. 174
A.32 Scatter plots showing the measured and calculated AMS fragment concentrations using MIR FG concentrations and vice versa. 174
A.33 Scatter plots showing fitted (PLSR models) normalized functional group composition using AMS mass spectra against their values from the MIR peak fitting. . 175
A.34 Plots showing the root mean squared error (RMSE) of predictions in the 10-fold cross validation for different number of latent variables in the PLSR models. . . 175
A.35 Time series of high-resolution stacked absolute functional group concentrations calculated using the AMS mass spectra. 176
A.36 Time series of high-resolution normalized functional group concentrations calculated using the AMS mass spectra. 176
A.37 Comparison between van Krevelen trajectories from the AMS elemental ratios (circle) and elemental ratios reconstructed from the interpolated FGs (cross) for a WB and a CC experiment. 177
A.38 AMS OA concentration and apparent POA concentration using a first-order wall loss for Exp. 5 (a). Residual FTIR spectrum based on the wall loss rate estimated from AMS OA (b). 179
A.39 AMS OA concentration and the first-order fit for Exp. 1 (reference) (a). AMS size-resolved organic mass. The results are based on particle time-of-flight data averaged from -2 to 2 h (b). 180
A.40 (a) OA mass concentration measured by FTIR (functional groups are color codded) and AMS (averaged over filter sampling periods). (b) Scatter plot comparing OA concentrations measured by AMS and FTIR. (c) OM:OC ratios measured by AMS and FTIR. 180
A.41 FTIR spectra of primary wood burning (WB) and pellet burning (PB) aerosols. . 181
A.42 Smoothed time series of AMS OA concentration, and different tracer fragments in a UV experiment (Exp. 4). 181
A.43 Smoothed time series of AMS OA concentration (green) and the percentage of POA transformed (black) in a UV experiment (Exp. 4). 182
A.44 Normalized concentrations of AMS fragments in a reference experiment (Exp. 3). 183
A.45 Smoothed time series of normalized AMS OA and sulfate concentration, and mass fragments attributed to certain tracers for a dark aging (nitrate) experiment in dry conditions. 183
A.46 Normalized concentrations of individual AMS fragments after passing chamber OA through a thermodenuder. 184

A.47 Ammonium-subtracted residual FTIR spectra of wood burning OA aged with UV and nitrate radical (Exp. 4, and 8, respectively). 185
A.48 Time series of OA, OC concentrations, percentage of POA aged, and f_{44}:f_{43} for wood burning emissions in different aging scenarios. 186
A.49 Loadings of the first three PCs. 187
A.50 An example of atmospheric smoke impacted $PM_{2.5}$ samples (prescribed burning in the eastern US 2013, validated by satellite observations; Yazdani et al., 2021e) with high organic loading and high acid signatures and very weak levoglucosan and invisible lignin-like signatures. 187
A.51 Scatter plot of TOR OC and WSOC. The dashed line shows the regression line. Red points are not considered in the regression. 189
A.52 Scatter plot of the carbonyl peak absorbance in the FTIR spectra against WSOC from the carbon analyzer instrument. 190
A.53 Baseline-corrected and blank-subtracted FTIR spectra of laboratory standards of aerosolized (atomized) levoglucosan collected on PTFE filters. 190
A.54 Analysis of interference of the fingerprint signatures of levoglucosan. 191
A.55 Normalized spectra of samples with true positive, false positive, and false negative results from FTIR smoke detection compared to satellite observations. . . . 193
A.56 Levoglucosan region before and after subtraction of inorganic interference in a smoke-impacted sample. 193
A.57 Levoglucosan region in two samples with low levoglucosan loadings. Vertical lines show the expected location of levoglucosan peaks. 194
A.58 FTIR spectrum of an atmospheric sample with visible lignin-like compounds signature, while levoglucosan signatures are not detected due to the low concentration and/or existence of interfering peaks. TOR OC, WSOC, and concentration of levoglucosan from HPAEC-PAD are shown. 194
A.59 HPAEC-PAD sum of anhydrosaccharides versus fitted values using PLSR models on FTIR spectra (a) and regression coefficients of the wavenumbers in the levoglucosan fingerprint region calculated using PLSR. 196
A.60 Comparison of the performance of the smoke classifier that uses FTIR smoke signatures and WSOC on training and test sets. B is defined in Eq. (A.14). 197
A.61 Comparison of the performances of smoke classifiers that use FTIR smoke signatures and WSOC (c–d) or the carbonyl peak (a–b) on training and test sets. . 198
A.62 Scatter plot showing how the smoke classifier classifies samples based on their levoglucosan and lignin signatures and the selected threshold. 199
A.63 Logistic regression for smoke classification. 199
A.64 Monthly maximum absorbance of lignin-like compounds in the FTIR spectra from IMPROVE filter samples in 2015. 200
A.65 The fraction of smoke-impacted samples separated by season. 201
A.66 Kernel density estimates of OM concentrations (a) and OM / $PM_{2.5}$ (b) for smoke-impacted (red) and normal samples (green). 201

A.67 FTIR spectra of fresh wood burning fine aerosols (high LG/OM ratio) from the IMPROVE network (color) and environmental chamber experiments of Yazdani et al. (2021c) (black). 202
A.68 Comparison of OM (a) and bbPOA concentrations (b) for two sites impacted by smoke and their FTIR spectra (c). 202
A.69 Schematic of WBT protocol for each fuel. 203
A.70 The emissions factors, EC/TC, and modified combustion efficiency (MCE; $(\Delta CO_2+\Delta CO)/\Delta CO_2$) separated by fuel type, test phase, and cookstove. 204
A.71 Scatter plot comparing $PM_{2.5}$ (gravimetric) measurements with the sum of artifact-corrected TOT OM (OM/OC × OC, where OM/OC is calculated from FTIR) and EC on quartz fiber filters. 205
A.72 Spectra of three samples from charcoal particulate emissions with significant contributions of inorganic nitrate. 205
A.73 Average mid-infrared spectra of particulate emissions separated by source and phase and stove (normalized by energy delivered to the pot). 206
A.74 Elemental composition of unburned fuels. 207
A.75 Scatter plot of aromatic CH stretching at 3050 cm^{-1} and out-of plane absorbances at 750 cm^{-1}. 209
A.76 Scatter plot of aromatic C=C stretching at 1600 cm^{-1} and out-of plane absorbances at 750 cm^{-1}. 210
A.77 Scatter plot of aromatic CH concentration estimated using the peak at 750 cm^{-1} and GC/MS total PAH concentration. 210
A.78 VIP scores of GC-MS sum of PAHs regressed against FTIR spectra. 211
A.79 Scatter plot of aromatic CH concentration estimated using the peak at 750 cm^{-1} and TOT EC concentration. 211
A.80 Baseline at 4000 cm^{-1} in the FTIR spectra regressed against EC and OC measured by TOT. 212
A.81 The single-bounce ATR-FTIR spectrum of atmospheric fine PM collected on a quartz fiber filter. 213

List of Tables

2.1 Chemicals used in the calibration set to analyze the effect of different physical/chemical properties of organic molecules on aliphatic C−H absorbance profile. 19
2.2 Importance of the principal components 27

3.1 Importance of the principal components 52

4.1 Summary of important fragment ions for each functional group based on the analysis method. Important fragments are shown in blue. 76

5.1 Description of experiments and initial O_3 and NO_2 concentrations. 91
5.2 Summary of wall loss rates for each experiment. 94

6.1 Confusion matrices of FTIR and satellite smoke detection methods for prescribed burns (a) and RIM fire (b) in 2013. 116

A.1 Predicted molecular weight, carbon number and OM/OC ratio for individual samples having similar OM/OC ratio. 155
A.2 Predicted molecular weight, carbon number and OM/OC ratio for individual samples having similar carbon number. 156
A.3 Table comparing satellite and FTIR smoke detection methods for the period of prescribed burns and RIM Fire in 2013. 192
A.4 Summary of alternative classifiers. 196
A.5 Original (pre-burn) fuel analysis. 208

1 Introduction

1.1 Atmospheric aerosols

Aerosol is a system including solid or liquid or multi-phase particles suspended in a gas, i.e., the continuum phase (Seinfeld and Pandis, 2016; Hinds, 1999). However, the word aerosol often only refers to the particles. The atmosphere even in remote areas contains high concentrations of aerosols (up to 10^7 cm^{-3}) with the diameter in the range of few nanometers to 100 µm (Seinfeld and Pandis, 2016). Atmospheric aerosols are believed to cause adverse health effects for example through causing oxidative stress (Bond et al., 2013; Shiraiwa et al., 2017b), reduce visibility by light scattering, and affect the climate and Earth's atmosphere radiative balance directly through light scattering and absorption and indirectly by affecting cloud formation (Seinfeld and Pandis, 2016; Hallquist et al., 2009).

Tropospheric aerosols are composed of sulfate, nitrate, ammonium, sodium, chloride, trace metals, water, crustal elements, and carbonaceous material. Carbonaceous material is composed of graphitic (elemental) carbon (EC) (mainly from combustion) and organic carbon (OC) (Seinfeld and Pandis, 2016). Particles with diameter smaller and larger than 2.5 µm are usually referred to as fine and coarse particles, respectively. Fine and coarse particle often have different formation, and removal mechanisms. Among the species found in atmospheric aerosols, sulfate, EC, and OC are mainly found in fine particles.

1.1.1 Atmospheric organic aerosols

Organic matter (OM) constitutes 20–90 % of the total fine atmospheric aerosol mass and has a very complex chemical composition (Kanakidou et al., 2005). Primary organic aerosols (POAs) are organic aerosols emitted directly either from anthropogenic or biogenic sources. Secondary organic aerosols (SOAs), on the other hand, are produced via oxidation of volatile organic compounds (VOCs) and condensation of low-volatility oxidation products. POA species have also been found to be sufficiently volatile that after dilution in the atmosphere evaporate, react with radicals in the gas phase, and recondense to the particulate phase

(Robinson et al., 2007; Seinfeld and Pandis, 2016).

VOCs in the troposphere react with oxidants (OH and NO_3 radicals, O_3, and Cl) or are transformed by photolysis (Ziemann and Atkinson, 2012). Although the exact mechanisms responsible for the formation of low-volatility organics in the atmosphere are not well understood, generally, these reactions include the formation of alkyl radicals (R·). Alkyl radicals then react to form intermediate species like alkoxy (RO·) and organic peroxy (ROO·) radicals and product species such as hydroperoxides (ROOH), carboxylic acids (RC(O)OH), carbonyls (RC(O)R), alcohols (ROH), organonitrates ($RONO_2$), peroxyacids (RC(O)OOH), and peroxynitrates ($ROONO_2$) (Fig. 1.1; Ziemann and Atkinson, 2012). The addition of oxygenated functional groups leads to the formation of species with a sufficiently low vapor pressure that can potentially form SOA (Kroll and Seinfeld, 2008).

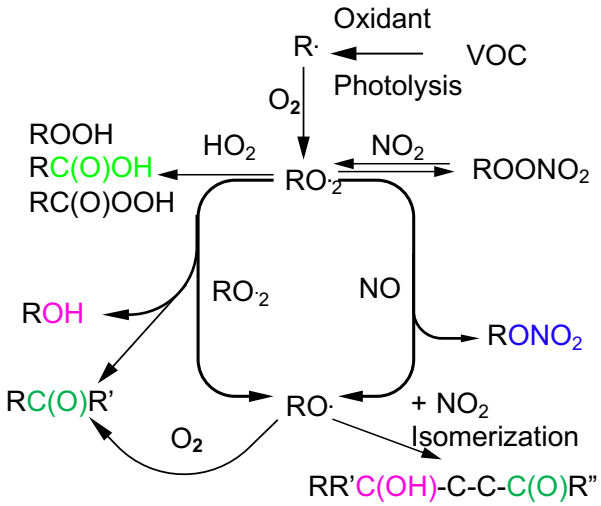

Figure 1.1 – The mechanisms of VOC degradation in the atmosphere, adapted from Ziemann and Atkinson (2012). Functional groups that are studied in this work are shown in color.

1.1.2 Biomass burning

Biomass burning (BB) is the burning of dead or living vegetation (grassland, forest, agricultural waste, or as fuel), which can occur naturally or can be man-made in the form of open fires or residential burning (Yadav and Devi, 2019). BB emissions affect air pollution, visibility and human health (Ford et al., 2018; IPCC, 2013). BB is an important source of POA, black carbon (BC), and brown carbon (BrC) in the atmosphere (Puxbaum et al., 2007; Bond et al., 2013; Andreae and Gelencsér, 2006; Lack et al., 2012). Puxbaum et al. (2007) estimated the contribution of biomass burning POA to total atmospheric OM to be up to 68 % in Europe. BB also emits VOCs (Karl et al., 2007) that can react with atmospheric oxidants to form lower-volatility compounds that potentially condense and form SOA. BB is also believed to be responsible for a large fraction of SOA (up to 70 %), especially in winter when biogenic emissions are absent (Qi et al., 2019; Lanz et al., 2010; Paglione et al., 2020). Biomass burning organic aerosols (bbOAs)

have also been found to have a significant potential for the production reactive oxygen species (ROS), which cause oxidative stress and adverse health effects (Verma et al., 2015).

The frequency of large wildfires have been increasing in recent years partly due to global warming and this trend is expected to continue in the future (O'Neill et al., 2021; Westerling, 2016; Ford et al., 2018). For example, emissions from wildfires in California are predicted to increase by up to 100 % by 2100 above the baseline period (1961–1990) (Hurteau et al., 2014). As a results, BB is expected to become an increasingly important contributor to atmospheric particulate matter (PM), while emissions from other sources are decreasing (Ford et al., 2018).

In order to fully understand the impacts of BB on climate, air pollution, and human health, both primary and secondary bbOAs should be characterized and quantified properly. Primary bbOA has been frequently identified in the atmosphere by using certain organic and inorganic molecular markers like anhydrosaccharides and potassium (Sullivan et al., 2008; Schneider et al., 2006; Ramadan et al., 2000). There are, however, known issues about the stability and specificity these markers. The identification and quantification of secondary bbOA in the atmosphere is even more challenging due to having a different chemical composition compared to the primary bbOA and the lack of complete knowledge about its markers (Nozière et al., 2015).

1.2 Measurement techniques for organic matter

Unlike other species in atmospheric aerosols, organics contain thousands of distinct compounds. This sheer complexity along with the low concentrations of individual compounds and challenges regarding the sampling makes the full characterization of organic aerosols and processes involving them difficult (Hallquist et al., 2009). In order to tackle this problem, analytical techniques from other disciplines have been modified and new instruments have been developed in the last 30–40 years (Nozière et al., 2015). Analytical techniques used for characterization of atmospheric OM include optical (e.g., Fourier transform infrared spectroscopy, FTIR, and UV-visible; Maria et al.,2002) and magnetic spectroscopy (e.g., ^{1}H nuclear magnetic resonance; Paglione et al., 2014), mass spectrometric techniques with soft (e.g., electrospray; Lopez-Hilfiker et al., 2019) or hard (e.g., electron impact; DeCarlo et al., 2006) ionization. These techniques can be combined with separation techniques such as gas chromatography (e.g., thermal desorption aerosol gas chromatograph; Williams et al., 2014) or liquid chromatography (including ion exchange and size exclusion chromatography) to form multidimensional techniques that provide a better characterization of OM. Tandem mass spectrometry (MS^n) is another multidimensional technique to obtain structural information about OM (Nizkorodov et al., 2011). Multidimensional instruments are more suitable for laboratory due to their complexity (Nozière et al., 2015), whereas some one dimensional techniques have been used in online, filed-deployable instruments (e.g., DeCarlo et al., 2006). There is also a class of instruments such as thermal-optical reflectance or transmittance (TOR and TOT, respectively), and water-soluble organic carbon (WSOC) analyzers, which quantify OC or

WSOC in the particle phase without further characterization of their constituents (Hallquist et al., 2009).

1.2.1 Comparison of analytical techniques

Analytical methods are usually classified based on the fraction of OM mass they can analyze, their capability to determine the molecular structure of OM, and the size or time resolution they offer (Hallquist et al., 2009; Nozière et al., 2015). FTIR, for example, is capable of analyzing most (around 80 %) of OM mass (Russell et al., 2009a; Boris et al., 2019). The chemical selectivity of FTIR, on the other hand, especially in the condensed phase is considered to be only on the order of function group (FG) determination (Hallquist et al., 2009). The size and time resolution of FTIR depends mostly on the collection technique but it is often performed offline on filter samples (e.g., Maria et al., 2002; Reggente et al., 2016; Faber et al., 2017; Boris et al., 2019; Debus et al., 2019). NMR is also performed offline and provides information about the FG structure of OM (Decesari et al., 2007), but it is limited to water-soluble fraction (Nozière et al., 2015). The selectivity of NMR analysis is relatively higher than FTIR even in complex atmospheric mixtures (Nozière et al., 2015). Instruments based on mass spectrometric techniques with hard ionization such techniques as aerosol mass spectrometry (AMS) and aerosol chemical speciation monitor (ACSM) offer limited chemical specificity in terms of mass fragments that are common to a broad class of compounds (e.g., oxygenated organics) and are rarely specific to certain molecules. These techniques can analyze most of the non-refractory OM mass with a time resolution on the order of seconds and provide aerosol size information (in case of AMS) (Canagaratna et al., 2007; Ng et al., 2011b). Instruments with soft ionization techniques such as chemical and electrospray ionization reach a higher identification power due to a lower extent of fragmentation of organic molecules but suffer from low sensitivity, are more efficient for ionization of certain compounds, and have variable ionization efficiency based on the sample composition (Nozière et al., 2015; Iyer et al., 2016; Hermans et al., 2017). Thus, these instruments are not suitable for a quantitative and reproducible characterization of total OM mass. Moreover, even soft ionization techniques cannot reach an almost full characterization of OM unless coupled with separation techniques (Nozière et al., 2015). Techniques like gas chromatography-mass spectrometry (GC-MS) provide molecular speciation only for a small percentage of OM mass (on the order of 10 %) (Hallquist et al., 2009).

In summary, a perfect analytical method for the characterization of OM might not exist but a combination of different methods usually offers the desired results. This combination can also help us better interpret measurements of each instrument when they are used alone (e.g., functional group mass fragment relationships, Faber et al., 2017). The combination of techniques can also provide information that is not accessible using each technique alone (e.g., high-time-resolution functional group composition; Yazdani et al., 2021b). Moreover, the full characterization of OM (i.e., identification of all individual species) might not be necessary for the majority of purposes. For example, elemental ratios (H:C, O:C, N:C, and S:C) are commonly used to characterize OM and provide information that allows us to understand

aerosol sources, processes, impacts, and fates (Canagaratna et al., 2015). However, identification and quantification of certain molecules or class of molecules might be necessary for the source apportionment of aerosols (e.g., anhydrosaccharides for biomass burning) or to better understand their health impacts (e.g., mutagenic and carcinogenic polycyclic aromatic hydrocarbons, PAHs, from solid fuel emissions) (Shen et al., 2017b; Sullivan et al., 2008).

1.2.2 FTIR spectroscopy

Most molecules (organic or inorganic) with covalent bonds absorb some wavelengths of electromagnetic wave. The absorption of mid-infrared radiation in the 2.5–25 µm (4000–400 cm^{-1}) range, which is a quantized process, corresponds to stretching (symmetric and asymmetric) and bending (scissoring, rocking, wagging, and twisting) vibrational motions of most covalent bonds. With the absorption of electromagnetic wave with the frequency that matches that of the vibrational motion of the bond, the amplitude of the vibration motion increases. However, only bonds with varying dipole moment with time are able to absorb the corresponding infrared radiation (Pavia et al., 2008). Figure 1.2 shows the typical absorption frequencies of different bonds. The 1300–900 cm^{-1} region, also known as fingerprint region, is usually more complex and has patterns specific to certain molecules (Nozière et al., 2015). The absorption profile of the same bond can also vary to some extent based on the inter- and intra-molecular interactions and the phase state of the compound (Pavia et al., 2008; McHale, 2017). The two aforementioned aspects of FTIR spectroscopy have not received much attention in the study of atmospheric aerosols. In the infrared spectra, fundamental absorptions (with the frequency $\tilde{\nu}$) are produced by excitation of the bond to the excited state with the lowest energy. There are also other absorptions such as overtones (excitation to a higher-energy excited state instead of the lowest one; $n\tilde{\nu}$), combination bands (excitation of a combination of vibrational motions; $\tilde{\nu}_1 + \tilde{\nu}_2$), difference bands (resulting from the difference of two interacting bands; $\tilde{\nu}_1 - \tilde{\nu}_2$), and Fermi resonance (coupling of a fundamental vibration with an overtone or combination band with similar energies) (Pavia et al., 2008). The mentioned bands (except fundamental absorptions) are often weak but can still be informative. For example, Fermi resonance of aliphatic CH is used to identify aldehydes.

Generally, the infrared spectrophotometers used for obtaining the absorption spectra are either based on dispersive (emerged in 1940s) or Fourier transform (FT). FTIR instruments, which became commercially available in 1969, offer a greater speed and sensitivity. Thus, they are more commonly used in today's research (Griffiths and Haseth, 2007).

Since 1950s, infrared spectroscopy has been used to analyze single aerosols or the ensemble of aerosols collected on different substrates (directly or after extraction). Mader et al. (1952) used infrared spectroscopy for the first time to analyze atmospheric aerosols and those generated from gasoline in a plastic chamber exposed to sunlight and identified significant amounts of oxygenated functional groups. In this work, which was published before the advent of FTIR instruments, ether-soluble aerosols were extracted from high-volume filters.

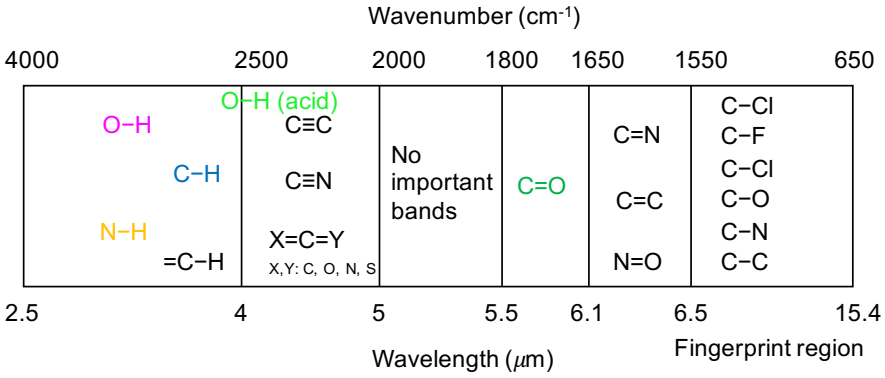

Figure 1.2 – Absorption wavelengths of different covalent bonds in the infrared spectrum, adapted from Pavia et al. (2008) and modified. Functional groups are highlighted with the color scheme used in this study.

Infrared spectra of a single micrometer-sized ammonium sulfate levitated particle was first obtained by Arnold and Pluchino (1982) and Arnold et al. (1984) with a sensitivity in the picogram range. The spectra of single micrometer-sized particles deposited on a substrate (e.g., KBr or ZnSe) have also been obtained using IR microscopy (Allen and Palen, 1989). In this method the extent to which the IR beam can be focused is limited and contaminants on the substrate may cause interference in the spectra (Allen and Palen, 1989). IR microscopy have also been used for the ordinary analysis (not single particle) of atmospheric or smog chamber samples collected directly on KBr or ZnSe impaction disks with detection limits in the low nanogram to high picogram range (Kellner and Malissa, 1989; Allen and Palen, 1989; Palen et al., 1992, 1993). These analyses, however, were limited to determining the mole fraction of functional groups (organonitrate, aldehyde, alcohol, acid, ketone, and aliphatic CH) and were unable to estimate the total mass loading due to an aperture smaller than a typical impactor deposit (Palen et al., 1992).

Allen et al. (1994) analyzed organic and inorganic functional groups in size-segregated ambient aerosols collected on ZnSe impaction surfaces of a Hering low pressure impactor (LPI) directly with transmission mode FTIR. The absorbances corresponding to organic and inorganic functional groups (e.g., sulfate, nitrate, and ammonium) were then integrated and related to the loadings of functional groups using the Beer-Lambert law and calibration constants (absorption coefficients) derived from laboratory standards. The variability of the absorption coefficients (determining the absorbance area per mole of the functional group) of the same functional group from compound to compound has been mentioned to cause uncertainties for the quantitative FTIR analysis, regardless of the method (Allen and Palen, 1989; Reggente et al., 2019a; Takahama et al., 2013).

Attenuated total reflectance (ATR) spectroscopy is another method of infrared spectroscopy used for the analysis of thin layers of aerosols with high sensitivity and the potential of implementation for the real-time analysis. Johnson et al. (1982) used ATR for the characterization

of ambient aerosols. Arangio et al. (2019) coupled electrospray with ATR and achieved a low detection limit (lower than 100 ng) and a linear response of absorbance versus deposited mass for low film thicknesses of dissolved organic and inorganic materials. This method is suitable for FTIR analysis on extracted organics from infrared-opaque filters (e.g., quartz fiber filters).

FTIR has also been used to detect and quantify organic and inorganic species collected on filters. The direct FTIR analysis on filters is nondestructive and does not need prior sample preparation. However, it has certain disadvantages due to absorption and scattering of the filter substrate in the mid-infrared range and sampling artifacts (Subramanian et al., 2004). Quartz fiber filters are often considered unsuitable for FTIR analysis in transmission mode, due to their strong absorption and scattering in the mid-infrared range (Allen and Palen, 1989). Nonetheless, quartz filters have been quantitatively analyzed directly for organics using the ATR configuration (e.g., Pereira et al., 2019) and for elemental carbon using diffuse reflection (DF) FTIR (Parks et al., 2021). FTIR analysis on polytetrafluoroethylene (PTFE) filters often poses certain challenges as C−F stretching and bending bands of the filter substrate often dominate the 1150–1250 cm^{-1} range and around 640 cm^{-1}, respectively. In addition, there is an increasing baseline with wavenumber do to filter membrane light scattering (Mcclenny et al., 1985). As a result, baseline correction and blank subtraction are often required to render the spectra suitable for quantitative analysis. Mcclenny et al. (1985) and Pollard and Jaklevic (1988) used FTIR to quantify the concentrations of ammonium, sulfate, and nitrate in the atmospheric aerosols collected on PTFE filters and showed good agreement with X-ray fluorescence (XRF) measurements. Maria et al. (2002) quantified OC collected on PTFE filters directly for the first time using FTIR absorbances of aliphatic CH and carbonyl groups. Maria et al. (2003) expanded the set of quantified functional groups from the previous work and found good agreement with collocated thermal-optical measurements. Russell et al. (2009b) used an algorithm to perform baseline correction and peak fitting with a revised an automated version of the approaches of Maria et al. (2002, 2003) and Maria and Russell (2005). Takahama et al. (2013) described the algorithm for the quantification of carboxylic COOH and carbonyl CO from PTFE filters. Kuzmiakova et al. (2016) described a new, automated baseline correction protocol for FTIR analysis on PTFE filters using smoothing splines. The development of baseline correction, peak fitting, and integration algorithms helped improve the reproducibility of FTIR analyses for aerosols collected on PTFE filters. The mentioned algorithms have been applied to the FTIR analysis of aerosols collected on PTFE filter in several campaigns and smog chamber experiments (Gilardoni et al., 2009; Hawkins et al., 2010; Chhabra et al., 2011a; Takahama et al., 2011; Liu et al., 2011). The AIRSpec open platform is the most recent development with this regard that includes several chemometric methods developed for the FTIR analysis of atmospheric aerosols (Reggente et al., 2019b).

Recently, more advanced statistical approaches have been used to analyze FTIR spectra. For example, partial least squares regression (PLSR) have been applied to the FTIR analysis of $PM_{2.5}$ samples collected on PTFE filters in several works, enabling the quantitative analysis without baseline correction or peak fitting (e.g., Ruthenburg et al., 2014; Reggente et al., 2016; Boris et al., 2019; Takahama et al., 2019). Bürki et al. (2020) presented a probabilistic framework

(Bayesian modeling) that enables the estimation of collocated TOR OC measurements from FTIR spectra of filter samples, incorporating prior knowledge of model parameters (e.g., the fraction of carbon associated with each FG) from laboratory standards and atmospherically relevant molecules. In addition to the functional group quantification, positive matrix factorization (Hawkins and Russell, 2010; Russell et al., 2011; Takahama et al., 2011, PMF;) and cluster analysis (Corrigan et al., 2013; Bürki et al., 2020) have been frequently used to identify and quantify the contribution of different sources to atmospheric aerosols collected on PTFE filters (e.g., biomass burning, biogenic, fossil fuel combustion, and marine). However, the FTIR signatures of marker molecules for identification of aerosol sources have not received enough attention in previous studies.

FTIR has also been combined and compared with other analytical techniques in a few studies (predominantly with AMS) for the chemical characterization of atmospheric and smog chamber (also known as environmental simulation chamber)aerosols showing promising results (Gilardoni et al., 2009; Russell et al., 2009b; Frossard et al., 2011; Liu et al., 2011; Corrigan et al., 2013; Frossard et al., 2014; Faber et al., 2017; Russell et al., 2009a).

1.3 Research objectives and dissertation structure

The main objectives of this work are chemical characterization of primary and secondary organic aerosols from different sources especially biomass burning, understating their evolution during the aging process in the atmosphere, and identification and quantification of these aerosols in the atmosphere. This thesis is composed of six different projects, which have been presented in chronological order. In these projects, FTIR spectroscopy is used to analyze organic aerosols collected on PTFE filters from a variety of sources: atmospheric samples, and aerosols from combustion sources collected in environmental simulation chambers and cookstove test facilities. FTIR measurements are compared and (statistically) combined with various collocated analytical techniques (e.g., AMS, IC, thermal optical OC, GC-MS) to better characterize the chemical composition of OM and its formation mechanisms. FTIR has been used extensively in the past to quantify the functional group composition of OM. Although functional group characterization is still an important part of this work, FTIR has also been used to extract more detailed structural information about OM and also to detect and quantify certain marker molecules related to biomass burning. One of the key advantages of FTIR that is its scalability to large air pollution monitoring networks with hundreds of sites has been shown in this work.

Chapter 2: Estimating mean molecular weight, carbon number, and OM/OC with mid-infrared spectroscopy in organic particulate matter samples from a monitoring network

In this chapter, a spectroscopic method is developed that allows us to estimate some mixture-averaged properties of OM (e.g., mean molecular weight, carbon number, and the OM:OC ratio) that are relevant for understanding different atmospheric processes and aerosol sources. The calibration models are built on laboratory standards of atmospherically-relevant organic

molecules and are applied to filter samples of $PM_{2.5}$ collected in an existing air pollution monitoring network. This method identifies distinct mean molecular weight and carbon number distributions for urban, rural, and smoke-impacted atmospheric samples.

Chapter 3: Characterization of primary and aged wood burning and coal combustion organic aerosols in an environmental chamber and its implications for atmospheric aerosols

In this chapter, the functional group composition of fresh and oxidized aerosols from wood burning and coal combustion are characterized in an environmental chamber. The corresponding FTIR spectra are related to the unburned fuel composition and oxidation products from the major volatile organic compounds know for each source. By comparing AMS and FTIR spectra, we find that AMS spectra of primary and aged burning aerosols in the chamber contain functional group information that is consistent with collocated FTIR measurements. Infrared spectra of wood burning aerosols are shown to have striking similarities with those of ambient biomass burning aerosols regarding the spectral profile and the fingerprint absorptions of biomass burning markers.

Chapter 4: Fragment ion-functional group relationships in organic aerosols using aerosol mass spectrometry and mid-infrared spectroscopy

AMS provides high-time-resolution characterization of OM. However, the extensive fragmentation of molecules poses challenges toward detailed understanding of molecular structure of burning-associated aerosols. This chapter details various multi- and univariate statistical techniques for understanding the composition of organic aerosols through the combination of collocated AMS and FTIR spectroscopy measurements in an environmental chamber. Our analysis reveals that functional group information is retained in AMS spectra in terms of a combination of small (low m/z) and large (high m/z, source specific) fragments. AMS and FTIR measurements are combined statistically and predictive models are developed that can estimate high-time-resolution functional group compositions from the AMS spectra. These models can elucidate the contribution of functional groups and their evolution during the course of aging.

Chapter 5: Chemical evolution of primary and formation of secondary biomass burning aerosols during daytime and nighttime

In this chapter, AMS and FTIR are used in tandem to better understand and quantify the evolution of primary and aged bbOA, two major sources of atmospheric OM, during chemical aging. We adopt a particle wall loss correction method based on AMS organic measurements and develop a procedure to quantify the changes in the composition of primary and aged bbOA. We find that 15 % of primary bbOA mass transforms at relatively short timescales on the order of a day in the atmosphere. The results of this study allow us to evaluate the stability of fresh bbOA molecules (including bbOA markers) and the importance of aging mechanisms other than the homogeneous gas-phase oxidation.

Chapter 6: Identification of smoke-impacted $PM_{2.5}$ samples with mid-infrared spectroscopy in a monitoring network

In this chapter, the FTIR biomass burning marker method introduced in Chapter 3 is developed further and evaluated against satellite detection of fire and smoke. In addition, the ability of FTIR to quantify anhydrosugars, well-known biomass burning markers, is compared to a sensitive chromatographic method (high performance anion exchange chromatography, HPAEC). A smoke classifier model is developed that only relies on FTIR spectra to detect the existence of wood burning smoke in atmospheric samples. The developed method is used to study the concentrations of biomass burning aerosols across the US in 2015 by analyzing approximately 20,000 PTFE samples of $PM_{2.5}$ collected at 160 sites of the IMPROVE network. This work showcases the scalability of FTIR to large air pollution networks .

Chapter 7: Quantified functional group compositions in household fuel burn emissions using FTIR

Billions of people burn solid fuels indoors for cooking and heating around the globe, which contributes to global EC emissions and premature deaths. In this chapter, we use FTIR to analyze $PM_{2.5}$ emissions collected on PTFE filters from fifteen cookstove types and five fuel types. FTIR spectra are found to be distinct for particulate emissions from combustion of each fuel and to have similarities to those of unburned fuels. A multi-variate statistical method is used to highlight the functional groups associated the most with OC from a collocated thermal optical transmittance (TOT) carbon analyzer. Our results show that unlike other samples analyzed in this thesis, aromatics and polycyclic aromatic hydrocarbons (PAHs) are major constituents of particulate emissions from cookstoves. We show that FTIR is a fast and non-destructive technique that provides information about OC that is in agreement with and complementary to thermal-optical measurements.

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

Published in the Journal of Atmospheric Measurement Techniques on 8 Jul. 2021

Authors : Yazdani, Amir; Dillner, Ann. M.; Takahama, Satoshi

Contributions: AY and ST conceived the project. AY prepared laboratory standards, performed the calibrations, and analyzed results. AY wrote the manuscript; ST and AMD provided regular input on the analysis and further editing of the manuscript. AMD provided laboratory and the ambient sample spectra and ST provided overall supervision of the project.

Reference: Yazdani et al. (2021a)

Abstract Organic matter (OM) is a major constituent of fine particulate matter which contributes significantly to degradation of visibility, radiative forcing, and causes adverse health effects. However, due to its sheer compositional complexity, OM is difficult to characterize in its entirety. Mid-infrared spectroscopy has previously proven useful in the study of OM by providing extensive information about functional group composition with high mass recovery. Herein, we introduce a new method for obtaining additional characteristics such as mean carbon number and molecular weight of these complex organic mixtures using the aliphatic C−H absorbance profile in mid-infrared spectrum. We apply this technique to spectra acquired non-destructively from Teflon filters used for fine particulate matter quantification at selected sites of Inter-agency Monitoring of PROtected Visual Environments (IMPROVE) network. Since carbon number and molecular weight are important characteristics used by recent conceptual models to describe evolution in OM composition, this technique can provide semi-quantitative, observational constraints of these variables at the scale of the network. For this task, multivariate statistical models are trained on calibration spectra prepared from atmospherically relevant laboratory standards and are applied to ambient samples. Then, the physical basis linking the absorbance profile of this relatively narrow region in the mid-infrared spectrum to the molecular structure is investigated using a classification approach. The multivariate statistical models predict mean carbon number and molecular weight that are consistent with previous values of organic-mass-to-organic-carbon (OM/OC) ratios estimated for the network using different approaches. The results are also consistent with temporal and spatial variations in these quantities associated with aging processes, and different source classes (anthropogenic, biogenic, and burning sources). For instance, the statistical models estimate higher mean carbon number for urban samples and smaller, more fragmented molecules for samples in which substantial aging is anticipated.

2.1 introduction

2.1.1 Organic aerosols and measurement methods

Organic mass is known to be an important constituent of fine particulate matter (PM). It is estimated to constitute 20–50 % of the total fine PM at mid-latitudes and up to 90 % in tropical forests (Kanakidou et al., 2005). This organic fraction contributes significantly to aerosol-related phenomena such as visibility and climate change, through radiative forcing and affecting cloud formation, and causes adverse health effects (Shiraiwa et al., 2017b; Hallquist et al., 2009). Such effects underscore the importance of better quantification of organic fraction in particulate matter which is a complex mixture of multitude of compounds whose compositions, concentrations, and formation mechanisms are not yet completely understood (Turpin et al., 2000; Nozière et al., 2015).

The determination of organic aerosol composition involves a large range of analytical and computational techniques. Among the widely known techniques are gas chromatography-mass spectrometry (GC-MS), mid-infrared spectroscopy – often referred to as Fourier transform

infrared spectroscopy (FTIR) – and aerosol mass spectrometry (AMS). GC-MS provides molecular speciation information but is limited to a small mass fraction of the organic aerosols as low as 10 % (Hallquist et al., 2009). AMS and FTIR, however, can be used to analyze most of the organic mass in addition to providing information about either chemical classes or functional groups (Hallquist et al., 2009). AMS is an on-line technique with a relatively high size and time resolution. Nevertheless, the extensive fragmentation caused by commonly used ionization method in AMS, i.e. electron impact (EI) ionization, makes the identification of original species difficult (Canagaratna et al., 2007; Faber et al., 2017). In recent years, soft ionization methods such as electrospray ionization (ESI), photoionization (PI), and chemical ionization (CI) have been used frequently for predicting physicochemical properties of OA, e.g. volatility (Li et al., 2016; Xie et al., 2020), phase state, and viscosity (Li et al., 2020b; DeRieux et al., 2018; Shiraiwa et al., 2017a) as a function of measured elemental composition and molecular weight. These methods minimize analyte fragmentation, providing better estimates of molar mass of individual molecules but often have other shortcomings such as ionization efficiency, which varies by molecule (Nozière et al., 2015; Iyer et al., 2016; Hermans et al., 2017; Lopez-Hilfiker et al., 2019).

In mid-infrared spectroscopy, the vibrational modes of organic molecules, whose frequencies fall in the range of mid-infrared electromagnetic radiation, are excited. The advantages of mid-infrared spectroscopy over other common techniques of quantifying OM are providing direct information on functional groups, while minimizing sample alteration during the analysis and having low sampling and analytical cost (Ruthenburg et al., 2014). However, this method only provides bulk functional group (FG) information and has uncertainties regarding the absorption coefficient for group frequencies (although this coefficient is roughly similar across different compounds; Hastings et al., 1952). Moreover, interpretation of mid-infrared spectrum is often complicated due to presence of overlapping peaks. In previous studies, different statistical methods were used to connect mid-infrared absorbances to molar abundance of different functional groups, from which OM, OC (organic carbon), and the OM/OC ratio were calculated with minimal assumptions (Coury and Dillner, 2008; Ruthenburg et al., 2014; Takahama et al., 2016; Boris et al., 2019). These studies showed good agreement between FTIR measurements and other methods of OM characterization. For example, Boris et al. (2019) showed that OC measured by FTIR is around 80 % of OC from thermal optical reflectance (TOR) measurements.

In addition to the abundance of organic functional groups, mid-infrared spectroscopy is informative about the environment in which organic bonds are vibrating (e.g., degree of hydrogen bonding; Pavia et al., 2008), therefore can be used to extract more detailed structural information about OM. This ability of mid-infrared spectroscopy has been investigated to a lesser extent in the context of atmospheric OM. In this work, we used this aspect to investigate two important structural parameters in OM, i.e. mean molecular weight, and mean carbon number. These two parameters are important characteristics used by recent conceptual models and parametrizations to describe evolution in atmospheric OM, in terms of its volatility and phase state (Shiraiwa et al., 2017a; Pankow and Barsanti, 2009; Kroll et al., 2011; Donahue

et al., 2011). Moreover, inspecting the spatial and temporal variations of these parameters helps us understand the processes involved in aerosol aging, especially fragmentation (Murphy et al., 2012), and can be useful for identification of the dominant sources (Price et al., 2017; Gentner et al., 2012).

In this paper, the mean molecular weight, carbon number, and OM/OC ratio of ambient aerosols, which were collected on polytetrafluoroethylene (PTFE) filters at selected IMPROVE sites, were estimated using FTIR spectroscopy. First, the aliphatic C–H region (2800–3000 cm^{-1}) was extracted from the baseline-corrected spectra of laboratory standards. The C–H spectral bands were then normalized to eliminate abundance information. Then, partial least squares regression (PLSR) was used to develop models on the high-dimensional and collinear spectral data. Thereafter, the derived statistical models were used to estimate the mean properties of ambient samples. Finally, a classification algorithm was applied to the PLSR model estimates to provide a better understanding of how they function.

2.1.2 Aliphatic C–H absorption and the molecular structure

We have used the aliphatic C–H region (2800–3000 cm^{-1}) in mid-infrared spectrum to build statistical models for estimating molecular weight and carbon number. This section describes the connection of that region of the spectrum with the molecular structure of organic aerosols and compares the approach used in this work with previous studies.

Recent studies using FTIR and AMS have shown that the aliphatic C–H is the most abundant functional group in organic aerosols (Russell et al., 2009b; Ruthenburg et al., 2014; Zhang et al., 2007) highlighting its importance in OM. This functional group also exhibits characteristics of "good group" frequencies in mid-infrared stretch region (Mayo et al., 2004). Since the hydrogen atom is much lighter than the carbon atom, most of the displacement during oscillation is related to the hydrogen, thereby the carbon atom and consequently its connection to the rest of the molecule is involved to a much lesser extent in the stretch (Mayo et al., 2004). This phenomenon results in a fairly consistent profile for C–H absorption band among different molecules containing this functional group and makes it possible to reduce the dimensionality of spectrum to few independent variables describing the band profile (advantageous when constructing statistical models using a limited number of samples). The light hydrogen atom also causes the aliphatic C–H functional group to absorb at a relatively high stretch frequency, making it isolated from most of other absorbing bonds (Mayo et al., 2004) except the broad carboxylic acid O–H stretch, which absorbs in the 2400–3400 cm^{-1} range and the ammonium N–H stretch (Pavia et al., 2008). These broad absorption profiles can be separated from the narrow aliphatic C–H bands by baseline correction. The unsaturated and aromatic C–H bonds, which absorb at a slightly higher frequency than aliphatic C–H, were not considered in this work. These bonds are not prevalent in atmospheric samples (Russell et al., 2011; Decesari et al., 2000) and their absorption usually falls below the FTIR detection limit (Russell et al., 2009b). The absorption bands attributed to unsaturated and aromatic C–H were not visible

in mid-infrared spectra of atmospheric samples of this study.

The aliphatic C−H (sp^3-hybridized) stretching band in mid-infrared spectrum is composed of four absorption peaks (two doublets) that are attributed to CH_2 (methylene) and CH_3 (methyl) symmetric and asymmetric stretches (Mayo et al., 2004). Methine (tertiary CH) also absorbs in this region, but has a very weak absorption compared to methyl and methylene (Pavia et al., 2008). The profile of these four peaks (characterized by peak frequency, intensity, and width) is affected by the structure of the molecule, inter- and intra-molecular interactions that change electron distribution, and the equilibrium geometry of the molecule (Atkins et al., 2017; Kelly, 2013) as discussed below.

Group vibrational modes in a molecule are not completely decoupled from the rest of the molecule (McHale, 2017). Equation (2.1) describes a 2-body harmonic oscillator model of molecular vibration (in a classical point of view), for which $\bar{\nu}$ is the fundamental wavenumber at which the bond vibrates, c is the speed of light, K is the spring constant of the chemical bond, m_H is mass of hydrogen atom and m_M is the mass of the rest of the molecule (assuming the rest of the molecule is stiff). The reduced mass of the system, μ, increases with increasing the molecular weight (Eq. 2.1), resulting in a decreased vibrational frequency (wavenumber). There are also effects that change the vibrational frequency through changing the bond strength. For example, electron-withdrawing effect of neighboring polar groups and ring structure strain elevate the absorption frequency of the oscillator by increasing the equivalent spring constant (Pavia et al., 2008). The Bohlmann effect, in which electron density is transferred from the lone pair of a neighboring nitrogen or oxygen into the C−H antibonding orbital, decreases the frequency by weakening the C−H bond (Lii et al., 2004). Hydrogen-bonding interactions and phase state can also affect absorption frequency and intensity of bands corresponding to vibrational modes (Fornaro et al., 2015; Kelly, 2013).

$$\bar{\nu} = \frac{1}{2\pi c}\sqrt{\frac{K}{\mu}}, \quad \text{where } \mu = \frac{m_H m_M}{m_H + m_M}. \tag{2.1}$$

The environment in which the molecules vibrate can effect the absorption peak width through different homogeneous and inhomogeneous broadening mechanisms. Slightly different interaction of molecules in liquids and amorphous solids (to a lesser extent in crystals) is the basis of inhomogeneous broadening (Kelly, 2013). This phenomenon determines the change in peak width due to phase state by changing the level of interaction between the molecules. Hydrogen bonding can also cause inhomogeneous broadening due to enhanced anharmonicity (Thomas et al., 2013). The weak hydrogen bond, which can exists for aliphatic C−H functional group (Desiraju and Steiner, 2001), broadens its absorption band slightly and shifts its absorption frequency.

The peak height ratios in aliphatic C−H region are also indicators of some structural fea-

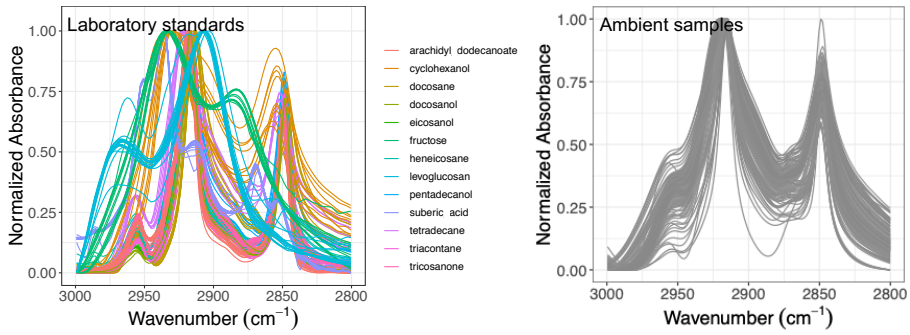

Figure 2.1 – Normalized aliphatic C–H spectra of the laboratory standards (left) and several atmospheric samples (right). This figure shows variation in absorbance profile among the standards and atmospheric samples.

tures of the molecule. For example, the ratio of peak heights of asymmetric CH_3 stretching to asymmetric CH_2 stretching shows the relative abundance of these groups in the sample (Orthous-Daunay et al., 2013). For straight-chain alkanes and some polymers, this ratio is directly related to the chain length and can be used to estimate the carbon number of a molecule (Lipp, 1986; Mayo et al., 2004). This ratio as well as the tertiary C–H absorption are informative about the degree of branching in the molecule. The ratio of symmetric to asymmetric CH_2 peak heights is an the indicator of rotational and conformational order in a molecule, and is related to chain length and phase state (Hähner et al., 2005; Corsetti et al., 2017; Orendorff et al., 2002). Price et al. (2017) compared that ratio between mid-infrared spectra of emissions under different engine conditions for ultra-low sulfur diesel (ULSD) and hydrogenation derived renewable diesel (HDRD) fuels and observed a slightly greater ratio for the ULSD emissions and suggested this was due to the differences in the carbon number distribution of the two fuel emissions. In addition, some other vibrational bands can affect this region through forming overtones and combination bands (Thomas, 2017). Overall, the absorbance profile in the aliphatic C–H region contains direct and indirect information about carbon number and molecular weight and shows significant variation in laboratory standards and atmospheric samples (Fig. 2.1) related to their molecular structure. In this work, we adopt a new approach for using mid-infrared spectra to characterize OM. We use the variations in the aliphatic C–H region to estimate mean carbon number and mean molecular weight of atmospheric samples. In previous studies on the mid-infrared spectrum of atmospheric aerosols, functional group molar abundance in laboratory standards or total OC from other methods such as TOR were considered as the response variable, while non-normalized absorbances were considered as independent variables (Takahama et al., 2013; Ruthenburg et al., 2014; Reggente et al., 2016). In this manner, linear models resembling the Bougher-Lambert-Beer law were developed. In this study, however, molecular weight and carbon number statistical models were developed using chemical formulas of the laboratory standards (no molar abundance information) and their normalized aliphatic C–H absorbances as independent variables. The current approach extracts detailed information from the mid-infrared spectrum complementary to previous approaches (Fig. 2.2).

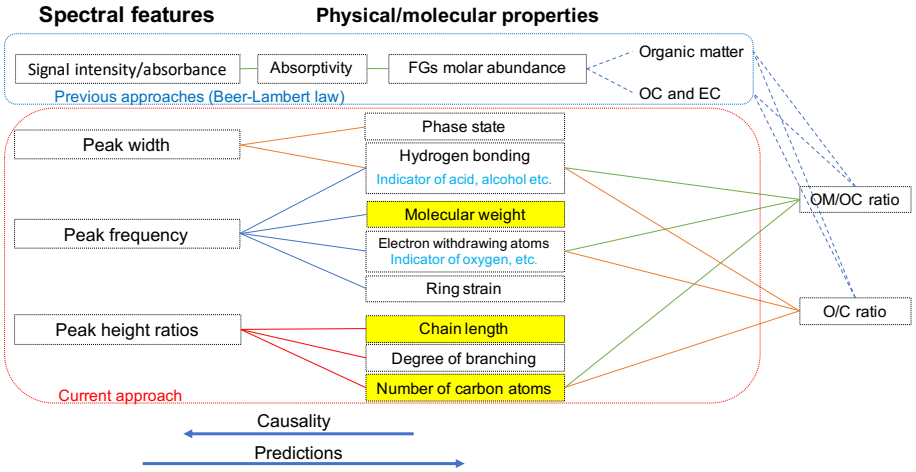

Figure 2.2 – Diagram showing the relation between spectral features and molecular/physical properties. The way previous approaches (e.g. Ruthenburg et al., 2014; Takahama et al., 2013) and the current approach use mid-infrared spectrum to estimate different parameters is shown in blue and red boxes, respectively. Highlighted molecular properties can only be estimated using the current approach.

2.2 Methods

We will describe the atmospheric samples as well as the laboratory standards for the calibration and test set in Sects. 2.2.1 and 2.2.2. Thereafter, the methodology for data analysis and interpretation will be discussed in Sects. 2.2.3, 2.2.4, and 2.2.5.

2.2.1 IMPROVE network monitoring sites (sampling and analysis)

Particulate matter with diameter less than 2.5 µm ($PM_{2.5}$) was collected on PTFE filters (25 mm diameter Teflo® membrane, Pall Corporation) every third day for 24 hours, midnight to midnight, at nominal flow rate of 22.8 L min^{-1} during 2011 and 2013 at selected sites in the Inter-agency Monitoring of PROtected Visual Environments (IMPROVE) network (http://vista.cira.colostate.edu/improve/,lastaccess:2020-10-08). There are, in total, 814 samples collected at 7 sites in the USA in year 2011 and 2161 samples collected at 16 different sites in the USA 2013 (see Fig. 2.3). 1 out 7 sites in 2011 and 4 out of 16 of sites in 2013 are urban sites and the rest are rural. FTIR analysis was performed on the PTFE filters using a Bruker-Tensor 27 FTIR spectrometer equipped with a liquid nitrogen-cooled, wideband mercury-cadmium-telluride (MCT) detector, and at a resolution of 4 cm^{-1} (data intervals of 1.93 cm^{-1}; Nyquist sampling). For samples with low molar abundance of organic compounds, especially aliphatic C−H, baseline correction could not be done properly in the aliphatic C−H region resulting in irregular and negative absorbance profiles. These samples were omitted from further analysis and only 798 were analyzed in this work. As can be seen from Fig. 2.4, data recovery is higher in urban sites than rural sites due to having a usually more prominent aliphatic C−H peak.

Due to this under-sampling, generalizing the results of this work to the whole of rural samples should be done with caution.

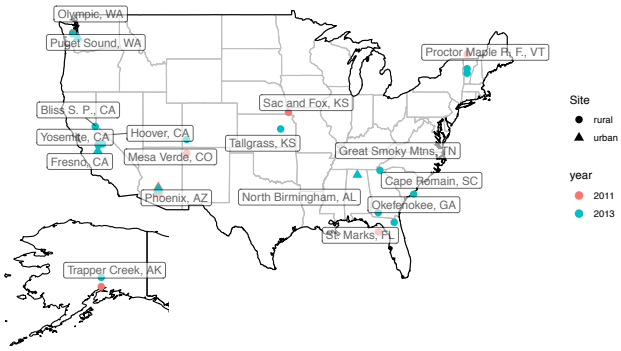

Figure 2.3 – The location of IMPROVE sites used for this work (the USA and Alaska); the year at which samples are taken is differentiated by color and the type of the site by point shape.

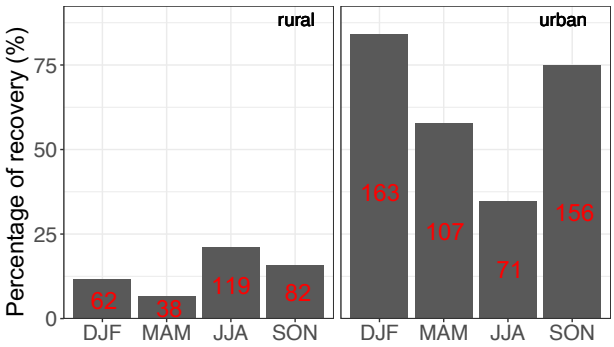

Figure 2.4 – Percentage of the samples which were recovered from each category (sample type and season) after baseline correction. The number of samples in each category is shown in red.

2.2.2 Laboratory standards (sampling and analysis)

Compounds containing relevant functional groups to atmospheric OM such as aliphatic C−H, alcohol and acid O−H, carbonyl C=O, and with different structures (straight-chain and cyclic) and various chain lengths were used to produce laboratory standards (Table 2.1). All compounds used for creating the standards contained aliphatic C−H, which is the main focus of this study. Five of these compounds were alkanes, just containing aliphatic C−H. Three were straight-chain alcohols containing alcohol O−H as well. One was cyclic alcohol and one was a cyclic ketone having carbonyl C=O; two were cyclic (not aromatic) sugar derivatives containing several O−H groups. The calibration set also contained an ester, a ketone and one dicaboxcylic acid. In addition to relevance to atmospheric OM, these standards were selected based on the availability of spectroscopic data and their suitability for atomization. These compounds had comparable absorption coefficients for aliphatic C−H and the effect of other functional

Table 2.1 – Chemicals used in the calibration set to analyze the effect of different physical/-chemical properties of organic molecules on aliphatic C−H absorbance profile.

Compound Name	Formula	Class	Phase State at 25°C	Molecular Weight ($g\,mol^{-1}$)	OM/OC
Tetradecane	$C_{14}H_{30}$	alkane	liquid	198.4	1.18
Hexadecane	$C_{16}H_{34}$	alkane	liquid	226.4	1.18
Heneicosane	$C_{21}H_{44}$	alkane	solid	296.6	1.18
Docosane	$C_{22}H_{46}$	alkane	solid	310.6	1.18
Triacontane	$C_{30}H_{62}$	alkane	solid	422.8	1.17
1-Pentadecanol	$C_{15}H_{32}O$	alkanol	solid	228.4	1.27
1-Eicosanol	$C_{20}H_{42}O$	alkanol	solid	298.6	1.24
1-Docosanol	$C_{22}H_{46}O$	alkanol	solid	326.6	1.24
Cyclohexanol	$C_6H_{12}O$	cyclic alcohol	liquid	100.2	1.39
Cyclohexanone	$C_6H_{10}O$	cyclic ketone	liquid	98.1	1.36
Fructose	$C_6H_{12}O_6$	Sugars and their derivatives	solid	180.2	2.50
Levoglucosan	$C_6H_{12}O_5$	Sugars and their derivatives	solid	162.1	2.25
Suberic acid	$C_8H_{14}O_4$	dicarboxcylic acid	solid	174.2	1.81
Arachdyl dodecanoate	$C_{32}H_{64}O_2$	ester	solid	480.9	1.25
12-Tricosanone	$C_{23}H_{46}O$	ketone	solid	338.7	1.23

groups, heteroatoms, and the molecular structure was analyzed indirectly via the change in the aliphatic C−H absorbance profile. Some of the laboratory standards and their resulting spectra were taken from Ruthenburg et al. (2014). The rest were created (using a similar protocol) from methanolic solutions with a concentration of 0.1 $g\,L^{-1}$ and analyzed by FTIR as follows. Atomized aerosols of the desired compounds were first generated by a TSI Model 3076 Aerosol Generator using the methanolic solutions. Then these particles were conducted by the flow system towards a 47 mm PTFE filter (Teflo® membrane, Pall Corporation), where they were collected. The flow system was composed of a silica gel dryer (for drying the aerosols before collection), a sharp-cut-off 1 µm cyclone and a diluter system (which facilitated the adjustment of aerosol concentration in the line). The pressure drop needed for the flow through the filter was provided by a rotary vacuum pump (Gast 0523-101Q-G588NDX) and the filter flow was controlled by a gas-flow controller (Alicat MCR-100-SLPM-D/5M). The mass on the filters ranged from few micro-grams to tens of micro-grams. After collecting the aerosols on the filters, FTIR analysis was performed on the PTFE filters using a Bruker-Vertex 80 FTIR spectrometer equipped with a deuterated lanthanum α alanine doped triglycine sulfate (DLaTGS) detector, with the same spectral resolution as the spectra of the ambient samples.

In total, 168 laboratory samples with different composition and molar abundance (absorption amplitude ranging from 0.001 to 2 before normalization) were used from which a subset of 43 samples was kept as a test set and the rest were used as the calibration set. The test set was used solely for the purpose of evaluation of the statistical models developed using the calibration set. However, the final statistical models, which were applied to ambient samples, were developed using all 168 laboratory standards to increase the precision.

2.2.3 Baseline correction and normalization

The baseline removal is often a useful step in mid-infrared spectroscopy on PTFE filters, like in other methods of spectroscopy. The baseline arises from light scattering by the filter membrane (Mcclenny et al., 1985) and particles collected on the filter as well as electronic transitions of some carbonaceous materials (Russo et al., 2014; Parks et al., 2019). For baseline removal, we used the smoothing spline method on 1500–4000 cm^{-1} region, where PTFE filter does not absorb, with parameter selection criteria similar to the approach taken by Kuzmiakova et al. (2016). Briefly, a cubic smoothing spline was fitted to the spectrum and then was subtracted from the raw spectrum to obtain the pure contribution of functional groups at each wavelength. The analyte region (the aliphatic C−H absorption region, 2800–3000 cm^{-1}) was manually excluded from the baseline by setting the weights in this region to zero in the the smoothing spline objective function (refer to Kuzmiakova et al., 2016). The rest of the spectrum between 1500–4000 cm^{-1} was included in the baseline by setting the weights one. After baseline correction, the aliphatic C−H absorbances were scaled between zero and one (Fig. 2.1) for all spectra so that the absorbance profiles were comparable regardless of the absorbance intensity (functional group abundance).

2.2.4 Building the calibration models

In order to estimate molecular weight and carbon number from the normalized aliphatic C−H absorbances in the mid-infrared spectra, we seek the solution of the following linear equation for the calibration models:

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{b} + \boldsymbol{e}, \tag{2.2}$$

where $\mathbf{X}$ is the normalized spectra matrix (the aliphatic C−H absorption region, 2800–3000 cm^{-1}), $\boldsymbol{y}$ is the vector of response variable (molecular weight or carbon number) and $\boldsymbol{e}$ is a vector of residuals ($\boldsymbol{y}$ and $\mathbf{X}$ are assumed to be centered). In spectroscopic applications, due to indeterminacy (more independent variables than the number of samples) and collinearity (inter-correlation between independent variable) the ordinary least squares (OLS) method is not applicable or is not robust unless regularized. Among the common methods developed for treating such a data structure, we chose univariate ($\boldsymbol{y}$ is a vector, i.e. has one variable) partial least squares regression (PLSR) for this work (Wold et al., 1983). Univariate PLSR projects $\mathbf{X}$ onto $\mathbf{P}$ basis with orthogonal scores $\mathbf{T}$ and residual matrix $\mathbf{E}$ (Eq. 2.3) such that the covariance between each score column and $\boldsymbol{y}$ is maximized (in each step of deflation). Thereafter, the response variable $\boldsymbol{y}$ is regressed linearly against the scores (Eq. 2.4). In Eq. (2.4), $\boldsymbol{c}$ is the regression coefficient of $\boldsymbol{y}$ as a function of scores ($\mathbf{T}$) and $\boldsymbol{f}$ is the vector of residuals.

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\top} + \mathbf{E}, \tag{2.3}$$

$$\boldsymbol{y} = \mathbf{T}\boldsymbol{c} + \boldsymbol{f}. \tag{2.4}$$

Determining the optimum number of latent variable (LVs), which are linear combinations of original wavenumbers in this study, is an essential step for developing calibration models with

predictive capability. After solving the PLSR problem for calibration models with different number of LVs, we ran a repeated 10-fold cross validation on the calibration models and calculated the root mean square error (RMSE) of predictions (for the calibration set) for each model. Thereafter, the model whose RMSE was within one standard error from the calibration model with minimum RMSE and had fewer LVs (i.e., a simpler model) was chosen (Hastie et al., 2009). Based on the above-mentioned procedure, the optimal number of LVs for molecular weight and carbon number calibration models was found to be 19 and 20, respectively.

2.2.5 Interpreting the calibration models using the basic spectral features

Although the PLSR models have considerably fewer LVs (approximately 20) than the original wavenumbers (105), the lack of physical interpretability and remaining number of LVs still hinders their physical interpretation. Therefore, we first analyze the basic (physically interpretable) features of the mid-infrared spectrum – peak frequencies, widths and ratios in the aliphatic C−H region – for the calibration set and their relation with carbon number and molecular weight (Sect. 2.3.1). Spatial and temporal variation of these patterns in the atmospheric samples are also analyzed and related to similar patterns in the laboratory standards.

The four basic features of the ambient sample spectra were used to build a classification and regression trees (CART) (Breiman et al., 1983) to approximate the PLSR predictions of mean molecular weight and carbon number and to better understand their connection with the underlying spectral absorption characteristics. In this approach, binary decision trees are generated to classify the PLSR estimates based on partitioned domains of their basic spectral features. The CART algorithm expands the trees in the order of decreasing explanatory power until certain stopping conditions (e.g., minimum number of observations in terminal nodes or minimum improvement of explanatory power at each step of splitting) are satisfied.

2.3 Results and discussions

First, the basic features of the aliphatic C−H profile are discussed in the atmospheric and the laboratory samples followed by a similarity check between the two (Sect. 2.3.1). Then, development of calibration models for predicting molecular weight and carbon is described, followed by investigation of their performance in the calibration and test (Sect. 2.3.2). Thereafter, the model estimates are discussed for atmospheric samples and compared to the results reported in literature (Sect. 2.3.3). Finally, the basic features introduced earlier are used to classify the results of the sophisticated (PLSR) models in order to obtain a better understanding of the way they function (Sect. 2.3.4).

2.3.1 Basic features

Basic features of the spectrum in the aliphatic C−H region were calculated for atmospheric samples and laboratory standards to study their temporal and spatial variation and their relation with molecular properties such as molecular weight, carbon number, and the OM/OC ratio. These variables, although few, can give a good estimate of the absorbance profile and make it more interpretable.

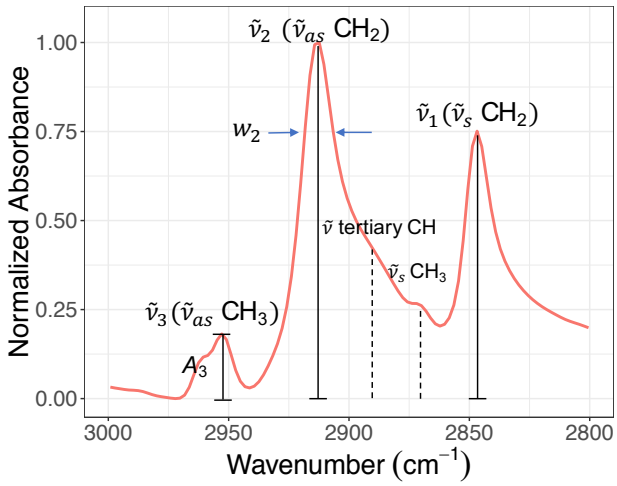

Figure 2.5 – A sample C−H spectrum showing the convention of peak parameters used in this study. The symmetric CH_2 ($\tilde{\nu}_s$ CH_2) wavenumber is denoted by $\tilde{\nu}_1$. The asymmetric CH_2 ($\tilde{\nu}_{as}$ CH_2) wavenumber is denoted by $\tilde{\nu}_2$ and the asymmetric CH_3 ($\tilde{\nu}_{as}$ CH_3) wavenumber by $\tilde{\nu}_3$. Absorbance and width of the ith peak are also denoted by A_i and w_i, respectively.

Figure 2.5 shows the convention of spectral features in the aliphatic C−H (2800–3000 cm^{-1}) region used in this study. Apart from methine group (tertiary C−H), which has a very weak absorption (Pavia et al., 2008), there are two doublets in this region corresponding to CH_2 and CH_3 symmetric and asymmetric stretching vibrations. The CH_3 symmetric peak is typically suppressed by the surrounding peaks and is not completely distinguishable. Among the remaining peaks, the symmetric CH_2 ($\tilde{\nu}_s$ CH_2) wavenumber is denoted by $\tilde{\nu}_1$. Likewise, the asymmetric CH_2 ($\tilde{\nu}_{as}$ CH_2) wavenumber is denoted by $\tilde{\nu}_2$ and the asymmetric CH_3 ($\tilde{\nu}_{as}$ CH_3) wavenumber by $\tilde{\nu}_3$. Absorbance and peak width of the ith peak are also denoted by A_i and w_i, respectively.

In the next subsections, the variation of the mentioned spectral features are studied in the laboratory standards and atmospheric samples. For this purpose, the atmospheric samples are separated into urban, rural and burning categories. The burning category constitutes 95 samples of urban or rural sites and is taken from clusters 9a, 9b and 10 of Bürki et al. (2020) based on their spectral similarity. These samples are believed to be influenced by residential wood burning or wildfires since they were usually collected during a known fire period (Rim Fire in California in 2013) or in Phoenix, AZ, during winter months when residential wood

burning typically occurs (Pope et al., 2017).

Asymmetric CH_2 peak wavenumber ($\tilde{\nu}_2$)

We calculated the second peak wavenumber ($\tilde{\nu}_2$) for the laboratory standards and atmospheric samples using a simple peak finding algorithm based on the first and second numerical derivatives of the spectrum. For the laboratory standards, the frequency generally decreases with increasing molecular weight until it reaches an asymptotic state after 200 g mol^{-1} (Fig. 2.6). The curve in Fig. 2.6 shows the theoretical peak frequency of the aliphatic C−H when the bond spring constant is assumed to be 10^3 N m^{-1} (Pavia et al., 2008), and the reduced mass is calculated based on a ball-and-string assumption composed of the hydrogen atom (first "ball") and the rest of molecule (second "ball"). The only effect considered in this model is the variation of the reduced mass of the oscillator. The fact that the less-oxygenated laboratory samples follow the theoretical line closely implies that the value of the spring constant considered here is, on average, a good approximation. However, especially for highly oxygenated (high OM/OC ratio) molecules and those in liquid phase (which have a lower molecular weight), the absorption frequency deviates from the theoretical line (higher frequency) due to higher levels of inter-molecular interaction.

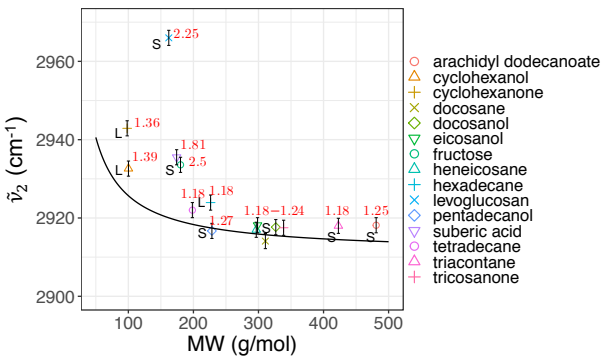

Figure 2.6 – Scatter plot showing the variation of the second peak wavenumber ($\tilde{\nu}_2$) with molecular weight (MW) in the calibration set, affected by the OM/OC ratio and phase state (S: solid, L: liquid). The black line shows the theoretical frequency with a spring constant equal to 10^3 N m^{-1} for all C−H bonds. The OM/OC ratio and phase state are shown for the samples. The error bars show uncertainty in calculated peak frequency due to FTIR scan resolution.

Regarding the atmospheric samples, most of categories have a peak density in 2915–2925 cm^{-1}, close to that of straight-chain molecules of the laboratory standards (Fig. 2.7, first row). Urban samples have a wider shoulder on the right side (around 2925 cm^{-1}) in summer when the samples are expected to be more aged. Other variations are believed to be insignificant considering the scan resolution of the FTIR instrument.

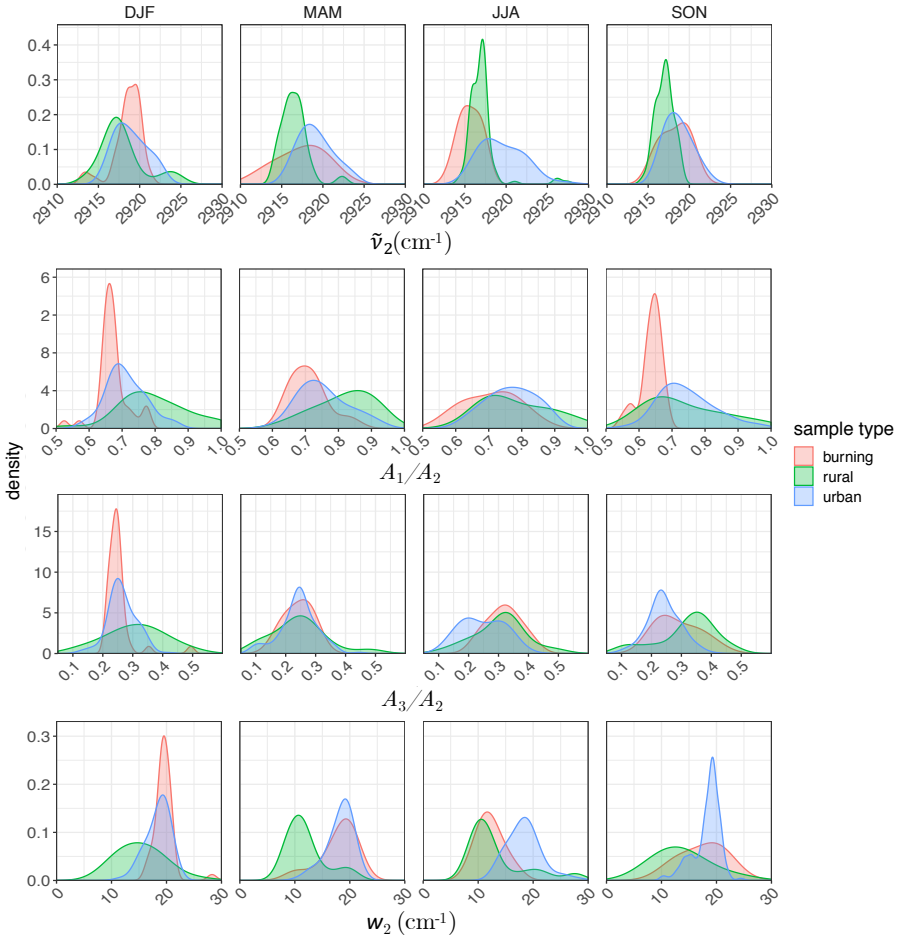

Figure 2.7 – Kernel density estimate of second peak wavenumber ($\tilde{\nu}_2$), the ratio of peak heights of symmetric CH_2 to asymmetric CH_2 stretching (A_1/A_2), the ratio of peak heights of asymmetric CH_3 to asymmetric CH_2 stretching (A_3/A_2), and the second peak width (w_2) of the aliphatic C−H band in mid-infrared spectra of the atmospheric samples segregated based on sample type and season.

Peak height ratios (A_i/A_2)

Analyzing the laboratory standards shows that a relatively linear but scattered relation exists between carbon number and the A_1/A_2 ratio in the calibration set (Fig. 2.8, upper panel). Suberic acid, that is the only dicarboxylic acid in the laboratory standards, does not follow the general trend, probably due to strong dimerization. As mentioned in Sect. 2.1.2, the A_1/A_2 ratio compares symmetric and asymmetric absorbance of methylene and its connection with carbon number has already been highlighted in FTIR analysis of some types of diesel fuels (Price et al., 2017). Increase in A_1/A_2 is also observed between solid and liquids, consistent with the work of Corsetti et al. (2017). We also observe a nonlinear relation between the A_3/A_2 ratio and carbon number with different levels based on branching and terminal functionalization (Fig. 2.8, lower panel). This ratio is equal to zero for molecules lacking methyl group such

as simple cyclic molecules while increasing as the number of branches containing terminal methyl increases.

Results show a clear separation in atmospheric samples regarding the sample type and season for both A_1/A_2 and A_3/A_2 ratios (Fig. 2.7, second and third row). The samples influenced by burning usually have the lowest A_1/A_2 ratio (Fig. 2.7, second row). This observation is consistent with the presence of molecules with longer chains, as observed for laboratory samples. Bürki et al. (2020) showed that the urban samples (in the same dataset) have their highest average OM/OC ratio in summer which is concurrent with the their highest A_1/A_2 ratio which suggests shorter chain length. The highest A_1/A_2 ratio for rural samples is observed in spring when the aerosols are highly oxidized (Bürki et al., 2020). This suggests that aged aerosols have lower carbon number probably due to the fragmentation process. The measured A_1/A_2 ratio for majority of the atmospheric samples ranges between 0.6 to 0.8, which is consistent with the value for laboratory standards. Results also show that the A_3/A_2 ratio is higher in rural samples compared to urban samples (with the exception of spring) suggesting a higher CH_3 to CH_2 abundance in those samples. This observation can be due to lower carbon number or higher number branches containing CH_3. Like the A_1/A_2 ratio, we observe fewer samples with low A_3/A_2 ratios in urban sites in summertime. The A_3/A_2 ratio falls between 0.1–0.4 for majority of the atmospheric samples, which is consistent with the value for laboratory standards. It is worth noting that peaks in atmospheric samples are more overlapped than laboratory standards, which makes calculation of peak ratios based on extrema of the original spectra imprecise. As a result, a peak fitting method based on Gaussian peaks was applied to atmospheric samples in order to obtain the peak ratios more precisely.

Peak width (w_i)

We observe a clear correlation between w_2 and the OM/OC ratio in the calibration set when solid and liquid phases are considered separately (Fig. 2.9). As mentioned in Sect. 2.1.2, hydrogen bonding increases the peak width, and the extent of hydrogen bonding is usually a good indicator of the OM/OC ratio. This is because hydroxyl, hydroperoxyl, and carboxyl groups, which form hydrogen bonds, are among the most effective functional groups in $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ formation due to the significant vapor pressure reduction they cause (Seinfeld and Pandis, 2016). In this study, w_2 is defined as the peak width at 75 % of the maximum amplitude. This position is chosen for robustness of the measurement algorithm (to avoid interference with other peaks); however, it can be converted to full width at half maximum (FWHM) assuming the proper peak profile (w_2 is 65 % of FWHM for a Gaussian peak). In addition to hydrogen bonding and phase state, superposition of a multitude of peaks with slightly different profiles can also have a statistical positive or negative effect on the peak width in mixtures (see Appendix A.1.4). The observed peak width in mid-infrared spectra of the atmospheric samples is the result of all above-mentioned factors. However, since all laboratory standards are produced with pure compounds, the significance of the mixture effect cannot be evaluated.

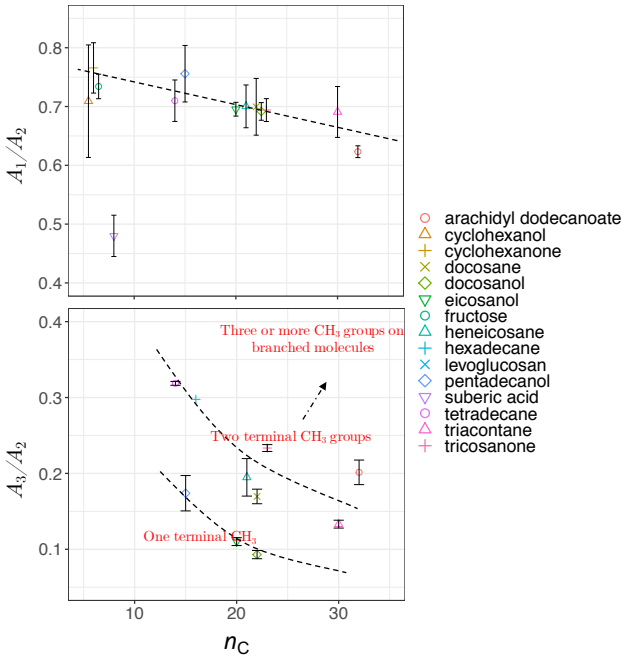

Figure 2.8 – Scatter plots showing the relation between carbon number (n_C) and the ratio of peak heights of symmetric CH_2 to asymmetric CH_2 stretching (A_1/A_2, upper panel), and the ratio of peak heights of asymmetric CH_3 stretching to asymmetric CH_2 stretching (A_3/A_2, lower panel), averaged for each substance in laboratory standards. Error bars show ± one standards error from the average and dashed lines are visual guides for the trends and levels.

Figure 2.7 (fourth row) shows a distinct distribution of w_2 considering spatial and temporal variations as well as sample category. Rural samples have a smaller value of w_2 compared to urban and burning samples, although the former are usually more oxidized (have higher OM/OC ratio). This observation suggests that other factors such as phase state and statistical effects likely outweigh the oxygenation effect on absorption peak width.

Spectral similarity (dimension reduction)

In previous sections, the basic features of spectra in the aliphatic C−H region were presented and discussed for the atmospheric samples and laboratory standards. Here, we check the spectral similarity between atmospheric complex mixtures and laboratory pure standards by means of principal component analysis (PCA), before developing calibration models.

The spectral data of laboratory standards are highly collinear as can be seen from their correlation matrix heat map (Fig. A.1). In this case, PCA is efficient for reducing the data dimension such that only the first six principal components (PCs) explain around 99 % of variance in the spectra (Table 2.2). For the sake of comparison, we have projected the spectra of atmospheric samples onto the six PCs. The results show that their scores, when projected onto laboratory PCs, are surrounded by laboratory standards. Many spectra, particularly urban ones, are

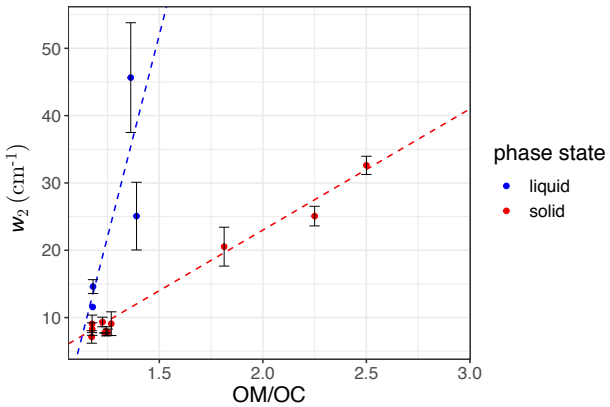

Figure 2.9 – The average value of second peak width (w_2) measured for each compound in the calibration set versus the OM/OC ratio, colored based on compound phase state at laboratory condition (25 °C). Error bars show ± one standards error from the average and dashed lines are visual guides.

clustered close to tetradecane for the first four PCs (Fig. 2.10); greater differentiation is found among the higher PCs. This observation suggests that the laboratory standards are able to capture the main variations in the spectra of atmospheric samples, which have a more regular aliphatic C−H profile close to that of straight-chain alkanes. We also found that PC3 appears to capture phase state information (see Appendix A.1.5).

Table 2.2 – Importance of the first six principal components in the laboratory standards.

	PC1	PC2	PC3	PC4	PC5	PC6
Standard Deviation	1.414	0.668	0.647	0.332	0.203	0.133
Proportion of Variance	0.651	0.145	0.136	0.036	0.014	0.006
Cumulative Proportion	0.651	0.796	0.932	0.968	0.982	0.988

2.3.2 Developing and evaluating the calibration models

PLSR with cross validation was used to develop quantitative models for molecular weight (MW) and carbon number (n_C) with the calibration set composed of 143 samples including all compounds over the available mass range. The OM/OC ratio was then calculated from these two parameters (OM/OC = $\frac{\text{MW}}{12.01 n_C}$). The developed PLSR models gave reasonably good fit results (r^2 ranging from 0.94 to 0.99) for molecular weight, carbon number, and indirect OM/OC ratio in the calibration set (Figure 2.11).

The prediction ability of the PLSR models was then evaluated using a test set composed of 43 samples which were not used for developing the models. The PLSR models also performed reasonably well in predicting molecular weight, carbon number and OM/OC ratio in the test set with r^2 ranging from 0.92 to 0.98 (Fig. 2.11). The predictions with high relative error were attributed to laboratory samples with low molar abundance (low signal-to-noise ratio), for which the baseline correction had the highest uncertainty. This is not a concern when

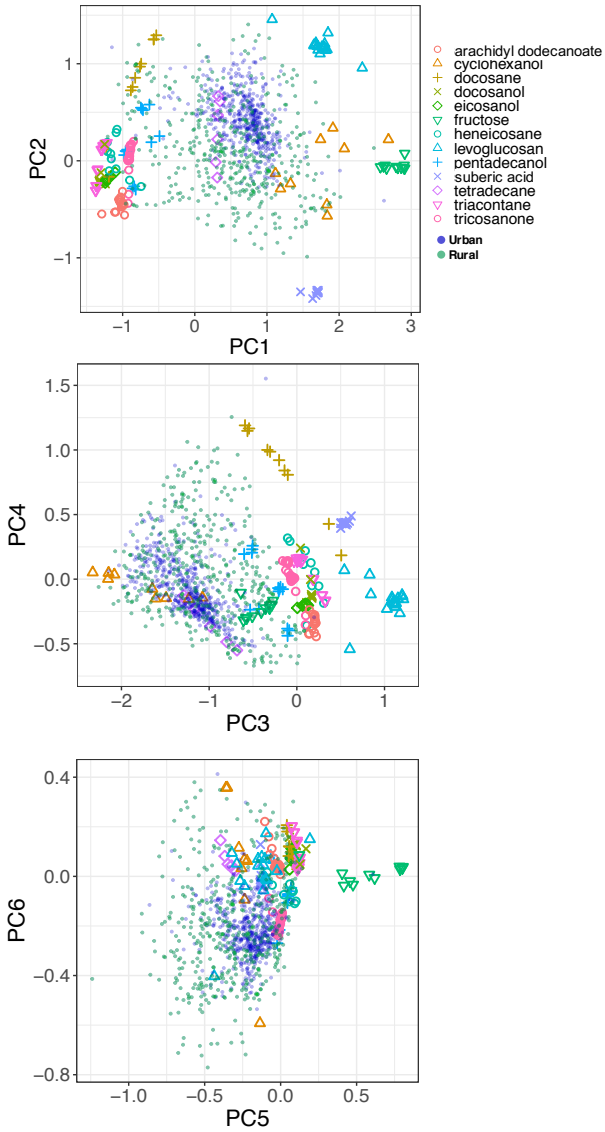

Figure 2.10 – Bi-plots showing the scores of normalized spectra of laboratory standards (color) and normalized spectra of atmospheric samples (filled circles) projected onto the first six principal components calculated for laboratory standards and listed in Table 2.2.

applying the PLSR models to atmospheric samples since the atmospheric samples with low signal-to-noise ratio were omitted in the first step (Sect. 2.2.1).

2.3.3 Applying the calibration models to atmospheric samples

After checking the performance of the PLSR models on the calibration and test set, all laboratory standards were used to build calibration models that were applied to the ambient samples. In the following sections, the estimates of OM/OC, mean molecular weight, and

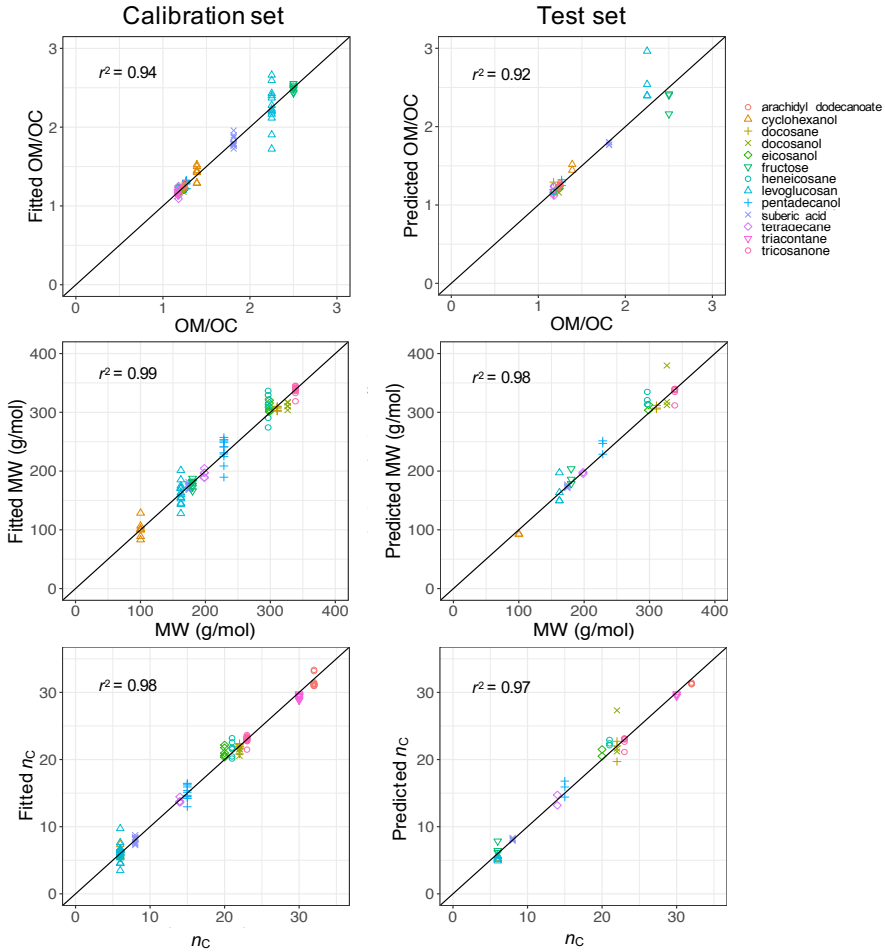

Figure 2.11 – Scatter plot of fitted (predicted) indirect OM/OC ratio, molecular weight (MW), and carbon number (n_C) against the values from chemical formula of the calibration set (test set). The diagonal black lines indicate the perfect fit (1:1).

mean carbon number for the ambient samples are shown in different categories based on season and sample type (rural, urban and burning) after omitting the physically unreasonable values. Thereafter, the trends and absolute values are compared to previous studies (when available) and our expectations based on aging process and aerosol emission sources.

In this work, we have assumed that we can obtain mean mixture (atmospheric samples) properties from the normalized spectrum of a mixture using the calibration models developed for pure compounds (laboratory standards). This assumption relies on the linearity of the property estimation models (which is consistent with our calibrations, Eq. 4.3), and equality of the absorption coefficients of the compounds existing in the mixture (see Appendix A.1.2 for more information). Thus, the absorption coefficient of aliphatic C−H has been assumed to be relatively similar between the compounds existing in atmospheric samples. Although the aliphatic C−H absorption coefficients of the laboratory standards were similar in this study,

the variability of this absorption coefficient is relatively less-studied for compounds existing in the atmospheric OM and needs to be addressed in the future. This assumption is a potential source of error that may change the accuracy of the results, but the estimates for atmospheric samples shown in the following sections suggest that this assumption does not overwhelm the findings.

OM/OC ratio

The OM/OC ratio is the first parameter that we investigate here since it has been studied extensively in atmospheric aerosols (Bürki et al., 2020; Hand et al., 2019; Ruthenburg et al., 2014; Takahama et al., 2011; Simon et al., 2011; Aiken et al., 2008). Moreover, it can be used as an indirect evaluation for mean molecular weight and mean carbon number estimates as the indirect OM/OC ratio is calculated from those two. An indirect OM/OC estimate that is consistent with previous studies implies that estimates of molecular weight to carbon number are also likely to be reasonable.

The OM/OC ratio is estimated to be generally lower for urban samples (≈ 1.5) than rural samples (≈ 1.8; Fig. 2.14, first row). The lower OM/OC ratio in urban sites is thought to be related to emission sources that are generally hydrocarbon, with low OM/OC ratio emitted from gasoline and diesel vehicles (fuel combustion and unburned motor oil) as a major part of anthropogenic $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ precursors (Gentner et al., 2012) as well as cooking. These organic molecules do not undergo significant oxidation and aging as the monitoring sites are generally close to the emission sources. In contrast, organic aerosols usually undergo several steps of oxidation and receive substantial condensation of oxidized vapors, which results in higher OM/OC ratio at rural and remote sites. Previous studies using several different methods (including FTIR and AMS) show the same trend in urban and rural sites (Ruthenburg et al., 2014; Zhang et al., 2007; Simon et al., 2011; Bürki et al., 2020). In addition, the majority of the samples are in the range that is usually considered for OM/OC ratio, i.e.,1.4–1.7 (Russell, 2003). We also observe that samples influenced by burning, especially residential wood burning, have lower OM/OC ratio (≈ 1.4) than those associated with more oxidized aerosol such as rural site, consistent with OM/OC estimates of Bürki et al. (2020).

The OM/OC ratio in urban sites is estimated to be higher in summer compared to other seasons, especially winter (Fig. 2.14, first row) which is believed to be caused by more intense photochemical aging in summertime (Kroll and Seinfeld, 2008). In rural sites, the trend becomes more complicated as vegetation, as major biogenic $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ emission sources, is more active in summer time (Yuan et al., 2018; Seinfeld and Pandis, 2016). Samples influenced by burning are also estimated to have higher OM/OC in summer when samples are affected by wildfires compared to winter when burning samples are mostly affected by residential wood burning. However, the contribution of photooxidation relative to emission sources is not clear in this case as they are coupled in these observations (Bürki et al., 2020).

In order to have a direct comparison with other methods, we chose the Phoenix, AZ, monitor-

ing site, for which recovery percentage of the baseline correction method is close to 100 %, and compared our indirect OM/OC ratio estimates to the corresponding ones calculated by Bürki et al. (2020). The latter method uses molar abundance information of functional groups in laboratory standards in addition to a much wider region of non-normalized mid-infrared spectrum (1500–4000 cm^{-1}). The median seasonal OM/OC ratios of this study underpredict that of Bürki et al. (2020) by 0.12 on average, while reproducing the same temporal trends. Some of the discrepancies may be due to insensitivity of spectral features to molecular characteristics in certain domains – for instance, the variation of peak frequency $\tilde{\nu}_2$ diminishes with increasing molecular weight (Sect. 2.3.1). However, the overall agreement between the two methods is reasonable considering the indirect nature of estimates in our work (Fig. 2.12).

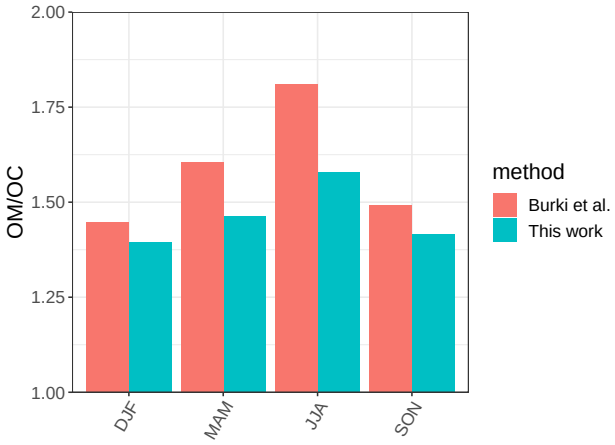

Figure 2.12 – Bar chart showing median OM/OC ratio calculated for each season based on samples collected in the Phoenix, AZ, monitoring site using our method and the one used by Bürki et al. (2020).

Molecular weight (MW)

The PLSR model estimates the mean molecular weight to range between 100–350 g mol^{-1} for majority of the samples (Fig. 2.14, second row). To the best of authors' knowledge no extensive study has been performed on mean molecular weight of ambient organic aerosol constituents. Nevertheless, the estimated range is reasonably close to that of the studies that have been done. Those studies measured molecular weights up to 200 g mol^{-1} for $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ constituents using GC-MS and ion chromatography (Cocker III et al., 2001; Jang and Kamens, 2001b; Kalberer, 2004), an average molecular weight between 200–300 g mol^{-1} for atmospheric HUmic-LIke Substances (HULIS) using electro-spray ionization (ESI) (Graber and Rudich, 2006), and an average molecular weight between 300–450 g mol^{-1} for oligomers formed in a smog chamber, measured using laser desorption/ionization mass spectrometry (LDI-MS) (Kalberer et al., 2006). Although particle-phase oligomerization processes result in high-MW compounds (Jang and Kamens, 2001a; Tolocka et al., 2004; Shiraiwa et al., 2014), the

abundance of these compounds is usually debated since the available experimental results regarding the reversibility of accretion reactions are contradictory (Kroll and Seinfeld, 2008). Moreover, oligomer formation may be overestimated in laboratory conditions compared to atmospheric particles (Kroll and Seinfeld, 2008; Kalberer, 2004; Trump and Donahue, 2014).

The PLSR molecular weight model estimates a lower mean molecular weight for rural samples ($\approx 200\,\mathrm{g\,mol^{-1}}$) compared to urban ones ($\approx 240\,\mathrm{g\,mol^{-1}}$), while burning samples are estimated to constitute the heaviest molecules ($\approx 290\,\mathrm{g\,mol^{-1}}$). This observation is consistent with our knowledge of emission sources. Emissions in urban areas are influenced by long-chain hydrocarbons from combustion products and motor oil (Gentner et al., 2012), while biomass burning is accepted to be the primary source of high-MW HULIS (Li et al., 2019). We also observe a decrease in mean molecular weight peak density in urban samples from winter to summer that is believed to be attributed to fragmentation during more intense photooxidation in summer (Hand et al., 2019; Jimenez et al., 2009), for emission sources that do not change drastically between the two seasons. The same phenomenon is observed in LDI mass-spectra of some urban samples in summer and winter reported by Kalberer et al. (2006). Although the reduction in mean molecular weight due to fragmentation can be compensated for by addition of heavy atoms to the molecule during oxidation, our results suggest that the overall direction of photooxidation in urban sites is reduction of the mean molecular weight.

Carbon number (n_C)

The PLSR carbon number model estimates that the recovered rural samples usually have lower mean carbon number compared to urban samples and the samples influenced by burning (Figure 2.14, third row). Higher mean carbon number estimates in urban sites (highest probability density around 16), which are coincident with high elemental carbon (EC) values from TOR measurements (Fig. A.2), can be attributed to major EC sources such as combustion of fossil fuel and biomass. This is also consistent with high $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ formation potential of molecules with 15–25 carbon in diesel fuel shown by Gentner et al. (2012). Samples affected by burning are estimated to have the highest mean carbon number among all samples. This observation is consistent with the emissions of plant cuticle waxes, mainly composed of straight-chain hydrocarbons, observed during biomass burning (Hawkins and Russell, 2010) as well as HULIS (Graber and Rudich, 2006). We also observe a decrease in estimated mean carbon number of urban samples from winter to summer suggesting fragmentation during aging and photooxidation processes.

The carbon-oxygen estimates of the PLSR models are consistent with the existing numerical simulation. We compared our estimates with the numerical simulations by Jathar et al. (2015). Multi-generational oxidation model used by Jathar et al. (2015) (Statistical Oxidation Model, SOM, in a 3-D air quality model) for simulating $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ in Los Angeles and Atlanta (two urban locations) shows that carbon number in $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ ranges from 3 to 15 with the concentration peaks around 7, 10 and 15 (Fig. 2.13). For this comparison, we calculated the carbon-oxygen grid from our molecular weight and carbon number estimates, assuming the

organic molecules have a chemical formula of $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ (a common assumption and one used by Jathar et al., 2015). Our PLSR models for the IMPROVE network estimate mean carbon number peaks (number density) for rural, urban, and burning samples to be around 8, 16 and 18 respectively, while the total range is limited to 3–19 (Fig. 2.13). We also estimate the oxygen number to range from 2 to 6 for the majority of the samples. It should be noted that this is as an order of magnitude comparison since the time frame and the location of the two studies are different and the numerical simulation by Jathar et al. (2015) only considers $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$.

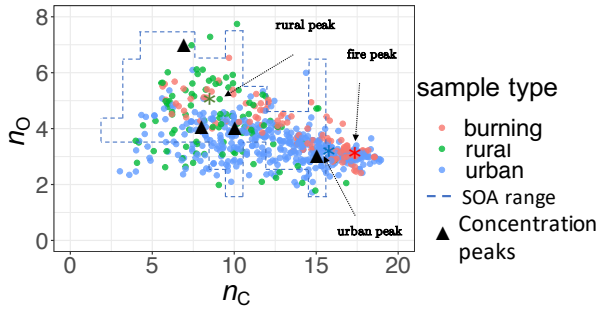

Figure 2.13 – Comparison between carbon-oxygen grid simulated by Jathar et al. (2015) for Atlanta and Los Angeles with sample points estimated for IMPROVE network (2011 and 2013) from the molecular weight and carbon number estimates of this study. The dashed lines show the range of simulated carbon and oxygen and the triangles indicate the location of the highest $C_{N_c}H_{2N_c+2-N_o}O_{N_o}$ concentrations for the simulations of Jathar et al. (2015).

2.3.4 Calibration model interpretation

Reducing the spectrum to four basic features introduced in Sect. 2.3.1 ($\tilde{\nu}_2$, A_3/A_2, A_1/A_2, w_2) is a manual data compression onto a basis set of interpretable variables. Though information loss is inevitable, it was shown in Sect. 2.3.1 that these basic features are still sufficient for qualitative explanation of spectral variations associated with different emission source and aerosol aging process. In this section, predictions made by the PLSR models on the ambient samples are grouped based on the four basic features using CART (Fig. 2.15) in order to form a better understanding of how the sophisticated PLSR models function.

The regression trees show that the peak ratios are observed to be the main grouping parameter for both carbon number and molecular weight (Fig. 2.15). The inverse relation of peak ratios with carbon number appears in most of the splitting nodes of carbon number and molecular weight regression trees (Fig. 2.15). This is consistent with the observed relation between carbon number and peak ratios in the calibration set (Fig. 2.8). Assuming that molecular weight is highly correlated with carbon number, the classification of molecular weight based on peak ratios is also expected. The peak frequency ($\tilde{\nu}_2$) appears once as a node in molecular weight tree and classifies the estimates based on the same trend that was observed in the calibration set (Fig. 2.6). The second peak width (w_2) also appears few times in the nodes

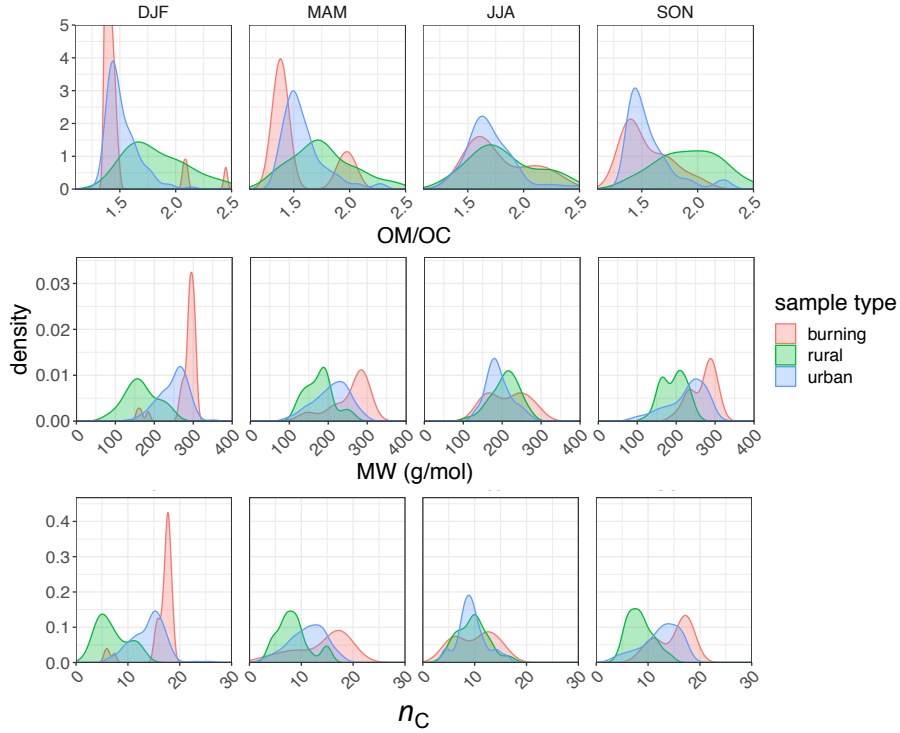

Figure 2.14 – Kernel density estimates of indirect OM/OC ratio, molecular weight (MW) and carbon number (n_C) estimated from normalized aliphatic C–H mid-infrared absorbances by PLSR models (segregated by sample type and season).

probably adding information about the OM/OC ratio and phase state. The two trees shown in Fig. 2.15 explain only around 50 % of the variation of estimates made by the PLSR models. The explained variation can be increased to an arbitrarily high number through the use of more branches in the fitting data set, but the predictive capability of regression trees for new samples depends highly on their similarity to the training set.

In summary, regression trees show that the predictions of the PLSR models are generally consistent with the observed trends of the basic features in the calibration set (Appendix A.1.6 supports this conclusion for individual spectra for which the PLSR models estimate quite different parameters). This observation implies that the PLSR predictions of carbon number and molecular weight are not independent of these basic features. However, the sophisticated PLSR models use other fine features in addition to the mentioned basic features to extract more detailed information and to reduce variabilities stemming from different sources such as baseline correction.

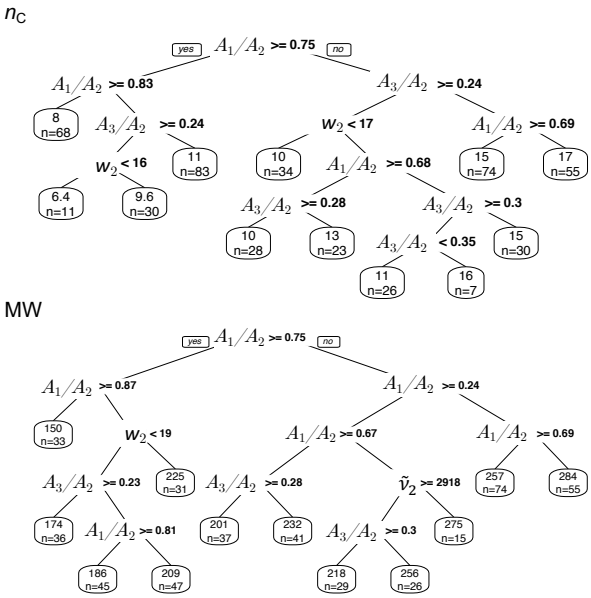

Figure 2.15 – Regression tree of molecular weight (MW) and carbon number (n_C) estimates in atmospheric samples based on the basic spectral features: second peak frequency ($\tilde{\nu}_2$), the ratio of peak heights of symmetric CH_2 stretching to asymmetric CH_2 stretching (A_1/A_2), the ratio of peak heights of asymmetric CH_3 to asymmetric CH_2 stretching (A_3/A_2) and second peak width (w_2) of aliphatic C−H band.

2.4 Concluding remarks

Normalized aliphatic C−H absorbances in mid-infrared spectrum were used in this study to estimate mean carbon number and molecular weight of the atmospheric OM. First, it was shown that the spectral features, such as peak frequencies and peak height ratios are correlated with carbon number, molecular weight, and the OM/OC ratio for laboratory standards. We also observed a meaningful temporal and spatial variation of those features in atmospheric aerosol samples. Thereafter, PLSR models were developed on laboratory standards to estimate the mentioned parameters in the atmospheric aerosol samples from the IMPROVE network. The estimated molecular weight and carbon number reconstruct the OM/OC values in the atmospheric aerosols that are consistent with previous studies with a reasonable difference (an average undeprediction of 0.12). These new statistical models estimate lower mean carbon number and mean molecular weight in more aged aerosols of similar sources, highlighting the fragmentation role in the aging process (Murphy et al., 2012). Moreover, they estimate relatively less oxidized, heavier molecules with higher carbon number for samples influenced by burning. The findings show that the new technique can help us better understand characteristics of OM due to emission sources and atmospheric processes. In addition, since carbon number and molecular weight are important characteristics used by recent conceptual models or parametrizations (e.g., Shiraiwa et al., 2017a; Li et al., 2016; Pankow and Barsanti, 2009; Kroll et al., 2011; Donahue et al., 2011) to describe the evolution in the OM composition, this

technique can provide semi-quantitative, observational constraints on these variations at the scale of the network as well as for laboratory experiments. We also found that the phase state of the laboratory standards clearly affects their spectroscopic features. These features can be used to develop predictive models that can estimate the phase state of atmospheric OM.

Only around 27 % of the existing samples could be analyzed with our approach due baseline correction limitations posed by low OM mass (compared to inorganic mass) on the filters. Under-sampling is more severe in rural sites although expected trends (such as higher OM/OC ratio) are observed even in the current subset. As a result, one should be cautious when extending the results of this study to draw general trends. Although some inaccuracy in the results is likely due to extrapolating from laboratory standards and the indirect nature of the approach (for which more research is needed), the estimates of molecular weight, carbon number, and the OM/OC ratio were shown to be reasonable. Further evaluation with different molecules and molecular mixtures can better constrain these estimates.

3 Characterization of primary and aged wood burning and coal combustion organic aerosols in an environmental chamber and its implications for atmospheric aerosols

Published in the Journal of Atmospheric Chemistry and Physics on 8 Jul. 2021

Authors : Yazdani, Amir; Dudani, Nikunj; Takahama, Satoshi; Bertrand, Amelie; Prévôt, André S. H.; El Haddad, Imad; Dillner, Ann M.

Contributions: IEH, ST and AY conceived the project and manuscript. AB and IEH performed the chamber experiments. AB provided AMS spectra. ND prepared and assembled the filter sampling set-up and took the FTIR spectra. AMD provided laboratory standard and ambient sample spectra. AY wrote the code for data analysis and post processing, performed the data analysis, prepared laboratory standards and analyzed samples, and wrote the manuscript. ST edited the manuscript and provided regular feedback on the analysis. IEH, ASHP, AB, AMD, and ND provided input on the analysis and the further editing of the manuscript. ST and IEH provided overall supervision of the project.

Reference: Yazdani et al. (2021c)

Abstract Particulate matter (PM) affects visibility, climate, and public health. Organic matter (OM), a uniquely complex portion of PM, can make up more than half of total atmospheric fine PM. We investigated the effect of aging on the secondary organic aerosol (SOA) concentration and composition formed from wood burning (WB) and coal combustion (CC) emissions, two major atmospheric OM sources, using mid-infrared (MIR) spectroscopy and aerosol mass spectrometry (AMS). For this purpose, primary aerosols were injected into an environmental simulation chamber and aged using hydroxyl (diurnal aging) and nitrate (nocturnal aging) radicals to reach an atmospherically-relevant oxidative age. A time-of-flight AMS instrument was used to measure the high-time-resolution composition of non-refractory fine PM, while fine PM was also collected on PTFE filters before and after aging for MIR analysis. AMS and MIR spectroscopy indicate an approximately three-fold enhancement of organic aerosol (OA) concentration after aging (even without wall-loss correction). The OM:OC ratios also agree closely between the two methods and increase, on average, from 1.6, before aging, to 2, during the course of aging. MIR spectroscopy, which is able to differentiate among oxygenated groups, shows a distinct functional group composition for aged WB (high abundance of carboxylic acids) and CC OAs (high abundance of non-acid carbonyls) and detects aromatics and polycyclic aromatic hydrocarbons (PAHs) in emissions of both sources. The MIR spectra of fresh WB and CC aerosols are reminiscent of their parent compounds with differences in specific oxygenated functional groups after aging, consistent with expected oxidation pathways for the major volatile organic compounds (VOCs) emitted from each source. The AMS mass spectra also show variations with source and aging that are consistent with the MIR functional group (FG) analysis. Finally, a comparison of the MIR spectra of primary and aged WB aerosols in the chamber with that of ambient samples affected by residential wood burning and wildfires reveals similarities regarding the high abundance of organics, especially acids, and visible signatures of lignin and levoglucosan. This finding is beneficial for the source identification of atmospheric aerosols and interpretation of their complex MIR spectra.

3.1 Introduction

Particulate matter (PM) affects visibility and climate (Hallquist et al., 2009). For example, fine PM can play the role of cloud condensation nuclei (CCN) impacting cloud formation (McFiggans et al., 2004). PM can also considerably perturb the transfer of different wavelengths of electromagnetic radiation by scattering or absorption phenomena (Seinfeld and Pandis, 2016). In addition, exposure to ambient fine PM is estimated to have caused 8.9 million premature deaths worldwide per year (in 2015; Burnett et al., 2018). Organic matter (OM), which constitutes up to 90 % of total fine atmospheric PM, is a key factor in aerosol-related phenomena (Russell, 2003; Shiraiwa et al., 2017b). However, its chemical composition and formation mechanisms have not yet been fully characterized due to its compositional complexity (Kanakidou et al., 2005; Turpin et al., 2000).

Biomass burning particulate emissions (including those from residential wood burning, prescribed burning, and wildfire) are major contributors to total atmospheric OM with an increas-

ing importance due to rising wildfire activities (Westerling, 2016; DeCarlo et al., 2008; Sullivan et al., 2008). Biomass burning primary organic aerosols (POAs) account for 16 % to 68 % of total OM mass in Europe (Puxbaum et al., 2007; Paglione et al., 2020) and more than 30 % of $PM_{2.5}$ mass in samples collected in Beijing (Srivastava et al., 2021). Coal combustion (for electricity and heat generation) is another major POA source in China and some regions of Europe (Haque et al., 2019; Junninen et al., 2009), emitting considerable amounts of carcinogenic and mutagenic polycyclic aromatic hydrocarbons (PAHs) (Sauvain et al., 2003). Approximately 40 % of the world's electricity (and up to 66 % in China) is generated in coal-fueled power plants (World Coal Association, 2020). Biomass burning and coal combustion are also believed to be responsible for a large fraction of secondary organic aerosol (SOA), especially in winter when biogenic emissions are largely absent (Qi et al., 2019; Lanz et al., 2010; Zhang et al., 2020; Paglione et al., 2020). Recent studies highlight the contribution of biomass burning by showing the predominance of carbon from non-fossil-fuel origins in SOA even in industrial regions (Haddad et al., 2013; Beekmann et al., 2015). To date, primary biomass burning emissions have been investigated in several works (e.g. Johansson et al., 2004; Bäfver et al., 2011; Alves et al., 2011). However, SOAs and their chemical compositions have not been studied extensively until recently (e.g. Bertrand et al., 2017; Tiitta et al., 2016; Bruns et al., 2015) due to sophisticated experimental set-up requirements.

The determination of organic aerosols (OA) and SOA chemical composition, involves a large range of analytical and computational techniques. Aerosol mass spectrometry (AMS) and mid-infrared (MIR) spectroscopy are two methods capable of analyzing most of OA mass in addition to providing information about chemical class or the functional groups (FGs) (Hallquist et al., 2009). In an AMS, non-refractory aerosol is first vaporized (normally at 600 °C). Thereafter, the vaporized fraction is turned to ionized fragments and is then detected by the mass spectrometer to obtain on-line atomic composition of non-refractory aerosol. This method is well characterized for its capability to estimate elemental composition with a high time resolution and detection limit (Canagaratna et al., 2007, 2015). There are, however, some known challenges: particle collection efficiency of aerodynamic lens section (Canagaratna et al., 2007); particle bounce back in the vaporizer (Kumar et al., 2018); potential reactions occurring in ion chamber (Faber et al., 2017); and, most importantly, extensive molecule fragmentation (Faber et al., 2017; Canagaratna et al., 2007) with the common ionization method (i.e. electron ionization, EI) that makes the interpretation of the AMS mass spectrum difficult.

MIR spectroscopy, which is commonly performed off-line on polytetrafluoroethylene (PTFE) filters, provides direct information about FG abundances in OA collected on the filters. This information can be converted to OM, OC (organic carbon) and the OM:OC ratio with the aid of statistical models and laboratory standards with a few assumptions (Boris et al., 2019; Ruthenburg et al., 2014; Reggente et al., 2016; Coury and Dillner, 2008). Recent studies show good agreement between MIR OM and OC estimates and the thermal optical reflectance (TOR) OC and the residual OM method for monitoring networks (Boris et al., 2019). The main advantage of MIR spectroscopy over other common methods is that it is relatively

fast, inexpensive, and non-destructive to the filter sample during analysis (Ruthenburg et al., 2014). However, sampling for MIR spectroscopy is usually performed for at least several minutes resulting in lower temporal resolution (Faber et al., 2017). Moreover, the presence of overlapping peaks complicates the interpretation of the MIR spectrum. Collocated AMS and MIR measurements can combine the advantages of both techniques and provide high-time-resolution measurements with FG quantification (e.g. Faber et al., 2017; Chen et al., 2016; Frossard et al., 2014; Russell et al., 2009a; Chhabra et al., 2011b). The uncertainties of OC and OA mass concentrations derived based on Fourier transform infrared spectroscopy (FTIR) and AMS have been reported to be within 35 % (Dillner and Takahama, 2015; Gilardoni et al., 2009; Russell et al., 2009b; Frossard et al., 2011; Liu et al., 2011; Corrigan et al., 2013; Frossard et al., 2014; Reggente et al., 2019a) and 25 %, respectively (Canagaratna et al., 2007). Precision for replicate measurements with the same instrument has been shown to be substantially higher (e.g., Debus et al., 2019).

This study is one of the few examples of AMS and MIR spectroscopy being combined to provide a superior chemical resolution for analyzing burning emissions in an environmental chamber and in the atmosphere. In this work, a series of wood burning (WB) and coal combustion (CC) experiments were conducted in an environmental chamber at the Paul Scherrer Institute (PSI). An AMS measured the chemical composition of OA throughout the aging process while fresh and aged fine aerosols were collected on separate PTFE filters, making it possible to combine the measurements of the two methods. We investigated the MIR spectra and FG composition of POA and SOA formed after diurnal and nocturnal aging processes in relation to their parent compounds and the oxidation products of their identified volatile organic compounds (VOCs). These results were combined with the high-resolution AMS mass spectra to evaluate the consistency of the two techniques and to better understand differences in the chemical composition caused by different emission sources and aging. Finally, the MIR spectra of the chamber biomass burning samples were compared to those of some atmospheric burning-influenced aerosols collected at the Interagency Monitoring of PROtected Visual Environments (IMPROVE) network (http://vista.cira.colostate.edu/improve/) to understand their similarities and to develop a method for the identification of atmospheric burning-influenced samples using MIR spectroscopy.

3.2 Methods

In the following sections, the experimental set-up (Sect. 3.2.1), on-line and off-line measurement techniques (Sects. 3.2.2 and 3.2.3), and atmospheric sample collection (Sect. 3.2.4) are described in detail. Thereafter, the statistical methods used for post-processing are discussed in Sects. 3.2.5 and 3.2.6.

3.2.1 Laboratory experimental set-up and procedure

Four wood burning (WB) and five coal combustion (CC) experiments were conducted in a collapsible Teflon chamber of 6 m^3 at the Paul Scherrer Institute (PSI). We studied the effects of fuel source and diurnal/nocturnal aging on the chemical composition of the emissions. The experimental set-up in this work was similar to that used by Bertrand et al. (2017) (Fig. A.9).

For the WB experiments, we have followed the procedures developed in Bertrand et al. (2017, 2018a), which favor smoldering-dominated wood fires. Three beech wood logs (approximately 2.5 kg) without bark, and an additional 300 g of kindling (beech) were burned in a modern wood stove (2010 model). The logs were ignited using three fire starters composed of wood shavings, paraffin and natural resins. The moisture content of the logs was measured to be around 11 %. Each burning experiment was started with a lighter followed by immediate closing of the burner door. Emissions past the ignition, in which kindling wood and starters were fully combusted, were injected into the chamber.

In the CC experiments, 300 g of bituminous coal from Inner Mongolia (63 % carbon content) was burned. First, the ash drawer of the stove was loaded with kindling wood (beech), which was ignited and served to ignite the coal. The wood was removed from the drawer after proper ignition of the coal. The emissions past the ignition phase were injected in the chamber via a single injection. Klein et al. (2018) have shown that the temperature in the stove at the starting phase significantly affects the total emission rates of SOA precursors and to a lesser extent their composition. Here, the ignition temperature spanned a similar range as in Klein et al. (2018). Control experiments were performed to evaluate the effect of kindling wood on the emissions, for which we followed a similar procedure as for the real experiments but without putting coal in the stove. The resulting emissions after removing the ignited kindling were not different from background for both particle and gas phases.

In both CC and WB experiments, the injection was continued (from 5 to 25 minutes) until the measured concentration of primary organic aerosol by the high-resolution time-of-flight (HR-TOF) AMS reached values of approximately 20 $\mu g\, m^{-3}$. WB and CC samples were extracted from the chimney and diluted using an ejector diluter (DI-1000, Dekati Ltd.) before being injected into the chamber. The lines from the chimney to the environmental chamber were heated to 413 K to limit semi-volatile compounds condensation in the lines. The average temperature and relative humidity of the chamber after injection were maintained at 293 K and 55–60 %, respectively. The emissions were left static for 30 minutes in the chamber after injection to ensure proper mixing and for sampling and characterizing the primary organic aerosol. Thereafter, emissions were aged using the hydroxyl or nitrate radical for simulating diurnal and nocturnal aging mechanisms.

The OH radical was produced by the photolysis of nitrous acid (HONO) continuously injected into the chamber, using UV lights (40 × 100 W, Cleo Performance, Philips). HONO was generated by reacting constant flows of diluted sulfuric acid (H_2SO_4) and sodium nitrite ($NaNO_2$) in a custom gas flask. Pure air passed through the flask and then a PTFE membrane filter

before being injected into the chamber to ensure particle-free flow of HONO (the procedure is explained by Taira and Kanda, 1990; Platt et al., 2013). Before aging, 1 µL of deuterated butanol-D9 (98%, Cambridge Isotope Laboratories) was injected into the chamber to measure the OH radical exposure (Barmet et al., 2012). The concentration of butanol-D9 was monitored by a proton transfer reaction time-of-flight mass spectrometer (PTR-ToF-MS 8000, Ionicon Analytik). Emissions were aged for around four hours to reach OH exposures of (2–3)$\times 10^7$ molec cm^{-3} h corresponding to 20–30 hours of aging in the atmosphere (assuming a 24-hour average OH concentration of 1×10^6 molec cm^{-3} in the atmosphere; Seinfeld and Pandis, 2016). For the nocturnal aging experiments, the NO_3 radical was produced by a single injection of O_3 and NO_2 in the chamber. The NO_2:O_3 ratio we have used is approximately 1, and their concentrations were approximately 50 ppb. The contribution of NO_x from combustion was less than 10 ppb. The concentration of NO_3 was inferred from the reactivity of phenol emitted from wood burning and coal combustion (k_{NO_3} = 3.9 × 10^{-12} cm^3 molec^{-1} s^{-1}). We calculated an initial concentration of NO_3 of 1–2.5 × 10^7 molec cm^{-3}. The effects of vapor wall losses of primary semi-volatile species and of oxidized vapors in our chamber are detailed in (Bertrand et al., 2018a) and in (Jiang et al., 2020), respectively, and are beyond the scope of this paper.

After each experiment, the chamber was cleaned by injecting O_3 for 1 hour and irradiating with a set of UV lights while flushing with pure air. Then, the chamber was flushed with pure air in the dark for at least 12 hours (similar to the procedure described by Bruns et al., 2015). The pure air injection system consists of a generator (Atlas Copco SF 1 oil-free scroll compressor with 270 L container, Atlas Copco AG, Switzerland) coupled to an air purifier (AADCO 250 series, AADCO Instruments, Inc., USA), which provides a hydrocarbon background of sub 10 ppbC. The background particle- and gas-phase concentrations were measured in the clean chamber before each experiment. Blank experiments were performed, in which the chamber was filled with either pure air or a mix of pure air and ambient air sampled through the heated sampling system and the lights were switched on. In these experiments, approximately 1 µg m^{-3} of organic aerosol was formed.

3.2.2 On-line PM measurement

After the primary emission injections, PM emissions in the chamber were monitored using two on-line techniques. Non-refractory particle composition was measured at a temporal resolution of 30 seconds by a HR-TOF AMS (Aerodyne Research Inc.; DeCarlo et al., 2006) operating in V mode (mass resolution $\Delta m/m = 2000$), with a vaporizer temperature of 600 °C and pressure of approximately 10^{-7} Torr, and EI operating at 70 eV, equipped with a 2.5 µm inlet aerodynamic lens. Data post processing was performed in Igor Pro 6.3 (Wave Metrics) using SQuirrel 1.57 and Pika 1.15Z routines. The elemental and OM:OC ratios were determined according to Aiken et al. (2008). The reported OA concentrations were not wall-loss corrected. The AMS ionization efficiency was calibrated using NH_4NO_3 particles. A condensation particle counter (CPC, 3025A TSI) measured total particle number concentrations and a scanning mobility particle sizer (SMPS, CPC 3022, TSI) measured particle size distribution. Particles

were dried (Nafion, Perma Pure LLC) upstream of the AMS, SMPS and CPC. The AMS collection efficiency was verified using SMPS and ranged between 0.7–1.1 and therefore was assumed to be approximately 1 for our conditions.

3.2.3 Off-line PM sampling and measurement

Primary and aged PM emissions were collected on separate PTFE filters (47 mm diameter Teflo® membrane, Pall Corporation) for 20 minutes after injection of primary emissions into the chamber and after four hours of aging. The aerosol collection area was limited to a circle with a diameter of 1 cm in the center of the filter using PTFE masking elements placed above and below the filters (Russell et al., 2009b). Sampling on PTFE filters was performed at a flow rate of 8 L min^{-1} using a flow system composed of a sharp-cut-off cyclone (1 µm at a nominal flow rate of 16 L min^{-1}) and a silica gel denuder. Hereafter, these PTFE filters are referred to by their fuel, and oxidant: e.g. WB_OH refers to the filters corresponding to the WB experiments aged with OH. After sampling, filters were immediately stored in filter petri dishes at 253 K before MIR analysis to minimize volatilization and chemical reactions. The PTFE filters were analyzed using a Bruker-Vertex 80 FT-IR instrument equipped with an α deuterated lanthanum alanine doped triglycine sulfate (DLaTGS) detector, at a resolution of 4 cm^{-1}. The FTIR sample chamber was continuously purged with dry air treated with a purge gas generator (Puregas GmbH) to minimize water vapor and carbon dioxide interferences.

3.2.4 Atmospheric samples (IMPROVE network)

Particulate matter with diameter less than 2.5 µm ($PM_{2.5}$) was collected on PTFE filters (25 mm diameter Teflo® membrane, Pall Corporation) every third day for 24 hours, midnight to midnight, at a nominal flow rate of 22.8 L min^{-1} during 2011 and 2013 at selected sites in the IMPROVE network (approximately 3050 samples from 1 urban and 6 rural sites in 2011 and 4 urban and 12 rural sites in 2013). The PTFE filters were analyzed using a Bruker-Tensor 27 FT-IR instrument equipped with a liquid nitrogen-cooled, wide-band mercury-cadmium-telluride (MCT) detector, at a resolution of 4 cm^{-1}. In this work, atmospheric samples were divided into four mutually exclusive sub-groups: urban, rural, residential wood burning, and wildfire (residential wood burning, and wildfire samples were identified by Bürki et al., 2020). Bürki et al. (2020) separated and identified the burning-influenced samples in the same dataset using cluster analysis and subsequent analysis of the clusters. They divided the burning-influenced samples into residential wood burning and wildfire sub-categories by extending further down the hierarchical tree. The residential wood burning sub-category was labeled according to its occurrence during winter months in Phoenix, AZ (Bürki et al., 2020), where residential wood burning commonly takes place (Pope et al., 2017). The wildfire sub-category was labeled due to its occurrence during a known fire event (Rim Fire, CA, 2013).

3.2.5 Post-processing of MIR spectra to identify and quantify functional groups in laboratory and IMPROVE samples

After obtaining the MIR spectra of the laboratory and ambient samples, they were processed the same. The post-processing enabled the quantification of alcohol COH (aCOH), carboxylic acid (COOH), aliphatic CH (aCH), and non-acid carbonyl (naCO) functional groups and the identification of PAHs, organonitrates, levoglucosan, inorganic sulfate, and nitrate. In this section, the methods used for spectral post-processing are described in detail.

Baseline correction

Baseline correction was performed to eliminate the contribution of background drift, light scattering by filter membrane and particles, and absorption by carbonaceous material due to electronic transitions (Russo et al., 2014; Parks et al., 2019). For this purpose, we used a smoothing spline method similar to the approach taken by Kuzmiakova et al. (2016). In this method, a cubic smoothing spline was fitted to the raw spectra (excluding organic FG bands) and then was subtracted from them to obtain the net absorption due to FG vibrations at each wavelength. The current version extends the baseline correction algorithm by Kuzmiakova et al. (2016) (limited to 1500–4000 cm^{-1}) to the 400–4000 cm^{-1} range.

Blank subtraction

Although PTFE filters are optically thin in the MIR range, they have several absorbing bands due to the C−F bond vibrational modes (e.g. at 1000–1320 cm^{-1}; Quarti et al., 2013). These bands overlap with some organic and inorganic function group bands and limit the information that can be extracted from the analysis of OA on PTFE filters. To mitigate this issue, a scaled version of a baseline-corrected blank filter spectrum (the contribution of filter membrane scattering was excluded, and only PTFE absorptions were maintained) was subtracted from the baseline-corrected sample spectra to retrieve some of the overlapping features (Fig. A.10). The scaling procedure compensates for the variation in blank filter absorbances due to factors, such as non-uniformity in PTFE membrane morphology due to manufacturing variability and/or aerosol loading differences between filters and within the same filter (Debus et al., 2019; Quarti et al., 2013) by scaling the C−F peak at 1210 cm^{-1}. This approach is different from those taken by Takahama et al. (2013) and Maria et al. (2003) that subtracted a scaled raw spectrum of a blank filter from sample spectra before baseline correction. The blank subtraction algorithm allowed us to identify bands related to aromatics and PAHs at 690–900 cm^{-1} (Centrone et al., 2005), organonitrates ($RONO_2$) at 850 cm^{-1}, alcohol CO stretching at 1050 cm^{-1}, levoglucosan bands at 860–1050 cm^{-1}, and inorganic sulfate and nitrate bands at 620 and 1400 cm^{-1}, and also it allowed us to better quantify the carbonyl absorbances at around 1700 cm^{-1}.

Quantifying organic functional groups

After baseline correction and blank subtraction, the multiple peak-fitting algorithm described by Takahama et al. (2013) and implemented by Reggente et al. (2019b), functioning based on non-linear least squares analysis, was applied to the spectra to obtain major FG abundances of aCOH, COOH, aCH, and naCO (Yazdani et al., 2021b). The $RONO_2$ group abundances were not quantified due to the extensive overlap of its absorbances with other compounds and in order to keep the MIR estimates consistent with those of AMS, for which only total (organic plus inorganic) nitrate was estimated. After estimating FG abundances, the O:C, H:C and OM:OC ratios were calculated with few a assumptions about the number of carbon atoms attached to each FG (refer to Chhabra et al., 2011b; Russell, 2003; Maria et al., 2002). A few other peaks attributed to aromatics and PAHs were analyzed qualitatively due to the lack of calibration models. In addition to the common FG analysis, we used MIR fingerprint features to identify relevant substances to biomass burning (e.g. levoglucosan and lignin).

3.2.6 Dimensionality reduction of AMS mass spectra

While AMS provides a well-characterized, time-resolved measurement of OA aging, the extensive fragmentation of molecules, high number of ion fragments, and collinearity of ion fragment intensities make the interpretation of AMS mass spectra complex (Faber et al., 2017; Canagaratna et al., 2007). We used principal component analysis (PCA; Hotelling, 1933) to reduce the dimensionality of the AMS mass spectra in order to identify the most important drivers of variability in the spectral data and their connection with the FG composition of OA. The advantage of PCA analysis over analysis of the normalized conventional mass fragments (e.g. f_{43} and f_{44} in Fig. 3.5) is that PCA loads the highly correlated fragment ions onto the same principal components (PCs) that are orthogonally oriented to each other, thereby reducing redundancy among dimensions. Furthermore, PCA describes the range of variability spanned specifically by this dataset, accentuating smaller variations that might be lost using the conventional ranges spanned by normalized mass fragment analysis.

Before applying PCA, mass spectra at each measurement were normalized using the corresponding OA concentrations to eliminate signal variability due to changes in OA concentration. The correlation matrix of the normalized AMS mass spectra (correlation of ion intensities at different m/z values) shows that signals at several m/z values are correlated (Fig. A.11). Thus, the dimensionality of the mass spectra can be reduced considerably without significant loss of information. Our data matrix, $\mathbf{X}$, is a $i \times j$ matrix with i observations (2979 AMS measurements) and j variables (335 fragment ions) with a rank l ($l \leq \min\{i, j\}$). The columns of $\mathbf{X}$ are centered to avoid intercepts in scores. The PCA calculation was performed by singular value decomposition of the centered data matrix (Abdi and Williams, 2010), $\mathbf{X} = \mathbf{P\Delta Q}^T$, where $\mathbf{P}$ is the $i \times j$ matrix of left singular vectors, $\mathbf{Q}$ is the $j \times l$ matrix of right singular vectors, known as loadings, and $\mathbf{\Delta}$ is the diagonal matrix of singular values ($\mathbf{\Delta}^2$ is the diagonal matrix of eigenvalues of covariance matrix, $\mathbf{X}^T\mathbf{X}$). The $i \times l$ matrix of factor scores, $\mathbf{F}$, is obtained as $\mathbf{F} = \mathbf{P\Delta}$.

3.3 Results and discussions

In this section, peak fitting is applied to MIR spectra to estimate major FG concentrations. Thereafter, relative FG abundances are compared in different OAs to understand compositional differences arising due to aerosol source and age (Sects. 3.3.1 and 3.3.3). In addition, the MIR spectra of primary and aged OA emissions are compared to those of the fuel sources for identifying the chemical similarities between POA and fuel sources and for better understanding the important oxidation pathways for SOA formation (Sects. 3.3.2 and 3.3.4). In Sec. 3.3.5, AMS mass spectra are analyzed using PCA to understand the factors driving the spectral variability and how they relate to FG composition. Finally, the MIR spectra of chamber WB aerosols are compared to those of previously identified atmospheric burning-influenced samples, and a new method for identifying biomass burning aerosols using MIR spectroscopy is proposed (Sect. 3.3.6).

3.3.1 Wood burning – functional group composition

The primary WB aerosols have a high abundance of the aCH group (Figs. 3.1a and A.14a). This FG, which absorbs in the 2800–3000 cm^{-1} region in the MIR spectrum (Fig. 3.2a), constitutes around 50 % of the total mass of fresh WB OA (Fig. 3.1a). The aCOH group, which appears as a broad peak around 3400 cm^{-1} (Fig. 3.2a), has the second highest concentration (around 30 % of the total mass; Figs. 3.1a). This FG is also ubiquitous in hemicellulose, cellulose, and lignin – three main components that constitute 20–40, 40–60, and 10–25 wt. % of lignocellulosic biomass, respectively (McKendry, 2002). The COOH group, which appears as a broad peak in the 2400–3400 cm^{-1} range and a sharp carbonyl peak at approximately 1700 cm^{-1} (Fig. 3.2a), is the third most abundant FG in fresh WB aerosols and constitutes 10–20 % of the OA mass (Figs. 3.1 and 3.2a). The primary WB samples have OM:OC ratios ranging from 1.6 to 1.8 (Figs. 3.2b and A.14b).

AMS and MIR estimates of OA concentration are highly correlated (R^2 = 0.92; refer to Yazdani et al., 2021b, for detailed comparison) and show an almost three-fold increase in OA mass concentration with aging. The relative aCH abundance decreases substantially (by up to 30 %) in WB aerosols with aging (Figs. 3.1a and A.14a). The relative decrease in the aCH abundance is less prominent when aerosols are aged with the nitrate radical, probably due to the fact that organonitrates are excluded from quantitative analysis, thus the total OA concentration being underestimated and also because only a limited number of precursors react with the nitrate radical, leading to different SOA species. The aCH profile changes from the superposition of sharp peaks (observed for long-chain hydrocarbons) in primary WB spectra to broad peaks (observed in the spectra of oxygenated species) in aged WB spectra (Fig. 3.2a; Yazdani et al., 2021a). The aCOH relative abundance in WB aerosols also decreases with aging (Fig. 3.1a). Relative abundances of COOH increase significantly in WB aerosol with aging, suggesting carboxylic acid formation to be the dominant VOC oxidation pathway for biomass burning (Fig. 3.1). The aged WB samples with high carboxylic acid concentration have a broad OH

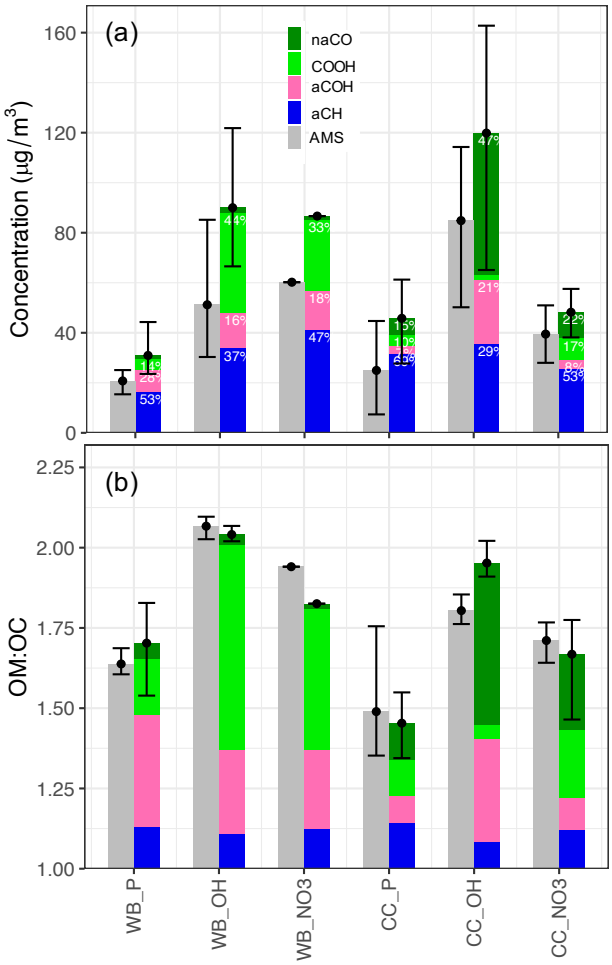

Figure 3.1 – (a) Bar plot of averaged MIR (separated by functional group) and AMS OA concentration estimates without wall loss correction. The contribution of each functional group (with contribution > 5%) to total OA, the type of aerosol (P: primary), emission source (WB and CC), and oxidant used for aging (OH: hydroxyl radical, NO_3: nitrate radical) are indicated for each category. (b) Bar plot of averaged MIR and AMS OM:OC estimates separated by the contribution of each functional group. The error bars show the maximum and minimum concentrations and OM:OC values for each category. For the estimates of each individual experiment refer to Fig. A.14.

peak ranging from 2400 to 3400 cm^{-1} and their carbonyl absorption frequency is on the lower end of its range (approximately 1708 cm^{-1} compared to 1715 cm^{-1} for ketone carbonyl; Fig. 3.2a) due to the weakening of the C=O bond in dimerized acids (Pavia et al., 2008). Fourier self-deconvolution (Kauppinen et al., 1981)helps identifying overlapping peaks in the complex 1500–1800 cm^{-1} region (Fig A.12). Phenol, methoxyphenols, and naphthalene are among the most important SOA precursors, based on their SOA yields, present in wood smoke reported by Bruns et al. (2016)and Stefenelli et al. (2019). The high abundance of COOH in the aged WB

OA of this study is consistent with the considerable carboxylic acid formation reported from these precursors (Chhabra et al., 2011a; Kautzman et al., 2010; George et al., 2015). The aged WB samples have the highest OM:OC ranging from 1.8 to 2.1 (Figs. 3.2b and A.14), with high concentrations of COOH.

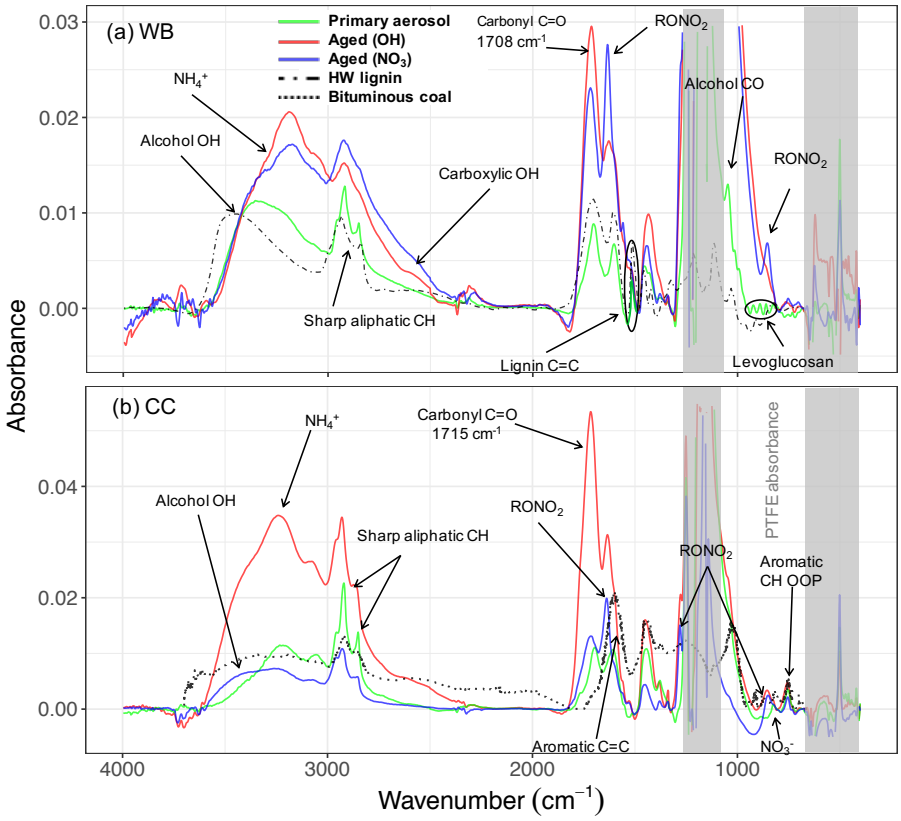

Figure 3.2 – The MIR spectra of primary and aged (with OH and NO_3) wood burning (WB) and coal combustion (CC) OAs and their parent compounds: lignin and bituminous coal. The dash-dotted line demonstrates the spectrum of solvent-extracted hardwood (HW) lignin (in KBr pellets) taken from Boeriu et al. (2004). The dotted line shows the spectrum of pulverized bituminous coal taken from Zhang et al. (2015).

MIR spectroscopy is able to distinguish between organic and inorganic nitrates due to differences in their absorption frequencies (Day et al., 2010). In this work, we have investigated the variations in the $RONO_2$ bands qualitatively by analyzing their absorbances at 1630, 1273, and 850 cm^{-1} (Day et al., 2010). These peaks are negligible in the primary WB aerosols. However, their absorbances (thus the abundance of organonitrates) increase in WB aerosols aged by both hydroxyl and nitrate radicals. with much more prominent contributions, when the nitrate radical is used (Fig. 3.2a). This suggests that the $RONO_2$ formation could be an important SOA formation pathway from WB emissions, which, nevertheless, we do not account for quantitatively. A relatively small peak at 1560 cm^{-1} is also observed in WB aerosols aged by the nitrate radical, which can be attributed to nitroaromatics (Pavia et al., 2008). Weak signatures

of aromatics or PAHs are visible at 690–900 cm^{-1} in the MIR spectra of both primary and aged WB aerosols. The absorption in this region appears as a single peak centered around 754 cm^{-1}. The intensity of this peak is correlated (R^2 = 0.70) to the concentration of fragments in the AMS mass spectra that were reported to be attributed to aromatics (Bruns et al., 2015; Pavia et al., 2008) (Fig. A.18). AMS results suggest aromatics and PAHs in WB aerosols constitute up to 8 % of total OA. However, underestimation is possible due to the incomplete list of ion fragments considered in this work and fragmentation of oxygenated aromatics and PAHs during EI ionization (Appendix A.2.6).

3.3.2 Wood burning – primary signatures

Using the MIR signatures of levoglucosan in primary WB aerosols (Fig. A.19), it is estimated that 22–48 % of the aCOH absorbance and 15–29 % of total OA mass is due to the presence of levoglucosan as a main product of high temperature pyrolysis of cellulose and an important biomass burning marker (Shen and Gu, 2009; Puxbaum et al., 2007; Hennigan et al., 2010). The reported range is consistent with measurements by a thermal desorption aerosol gas chromatograph coupled to a HR-TOF AMS (TAG-AMS) TAG-AMS obtained by Bertrand et al. (2018a). The abundance of the AMS m/z 60 fragment ion, related to levoglucosan fragmentation, (Schneider et al., 2006) and its MIR absorbances (used for the first time in this work) are highly correlated (R^2 = 0.76; Fig. A.20), both showing a consistent decrease in levoglucosan absolute concentration with aging regardless of the type of oxidant (Fig. A.20). The contribution of levoglucosan to total OA mass is estimated to decrease to less than 5 % after aging. This is consistent with the significant enhancement in OA mass with SOA formation and the degradation of levoglucosan through its reaction with OH and its loss to the chamber walls (Bertrand et al., 2018b; Hennigan et al., 2010; Zhao et al., 2014).

The MIR spectra of hard wood lignin and fresh WB aerosols are very similar, suggesting the presence of similar molecular structures between the primary WB aerosol composition and its parent compound, lignin (Fig. 3.2a). One specific aspect of this similarity is the high absorbances of the aCOH group in both samples. By comparing the MIR spectra of fresh WB aerosols with that of lignin (Appendix A.2.8), it can be inferred that the sharp peak at 1515 cm^{-1} and a part of the broader peak at 1600 cm^{-1} are attributed to aromatic rings in lignin structure (Hergert, 1960; Yang et al., 2007; Derkacheva and Sukhov, 2008) observed also in the spectrum of lignin monomers such as coniferyl alcohol (Bock and Gierlinger, 2019) and pyrolysis products of lignin with the same substitution pattern (Duarte et al., 2007; Simoneit et al., 1993). The peak at 1515 cm^{-1}, however, diminishes with aging, implying a change in aromatic ring substitution or ring opening by oxidant attack. The peak around 1600 cm^{-1} is suppressed by the $RONO_2$ peak around 1630 cm^{-1} after aging. Overall, our analysis shows that the large majority of the primary OA from WB emissions is composed of anhydrous sugars and lignin pyrolysis products.

3.3.3 Coal combustion – functional group composition

Peak fitting analysis suggests that the aCH group from all compounds containing this FG (alkanes and other compounds) constitutes around 60–80 wt. % of primary CC OA – a larger fraction compared to fresh WB OA (Figs. 3.1b and A.14b). The abundance of short and long-chain alkanes has also been reported in VOC emissions of bituminous coal (Liu et al., 2017; Klein et al., 2018). The naCO group is the second most abundant FG in the fresh CC OA, constituting, on average, 15 % of its mass (Fig. 3.1a), much higher than its contribution to WB emissions. This functional group has been reported to constitute 5–15 % wt. of bituminous coal VOC emissions (Liu et al., 2017; Klein et al., 2018). The concentration of the COOH group is usually low in the primary CC aerosols (around 10 %; Fig. 3.1a). The primary CC aerosols are estimated to have the lowest OM:OC ratios (1.35–1.5) among all samples, which is justified by their strong aCH absorbances (Fig. 3.2b). Inorganics (ammonium, sulfate, and nitrate) have prominent absorbances in the MIR spectra of fresh (and aged) CC aerosols, which have not been observed in the case of WB aerosols. The high abundance of inorganic salts can be attributed to the sulfur and nitrogen that are present in bituminous coal (Vasireddy et al., 2011).

In the aged CC aerosols, the relative abundance of the aCH group decreases drastically (on average 40 %), especially when the hydroxyl radical is used as oxidant, despite an almost three-fold increase in the OA mass concentration (without wall-loss correction; Figs. 3.1a and A.14a). On the other hand, the abundance of the naCO carbonyl group increases by up to 40 %, suggesting that carbonyl production is the dominant oxidation pathway for VOCs of CC emissions, predominantly when the hydroxyl radical is used (Fig. 3.1a). This observation is consistent with the high carbonyl abundance in SOA formed from the OH oxidation of alkane precursors, which are abundant in CC SOA, at close OH exposures to those in this work (Lambe et al., 2012; Lim and Ziemann, 2009). This does not, however, preclude the contribution of oxidation products of aromatic compounds, which are also abundant in CC VOCs (Liu et al., 2017; Klein et al., 2018), in aged CC OA. The absence of the broad carboxylic acid OH peak for these samples, and the peak location of the carbonyl group (1715 cm^{-1}) in their MIR spectra (Fig. 3.2b) suggest that the majority of the carbonyl group in the OH-aged CC aerosols is ketone (Pavia et al., 2008). An increase in the aCOH abundance is also observed in the CC aerosols with aging. This increase is more remarkable than that in the WB aerosols (Fig. 3.1).

The $RONO_2$ signature becomes clearly visible in aged CC aerosols, when the nitrate radical is used as oxidant (Fig. 3.2b) or when the hydroxyl radical is used in the presence of NO_x (> 50 ppb). This observation is consistent with VOC oxidation pathways proposed by Kroll and Seinfeld (2008) and measured by Ayres et al. (2015). In addition, a new peak at 1350 cm^{-1} that can be attributed to the S=O group in sulfonates (Pavia et al., 2008) appears in some aged CC aerosols (Fig. 3.2b).

The aromatic CH out-of-plane (OOP) absorption appears as a relatively stronger peak at 754 cm^{-1} compared to WB aerosols in the spectra of CC aerosols and bituminous coal (Fig.

3.2b; Sobkowiak and Painter, 1992; Sobkowiak and Painter, 1995). The normalized CH OOP absorbances by total OA mass are, on average, higher for CC aerosols compared WB, suggesting a higher contribution of this FG. The concentration of AMS fragment ions corresponding to PAHs suggest that aromatics and PAHs of CC aerosols account for up to 7 % of total OA mass, although some underestimation is possible due to the limited number of fragment ions considered and the fragmentation of oxygenated aromatics and PAHs upon ionization (Appendix A.2.6). These measurements suggest lower concentrations of aromatic compounds in CC OA compared to CC VOC emissions measured by Liu et al. (2017) and Klein et al. (2018).

The aged CC aerosols have slightly lower OM:OC ratios compared to the aged WB aerosols (Fig. 3.1b). For both emission sources, aerosols aged with the hydroxyl radial have higher OM:OC ratios (approximately 0.2) than those aged with the nitrate radical (Fig. 3.1b). This can be attributed to the different rate constants of VOC reactions with nitrate and hydroxyl radicals (Ziemann and Atkinson, 2012). Furthermore, organonitrates, which are abundant when aerosols are aged with the nitrate radical, are not considered either in AMS OA estimates or in MIR peak fitting. This exclusion causes the underprediction of both the OM:OC ratio and OA concentration estimates when the nitrate radical is used.

3.3.4 Coal combustion – primary signatures

Bituminous coal contains 69–86 wt. % carbon and its chemical structure is formed by highly substituted rings that are connected by alkyl or ether bridges (oxygen or sulfur) (Vasireddy et al., 2011). The MIR spectra of primary CC aerosols have weaker aromatic C=C absorbances at 1610 cm^{-1} than bituminous coal (Fig. 3.2b). This peak is suppressed by the $RONO_2$ peak at 1630 cm^{-1} with aging (Fig. 3.2b). The carbonyl peak (1710 cm^{-1}) observed in the spectra of the primary CC aerosols is absent in that of bituminous coal (Fig. 3.2b), suggesting carbonyl formation during coal combustion. Zhang et al. (2015) also reported a similar carbonyl generation for coal oxidation at high temperatures. We also observe that the aCH peaks are considerably sharper in the spectra of fresh CC aerosols than the pulverized coal (Fig. 3.2b), suggesting a higher relative contribution of this FG in primary CC OA than in bituminous coal.

The analyses of Sects. 3.3.1–3.3.4 suggest that aerosol source (WB and CC) and type (primary and aged with different oxidants) are both drivers of variability in the MIR spectra. These changes are substantially higher than uncertainties associated with OA mass concentrations derived based on FTIR and AMS. Furthermore, we found that aged aerosols of WB and CC, have distinguishable functional group compositions, suggesting that MIR spectra retain source class information at least up to the studied levels of aging.

3.3.5 Comparison between AMS mass spectra and FG composition of OA

In this section, we investigate the main differences in bulk chemical composition of the primary and aged emissions measured by the AMS and their FG content. We applied PCA to the

AMS mass spectra to facilitate their interpretation and to better understand the major sources of variability in the spectra and their connection with the FG composition of OA. Moreover, we projected AMS PMF (positive matrix factorization) factors obtained by Elser et al. (2016) and Aiken et al. (2009) (High Resolution AMS Spectral Database: http://cires.colorado.edu/jimenez-group/HRAMSsd/) onto the PC space to compare and contrast the AMS spectra obtained from the chamber experiments with AMS PMF factors of atmospheric samples. By applying PCA, the dimensionality of the normalized AMS mass spectra was reduced from 330 fragment ions to 3 PCs while explaining around 91 % of the variance in the mass spectra (Table 3.1).

Table 3.1 – Importance of the first three principal components.

	PC1	PC2	PC3
Standard Deviation	0.046	0.019	0.01461
Proportion of Variance	0.715	0.120	0.071
Cumulative Proportion	0.715	0.835	0.907

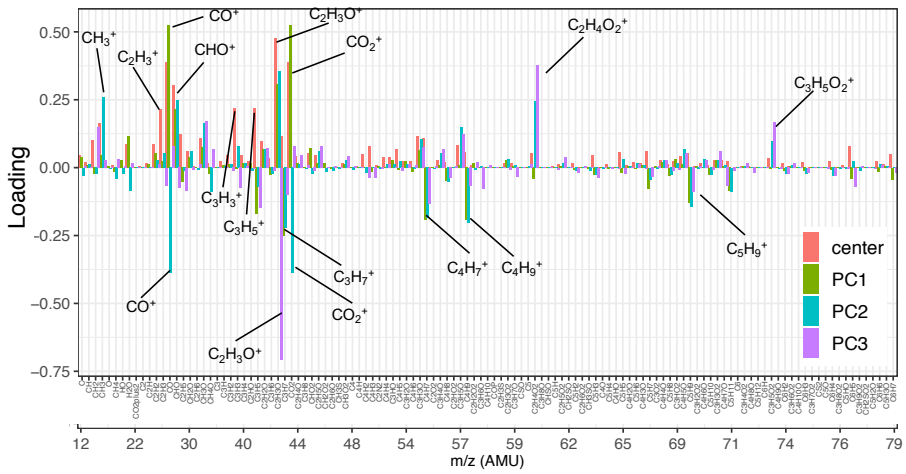

Figure 3.3 – Loadings of the first three principal components and the normalized mean AMS mass spectrum (shown up to m/z 80). Fragment ions with high positive/negative loadings are indicated by their formula. Refer to Fig. A.15 for the heavier fragments.

As can be seen from Fig. 3.3, the mean (center) mass spectrum, which represents the average of all mass spectra, has its highest values at a few ion fragments: $C_2H_3O^+$, CO^+, $CO_2{}^+$, CHO^+, $C_2H_3{}^+$, $C_3H_3{}^+$, and $C_3H_5{}^+$. The loadings, which are informative about the most important variations in the mass spectra, have also high values at a few mass fragments, making their interpretation simple. The first PC, which explains around 72 % of the variance alone, has high positive loadings of $CO_2{}^+$ and CO^+, CHO^+, and $C_2H_3O^+$ fragment ions (Fig. 3.3). The $CO_2{}^+$ fragment has been previously ascribed to mono- and dicarboxylic acids (Canagaratna et al., 2007; Faber et al., 2017; Russell et al., 2009a; Frossard et al., 2014). The CO^+ fragment is estimated directly from $CO_2{}^+$ (Aiken et al., 2008), and the other mentioned ion fragments represent other oxygenated species (Faber et al., 2017; Chhabra et al., 2011a). The high loadings of major oxygenated ion fragments and the empirical observation that primary and

aged samples are separated along the PC1 axis (Fig. 3.4) suggest that the first PC indicates the general direction of aging and the extent of oxidation, which appear to be the major sources of variation in the data (based on high explained variance by PC1). As can be seen from Fig. 3.4, PC1 scores are the lowest for primary CC and WB aerosols and increase with aging for both WB and CC samples. The higher oxidation state of aged WB aerosols compared to aged CC (i.e., higher PC1 score) aerosols, is consistent with the high abundance of COOH and naCO FGs in the aged WB and CC aerosols, respectively. The WB and CC samples that are aged with the hydroxyl radical have higher PC1 scores than those aged with the nitrate radical, indicating the former are more oxidized. The order of PC1 scores across the samples is reminiscent of the order of OM:OC ratios discussed earlier.

The second PC, which explains about 12 % of the variance, mainly contrasts primary WB (with high abundance of levoglucosan and lignin) with primary CC (with high abundance of aCH) aerosols by high positive loadings of $C_2H_4O_2{}^+$, representing levoglucosan fragmentation (Schneider et al., 2006), and $C_8H_9O_2{}^+$ (Fig. A.15), representing lignin fragmentation (Tolbert and Ragauskas, 2017; Saito et al., 2005), and negative loadings of $C_xH_y{}^+$ fragments, attributed to the aCH group (Faber et al., 2017). In addition, this PC has high positive loadings of $C_2H_3O^+$, CHO^+ mass fragments and high negative loadings CO^+ and $CO_2{}^+$ (Fig. 3.3), differentiating between the COOH group and other oxygenated FGs such as aCOH (Faber et al., 2017). The primary CC aerosols have the lowest PC2 scores due to high aCH content. Their PC2 scores increase with aging (Fig. 3.4a), indicating the production of CHO^+ and $C_2H_3O^+$ fragments (non-acid oxygenated FGs, such as the naCO and the aCOH groups) and decrease slightly before the end of aging probably due to the production of $CO_2{}^+$ (the COOH group) outweighing other oxygenated fragments. On the other hand, the primary WB aerosols have the highest PC2 scores (Fig. 3.4a) due to the high abundance aCOH, levoglucosan, and lignin. The PC2 scores reduce drastically with aging for WB aerosols, especially when the hydroxyl radical is used, suggesting the formation of the COOH group and degradation of levoglucosan and lignin (Fig. 3.4a). PC2 scores are higher for WB and CC samples aged with the nitrate radical compared to those aged with the hydroxyl radical (Fig. 3.4a) due to the lower concentration of the $CO_2{}^+$ fragment ion (Figs. 3.5 and A.16).

The first two principal components, which together explain 84 % of total variance, are able to separate aerosols according to their oxidation state (PC1) and their source (PC2) and show that the trajectories of CC and WB OA start to converge, especially when aging with the hydroxyl radical (Fig. 3.4a). This observation implies that the spectral differences between samples of the same or different categories decrease with aging despite the higher oxidation state (contribution to PC1) of the aged WB compared to the aged CC aerosols. The primary WB aerosols on PC1–PC2 biplot (Fig. 3.4a) are located close to biomass burning OA (BBOA) factors obtained by Aiken et al. (2009) and Elser et al. (2016), highlighting their chemical similarity. The same is observed for CC aerosols with coal combustion OA (CCOA) and hydrocarbon-like OA (HOA) factors. The OH-aged aerosols are, however, more similar to the semi-volatile oxygenated OA (SV-OOA) factor, representing less oxygenated OA (Fig. 3.4a) and have considerably lower PC1 compared to the aged OA (OOA) factor by Aiken et al. (2009), which represents aged SOA

(i.e., more aged than SV-OOA; not in the range of Fig. 3.4). The location of the samples in the f_{44}–f_{43} plot (far from the triangle vertex; Fig. 3.5) also suggests a moderate extent of oxidation for the aged WB and CC samples.

The third PC, which explains 7 % of the variance, mainly separates the aged aerosols based on the type of oxidant used and has high negative loading of $C_2H_3O^+$ and positive loadings of the $C_2H_4O_2{}^+$ and $C_3H_5O_2{}^+$ fragments (Fig. 3.3). PC3 scores have the highest values for the primary WB aerosols due to the abundance of levoglucosan and decrease with aging (Fig. 3.4b) due to the degradation of levoglucosan and the generation of the $C_2H_3O^+$ fragment (Fig. 3.5). PC3 scores are lower for the primary CC aerosols due to the absence of levoglucosan and decrease further with aging due generation of the $C_2H_3O^+$ fragment. For both sources, the aerosols aged with the nitrate radical have considerably lower PC3 scores (Fig. 3.4b) due to the high relative abundance of the $C_2H_3O^+$ fragment (Figs. 3.5 and S7). The aged CC aerosols with a high abundance of naCO also have relatively low PC3 scores (Fig. A.17b), suggesting a higher $C_2H_3O^+$ concentration for those samples (also observed from Fig. 3.5). This observation suggest that the species formed during aging with the nitrate radical and the naCO group produce higher concentrations of the $C_2H_3O^+$ fragment.

The PCA analysis shows that both aerosol source (WB and CC) and type (primary and aged with different oxidants) are responsible for variability in the AMS mass spectra (similar to the MIR spectra). We also found that the primary WB and CC aerosols have similar mass spectra to the BBOA and HOA factors, respectively, and OH-aged OAs of both sources are similar to the SV-OOA factor. Furthermore, the spectral variations are consistent with our functional group analysis via FTIR, suggesting that the AMS mass spectra maintain some functional group and source class information even after aging and in spite of the extensive fragmentation (discussed further by Yazdani et al., 2021b). However, even at the moderate levels of aging of this work, a part of this information only exists in higher PCs, which explain the lower variance in the data (e.g., PC2, explains 12 % of the variance across source classes and oxidative aging studied in this work and distinguishes OH-aged CC and NO_3-aged WB). These findings are consistent with past reports suggesting that AMS is most sensitive to aging Jimenez et al. (2009), and underscores the challenges in identifying source classes in highly aged atmospheric OA using AMS.

3.3.6 Atmospheric biomass burning-influenced aerosols

One of the main purposes of chamber experiments is understanding the characteristics of biomass burning aerosols in the atmosphere. We present a simplified comparison of the MIR spectra of WB aerosols in the chamber with the mean spectra of atmospheric samples affected by burning and other sources. Although the nitrate radical exposures in the chamber experiments were comparable to those of atmospheric samples, no comparable $RONO_2$ bands were observed in ambient samples. This observation implies that organonitrates potentially degrade later in the sequence of reactions due to thermal decomposition (Barnes et al., 1990;

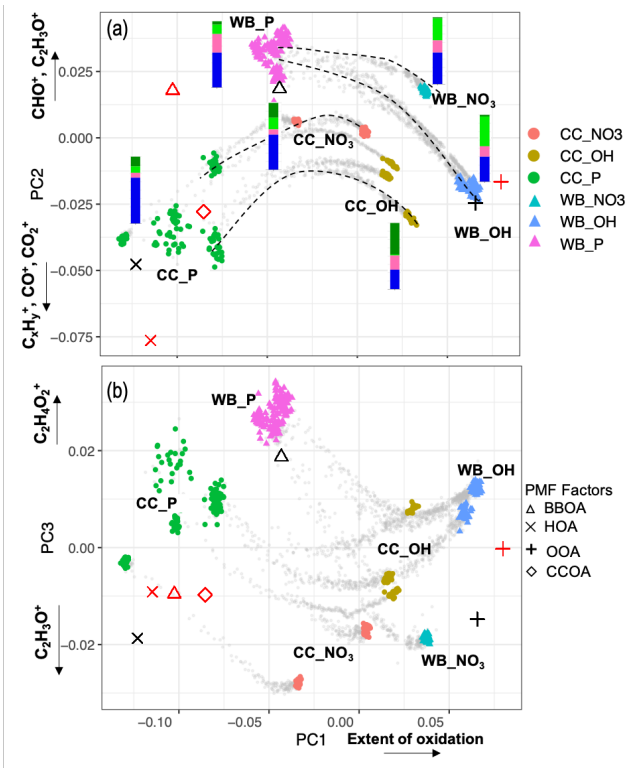

Figure 3.4 – Biplots of PC2–PC1 (a), and PC3–PC1 (b) scores. The AMS measurements corresponding to filter sampling periods are color-coded by category. The AMS measurements out of filter sampling periods are illustrated by gray dots and some oxidation trajectories are indicated by dashed curves. AMS BBOA, CCOA, HOA, and SV-OOA factors from Elser et al. (2016) (red) Aiken et al. (2009) (black) are projected onto PCs for comparison. Average FG composition for each category estimated from the MIR spectra is shown beside the category with the same color scheme as Fig. 3.1.

Kroll and Seinfeld, 2008) or due to hydrolysis in the particle phase (Day et al., 2010; Ng et al., 2017; Liu et al., 2012b). As a result, the chamber WB aerosols aged with the nitrate radical were excluded from this comparison due to their very prominent $RONO_2$ bands.

Mean spectra

The atmospheric samples were divided into four sub-groups: urban, rural, residential wood burning, and wildfire as described in Sec. 3.2.4. The individual spectra were baseline-corrected and blank-subtracted and then were normalized (Euclidean norm) based on their absorbances in the 1300–4000 cm^{-1} range. This procedure allowed us to make the spectra with different absorbance magnitudes (i.e., different aerosol mass concentrations) comparable. Thereafter, a single mean spectrum was calculated for each sub-group and used as its average representation. These spectra were compared to those of fresh and aged WB OAs in the chamber. Since major inorganic compounds of fine aerosols (ammonium, sulfate, and nitrate) are also

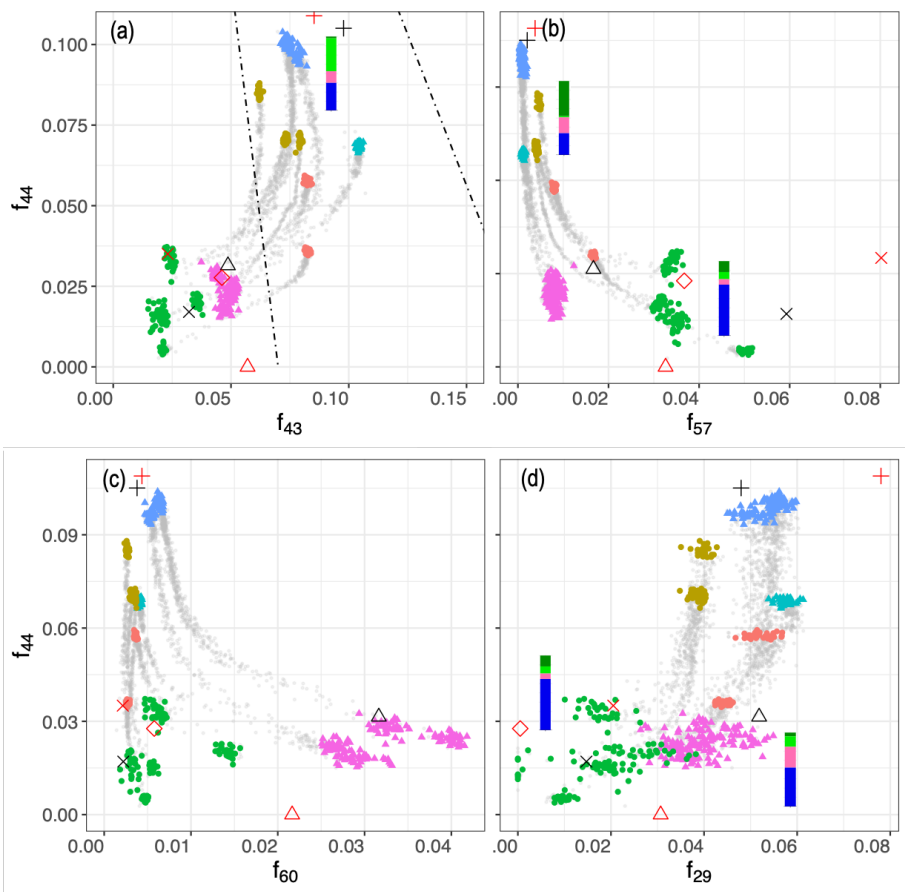

Figure 3.5 – Scatter plots of f_{44} against f_{43}, f_{57}, f_{60}, and f_{29}. Dashed lines show the outline of the triangle (Ng et al., 2010) and the black shapes show the location of PMF factors (Elser et al., 2016; Aiken et al., 2009; Ulbrich et al., 2009). Average FG composition for each category estimated from the MIR spectra are shown beside the category. Figure 3.4 legend is applicable here.

IR-active and their absorbances overlap with those of organic FGs (especially aCOH), the analysis of organic FGs in ambient samples is not always straightforward. To mitigate this problem, ammonium absorbances were subtracted from the mean spectra to obtain the pure contribution of organic compounds.

As can be seen from Fig. 3.6a, the mean spectrum of rural samples has strong absorptions at ammonium (doublets at 3200 and 3050 cm^{-1}), nitrate (1400 cm^{-1}), and sulfate (620 cm^{-1}) regions, while the peaks attributed to organic compounds are relatively weak (e.g. very weak aCH and carbonyl CO peaks and indistinguishable aCOH absorption due to strong inorganic absorptions; Fig. 3.6a). The urban mean spectrum also has strong ammonium and sulfate absorptions, suggesting the abundance of inorganic compounds in urban sites. The organic signatures in the urban spectrum, however, are slightly stronger than those of rural sites with the clear presence of broad carboxylic acid absorption (Fig. 3.6b). Neither rural, nor urban

mean spectra are similar to that of chamber WB spectra.

The mean spectrum of residential wood burning samples has prominent absorptions of ammonium, sulfate and also nitrate. Cold weather and high concentrations of nitric acid resulting from fossil fuel combustion and biomass burning are believed to be the primary reasons for the presence of ammonium nitrate on the filters in spite of being relatively volatile (Chow et al., 2005; Seinfeld and Pandis, 2016). This spectrum has considerably stronger signatures of organic compounds (Fig. 3.6b). Very sharp aCH peaks, strong acid COOH absorptions, and a visible aCOH absorption on the left shoulder of the ammonium peak can be seen in the mean spectrum of residential wood burning. Comparing this spectrum with that of aged WB in the chamber reveals their striking similarities. Both spectra have close inorganic-to-organic ratios that result in similar profiles in Fig. 3.6a. In addition, both spectra have visible alcohol and acid signatures, which are identified to be important in biomass burning aerosol composition (Corrigan et al., 2013; Takahama et al., 2011; Russell et al., 2011; Hawkins and Russell, 2010). Nevertheless, the aCH absorption in the mean spectrum of residential wood burning is stronger than that in the mean spectrum of aged WB in the chamber. This observation might be attributed to the long-chain hydrocarbons existing in the cuticle wax of vegetation detritus that is absent in these chamber WB experiments (Hawkins and Russell, 2010). The spectral comparison of residential wood burning aerosols and fresh WB aerosols in the chamber shows that fresh WB lacks inorganics compared to residential wood burning. Moreover, the relative abundance of aCOH is significantly higher in WB aerosol, indicating the residential wood burning samples are aged to some degree.

In addition to FG identification, we can discuss the presence of specific marker compounds. A weak signature of lignin C=C (at 1515 cm^{-1}) can be observed in the mean spectrum of residential wood burning (Fig. 3.6a). In addition, weak levoglucosan absorbances are visible in some residential spectra. Both mentioned signatures in burning-influenced atmospheric samples are stronger than that of aged chamber WB and weaker than that of fresh chamber WB aerosols, suggesting that most of these atmospheric samples have, on average, experienced aging within the range explored by our chamber experiments.

The mean spectrum of wildfire samples is also very similar to that of residential wood burning except having slightly weaker aCH absorbances. Consequently, the strong COOH absorption and the visible lignin and levoglucosan signatures are also the characteristics of the wildfire mean spectrum as they were for that of residential wood burning. The mean spectra of aged chamber WB, residential wood burning, and also wildfire aerosols (Fig. 3.6b) are similar to the biomass burning PMF factor obtained by Hawkins and Russell (2010).

Finally, the samples affected by wildfire are the only atmospheric samples in the IMPROVE network (2011 and 2013) that have visible (but weak) PAH signatures, possibly leading to a higher PAH-related toxicity. The areal density of atmospheric and chamber aerosols collected on PTFE filters is comparable in this work (approximately 10 $\mu g\ cm^{-2}$). As a result, the low absorbance of PAHs in the atmospheric samples implies that most of these compounds are

degraded in the atmosphere or during the filter transportation and storage.

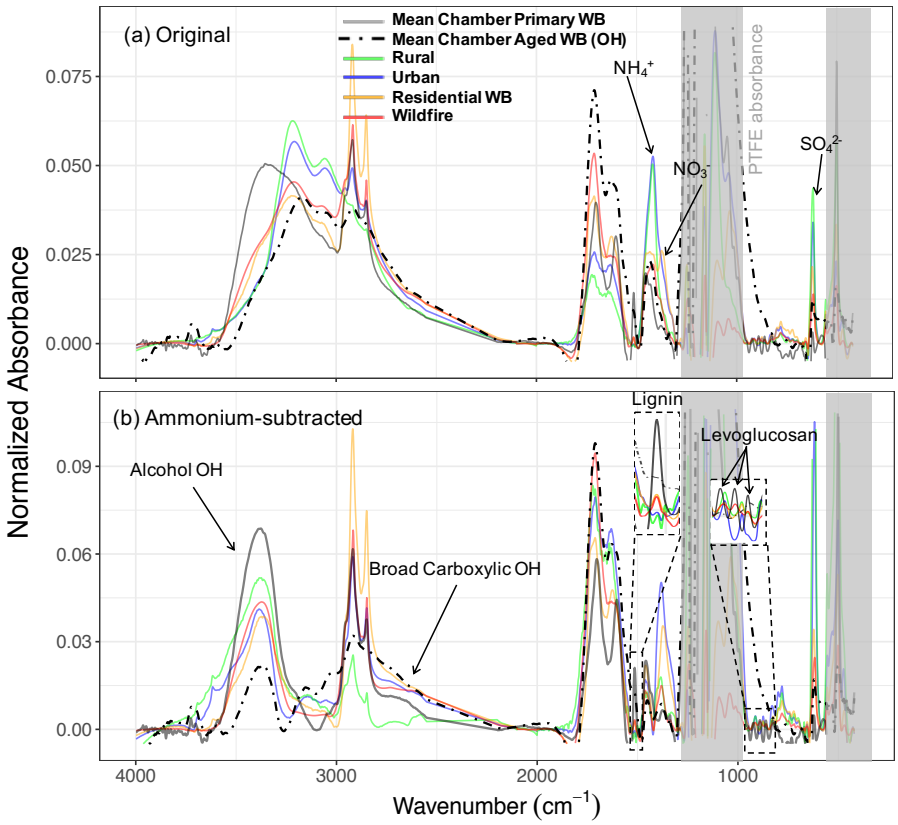

Figure 3.6 – Normalized (Euclidean norm) mean spectra of ambient aerosols (rural, urban, residential wood burning, and wildfire) and chamber WB aerosols before (a) and after (b) subtraction of ammonium profile. Inorganic nitrate and sulfate absorbances still exist in panel (b).

OM:OC ratios

The OM:OC ratio is an important parameter often reported for aerosols collected at monitoring networks (Boris et al., 2019; Ruthenburg et al., 2014; Bürki et al., 2020; Yazdani et al., 2021a; Hand et al., 2019; Reggente et al., 2019a). We compared OM:OC ratios from these chamber samples to burning samples identified by Bürki et al. (2020). They used a probabilistic framework on MIR spectra, which resulted in OC estimates that were consistent with those from collocated TOR measurements. The average OM:OC estimates for samples influenced by wildfires and residential wood burning were reported to be 1.65 and 1.45, respectively (Bürki et al., 2020) compared to 1.65 and 2 (AMS and MIR spectroscopy obtained close values) for primary and aged WB aerosols in this work. The average OM:OC estimates for these atmospheric burning-influenced samples are clearly closer to that of primary WB aerosols in chamber. Although some biases may exist between the two methods due to different estimation methods and calibration standards (Reggente et al., 2019a), this observation is believed to

be mainly due to two reasons: First, identified burning-influenced samples in the atmosphere are not as oxidized as the aged WB aerosols in the chamber are. Second, sharp aCH peaks, attributed to cuticle wax, are observed in the MIR spectra of ambient fire-influenced samples (Hawkins and Russell, 2010), while being absent in that of chamber WB aerosols (without bark). The high abundance of the aCH group might lower the OM:OC estimate of atmospheric burning-influenced samples.

Identification of burning-influenced samples

Biomass burning aerosol composition and its molecular markers continually evolve with aging as discussed in previous sections. This constant evolution poses a substantial challenge to identify the burning-influenced aerosols in the atmosphere and to quantify the contribution of biomass burning OM to the PM mass. In this section, we assess two different approaches for the identification of atmospheric burning-influenced aerosols using MIR spectroscopy.

While each FG is not a marker for any specific source, their proportions could possibly be informative when used with multivariate methods. Cluster analysis by Bürki et al. (2020), for example, identified 45 burning-influenced samples (in 3050 samples) based on their spectral similarity (i.e. FG proportions and organic-to-inorganic ratio), which were supported, to some extent, by matching their collection time and location with the known burning events (e.g., Rim Fire in California in 2013 and residential wood burning in Phoenix in winter). Nevertheless, the signatures of parent compounds (e.g. lignin) and specific biomass burning markers (e.g. levoglucosan) – direct identifiers of burning – were not considered in their approach. This is because the extended baseline correction and filter subtraction introduced in this work are necessary for identifying these signatures. Moreover, the contribution of these signatures to total variations in the MIR spectra is too minor to be featured for spectral separation when selecting only a limited number of clusters.

In contrast to FG proportions, tracers are less ambiguous for source identification, but they eventually degrade with aging as discussed in Sec. 3.3.2. In the tracer approach developed in this work, samples with lignin absorbances above a certain threshold were labeled burning-influenced (absorbances> 0.0006 in Euclidean normalized spectra to account for samples with a non-negligible contribution of lignin and absorbances> 0.0004 in non-normalized spectra to discard samples with low mass loadings and noisy spectra). Samples with levoglucosan mass (estimated from MIR spectroscopy) contributing more than 5 % to total OM (1.8 of collocated TOR OC estimates) were also labeled burning-influenced (ranging from 5 to 15 % of total OM). We identified in total 173 samples (out of 3050 samples) with the tracer approach, which included the majority of the previously identified burning sample using cluster analysis (38 out of 45) (Fig. 3.7), suggesting that most of the atmospheric samples with a MIR spectrum similar to that of burning aerosols have also visible tracer molecule signatures. There were, however, 135 samples with levoglucosan and/or lignin signatures above the defined threshold that were not identified by the cluster analysis. The relatively high number of false negatives (i.e. missing burning label when levoglucosan and lignin peaks were present) in the cluster

analysis suggests that uncertainties can be high with this approach. There were also 7 samples with negligible levoglucosan and lignin signatures (not identified by the tracer method) that were identified to be burning-influenced by the cluster analysis. The latter discrepancy might exist due to misidentification by the cluster analysis or because the discussed signatures diminish with extensive aging and cannot be detected by the tracer method. The same could be true for cluster 10 of Bürki et al. (2020), which was proposed to be possibly influenced by burning, but only 3 out of 28 samples have identifiable levoglucosan or lignin signatures.

The atmospheric burning-influenced samples identified by both methods have a high relative abundance of OM. OM constitutes more than 50 % of fine PM mass for the majority of these sample even at low PM loadings as observed in Fig. 3.7. However, a high OM:PM ratio cannot always be a reliable indicator of burning alone since it is not unique to biomass burning and there are many non-burning samples with a high OM:PM ratio (Fig. 3.7). Potassium is considered a good inorganic tracer of biomass burning (Sullivan et al., 2011b). Bürki et al. (2020) showed a higher-than-average K:$PM_{2.5}$ ratio for the residential wood burning samples identified in the IMPROVE network (2011 and 2013). However, as also mentioned by Sullivan et al. (2011b), high K is not observed in wildfire samples. In addition, there are some non–biomass-burning sources of potassium more likely to be found in urban areas such as incinerators and fly ash. The burning-influenced samples identified by the tracer method in this work have higher concentrations of potassium compared to the majority of other atmospheric samples (Fig. A.24). However, there are also some samples with relatively high K concentrations that are not impacted by burning and are most probably affected by mineral dust due to having a prominent Si−O−H peak above 3500 cm^{-1} (Bürki et al., 2020).

In summary, the tracer method appears to be able to identify the atmospheric burning-influenced samples including the majority of those identified by the cluster analysis. However, a definitive identification of burning periods is needed to better assess this new approach. Quantifying the contribution of biomass burning OM (fresh and aged) to the PM mass is another remaining challenge that needs to be addressed in the future work.

3.4 Concluding remarks

In this work, we used MIR spectroscopy and AMS to obtain complementary information about primary and secondary wood burning (WB) and coal combustion (CC) aerosols. MIR spectroscopy provides a detailed characterization of functional groups and implies that primary aerosols and their parent compounds have similar chemical compositions. This similarity diminishes after aging with the appearance of the peaks attributed to oxygenated species and disappearance of the peaks attributed to the parent compounds. We observed distinct FG compositions for OA based on the emission source (WB and CC), aerosol type (primary and aged), oxidant type, and NO_x concentration. The observed FG composition is informative about the dominant oxidation pathways of WB and CC VOCs and can be used to verify and improve the results of the chemically resolved SOA formation models.

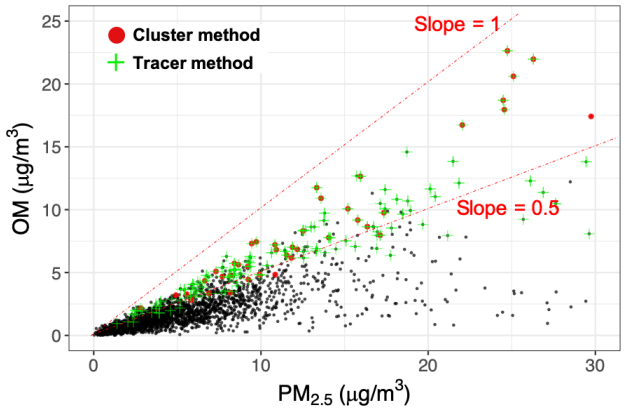

Figure 3.7 – Scatter plot comparing total $PM_{2.5}$ and OM mass in atmospheric samples. Red circles indicate residential wood burning and wildfire samples identified by cluster analysis (Bürki et al., 2020). Green crosses show burning samples identified based on lignin and levoglucosan signatures. Black circles indicate the existing atmospheric samples in the IMPROVE network (2011 and 2013; approximately 3050 samples). The dotted dashed lines delineate the range of OM mass fractions for samples designated as burning-influenced (the slope of 0.5 is arbitrarily chosen to guide the eye). OM was estimated by multiplying OC by 1.8 (assuming an average OM:OC ratio of 1.8). Refer to Fig. A.23 for a more detailed classification.

Dimensionality reduction of the AMS mass spectra reveals similarities between the AMS mass spectra of the primary CC OA and the HOA factor, the primary WB OA and the BBOA factor, and the OA aged with OH and the SV-OOA factor, respectively. In addition, this analysis suggests that both the aerosol source and type are major drivers of variability in the AMS mass spectra. These variations are also consistent with our FG analysis, implying that AMS mass spectra maintain some functional group information in spite of the extensive fragmentation. However, variations due to the change in the FG composition occasionally constitute a small fraction of the total variation (stored in higher PCs).

We also found that the MIR spectra of WB aerosols in the environmental chamber are similar to those of ambient samples affected by wildfires and residential wood burning. The high abundance organics (especially acids) and existence of peaks attributed to lignin and levoglucosan are the main aspects of this similarity. This result helped us better interpret the MIR spectra of atmospheric samples and was used to aid the identification of ambient burning-influenced aerosols.

Finally, it was found that PAHs and aromatics are quantifiable in chamber aerosols using MIR spectroscopy but are not visible in the spectra of atmospheric samples except for a few burning-influenced instances. Considering the fact that the areal density of aerosols collected on PTFE filters is similar in the atmospheric and chamber samples, this observation suggests that either the chamber conditions (e.g., fuel, aging, VOC and OA concentrations) are more conductive to PAHs or most aromatics and PAHs are degraded in the atmosphere or during the transportation and storage.

4 Fragment ion-functional group relationships in organic aerosols using aerosol mass spectrometry and mid-infrared spectroscopy

Accepted for publication in the Journal of Atmospheric Measurement Techniques Discussions on 8 Nov. 2021.

Authors : Yazdani, Amir; Dudani, Nikunj; Takahama, Satoshi; Bertrand, Amelie; Prévôt, André S. H.; El Haddad, Imad; Dillner, Ann M.

Contributions: IEH and ST and AY conceived the project and manuscript. AB and IEH performed the chamber experiments. AB provided AMS spectra. ND prepared and assembled the filter sampling set-up and took the FTIR spectra. AMD laboratory standard and ambient sample spectra. AY wrote the code for data analysis and post processing, performed the data analysis, prepared laboratory standards and analyzed samples, and wrote the manuscript. ST edited the manuscript and provided regular feedback on the analysis. IEH, ASHP, AB, AMD, and ND provided input on the analysis and the further editing of the manuscript. ST and IEH provided overall supervision of the project.

Reference: Yazdani et al. (2021b)

Abstract Aerosol mass spectrometry (AMS) and mid-infrared spectroscopy (MIR) are two analytical methods for characterizing the chemical composition of OM. While AMS provides high-temporal-resolution bulk measurements, the extensive fragmentation during the electron ionization (EI) makes the characterization of OM components limited. The analysis of aerosols collected on PTFE filters using MIR, on the other hand, provides functional group (FG) information with reduced sample alteration but results in a relatively low temporal resolution. In this work, we compared and combined MIR and AMS measurements for several environmental chamber experiments to achieve a better understanding of the AMS spectra and the OM chemical evolution by aging. Fresh emissions of wood and coal burning were injected into an environmental simulation chamber and aged with hydroxyl and nitrate radicals. A high-resolution time-of-flight (HR-TOF) AMS measured the bulk chemical composition of fine PM. Fine aerosols were also sampled on PTFE filters before and after aging for the offline MIR analysis. After comparing AMS and MIR bulk measurements, we used multivariate statistics to identify the influential functional groups contributing to AMS OM mass for different aerosol sources and aging processes. We also identified the key mass fragments resulting from each functional group for the complex OM generated from biomass and fossil fuel combustion. Finally, we developed a statistical model that enables the estimation of the high-time-resolution functional group composition of OM using collocated AMS and MIR measurements. Using this approach, AMS spectra can be used to interpolate the functional group measurements by MIR, allowing us to better understand the evolution of OM during the course of aging.

4.1 Introduction

Particulate matter (PM) impacts visibility, climate, and human health (Hallquist et al., 2009). Organic matter (OM), which accounts for an important fraction of total fine atmospheric PM mass, has profound effects on aerosol-related phenomena (Turpin and Lim, 2001; Russell, 2003; Shiraiwa et al., 2017b). Characterizing the organic fraction is necessary to reduce the uncertainties associated with the impact of PM and can in turn affect the policies related to climate change and air quality management (Zhang et al., 2011; Turpin and Lim, 2001). However, OM chemical composition and formation mechanisms have not yet been fully understood due to their complexity.

Different analytical and computational techniques exist for determination of the chemical composition of organic aerosols (OAs) and especially secondary organic aerosols (SOAs). Modeling all important SOA-related reaction and species is not feasible for large-scale models (Jathar et al., 2015) and simpler models often do not reproduce the measured concentrations and dynamics of SOA evolution in polluted regions (Volkamer et al., 2006). Among the analytical techniques, aerosol mass spectrometry (AMS) and mid-infrared (MIR) spectroscopy are able to provide bulk chemical information for most of OM mass (Hallquist et al., 2009).

AMS provides information about the chemical composition of OM and its temporal variations in terms of ensemble mass spectra acquired over short time intervals (Zhang et al., 2011).

Aerodyne Research aerosol mass spectrometer (used in this work and referred to as "AMS") is the most widely used thermal desorption-based mass spectrometers in aerosol research. AMS is capable of quantifying non-refractory species (e.g., sulfate, nitrate, ammonium, chloride, and OM) by thermal vaporization (typically at 600 °C) and electron ionization (EI; typically at 70 eV) (Canagaratna et al., 2007). In spite of the valuable information that AMS provides, the AMS OM fragment ions are not molecule-specific and are difficult to interpret due to the extensive fragmentation of molecules with the high-energy electron impact (EI) ionization and flash vaporization. This limits the level of molecular details that can be extracted from the AMS mass spectra (Canagaratna et al., 2007; Kumar et al., 2018; Faber et al., 2017; Chhabra et al., 2011a). Organic aerosol components can also undergo oxidation, dehydration and/or decarboxylation reactions inside the AMS ionization chamber (Canagaratna et al., 2015). In addition, uncertainties regarding the relative ionization efficiency (Xu et al., 2018), fragmentation tables (Aiken et al., 2008), gas-phase interference (Canagaratna et al., 2015), and collection efficiency (Frossard et al., 2014) have been reported. There are soft ionization methods, such as electrospray ionization (ESI) and chemical ionization (CI) that minimize the analyte fragmentation at the expense of variable ionization efficiency, signal-to-noise ratio, and quantifying bulk OM composition (Lopez-Hilfiker et al., 2019; Nozière et al., 2015; Iyer et al., 2016; Zahardis et al., 2011).

MIR spectroscopy, which is commonly performed off-line on polytetrafluoroethylene (PTFE) filters (Takahama et al., 2013; Ruthenburg et al., 2014), is used as a complementary method to AMS in this work. This non-destructive method gives direct functional group (FG) information; provides measurements consistent with commonly used instruments in monitoring networks (Boris et al., 2019); and is capable of differentiating between the composition of different oxidized OMs. For example, Liu et al. (2012a) observed very similar AMS mass spectra for several SOA samples, while the complementary MIR spectra indicated clear chemical differences for these aerosols. In addition to FG quantification, MIR spectroscopy has been recently used to quantify biomass burning markers in the atmospheric aerosols (e.g., levoglucosan and lignin; Yazdani et al., 2021c). However, MIR spectroscopy on filters has low temporal resolution compared to on-line instruments such as AMS (Faber et al., 2017; Yazdani et al., 2021c). Moreover, volatilization of volatile organic compounds from PTFE filters during or after sampling can affect OM mass and composition (Subramanian et al., 2004). Uncertainties regarding variable absorptivities by different organic molecules (Hastings et al., 1952), peak overlaps (Pavia et al., 2008), scattering, and PTFE interference (Takahama et al., 2013) have also been reported for this technique.

Past studies compared AMS and MIR OM, O:C, and positive matrix factorization (PMF) factors in field campaigns (Gilardoni et al., 2009; Russell et al., 2009b; Frossard et al., 2011; Liu et al., 2011; Corrigan et al., 2013; Frossard et al., 2014), reporting reasonable agreement between the instruments despite the aforementioned uncertainties. Two controlled laboratory studies tried to understand the relationship between fragment ions and functional groups (Faber et al., 2017; Russell et al., 2009a) using univariate correlations. In this work, we compare and combine the AMS and MIR measurements for the organic aerosols of moderate to high

complexity from biomass burning and coal combustion emissions, two major sources of atmospheric OM, in an environmental chamber. We apply additional uni- and multivariate techniques to further interpret the relationships between more than 300 AMS fragment ions and 4 MIR FGs, and provide a method to predict the high-time-resolution evolution of FGs using AMS.

4.2 Methods

In the following sections, the experimental set-up (Sect. 4.2.1), on-line and off-line measurement techniques (Sects. 4.2.2 and 4.2.3) are described. Thereafter, statistical methods used for combining AMS and MIR measurements are explained (Sects. 4.2.4–4.2.6). The experimental set-up, procedure, and data used in this work are the same as those reported by Yazdani et al. (2021c).

4.2.1 Laboratory experimental set-up and procedure

Briefly, we conducted four wood burning (WB) experiments with beech wood logs and five coal combustion (CC) experiments with bituminous coal using ordinary modern stoves (Bruns et al., 2015). The emissions were diluted and then injected into a 6 m^3 Teflon bag environmental chamber of at the Paul Scherrer Institute (PSI) in Villigen, Switzerland. The injections were continued until the concentration of PM and OM measured by the scanning mobility particle sizer (SMPS) and high-resolution time-of-flight (HR-ToF) AMS reached atmospherically-relevant values. The emissions were held in the chamber for 30 minutes after injection to improve mixing. Thereafter, primary emissions were chemically aged using the hydroxyl or nitrate radical in order to simulate daytime and nighttime aging mechanisms, respectively. For the diurnal aging simulations, the OH radical was produced by the photolysis of HONO and the OH exposures reached $(2–3)\times10^7$ molec cm^{-3}h (measured using butanol-d9; Barmet et al., 2012) corresponding to 20–30 hours of aging in the atmosphere. For the nocturnal aging experiments, the NO_3 radical was produced by a single injection of O_3 and NO_2 in the chamber. The nitrate radical concentration was estimated to be $(1.5–2.5)\times10^7$ molec cm^{-3} for the first hour of aging process based on the phenol concentration decay in the gas phase.

4.2.2 Online AMS PM measurements

Non-refractory particle composition was measured with a HR-ToF AMS operating in V mode (mass resolution $\Delta m/m$ = 2100 for m/z 200; DeCarlo et al., 2006) with a 2.5 µm inlet aerodynamic lens throughout the experiment. The raw signal was postprocessed in Igor Pro 6.3 (Wave Metrics) using SQUIRREL 1.57 and PIKA 1.15Z routines. Elemental ratios of OM were estimated following the approach of Canagaratna et al. (2015) (Fig. A.25). The AMS OM concentrations reported in this work are not corrected for the chamber wall losses and the measured nitrate is assumed to be inorganic for ease of comparison with MIR.

4.2.3 Offline MIR PM measurements

Two 47 mm PTFE filters (Pall corporation) were used for each experiment to sample the primary PM after its injection into the chamber and the aged PM after approximately three to four hours of aging. The filter holder was placed downstream of a sharp-cut-off cyclone and a silica gel denuder and the flow rate through the filter was maintained at 8 $\mathrm{L\,min^{-1}}$. We used a similar naming convention for the filters to that of Yazdani et al. (2021c). Filters were immediately stored in filter petri dishes at 253 K after sampling and before MIR analysis to minimize volatilization and chemical reactions. The PTFE filters were analyzed using a Bruker Vertex 80 Fourier transform infrared (FTIR) instrument equipped with an α deuterated lanthanum alanine doped triglycine sulfate (DLaTGS) detector, at a resolution of 4 $\mathrm{cm^{-1}}$. The spectra were averaged over 64 scans.

MIR Spectral postprocessing

The MIR spectra were baseline-corrected to eliminate the contribution of light scattering by filter membrane and particles as well as absorption by graphitic carbon (Parks et al., 2021). We used a smoothing spline described by Kuzmiakova et al. (2016). After baseline correction, the scaled and baseline-corrected spectrum of a blank filter was subtracted from the baseline-corrected sample spectra to minimized the interference of PTFE C – F bands. After baseline correction and blank subtraction, the multiple peak-fitting algorithm described by Takahama et al. (2013) was applied on the spectra to obtain FG abundances of alcohol (aCOH), carboxylic acid (COOH), alkane (aCH), non-acid carbonyl (naCO) (Supplement Fig. A.25). After obtaining FG abundances, the O:C, H:C, and OM:OC ratios (Fig. A.25) were calculated assuming 0.5 C atom for each of aCH and aCOH bonds (Chhabra et al., 2011b; Russell, 2003; Maria et al., 2002).

4.2.4 Identifying influential MIR absorbances for AMS OM

The AMS OM estimates and the MIR spectra are combined statistically to identify the functional groups that influence the former OM mass concentration the most. The results of this method, which are affected neither by uncertainties of MIR peak fitting nor by absorption coefficients, identify important predictors of OM estimated by more routinely applied methods prior to peak fitting. This technique can be applied even when absorption coefficient data are not available for all FGs, and help decide which FGs are needed to be included in the MIR peak fitting. The averaged AMS OM concentrations over the filter sampling periods were regressed against the corresponding MIR spectra using partial least squares regression (PLSR). Thereafter, the influential absorbances in the MIR spectra for the organic OM concentration were determined based on variable importance in projection (VIP) scores of the PLSR model (Fig. A.26). This procedure was applied separately for the primary and aged aerosols of each source to highlight the compositional differences. By regressing AMS OM concentration against MIR

absorbances, we seek a solution of the following linear equation for coefficients $\boldsymbol{a}$:

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{a} + \boldsymbol{e}, \tag{4.1}$$

where $\mathbf{X}$ $(n \times p)$ is the MIR spectra matrix with n samples and p independent variable (wavenumbers), $\boldsymbol{y}$ $(n \times 1)$ is the vector of response variable, AMS OM concentration, and $\boldsymbol{e}$ is the vector of residuals. In this work, the univariate partial least squares regression (Wold et al., 1983) is used. The univariate PLSR projects $\mathbf{X}$ onto $\mathbf{P}$ $(p \times h)$ (h is the number of latent variables) basis with orthogonal scores $\mathbf{T}$ $(n \times h)$, while maximizing the covariance between scores and the response variable, $\boldsymbol{y}$. In Eq. (4.3), $\boldsymbol{b}$ and $\boldsymbol{f}$ indicate the regression coefficients the vector of residuals, respectively.

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\top} + \mathbf{E}, \tag{4.2}$$

$$\boldsymbol{y} = \mathbf{T}\boldsymbol{b} + \boldsymbol{f}. \tag{4.3}$$

After solving the regression equation using PLSR for different number of latent variable (LVs), a repeated 10-fold cross validation was applied to find the optimal number of latent variables and avoid under/overfitting. Examining loadings and coefficients directly can be informative about the important absorbances. For instance, the first weight vector, $\boldsymbol{w}_1$, can be a good estimate of important bands but it is limited to the cases that signal is not dominated by other factors rather than the analyte, such as inorganics, and filter absorption (Haaland and Thomas, 1988). In this work, we used a more general method, VIP scores (Wold et al., 1993), to identify the important absorption bands. This metric is a root mean square of loading weights of all h latent variables used in the model weighted by their fraction of the captured response (Chong and Jun, 2005; Takahama et al., 2016). The VIP score of the jth wavenumber is calculated by considering all h latent variables in the model as shown in Eq. (4.4). Since the average of squared VIP scores is equal to one, generally, the wavenumbers with VIP score greater than one are considered influential due to higher-than-average contribution to estimating the response variable. In Eq. (4.4), $\boldsymbol{t}_k$ and $\boldsymbol{w}_k$ represent the kth columns of $\mathbf{T}$, score matrix, and $\mathbf{W}$ $(p \times h)$, weight matrix, respectively and the relationship between $\mathbf{T}$ and $\mathbf{W}$ is described by Eq. (4.5) (Helland, 1988).

$$\boldsymbol{VIP}_j = \sqrt{p\frac{\sum_{k=1}^{h} SS(b_k\boldsymbol{t}_k)(\boldsymbol{w}_{jk}/\|\boldsymbol{w}_k\|)^2}{\sum_{k=1}^{h} SS(b_k\boldsymbol{t}_k)}},$$
$$\text{where } SS(b_k\boldsymbol{t}_k) = b_k^2\boldsymbol{t}_k^{\top}\boldsymbol{t}_k. \tag{4.4}$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}(\mathbf{P}^{\top}\mathbf{W})^{-1}. \tag{4.5}$$

4.2.5 Identifying FG-ion fragment relationships

Covariance and correlation coefficients were used to understand the connection between fragment ions (up to m/z 202 for which the signal-to-noise ratios are still significant) and FGs (Fig. A.27a). We used normalized functional group abundances by MIR total OM and normalized fragment ion concentrations by AMS total OM (averaged over filter sampling periods) and calculated covariances and correlations between 4 FGs and more than 300 fragment ions. The major difference between the fragment-FG correlation and covariance is that the former is more informative about the fragments with low concentrations while the latter highlights the fragments with higher concentrations. Data normalization was performed to avoid correlations introduced by the changes in the total OM mass concentration as oxygenated fragments are highly correlated before normalization (Fig. A.29). In addition, negative correlations (anti-correlations) and covariances were omitted as they do not show the production of fragments by FGs. For example, often the aCH relative abundance decreases with aging as the relative concentrations of oxygenated FGs such as COOH and oxygenated fragment ions such as CO_2^+ increase, leading to a significant anti-correlation between the aCH FG and the CO_2^+ fragment. Russell et al. (2009a) and Faber et al. (2017) have already applied the univariate fragment-FG (correlation) analysis for different sources using unit-mass-resolution and HR AMS data, respectively. However, their analysis has been limited to only a few light fragment ions.

Univariate methods can be difficult to interpret when ion fragments are produced by multiple functional groups. Therefore, in addition to the univariate methods, the VIP scores method was used to highlight the influential spectral regions and FGs for major fragment ions with high concentrations (CO_2^+, CHO^+, $C_2H_3O^+$, $C_3H_5^+$) and for two biomass-burning-related fragment ions ($C_2H_4O_2^+$ for levoglucosan and $C_8H_9O_2^+$ for lignin). This multivariate approach, which is not affected by the MIR peak fitting uncertainties, is similar to identifying the influential MIR spectral regions for the AMS OM as discussed in Sect. 4.2.4 except that the concentrations of individual fragment ions are regressed against the MIR spectra (Fig. A.27b).

4.2.6 Interpolating FG abundances using AMS mass spectra

In order to estimate the high-time-resolution FG composition of OM, FG abundances for all filters (normalized by the MIR total OM mass concentration), which were calculated from peak fitting, were regressed against the AMS spectra (normalized by the AMS total OM mass concentration and averaged over the filter sampling periods) using PLSR. A repeated 10-fold cross validation was applied to indicate the optimal number of latent variables. These models were then used to interpolate (high-time-resolution) FG compositions using the AMS spectra and investigate the evolution OM during the course of oxidation when only AMS measurements existed (Fig. A.28). Thereafter, the high-resolution O:C ratios separated by FG contribution ($O:C = O_{COOH}:C_{total} + O_{naCO}:C_{total} + O_{aCOH}:C_{total}$) were calculated from the high-resolution resolution FG compositions following the same approach of Sect. 4.2.3.

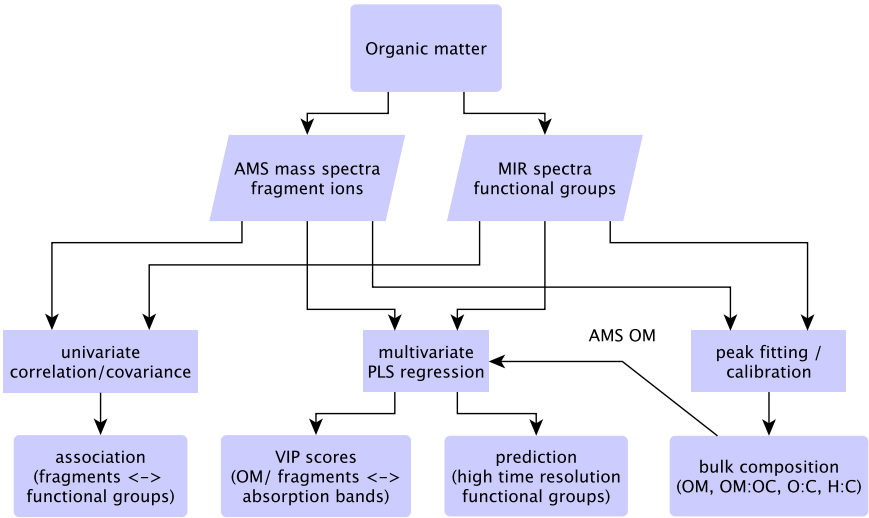

Figure 4.1 – Statistical relations and strategy for comparison of MIR and AMS measurements. The correlation/covariance analysis is performed between AMS mass fragments and MIR functional group abundances from peak fitting. The PLS regression is performed using the AMS total OM or individual fragment concentrations as the independent variable and the MIR absorbance spectra as the dependent variables.

4.3 Results and discussions

In the following subsections, bulk OM parameters from AMS and MIR are combined and compared in Sect. 4.3.1. There after, fragment ion-FG relationships are investigated in Sect. 4.3.2. Finally, PLSR models are developed to predict FG composition of OM using the AMS mass spectra (Sect. 4.3.3). Our approach for combining and comparing the AMS and MIR measurements is demonstrated in the diagram of Fig. 4.1.

4.3.1 Combination and comparison of OM measurements

Influential spectral regions of the MIR spectra and their corresponding FGs are determined for the AMS OM using VIP scores (Sect. 4.3.1). Thereafter, the OM mass concentration, OM:OC, O:C, and H:C ratios calculated using peak fitting to MIR spectra are compared to the average values from AMS for the primary and aged aerosols.

Influential group frequencies for total AMS OM

VIP scores of the PLSR models regressing MIR absorbances against AMS OM mass concentration highlight certain FGs to be important regarding the OM mass for each fuel burned and aerosol age. As can be seen from Fig. 4.2, carbonyl CO, aCOH have the highest VIP scores (greater than one) for primary WB aerosols, highlighting their importance in the primary WB

OM. The high VIP scores of aCOH is consistent with the fact that aCOH is a major part of wood constituents. Although $\nu(CH_2)$ and $\nu(CH_3)$ (stretching vibrational modes) at 2800–3000 cm^{-1} do not have high VIP scores for primary WB aerosols, the VIP scores for $\delta(CH_2)$ at 1470 cm^{-1} (bending vibrations) are high, suggesting the importance of aCH. The PLSR model probably use the information from the bending peak due to the lower overlap with other bands in the primary WB spectra. The peak around 1600 cm^{-1} which has a greater-than-one VIP score for primary WB is the result of several overlapping peaks attributed to the organonitrates, aromatic ring $\nu(C{=}C)$, amine $\delta(N-H)$, amide $\delta(N-H)$, and carboxylate $\nu(C{=}O)$ (Pavia et al., 2008). These overlapping absorbances make peak assignment in this region uncertain and complex. However, this peak is accompanied by the lignin-related $\nu(C{=}C)$ vibrations at 1515 cm^{-1} (Yazdani et al., 2021c), suggesting the abundance of lignin-like products in primary WB OM as also proposed by Bertrand et al. (2018a). For the aged WB aerosols, VIP scores are the highest for the broad carboxylic $\nu(OH)$ absorbances at 2400–3400 cm^{-1} and carbonyl $\nu(CO)$ (acid carbonyl) at 1700 cm^{-1}, suggesting carboxylic acids to be important contributors to the OM mass after SOA formation (Yazdani et al., 2021c). In contrast to the primary WB aerosols, aCOH does not have high VIP scores for the aged WB aerosol (Yazdani et al., 2021c).

For the primary CC aerosols, $\nu(CH_2)$ has the highest VIP scores, suggesting the abundance of hydrocarbons likely from volatile compounds of coal. The aromatic ring $\nu(C{=}C)$ peak at 1600 cm^{-1}, however, has relatively lower VIP scores, implying that the aromatic rings do not constitute the majority of primary CC OM although coal is mainly composed of highly substituted aromatics. The ammonium $\nu(N-H)$ peaks at 3200–3400 cm^{-1} have negative coefficients and high VIP scores, implying that the PLSR model compensates ammonium interference with organics by assigning negative coefficients to the region. For the aged CC aerosols, which are mostly composed of CC SOA, the VIP scores of $\nu(CH_2)$ are considerably lower. By contrast, carbonyl $\nu(CO)$ and aCOH regions (observed on the shoulder of $\nu(N-H)$ peaks) have the highest VIP scores, suggesting that CC SOA is mostly composed of carbonyls and alcohols. The out-of plane aromatic CH band, $\gamma(CH)$, at 750 cm^{-1}, in spite of being visible, does not have high VIP scores in spite of their positive coefficient, suggesting that aromatic CH (rCH) is not a major part of the aged CC. The $RONO_2$ absorption region at 1630 cm^{-1} does not have high VIP scores although this region is very prominent in the aerosols aged with the nitrate radical (Yazdani et al., 2021c). This is because the AMS OM concentrations used in this study do not consider the majority organonitrate mass as both ${NO_2}^+$ and NO^+ are attributed to inorganic nitrate. As a result, the organonitrate abundances do not affect the regression models.

In general, the important FGs highlighted by the VIP scores method are the same ones targeted in past studies of atmospheric aerosols (e.g., Ruthenburg et al., 2014; Russell et al., 2009b) and are consistent with our knowledge of POAs and SOAs related to combustion sources (e.g, Bertrand et al., 2018a, 2017; Yazdani et al., 2021c). In addition, they provide insight into the fraction of the combusted fuel that is important for OM formation (e.g., hydrocarbons in CC).

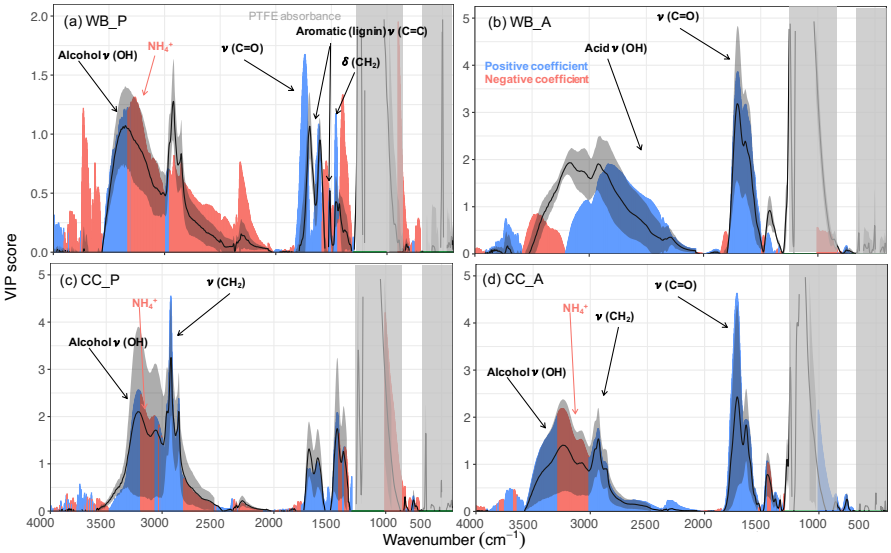

Figure 4.2 – VIP scores of the MIR absorbances regressed against the AMS OM concentration (averaged over the filter sampling periods) for (a) primary wood burning, (b) aged wood burning, (c) primary coal combustion, and (d) aged coal combustion aerosols. Blue/red regions correspond to wavenumbers with positive/negative regression coefficients in the PLSR models, respectively. Solid curves show the average spectrum (± one standard deviation shown by the shaded bands) for each category. Important FGs and their locations are indicated for each category. The PTFE C–F absorption regions are masked by gray rectangles.

AMS and MIR OM mass concentrations

It was shown in the last section that four FGs, aCH, COOH, aCOH, and naCO are the most influential functional groups regarding OM mass. The abundances of the mentioned FGs were estimated using peak fitting to the MIR spectra. The aromatic C=C group (rC=C) was not quantified due to the interference with other functional groups and the lack of absorption coefficient data. The peak-fitting results show that the OM concentration estimates from AMS and MIR are highly correlated (Fig. 4.3a, $R^2 = 0.92$). The slope of the MIR OM concentration versus that of AMS (not corrected for collection efficiency) is 1.3. This slope is within the previously reported range (Gilardoni et al., 2009; Russell et al., 2009a,b; Liu et al., 2011) considering collection efficiency of AMS (Yazdani et al., 2021c; Kumar et al., 2018; Canagaratna et al., 2007) and the aerosol volatilization artifacts from PTFE filters (Ruggeri, 2017; Subramanian et al., 2004). Yazdani et al. (2021c) reported the AMS the collection efficiency to range between 0.7 and 1.1 for the same experiments. The OM concentrations estimated by both methods indicate the significant enhancement with aging even without particle and vapor wall loss consideration (on average 2.4 and 2.7 times by AMS and MIR, respectively). The enhancement ratios are in the range of values that were previously reported for SOA formation from log-wood stoves (Bertrand et al., 2017; Tiitta et al., 2016; Grieshop et al., 2009; Heringa et al., 2011; Hennigan et al., 2010). Using different absorption coefficient values for MIR FGs (discussed by Reggente et al., 2019a) has little effect on the correlation coefficient.

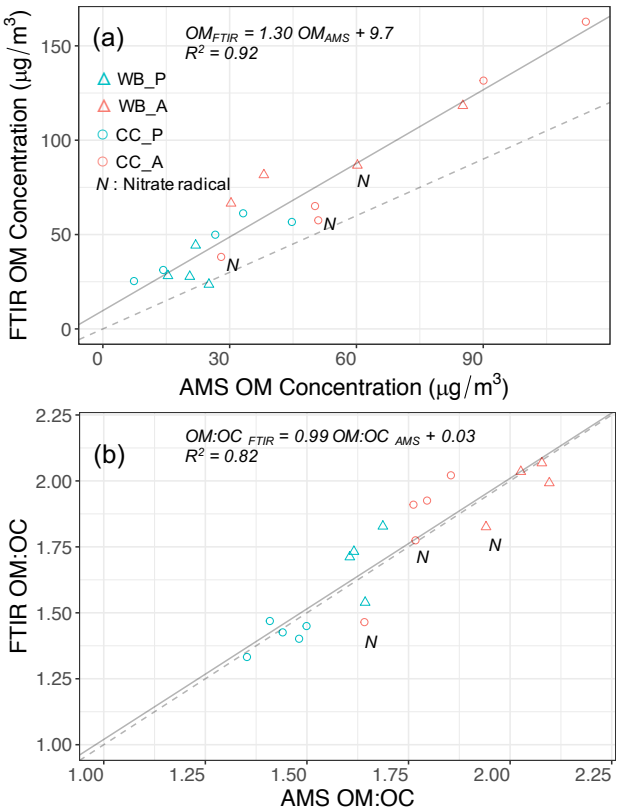

Figure 4.3 – (a) Scatter plot comparing OM concentration estimates by AMS and MIR for primary (P) and aged (A) aerosols of wood burning (WB) and coal combustion (CC) in the environmental chamber experiments. (b) Scatter plot comparing OM:OC estimates from AMS and MIR for primary and aged aerosols in each of the environmental chamber experiments. The fitted and 1:1 lines are solid and dashed, respectively.

AMS and MIR OM:OC ratios

The OM:OC ratios calculated from the AMS mass spectra were averaged over the filter sampling periods and compared to those measured from peak fitting to MIR spectra. The OM:OC estimates of these two methods agree very well (R^2 = 0.82 and slope = 0.99; Fig. 4.3b) with an average difference of less than 0.15 (Yazdani et al., 2021c). The correlation coefficient of OM:OC ratios is also found to be insensitive to the choice of absorption coefficients reported by Reggente et al. (2019a) for MIR spectroscopy. The fact that both methods capture similar OM:OC and mass concentration trends, suggests that a similar fraction of OM is monitored by both and the uncertainties associated with each method is less than variations due to fuel sources and aging processes. The primary CC aerosols are estimated to have the lowest OM:OC ratios (1.35–1.5), which is justified by their strong hydrocarbon (aCH) signatures (Fig. 4.2c). The primary WB samples have slightly higher OM:OC ratios (1.6–1.7 from AMS) primarily due to a relatively higher concentration of aCOH (Fig. 4.2a). Both instruments estimate that the aged aerosols of the two sources, regardless of the aging method, have higher OM:OC than

the primary ones (Fig. 4.3b). The aged WB aerosols have the highest OM:OC ranging from 1.9 to 2.1 (from AMS), with high concentrations of COOH (from MIR). The aged CC aerosols have lower average OM:OC ratios compared to the aged WB aerosols, ranging from 1.6 to 1.8. For both emission sources, AMS and MIR show that aerosols aged with the hydroxyl radical have higher OM:OC ratios than those aged with the nitrate radical (Fig. 4.3b). Attributing total AMS nitrate to organics to estimate an upper bound for OM:OC, makes this difference less prominent. However, the nitrate radical only reacts efficiently with certain precursors compared to the hydroxyl radical, resulting in different SOA composition that is reflected in both AMS and MIR measurements.

AMS and MIR van Krevelen trajectories

The local slope of the aging trajectory in the van Krevelen diagram is informative about the changes in the functionality of OM (Heald et al., 2010; Ziemann and Atkinson, 2012; Chhabra et al., 2011a), which is also directly measured with MIR spectroscopy. Figure 4.4 shows the van Krevelen diagram of the WB and CC OM in different experiments. In the WB experiments, AMS oxidation trajectories vary between a straight line and a convex (L-shaped) curve. In the first WB experiment with the hydroxyl radical (WB_OH_1), AMS aging trajectory is almost a straight line implying a monotonic change of FGs during the course of aging (Fig. 4.4a). In the second experiment (WB_OH_2), however, the trajectory is convex with a reduced slope toward the end of aging, implying an increase in the abundance of FGs resulting in a low trajectory slope, e.g. acids (Fig. 4.4b). This is supported by the high concentration of the COOH group observed in the MIR spectra of the corresponding aged aerosols (Yazdani et al., 2021c) and will be investigated further in Sect. 4.3.3. The WB experiment with the nitrate radical (WB_NO3_1) has a slope close to zero with a lower final O:C probably due to the exclusion of the organonitrate group and different SOA formation reactions of the nitrate radical. The relatively low decrease in H:C with aging is supported by a relatively lower decrease of aCH absorptions in the MIR spectra of WB aerosols that are aged with the nitrate radical (Yazdani et al., 2021c). The modest decrease in H:C with aging is observed to be a characteristic of aging with the nitrate radical regardless of the emission source (Fig. 4.4d, h, and i), suggesting a more effective H atom abstraction by OH.

The starting points of the WB oxidation trajectories (from AMS) have H:C ratios in the range of 1.6–1.8 and O:C ratios in the range of 0.3–0.4. The ending points have H:C ratios in the range 1.4–1.6 and O:C ratios in the range of 0.6–0.7. The observed values are close to that of OA measured by Chhabra et al. (2011a), the OA emissions of logwood combustions by Tiitta et al. (2016), and ambient OA reported by Heald et al. (2010). The average O:C and H:C trends calculated from MIR spectroscopy are generally consistent with that of AMS, showing a decline in H:C ratio and increase in O:C ratio. However, there is an offset in the absolute values; in general, H:C is estimated to be approximately 0.2 higher by MIR spectroscopy both for the primary and aged WB aerosols.

For the CC experiments, AMS oxidation trajectories usually start at H:C ratios around 1.7–1.9

(higher than that of WB supported by the strong aCH signature; Fig. 4.2c) and O:C ratios around 0.20–0.25 and end usually at H:C and O:C ratios around 1.5–1.7 and 0.35–0.55, respectively. The high H:C ratios before aging support the low amount of aromatics to aliphatic CH observed from MIR spectra. In most of the experiments, a positive curvature in trajectory is observed (Fig. 4.4e-i), implying a change in the type of FGs produced during the course of aging. The average slopes are close to −1 (from AMS) in the majority of experiments and are slightly higher than those of the WB experiments. The average oxidation slopes that are estimated from MIR spectroscopy are also higher for the CC experiments compared to WB. This observation is supported by the formation of SOAs with higher naCO abundances for CC (Yazdani et al., 2021c), resulting in higher trajectory slopes . However, MIR generally estimates higher O:C (by 0.05–0.1) and lower H:C ratios (by approximately 0.2) for the aged CC aerosols compared to AMS.

The observed deviations between the two instruments might stem from the low OM mass collected on the filters that increases the baseline correction and peak fitting uncertainties in MIR analysis, in addition to the existence of FGs that are not considered in the peak fitting algorithm (e.g., ethers, PAHs, rC=C and rCH). Sampling biases of semi-volatile compounds on PTFE filters (Subramanian et al., 2004) and the uncertainties of AMS elemental ratio calibrations (Canagaratna et al., 2015; Aiken et al., 2008) can also affect the results.

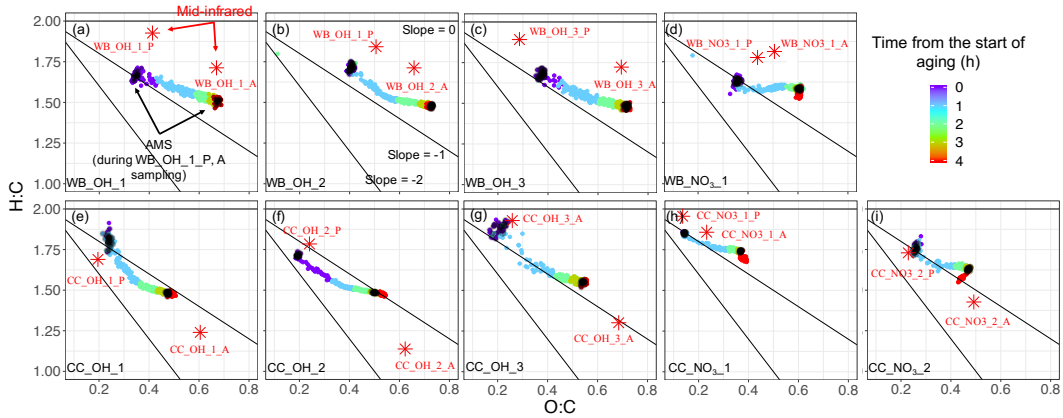

Figure 4.4 – Comparison of AMS van Krevelen (H:C vs O:C) aging trajectories (color circles) for wood burning (WB) and coal combustion (CC) experiments with MIR estimates for aerosols collected on PTFE filters before and after aging (red stars). Black circles in AMS trajectories correspond to the filter sampling periods for the primary and aged aerosols. The filter names are the same as Yazdani et al. (2021c). P: primary, A: aged.

4.3.2 AMS fragment ion-MIR FG relationships

In Sects. 4.3.2 and 4.3.2, the connection between the AMS fragment ions and MIR FGs is investigated using different statistical methods (covariance, correlation, and VIP scores). The combined summary of these analyses is shown in Table 4.1.

Table 4.1 – Summary of important fragment ions for each functional group based on the analysis method. Important fragments are shown in blue.

FGs	Covariance	Correlation	VIP scores	Multivariate regression CO_2^+, CHO^+, $C_2H_3O^+$
aCH	$C_xH_{2x\pm1}^+$ (e.g., $C_3H_7^+$, $C_4H_9^+$)	$C_xH_Y^+$ (e.g., $C_3H_5^+$, $C_3H_7^+$, $C_6H_9^+$, $C_8H_9^+$, heavy fragments: $C_8H_{17}^+$)	$C_3H_5^+$	–
aCOH	CHO^+, CH_3O^+, $C_2H_4O_2^+$	$C_2H_5O^+$, $C_3H_6O_2^+$, $C_5H_6^+$ (phenol)	CHO^+, $C_2H_4O_2^+$, $C_8H_9O_2^+$	CHO^+, $C_2H_3O^+$
COOH	CO_2^+, CO^+, $C_2H_3O^+$, CHO^+	$C_xH_yO_{>1}^+$ (e.g., $C_2H_3O_2^+$, $C_7H_5O_4^+$)	CO_2^+	CO_2^+, CHO^+
naCO	CO_2^+, CO^+, $C_2H_3O^+$	$C_xH_yO_1^+$ (e.g.,$C_7H_4O^+$)	$C_2H_3O^+$	$C_2H_3O^+$
$RONO_2$	–	$C_2H_3O^+$, $C_4H_9NO_3^+$, $C_6H_{10}NO_2^+$	CHO^+, $C_2H_3O^+$	–

Correlation and covariance analyses (univariate)

The aCH group has high covariance with C_xH_{2x-1} and C_xH_{2x+1} fragments ($C_3H_5^+$, $C_3H_7^+$, $C_4H_7^+$, $C_4H_9^+$, and $C_5H_9^+$; Fig. 4.5). The highest correlations are between the aCH group and $C_3H_5^+$, $C_3H_7^+$, $C_5H_7^+$, and $C_6H_9^+$ (Fig. 4.6). The relationship of heavier fragments such as $C_7H_{13}^+$ and $C_8H_{15}^+$ with aCH is more prominent in the correlation analysis. These fragments are especially abundant in the primary CC aerosols, suggesting these aerosols is composed of relatively longer chain hydrocarbons that even after fragmentation, produce relatively heavy fragments. This observation is also supported by the MIR spectra of these samples, which possess sharp CH_2 and weak CH_3 peaks (Yazdani et al., 2021a). Faber et al. (2017) have previously shown the relation between $C_4H_7^+$ and aCH. The m/z 57 signal in the unit-mass-resolution mass spectra (which also includes $C_4H_9^+$ signal) has been proposed to be a tracer of unburned fuel emissions (Schneider et al., 2006). However, its correlation coefficient with aCH has been shown to be quite variable and sometimes negative (Russell et al., 2009a). This discrepancy partly stems from the contribution of $C_3H_5O^+$ to m/z 57 (Faber et al., 2017) and partly from the fact that molecules with different chain-lengths, degrees of branching, and heteroatoms produce different and source-dependent fragmentation patterns for $C_xH_y^+$. In addition, the existence of several highly correlated ion fragments with aCH in this study suggests the superiority of a multi-variate approach to obtain information about aCH from the AMS mass spectra.

The COOH group has the highest covariance with CO_2^+, CO^+, $C_2H_3O^+$, and CHO^+. Since the CO^+ concentration is estimated from that of CO_2^+, the former fragment it not investigated separately. The highest correlations are, on the other hand, with $C_2H_3O_2^+$ and several heavier fragments with multiple oxygen atoms such as $C_7H_5O_4^+$, which are abundant in the aged WB aerosols. The high covariance with the CO_2^+ fragment is supported by the fact that CO_2^+ is produced from the fragmentation of mono- and dicarboxylic acids (Duplissy et al., 2011; Zhang et al., 2005). The heavier fragments with multiple oxygen atoms ($C_xH_yO_{z>1}^+$) are also indicative of the COOH FG as also reported by Lambe et al. (2012) and might be source-specific as, for example, $C_7H_5O_4^+$ is only detected in the WB aerosols. In the WB aerosols, the COOH group is correlated significantly ($r \sim 0.96$) with CHO_2^+, the fragment known to be produced from the α-cleavage of carboxylic acids (Pavia et al., 2008). The $C_2H_4O_2^+$ fragment is also known to be produced from acids having γ hydrogen through McLafferty rearrangement and its concentration increases with the extensive aging for the WB and CC aerosols (Yazdani et al., 2021c). However, the strong interference of levoglucosan fragmentation, abundant in WB emissions, with $C_2H_4O_2^+$ makes investigation of COOH-$C_2H_4O_2^+$ relation difficult. With the help of MIR, which does not suffer from the same interference, samples with negligible levoglucosan concentrations were separated. For these samples, a fairly strong correlation ($r \sim 0.82$) between COOH and $C_2H_4O_2^+$ was observed.

The aCOH group covaries most significantly with CHO^+, CH_3O^+, $C_3H_5O^+$ (which contributes to m/z 57 in the unit-mass-resolution spectra), $C_2H_3O^+$, and $C_2H_4O_2^+$. The CHO^+ fragment has been often interpreted as the tracer of esters, polyols and compounds with polyfunctional

groups without the carboxylic COOH (Canagaratna et al., 2015). This fragment, however, is also known to be produced by aldehydes although the aldehyde C–H band is not observed in the MIR spectra of the samples under study. Faber et al. (2017) also showed that the signal ratio of $C_2H_3O^+$ to $C_4H_7{}^+$ is linearly correlated with the molar ratio of aCOH to aCH. The connection of $C_2H_3O^+$ with alcohols, however, should be treated with caution as carbonyls can also produce the same mass fragment. The $C_2H_4O_2{}^+$ fragment appears to be important for aCOH and has been previously shown to be related to anhydrous sugars in the biomass burning smoke (Schneider et al., 2006). The CH_3O^+ fragment is produced from the α cleavage of alcohols (Pavia et al., 2008). The highest correlations in this work are, however, between the aCOH group and $C_2H_5O^+$ and $C_3H_6O_2{}^+$ fragments and some other fragments such as $C_5H_6{}^+$. The $C_2H_5O^+$ fragment is known to be produced from the α cleavage of alcohols (Pavia et al., 2008). The $C_5H_6{}^+$ fragment, correlated to a lower extent with aCOH, can be produced by phenol fragmentation after loosing CO, which is also abundant in WB emissions (Bruns et al., 2017).

The naCO group covaries most significantly with CO^+, $C_2H_3O^+$, and $CO_2{}^+$. Contrary to COOH, CHO^+ appears to have a low covariance with naCO. The $C_2H_3O^+$ fragment is known to be produced by aliphatic ketones and aldehydes (Pavia et al., 2008; Eadon et al., 1971). As discussed by Yazdani et al. (2021c), the naCO in the CC samples are believed to be mostly ketone based on their C=O frequency. The naCO group is highly correlated with some $C_xH_y{}^+$ fragments (e.g., $C_4H_3{}^+$, $C_6H_4{}^+$) and some single-oxygen fragments (e.g, C_5HO^+, $C_7H_4O^+$, and $C_7H_5O^+$). The $C_7H_5O^+$ fragment is known to be produced by aromatic ketones (Pavia et al., 2008) and $C_xH_yO_1{}^+$ has been attributed to carbonyls (Lambe et al., 2012). The $C_2H_3O^+$:$CO_2{}^+$ ratio is relatively higher in samples aged with the nitrate radical or samples that have considerable amounts of the naCO group (Fig. A.29), suggesting that $C_2H_3O^+$ is produced mainly by molecules possessing naCO or SOAs formed with the nitrate radical.

There are mid-infrared signatures attributed to levoglucosan and lignin that are prominent in the primary WB aerosols and diminish with aging. These signatures are important markers of biomass burning and have been used to identify atmospheric smoke-impacted samples (Yazdani et al., 2021c). Correlation analysis of these features with the AMS ion fragments (Fig. A.30) show that the $C_8H_9O^+$ fragment is related to lignin. In fact, one might attribute the m/z 121 fragment to two peaks $C_7H_5O_2{}^+$ and $C_8H_9O^+$ for hydroxyphenyl (H) lignin and $C_4H_9O_2{}^+$ of guaiacyl (G) lignin, respectively (Li et al., 2012; Tolbert and Ragauskas, 2017). The $C_2H_4O_2{}^+$ and $C_3H_5O_2{}^+$ fragments have high correlation with MIR levoglucosan signatures. The fragment at m/z 102, in spite of having generally a low concentration, has the highest correlation with the levoglucosan concentration measured from the MIR spectra. This fragment might be used alternatively in case the interference of other compounds (e.g., acids) for the lighter fragments related to levoglucosan is substantial.

In this work, organonitrates were not quantified. However, the MIR peak attributed to $RONO_2$ has high correlation coefficients with $C_2H_3O^+$ and several other oxygenated fragments such as $C_4H_7O^+$. Nitrogenated fragments containing the nitrate and nitro groups such as $C_4H_9NO_3{}^+$

and $C_6H_{10}NO_2^+$ also appear to have moderate correlation coefficients (approximately 0.6) with the $RONO_2$ peak in the MIR spectrum (Fig. A.31) although the quantification of nitrogenated fragments is often not accurate in the V mode due to peak overlapping..

To summarize, the high correlation coefficients of several fragments with each FG suggest that FG information is retained to a good degree in the AMS spectra. We also found that multiple FGs are correlated with each of the major oxygenated fragments, (CO^+, $C_2H_3O^+$, CO_2^+, and CHO^+). As a result, a multivariate approach should be taken to infer FG abundances from the AMS spectra.

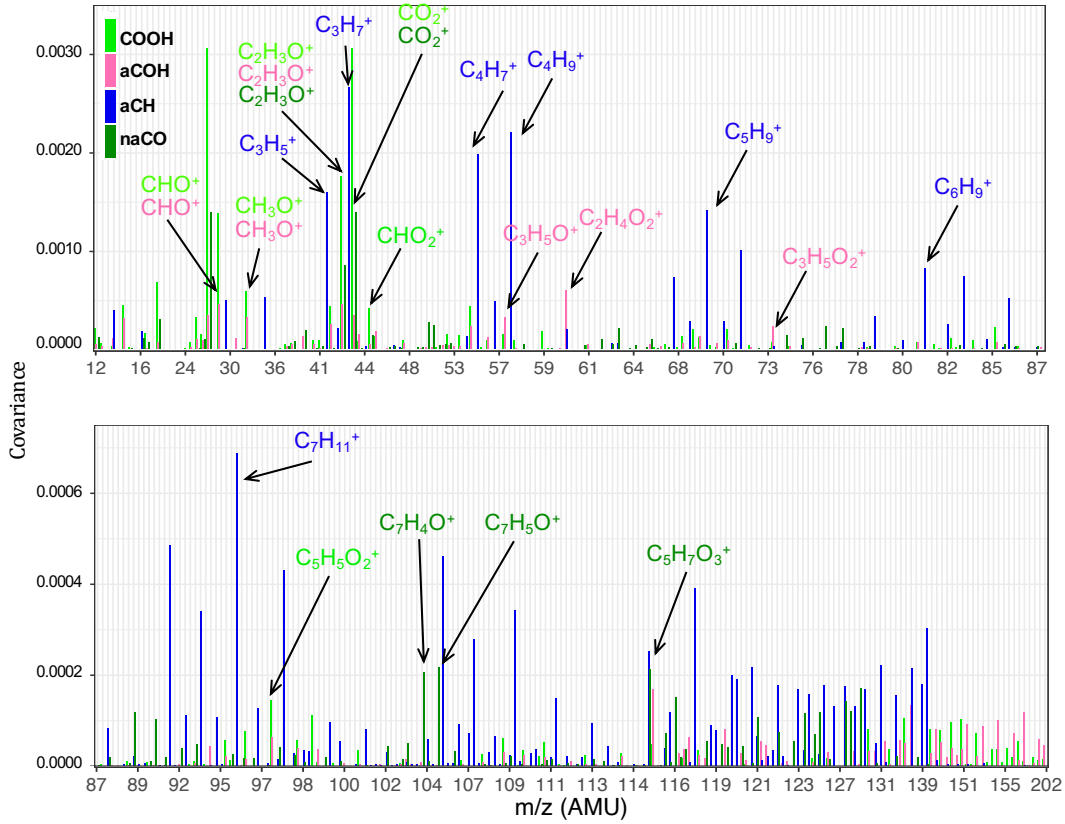

Figure 4.5 – Bar plots showing covariances of normalized AMS fragment ion concentrations and normalized FG abundances. Only positive covariances are shown.

VIP scores (multivariate)

As can be seen from Fig. 4.7, the CO_2^+ fragment has the highest VIP scores for the carbonyl ν(CO) and broad acid ν(OH) peaks from 2400 to 3400 cm^{-1}. This is consistent with previous studies (e.g., Zhang et al., 2005) and our univariate analyses (Sect. 4.3.2). On the other hand, the ν(CH) region (2800 to 3200 cm^{-1}), interfering with the broad acid OH stretching band, has high VIP scores with negative regression coefficient, showing that aCH relative concentration is anti-correlated with CO_2^+. Although some interference for the CO_2^+ fragment is expected

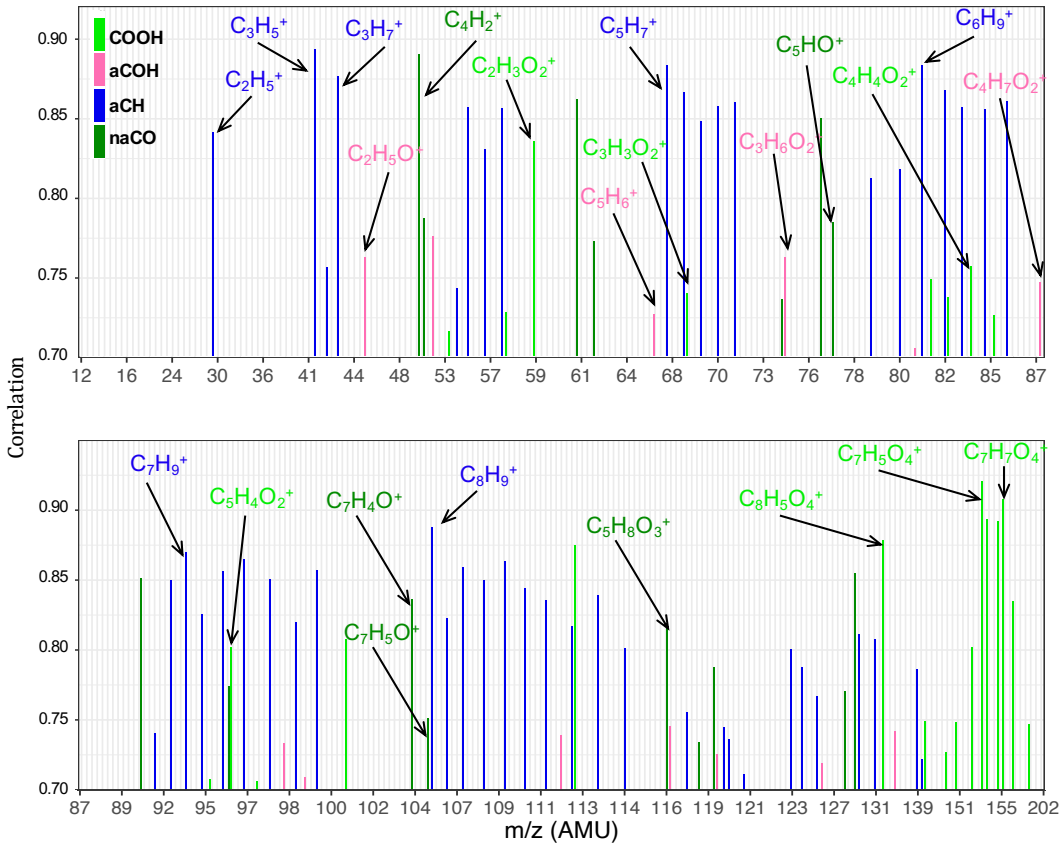

Figure 4.6 – Bar plots showing the Pearson correlation coefficients of normalized individual fragment ion concentrations and normalized FG abundances. Only values > 0.7 are shown.

from gas-phase CO_2 in the AMS spectra, our results show that the compensation method appears to eliminate this interference effectively.

The COH^+ has the highest VIP scores for the $RONO_2$ peaks and the broad alcohol ν(OH) at 3400 cm^{-1}. These results suggest that alcohols and the SOA species produced during the aging with the nitrate radical (that can also be alcohols) are mostly responsible for producing this mass fragment. Although the interference form the gas-phase $^{15}N\,^{14}N$ can be significant for CHO^+, our results show that CHO^+ appears to be meaningfully indicative of alcohols after compensation for the gas-phase interference.

The $C_2H_3O^+$ fragment has the highest VIP scores for the carbonyl CO and the $RONO_2$ peaks and also to a lesser extent for the broad acid ν(OH) peak. This observation suggests that $C_2H_3O^+$ is mostly produced by fragmentation of carbonyls and SOA species formed by aging with the nitrate radical and to a lesser extent carboxylic acids.

The $C_3H_5{}^+$ fragment was chosen for the VIP scores analysis due to having high concentrations for both CC and WB aerosols. This fragment has the highest VIP scores with positive regression coefficients for the ν(CH) (2800–3000 cm^{-1}) and δ(CH) (1300–1500 cm^{-1}) peaks, showing that

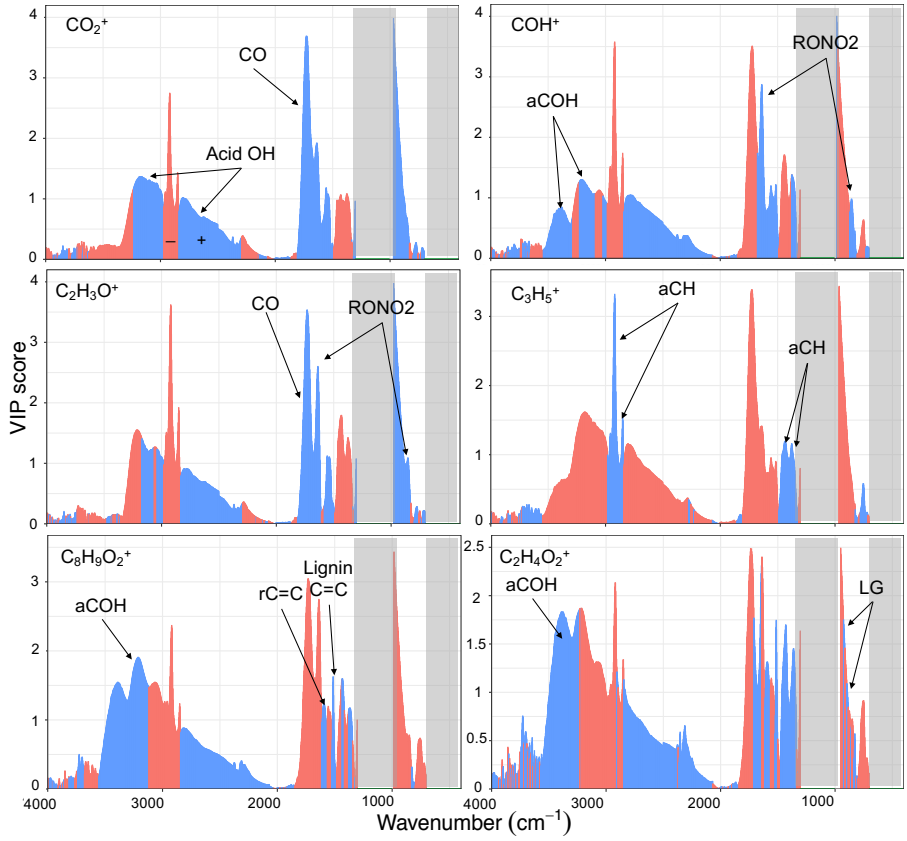

Figure 4.7 – VIP scores of MIR absorbances regressed against AMS fragment ion concentrations (averaged over the filter sampling periods). Blue and red scores correspond to wavenumbers with positive and negative regression coefficients in the PLSR models, respectively. Important FGs for each mass fragment are indicated and PTFE absorption regions are masked by gray rectangles.

this fragment is directly related to aCH for both sources. This result has been expected but also highlights the fact that $C_xH_y^+$ fragments should be chosen wisely based on the aerosol source to provide useful information about the aCH group.

Regarding the biomass burning markers, the $C_8H_9O_2^+$ fragment that is proposed to be produced by the fragmentation of lignin molecules (Li et al., 2012; Tolbert and Ragauskas, 2017) has the highest VIP scores with positive coefficients in the aromatic ν(C=C) (1515 and 1600 cm^{-1}) and aCOH regions, suggesting the connection of this fragment with aromatic compounds having a lignin-like substitution that generates the sharp peak at 1515 cm^{-1} and aCOH groups. The $C_2H_4O_2^+$ fragment that is proposed to be produced predominantly by fragmentation of levoglucosan molecules has the highest VIP scores with positive coefficients in the aCOH region (3400 cm^{-1}), suggesting the abundance of aCOH in molecules producing this fragment. In addition, high VIP scores with positive regression coefficients is observed in the 850–1000 cm^{-1} region, which was previously proposed to be related to levoglucosan

fingerprint absorbances (Yazdani et al., 2021c).

We also performed a simple multivariate linear regression between the oxygenated FGs (aCOH, naCO, and COOH) and major fragments (CO_2^+, CHO^+ and $C_2H_3O^+$; Fig A.32) for the experiments in which the hydroxyl radical was used as oxidant. As shown in Fig. A.32, regressing CO_2^+ against COOH, aCOH, and naCO results in the highest regression coefficient for the COOH group. In a similar regression for CHO^+, the relative contribution of aCOH increases (Fig. A.32). However, a high regression coefficient for COOH is still observed. The regression for $C_2H_3O^+$ highlights a relatively greater contribution of naCO (Fig. A.32). However, as for CHO^+, a high regression coefficient for COOH is also observed. As summarized in Table 4.1, different statistical methods suggest that the major fragments are usually produced more by a certain oxygenated FG, while interference from other FGs might also be significant. This motivates the use of multivariate methods for predicting FG abundances using fragment ion concentrations in the following section.

4.3.3 MIR FG interpolation using AMS mass spectra

We showed in previous sections that AMS and MIR measurements are consistent. We also found that FG information is maintained in the AMS mass spectra, which motivated the use of multivariate methods to access this information. For this purpose, normalized AMS spectra were regressed against normalized FG compositions from MIR peak fitting. The fit quality of the developed models is reasonable (Figs. A.33 and A.34) with their R^2 ranging from 0.71 to 0.94. These models use mass fragments to predict the FG compositions that were found to be important in previous sections. We used the developed PLSR models to interpolate the functional group composition of WB and CC OM between the filter sampling periods (primary and aged) using the AMS spectra (Fig. 4.8). These models are especially helpful as AMS has a considerably better time resolution and can be used to investigate the FG evolution of OM during the course of aging.

The interpolated functional group compositions (Figs. 4.8, A.35, and A.36) show different FG compositions and trends during the course of oxidation for WB and CC aerosols. This is predominantly seen in the fraction of oxygenated functional groups that emerge with aging. For the WB experiments, the aCH relative abundance falls steeply as aging with the hydroxyl radical starts (Fig. 4.8a). This is also observed for aCOH. On the other hand the COOH relative abundance increases significantly as soon as the aging starts and levels off after two hours of aging. The relative abundance of naCO, however, does not change significantly compared to other FGs for the WB experiment (Fig. 4.8a). There are slight differences between different experiments of WB aging with the hydroxyl radical (also observed in their van Krevelen trajectory in Fig. 4.4a–c). For instance, the relatively linear trajectory of the WB_OH_1 experiment (Fig. 4.8a) is concurrent with monotonic FG composition change (Fig. A.35a). On the other hand, the curved van Krevelen trajectories of WB_OH_2 and WB_OH_3 (Fig. A.36b–c) correspond to the consumption of naCO produced after the start of aging and the gradual increase of

the COOH relative abundance (Fig. A.36b–c). The different SOA species formed by oxidation with hydroxyl and nitrate radicals is also reflected in the evolution of OM FG composition. When aging with the nitrate radical, the decrease in the relative abundance of aCH is much less prominent after the start of aging compared to aging with the hydroxyl radical (Fig. 4.8b) although the OM mass enhancement is comparable between the two (Fig. A.35a–d). This observation suggests the formation of different SOA species with higher relative abundance of aCH when the nitrate radical is used. This is also supported by the horizontal trajectory in the van Krevelen diagram (no decrease in H:C) (Fig. 4.4d). No clear difference in the composition of oxygenated FGs (except organonitrate) is observed between aging with the hydroxyl and nitrate radicals. However, when the nitrate radical is used, the O:C ratio increases to lower levels and reaches a plateau faster (Fig. 4.8e–f; also true for the CC OM). This observation is consistent with the fact that the nitrate radical is produced with a single injection but the hydroxyl radical is generated continuously throughout the aging. It is observed that most prominent changes in the FG composition for both oxidants occur in the first hour of aging when the OM mass changes the most (Fig A.35), while small changes are observed toward the end of aging. Looking at the absolute abundances of functional groups, we observe that the mass concentrations of all FGs including aCH and aCOH increase during the course of aging (Fig A.35) and it is the different rate of their increase that changes their relative abundance as shown in Fig. 4.8.

We observe for the CC experiments a different composition of FGs emerging after the start of aging that also evolves differently as aging continues. Like for the WB experiments, the aCH relative abundance decreases drastically with aging (Fig. 4.8c), while its absolute concentration increases only slightly with aging (Fig. A.35). The decrease in the aCH relative abundance is, however, less prominent when the nitrate radical is used (Fig. 4.8d) as also supported by the lower decrease in the H:C shown the van Krevelen plots (Fig. 4.4h–i). Unlike the WB experiments, the relative abundance of the aCOH group increases slightly with aging when using both oxidants (Fig. 4.8c–d). The relative abundances of naCO and COOH show more complex behaviors. The relative abundance of naCO increases sharply and naCO becomes the major FG with the start of aging for both oxidant but decreases slightly after continued aging (Fig. 4.8c–d). The relative abundance of COOH decreases initially (Fig. 4.8c), however, after about one hour into the aging process (earlier with the nitrate radical), when there is no more significant OM enhancement, the COOH relative abundance starts to increase gradually. This observation is consistent with the ripening phenomenon (Wang et al., 2018b) in which the composition of the SOA keeps changing and becomes more oxidized, while the change in OM mass is minimal. This phenomenon is also observed in the L-shaped oxidation trajectories of CC OM in the van Krevelen plot of Figs. 4.4e–i for both oxidants.

To summarize, the interpolated FG compositions are supported by the van Krevelen trajectories, but provide insight into the oxidation pathways that cannot be independently obtained from the van Krevelen plots (e.g., several combinations of FGs can produce similar slopes). For two CC experiments, negative concentrations of COOH are predicted (Fig A.36e–f). These unphysical values are believed to resulted due to uncertainties of PLSR models and quantifi-

cation uncertainties for the COOH group from the MIR peak fitting. The predicted trends, however, are still informative.

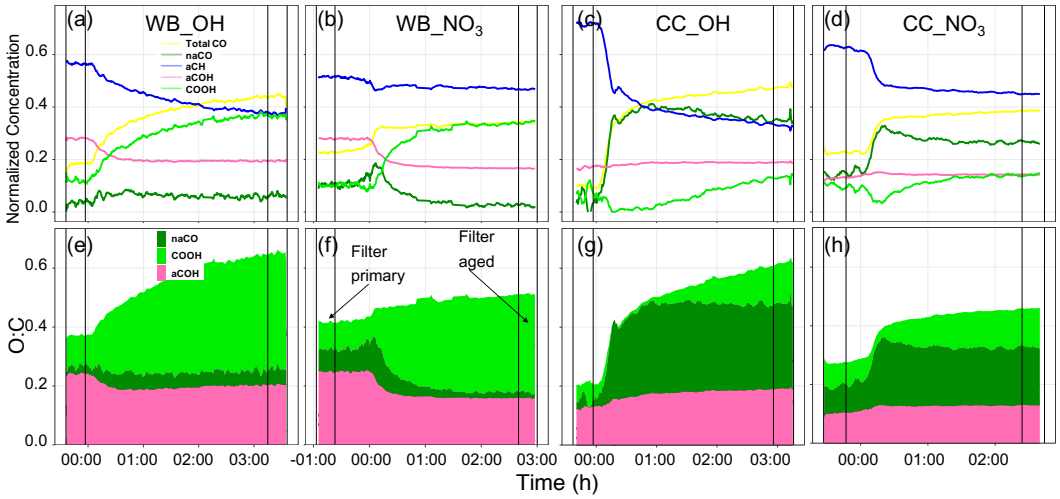

Figure 4.8 – Time series of normalized concentration of functional groups interpolated using AMS mass spectra (a–d) and time series of O:C ratios calculated from the interpolated functional groups (e–h). An example for each source (CC and WB) and oxidant (OH and NO_3) has been shown. The time zero indicates the start of aging (UV lights turned on or oxidant injected). The horizontal lines indicate the periods of filter sampling.

4.4 Concluding remarks

By comparing MIR and AMS measurements for a series of environmental chamber experiments, we found that the four MIR functional groups (FGs) highlighted by the VIP scores method (i.e., aCH, aCOH, COOH, and naCO) explain the OM mass, OM:OC, H:C, and O:C in good agreement with AMS measurements for the wood burning (WB) and coal combustion (CC) aerosols.

By using univariate and multivariate methods, we found that several small (low m/z) and large (m/z > 100 and generally source-specific) AMS fragment ions are informative about the FG composition of primary and secondary OMs. The heavy, source dependent fragments identified in the aged wood burning aerosols can potentially be used to identify and quantify the secondary biomass burning OM in the atmosphere.

The multivariate analysis indicates that when OMs with different proportions of oxygenated FGs (i.e., COOH, COH, and naCO) are fragmented in the AMS, they produce different proportions of the major oxygenated fragments (i.e., CO_2^+, CHO^+, $C_2H_3O^+$). A multivariate method was used to extract the high-time-resolution FG information from the AMS spectra to better understand the evolution of the OM composition during the course of aging. The results suggest the formation of moderately oxygenated FGs (e.g., naCO) soon after the start of aging and the transformation of moderately oxygenated FGs to the more oxygenated FG, COOH,

with continued aging.

5 Chemical evolution of primary and formation of secondary biomass burning aerosols during daytime and nighttime

Manuscript in preparation

Authors : Yazdani, Amir; Takahama, Satoshi; Kodros, Jack K.; Paglione, Marco; Masiol, Mauro; Squizzato, Stefania; Florou, Kalliopi; Pandis, Spyros N.; Nenes, Athanasios

Contributions: AN, SNP, ST, and AY conceived the project. AY prepared filter sampling set-up. JKK, MP, KF, MM, SS performed the chamber experiments. KF processed AMS row signals. AY processed FTIR spectra and performed the data analysis, and wrote the manuscript. ST, AN, and SNP provided regular input on the analysis and the further editing of the manuscript. ST and AN provided overall supervision of the project.

Reference: Yazdani et al. (2021d)

Abstract Primary emissions from wood and pellet stoves were aged in an atmospheric simulation chamber. Both daytime and nighttime conditions were simulated. Online AMS and offline FTIR measurements agreed reasonably well in terms of the organic aerosol (OA) mass concentration, OA:OC trends, and concentrations of biomass burning markers – lignin-like compounds and anhydrosugars. Around 15 % of the primary organic aerosol (POA) mass underwent some form of transformation (evaporation or oxidation) during the daytime oxidation at relatively short timescales (equivalent to 6–10 hours of atmospheric exposure) and to a lesser extent during the nighttime oxidation. The decay of certain semi-volatile (e.g., levoglucosan) and less volatile (e.g., lignin-like) POA species was substantial during aging, highlighting the role of gas-particle partitioning and heterogeneous reactions with oxidants. Lignin-like compounds were observed to degrade efficiently under both daytime and nighttime conditions, whereas anhydrosugars degraded efficiently only under daytime conditions. Among the marker mass fragments of primary biomass burning OA (bbPOA), large ones (higher m/z) were relatively more stable with regard to aging. The AMS and FTIR spectra of the formed biomass burning secondary OA (bbSOA) became more oxidized with continued aging and resembled those of aged organic aerosols in the atmosphere. The bbSOA formed during the daytime oxidation was dominated by acid contributions. Organonitrates, on the other hand, appeared to be an important product of nighttime reactions in both humid and dry conditions.

5.1 Introduction

Fine particulate matter (PM) in the atmosphere impacts climate and visibility (McFiggans et al., 2004; Hallquist et al., 2009) and is known to cause respiratory and cardiovascular diseases, leading to premature deaths (Pope et al., 2009; Shiraiwa et al., 2017b; Burnett et al., 2018). A major fraction (up to 90 %) of fine PM is organic. Organic aerosol (OA) has various sources and formation mechanisms in the atmosphere, resulting in its complex chemical composition (Russell, 2003; Kanakidou et al., 2005; Hallquist et al., 2009). Primary organic aerosols (POA) are emitted directly from their sources, whereas secondary organic aerosols (SOA) are formed through chemical reactions and condensation of volatile, intermediate-volatility, and semi-volatile organic compounds (VOCs, IVOCs, and SVOCs, respectively) (Seinfeld and Pandis, 2016). Oxygenated OA is often the most important part of atmospheric OA, highlighting the role of atmospheric chemistry and aging for OA (Zhang et al., 2007) as particles and gases are exposed to oxidants for days in the atmosphere (Wang et al., 2018b). Different types of aging have been proposed and investigated in previous studies: homogeneous gas-phase oxidation (Donahue et al., 2012), oligomerization (Kalberer et al., 2006), heterogeneous reactions with oxidants (Robinson et al., 2006a; George et al., 2008), and photolysis (Bateman et al., 2011; Henry and Donahue, 2012). Homogeneous gas-phase reactions are generally believed to dominate (Henry and Donahue, 2012), while other mechanisms may be important under different conditions (e.g., Hearn et al., 2005; Hung et al., 2005; Nah et al., 2014).

Biomass burning (BB) contributes significantly to atmospheric primary and secondary OA

(POA and SOA) (Puxbaum et al., 2007; Qi et al., 2019; Lanz et al., 2010) and black carbon (BC) (Bond et al., 2013). BB is expected to have an increasing contribution to $PM_{2.5}$ in the foreseeable future (Ford et al., 2018). Primary biomass burning OA (bbPOA) and the secondary OA (bbSOA) formed using hydroxyl and nitrate radicals have been investigated in several environmental chamber studies (Johansson et al., 2004; Bäfver et al., 2011; Alves et al., 2011; Hennigan et al., 2011; Tiitta et al., 2016; Bruns et al., 2015; Bertrand et al., 2017, 2018a; Yazdani et al., 2021a). These studies highlighted a net enhancement in the OA concentration with aging and different chemical compositions for bbSOA formed using different oxidants. In order to estimate the contribution of bbPOA after aging in the chamber experiments or in field measurements, it has often been assumed implicitly or explicitly to be stable during the course of aging (e.g., Robinson et al., 2007; Grieshop et al., 2009; Tiitta et al., 2016; Kodros et al., 2020), neglecting the effects of heterogeneous reactions, photolysis, and gas-particle partitioning of its component. This is while several studies have reported significant degradation of bbPOA markers like anhydrosugars and methoxyphenols (Hennigan et al., 2010, 2011; Slade and Knopf, 2013; Bertrand et al., 2018a; Yazdani et al., 2021c). These compounds constitute a significant fraction (up to 50 %) of the bbPOA mass (Fine et al., 2002; Bertrand et al., 2018a; Yazdani et al., 2021c) and their chemical processing can impact the bbPOA mass and composition significantly. Nonetheless, the overall change in the bbPOA mass and composition during the course of aging with different oxidants is not well understood.

Recent efforts using the volatility basis set (VBS; Donahue et al., 2006) address the volatility and gas-particle partitioning of POA including primary bbOA in simulations (Robinson et al., 2007; Theodoritsi et al., 2020). The chemical processing of bbPOA remains uncertain and heterogeneous reactions of bbOA compounds are not included in most models. Moreover, chemical transport models (CTMs) have difficulties reproducing the OA in areas or periods affected by fires (Fountoukis et al., 2014; Theodoritsi et al., 2020).

Aerosol mass spectrometry (AMS) and mid-infrared (MIR) spectroscopy are two methods used in this study that provide information about the bulk composition of the OA (Nozière et al., 2015) and certain bbOA markers (Schneider et al., 2006; Yazdani et al., 2021c). AMS has some known limitations due to the extensive molecule fragmentation and the variability of particle collection efficiency (Canagaratna et al., 2007; Faber et al., 2017; Kumar et al., 2018). Fourier transform infrared spectroscopy (FTIR) is a non-destructive method that measures the abundance of certain functional groups but with a limited temporal resolution. Functional group abundances are then used to estimate the OA mass concentration and elemental ratios (Coury and Dillner, 2008; Ruthenburg et al., 2014; Reggente et al., 2016; Boris et al., 2019). A recent study also shows the ability of FTIR to quantify bbOA marker molecules (Yazdani et al., 2021c).

In this work, primary biomass burning emissions are injected into an environmental simulation chamber and aged. AMS and FTIR are used in tandem to better understand and quantify the evolution of primary and aged bbOA during the chemical aging. We adopt a particle wall loss correction method based on AMS organic measurements and develop a

procedure to quantify the overall changes in the composition of primary and aged bbOA. The results of this study allow us to evaluate the stability of bbPOA (including its markers) and the importance of other aging mechanisms compared to the homogeneous gas-phase oxidation under atmospherically relevant conditions.

5.2 Methods

5.2.1 Experimental set-up and procedure

Primary emissions from common wood and pellet stoves were diluted and injected for 30–40 minutes into a 10 m^3 Teflon atmospheric simulation chamber located at the Foundation for Research and Technology-Hellas (FORTH), Greece. Olive wood logs with bark and EN*plus*® A1 pellets were used as fuel. Details of the chamber and combustion facilities have been discussed elsewhere (Kaltsonoudis et al., 2017; Kodros et al., 2020). The fuels and stoves used in this work are commonly used in the region. Primary emissions were held in the chamber after the injection for around two hours to ensure proper mixing and to characterize chamber wall losses.

Nine experiments were conducted in this work. For the three reference experiments (Table 5.1), emissions were held in the dark chamber without addition of any oxidants. For the two experiments simulating the daytime aging, reactions were initiated by turning on UV lights (J_{NO_2} = 0.59 min^{-1}) for 2 h and the subsequent generation of the hydroxyl radical via ozone photolysis in the presence of water vapor. For these experiments the average RH was roughly 50 % and the average OH concentration was (3–5)$\times 10^6$ molec cm^{-3}. This corresponds to 6–10 h of aging in the atmosphere assuming an average OH concentration of 10^6 molec cm^{-3} (Seinfeld and Pandis, 2016). The OH radical concentration was estimated using a proton-transfer-reaction mass spectrometer (PTRMS) monitoring the concentration of 1-butanol-d9 (Barmet et al., 2012). The four nocturnal aging experiments were conducted under two different RH regimes: dry (8–10 %) and humid (60–80 %). For these experiments, roughly 100 ppb of NO_2 was injected into the chamber before the injection of primary biomass burning emissions. Around two hours after the injection of primary emissions, aging was initiated by injection of O_3 (roughly 100 ppb) and production of the NO_3 radical via the reaction of ozone with NO_2.

5.2.2 On- and off-line PM measurements

The composition of non-refractory aerosols in the chamber was measured by a HR-ToF AMS (Aerodyne Research Inc.) operating in V mode with a 1 µm inlet aerodynamic lens. For certain experiments, AMS also measured the composition of chamber aerosols after being passed through a thermodenuder to study their volatility. A scanning mobility particle sizer (classifier model 3080, DMA model 3081, CPC model 3787, TSI) was used with adjusted flow rates to measure particle number size distribution in the 15–700 nm range. Primary and aged

Table 5.1 – Description of experiments and initial O_3 and NO_2 concentrations.

Experiment	Type	Fuel	NO_2 (ppb)	O_3 (ppb)	RH (%)
Exp. 1	Reference (no oxidant)	Wood	–	–	10
Exp. 2	Reference (no oxidant)	Pellet	–	–	10
Exp. 3	Reference (no oxidant)	Wood	–	–	50
Exp. 4	UV + 30 pbb SO_2	Wood	–	–	50
Exp. 5	UV + 80 pbb SO_2	Wood	–	–	50
Exp. 6	Dark and dry	Wood	100	100	10
Exp. 7	Dark and dry	Pellet	100	100	10
Exp. 8	Dark and humid	Wood	100	100	80
Exp. 9	Dark and humid	Pellet	100	100	60

PM_1 were collected on 47 mm PTFE filters (Pall corporation, 1 cm diameter of the collection surface). Sampling on PTFE filters was performed at a flow rate of 8 L min^{-1} for 20 min using a flow system composed of a silica gel denuder to minimize aerosol water and a sharp-cut-off cyclone. Filters were immediately stored in filter petri dishes at 253 K after sampling to minimize volatilization of aerosols and chemical reactions. Filters were analyzed using a Bruker Vertex 80 FTIR instrument equipped with an α deuterated lanthanum alanine doped triglycine sulfate (DLaTGS) detector and a custom-made filter mini-chamber to minimize water vapor and CO_2 interferences. The spectra were obtained at a resolution of 4 cm^{-1} and were averaged over 128 scans.

5.2.3 Data analysis

Initial spectral post-processing

The AMS raw signal was post-processed using the AMS software toolkits SeQUential Igor data RetRiEvaL (SQUIRREL) v1.57 and the Peak Integration by Key Analysis (PIKA) v1.16. The elemental and OA:OC ratios were calculated using the approach of Canagaratna et al. (2015). Organic nitrate concentrations were calculated based on NO^+ and $NO_2{}^+$ peak ratios in the AMS mass spectra following the approach of Farmer et al. (2010).

Baseline correction was performed on the FTIR spectra to eliminate the contribution of light scattering from the spectra (Russo et al., 2014; Parks et al., 2019) using smoothing splines. After baseline correction, blank subtraction was done to recover some of the overlapping features with PTFE peaks (e.g., levoglucosan fingerprint bands 860–1050 cm^{-1}). A multiple peak-fitting algorithm was applied on the FTIR spectra to estimate abundances of alcohol (referred to as aCOH), carboxylic acid (COOH), alkane (aCH), and non-acid carbonyl (naCO) groups (Takahama et al., 2013). The ratio of fingerprint absorbances related to levoglucosan and similar anhydrosugars (multiple peaks in the 860–1050 cm^{-1} range; Yazdani et al., 2021c) and lignin-like compounds (single sharp peak at 1515 cm^{-1}) were compared between primary and aged aerosols to estimate the diminution of these two biomass burning markers with

aging. Absolute levoglucosan concentrations were estimated using the mentioned fingerprint absorbances and the absorption coefficients calculated from aerosolized levoglucosan standards.

Wall loss correction and subsequent analysis

Particle wall loss correction was carried out assuming a first-order, time independent wall loss rate for the OA mass concentration (Pathak et al., 2007):

$$\ln C_{\mathrm{OA}}(t) = -k_{\mathrm{OA}}\, t + \ln C_{\mathrm{OA}}(0), \tag{5.1}$$

where k_{OA} is the wall loss rate constant and C_{OA} is the AMS OA mass concentration. The rate constant for each experiment was calculated using OA concentrations from 0.5–1 h after the injection of primary emissions into the chamber (to ensure proper mixing) up to the start of chemical aging or until the end for reference experiments . After calculating wall loss rate constants, wall-loss-corrected concentrations of individual AMS fragments were calculated by

$$C_i^{\mathrm{cor}}(t) = C_i^{\mathrm{obs}}(t) + k_{\mathrm{OA}} \int_0^t C_i^{\mathrm{obs}}(t)\, dt, \tag{5.2}$$

where $C_i^{\mathrm{obs}}(t)$ is the measured concentration of fragment i at time t, and $C_i^{\mathrm{cor}}(t)$ is its wall-loss-corrected concentration at time t. The fragments that fall below their initial concentration indicate POA transformation. The difference between the wall-loss corrected and initial concentrations of diminishing fragments (those with $C_i^{\mathrm{cor}}(t) - C_i^{\mathrm{cor}}(0) < 0$) was used to quantify the extent of POA transformation at each time after the start of aging:

$$\frac{\sum (C_i^{\mathrm{cor}}(t) - C_i^{\mathrm{cor}}(0))}{C_{\mathrm{POA}}(0)}. \tag{5.3}$$

The apparent contribution of the fresh POA to aged OA spectra (FTIR or AMS) was calculated using Eq. (5.4) assuming no evaporation, condensation, or heterogeneous reactions,

$$\boldsymbol{s}_{\mathrm{POA}}(t) = \boldsymbol{s}_{\mathrm{POA}}(0) \exp(-k_{\mathrm{OA}}\, t), \tag{5.4}$$

where $\boldsymbol{s}_{\mathrm{POA}}(t)$ is the apparent POA spectrum at time t, and $\boldsymbol{s}_{\mathrm{POA}}(0)$ is the POA spectrum at the start of aging. The residual spectrum at time t, $\boldsymbol{s}_{\mathrm{res}}(0)$, was then calculated by subtracting the apparent POA spectrum at time t from that of the aged OA (Appendix A.4.1):

$$\boldsymbol{s}_{\mathrm{res}}(t) = \boldsymbol{s}_{\mathrm{OA}}(t) - \boldsymbol{s}_{\mathrm{POA}}(t). \tag{5.5}$$

The residual spectrum is composed of positive and negative elements. The positive elements indicate the SOA species or oxidation products of fresh POA (e.g., due to heterogeneous oxidant reactions). The negative elements, on the other hand, indicate the loss of fresh POA species.

Wall loss rates calculated from AMS OA (k_{OA}) were compared to those calculated from AMS sulfate concentrations k_{SO_4} and the size-resolved coagulation-corrected rates from SMPS (k_c) (Wang et al., 2018a) for consistency check. The average k_c values are reported in the Table 5.2 based on the size distribution of organic particles.

Dimension reduction of AMS spectra

We also used principal component analysis (PCA; Hotelling, 1933) to simplify the high-dimensional, inter-correlated AMS spectra. This allowed us to better understand the evolution of bbOA during the course of aging in the chamber and to compare it to atmospheric aerosols. PCA calculations were performed on the normalized and uncentered AMS spectra (by total AMS OA concentration) from the chamber experiments (1600 spectra with 800 fragments) using singular value decomposition (Abdi and Williams, 2010). Thereafter, positive matrix factorization (PMF) factors for atmospheric OA from previous studies (e.g., Aiken et al., 2009) were projected onto the PC space for comparison.

5.3 Results

The wall loss rates calculated from the AMS OA are generally close to those calculated from the AMS sulfate or the size-resolved coagulation-corrected rates from SMPS (Table 5.2). The wall loss rates based on AMS OA are able to explain well the overall variations in the OA concentration for the reference experiments and before the initiation of aging for other experiments (Figs. A.38a and A.39a). Small but systematic differences between loss rates of different fragments was observed (Fig. 5.1a) that is investigated further in Sect. 5.3.1. The decrease in the OA concentration due to wall losses estimated by FTIR peak fitting for the reference experiments closely matched (less than 10 % difference) that by AMS. The latter supports the wall loss rate constants derived from AMS OA. In general, a reasonable agreement observed between OA mass concentrations estimated by AMS and FTIR (Appendix A.4.2).

5.3.1 Evolution of primary biomass burning aerosols

The chemical composition of primary wood burning (WB) aerosols has been characterized elsewhere (e.g., Bertrand et al., 2017; Yazdani et al., 2021c) and is only treated here briefly. The FTIR spectra of primary aerosols of WB largely resemble that of wood constituents such as lignin (Yazdani et al., 2021c). The aCOH FG is additionally very prominent for these aerosols and signatures of levoglucosan are observed in their FTIR spectra (Fig. A.41). The aCH peaks are more prominent in the FTIR spectra of WB aerosols of this work compared to those of Yazdani et al. (2021c) likely due to the combustion of wood with bark. Primary pellet burning (PB) aerosols also have similar FTIR spectra to that of wood (Pandey, 1999) and have higher relative concentrations of aCOH and levoglucosan compared to WB aerosols (Fig. A.41). For different fuels (wood and pellet), different ratios of lignin-related fragments (e.g., $C_8H_9O_2^+$,

Table 5.2 – Summary of wall loss rates for each experiment. The ratio of certain markers before aging (first PTFE filter) and after aging (second PTFE filter) are compared to that of apparent POA concentration.

	Reference (no oxidant)			UV		Dark (NO_3) and dry		Dark (NO_3) and humid	
	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6[b]	Exp. 7	Exp. 8	Exp. 9
AMS OA wall loss rate, k_{OA} (h^{-1})	0.057	0.063	0.082	0.103	0.099	0.071	0.088	0.112	0.097
AMS SO_4 wall loss rate, k_{SO_4} (h^{-1})	0.052	0.037	0.074	0.087	0.093	0.063	0.077	0.084	0.076
SMPS wall loss rate k_c (h^{-1})	0.10±0.06	0.07±0.03	0.10±0.06	0.10±0.05	0.10±0.03	0.07±0.02	0.07±0.02	0.12±0.04	0.09±0.05
POA ($C_{POA}(t)/C_{POA}(0)$)	0.90	0.85	0.73	0.63	0.68	0.77	0.73	0.64	0.68
Levoglucosan FTIR	0.88	0.83	0.54	0.27	0.40	–	0.88	0.35	0.65
$C_2H_4O_2^+$	0.88	0.91	0.52	0.27	0.36	0.90	0.78	0.40	0.54
$C_3H_5O_2^+$	0.96	0.91	0.67	0.36	0.45	0.90	0.79	0.51	0.58
Lignin FTIR	1.01	0.83	0.84	0.10	0.24	–	0.89	0.34	0.10
$C_8H_9O_2^+$	0.94	0.79	0.73	0.13	0.18	0.65	0.69	0.20	0.25
$C_9H_{11}O_3^+$	0.95	0.70	0.78	0.10	0.11	0.53	0.68	0.18	0.50
$C_{10}H_{13}O_3^+$	0.94	0.82	0.77	0.40	0.56	0.90	0.72	0.48	0.50
aCH FTIR[a]	>1	>1	>1	0.50	0.45	–	0.45	>1	>1
$C_4H_9^+$	0.91	0.83	0.75	0.35	0.38	0.75	0.60	0.59	0.50
$C_5H_{11}^+$	0.90	0.84	0.75	0.35	0.48	0.80	0.55	0.60	0.52

[a] Approximate values by which FTIR POA spectra should be scaled to avoid inverted local CH peaks when subtracting from aged OA spectra.
[b] FTIR measurements were discarded for Exp. 6 due to the unusually low OA mass sampled on the filter.

$C_9H_{11}O_3^+$, $C_{10}H_{13}O_3^+$; Tolbert and Ragauskas, 2017; Bertrand et al., 2017; Li et al., 2012) are observed, reflecting different composition of the fuels. We also observe that the relative abundance of lignin- and levoglucosan-related fragments and their FTIR absorbances vary among different experiments with the same fuel probably due to slightly different burning conditions.

We focus the rest of this section on the evolution of primary bbOA with aging. For this purpose, all fragments were corrected for wall losses and normalized by their concentrations at the start of aging. The trends were observed to be similar within each aging scenario regardless of the fuel burned. An experiment representing each category is shown in Fig. 5.1. The increase in the normalized concentration of fragments (usually oxygenated fragments) indicates the appearance of new species, for example, through SOA formation. A decrease in the normalized concentrations suggests that certain species diminish with aging either due to heterogeneous reactions or evaporation. As can be seen from Fig. 5.1a, in a reference experiments (Exp. 38), fragments follow the general trend of the total OA with minor deviations. This observation suggests that the composition of OA does not change noticeably in the absence of oxidants. The same is concluded from the similar FTIR spectral profiles of OA in the beginning and at the end of the reference experiments (Fig. 5.2a). A slight divergence, however, is observed in the time series after around two hours of leaving aerosols in the chamber (Fig. 5.1a). This diverging trend indicates minor differences in the effective loss rates of different compositions probably related to gas-particle partitioning of certain semi-volatile species. The fragment attributed to levoglucosan (and anhydrosugars) fragmentation ($C_2H_4O_2^+$) and light $C_xH_y^+$ fragments (e.g., CH_2^+, $C_2H_2^+$, $C_4H_6^+$) appear to have a higher-than-average loss rates, while the fragment attributed to lignin-related compounds ($C_9H_{11}O_3^+$) appears to diminish less. Similar trends are observed for all WB experiments during the period before the start of aging, implying slight but systematic differences between the loss rates of different OA species.

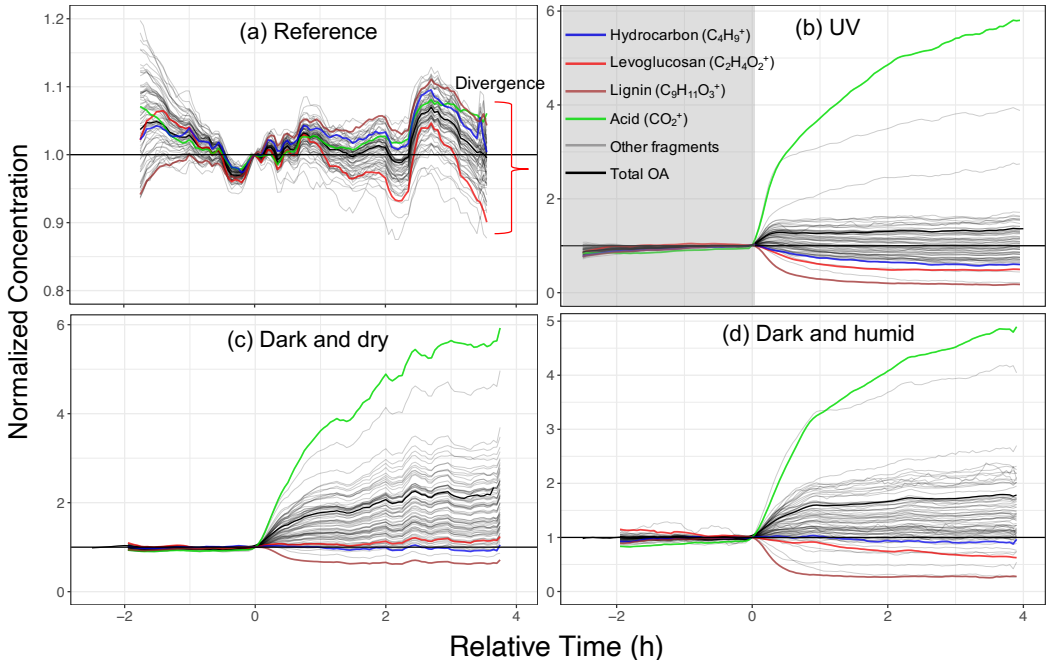

Figure 5.1 – Wall-loss-corrected and normalized concentrations of individual fragments ions in different aging scenarios. Only mass fragments contributing more than 0.25 % to total OA mass concentration are shown. Panels a–d represent Exp. 1, 4,6, and 8, respectively. Time zero indicates the start of aging.

As soon as the aging starts in the UV experiments (e.g., Exp. 4; Fig. 5.1b), several oxygenated fragment ions increase in concentration, among those, CO_2^+ has the most prominent growth. At the same time, the wall-loss-corrected concentrations of several mass fragments decrease by more than 50 % (Fig. 5.1b). Among these fragments, $C_9H_{11}O_3^+$, attributed to lignin-like compounds, and $C_2H_4O_2^+$, related to levoglucosan, decrease the most significantly. The $C_{10}H_{13}O_3^+$ fragment – a larger and less abundant fragment attributed to lignin-like compounds– appears to be more stable with regards to aging (Fig. A.42). The FTIR signatures of levoglucosan and lignin-like compounds also decrease significantly with aging with UV (Table 5.2), causing negative peaks in the residual spectra (Fig. 5.2b). The extent of the diminution for the mentioned species is significantly higher than what can be attributed to wall loss rate uncertainties (standard deviation = 0.016 h^{-1} across all experiments). In addition, the diminution is captured in all UV experiments by both AMS and FTIR, suggesting the insubstantial role of sampling artifacts for FTIR and variable collection efficiency (CE) for AMS. In addition to the mentioned fragments, several non-oxygenated fragments with the $C_xH_y^+$ formula, related to hydrocarbons and aromatics, are also lost with aging (e.g., $C_4H_9^+$, $C_5H_{11}^+$, $C_6H_6^+$, $C_8H_7^+$, and $C_8H_9^+$; Table 5.2). A locally inverted aCH peaks in the FTIR residual spectra of Exp. 4 (Fig. 5.2b) supports the hypothesis that hydrocarbons are lost with aging. The aforementioned observations suggest a significant change in certain POA components during aging with UV lights. There are also fragments (with single oxygen like $C_3H_4O^+$ or related to aromatics like $C_7H_7^+$) that increase in concentration briefly after the start of aging (around 30 min) and fall

below their initial concentration with continued aging. These fragments might indicate the moderately oxygenated intermediate species or those that partition to the particle phase with increased OA loading but their decay becomes more apparent with continued aging.

When emissions are aged with the nitrate radical in dry conditions (e.g., Exp. 6; Fig. 5.1c), the CO_2^+ fragment has the highest increase. Contrary to the UV experiments, the mass fragments related to levoglucosan fragmentation do not diminish more than total OA. The same is inferred from the FTIR spectra (Table 5.2). This implies that levoglucosan is relatively less reactive toward the nitrate radical. Only few fragments decrease more than average, among those, the lignin-related fragments, $C_9H_{11}O_3^+$, and $C_8H_9O_2^+$, decay the most. When emissions are aged with the nitrate radical oxidation in humid conditions (e.g., Exp. 8; Fig. 5.1d), CO_2^+ has the highest increase as in other cases. However, like the UV experiments, several mass fragments decrease more than what can be attributed to wall losses. The majority of the diminishing fragments are hydrocarbons ($C_xH_y^+$) or the fragments related to lignin-like compounds. The same is observed from the FTIR spectra (Table 5.2). We observe that the lignin- and levoglucosan-related fragments decay more prominently in humid compared to dry conditions. However, the decreasing trend for the levoglucosan-related fragment is also observed before the initiation of aging and it is not affected by it. This observation suggests that the nitrate radical reactions are not responsible for the decay of levoglucosan. In this case, other factors such as acid catalyzed levoglucosan reactions in the aqueous phase (Holmes and Petrucci, 2006) or a more efficient removal of gas-phase levoglucosan by chamber walls might play a role.

Our analysis shows that in the reference experiments around only up to 2 % of POA mass is transformed after leaving the emissions in the chamber for around 4 h. This value is considered as a baseline for other aging scenarios. By aging with UV lights, 10–15 % of total POA mass undergoes some form of transformation (oxidation and/or evaporation) (Fig. A.43). Aging with the nitrate radical in dry conditions results in negligible POA transformation, close to what observed in the reference experiments. Aging with the nitrate radical in humid conditions, however, results in a slightly higher POA transformation (approximately 5 %). The values obtained using our method should be considered as a lower bound estimate for the chemical processing of primary bbOA. This is because the increase in the concentration of some oxygenated fragments owing to SOA condensation might outweigh the decay for the same fragments caused by POA oxidation or evaporation. This can especially render the loss of light oxygenated fragments that are not specific to one species undetectable by the method. For instance in a UV experiment (Exp. 4), the most important levoglucosan-marker fragments in the AMS spectra ($C_2H_4O_2^+$, and $C_3H_5O_2^+$), which decrease the most significantly with aging, constitute only around 4 % of POA mass at the start of aging. Unlike these fragments, CHO^+, which is also produced in considerable amounts by levoglucosan fragmentation does not diminish during the course of aging due to the effect of SOA condensation on this fragment. In this experiment, we estimate that levoglucosan constitutes around 20 % of POA mass at the start of aging using FTIR. This value is consistent with previous studies determining the contribution of levoglucosan to primary bbOA (Bertrand et al., 2018a). We also estimate

that the loss of levoglucosan alone during the course of aging is equivalent to more than 10 % of POA mass. The inverted aCH in the FTIR residual spectrum (Fig. 5.2c) accounts for around 4 % of POA mass (10 % aCH mass). The decay of lignin-like compounds, however, was not quantified due to the lack of absorption coefficient and knowledge about their detailed molecular structure.

Gas-particle partitioning, heterogeneous oxidant reactions, and photolysis can play a role in the diminution of the mentioned species in the particle phase. Bertrand et al. (2018b) proposed the prominent role of gas-particle partitioning and the subsequent vapor loss to the chamber walls for the depletion of levoglucosan. Comparing the trends of levoglucosan-related fragments between reference and UV experiments of this work, the reaction of levoglucosan with the hydroxyl radical appears to be the dominant factor for its depletion in the particle phase. For semi-volatile compounds like levoglucosan, particle-phase depletion can be the result of gas-phase depletion and the subsequent gas-particle partitioning. However, the fast diminution of fragments that are produced by less volatile species (e.g., heavy lignin-like compounds) highlights the role of photolysis and heterogeneous reactions. In addition, a similar fast depletion of $C_xH_y^+$ fragments merely due to heterogeneous reactions with the hydroxyl radical was also reported by George et al. (2008).

The FTIR signature (sharp peak at 1515 cm^{-1}) used for identification of lignin-like compounds in this work has also been observed in the FTIR spectra of the HUmic LIke Substance (HULIS) isolated from aqueous extracts of atmospheric aerosols (Graber and Rudich, 2006). As discussed by Yazdani et al. (2021c), this peak can also be produced by small and relatively volatile molecules with a similar aromatic ring substitution to lignin (e.g., methoxyphenols and substituted syringols). The $C_9H_{11}O_3^+$ and $C_{10}H_{13}O_3^+$ fragments in the AMS spectra might not be exclusive to non-volatile lignin and can be produced by smaller, more volatile molecules in bbOA resulting from lignin pyrolysis. There are, however, three observations that suggest the lignin-related fragments in this work can be attributed to compounds with at a lower volatility than levoglucosan that exist predominantly in the particle phase under the conditions of the experiments (e.g., OA loading and temperature). First, the thermodenuder data suggest the lower volatility of compounds producing $C_9H_{11}O_3^+$ and $C_{10}H_{13}O_3^+$ compared to species producing $C_2H_4O_2^+$, mainly levoglucosan (Fig. A.46). These compounds, however, appear to evaporate at higher temperatures and are still categorized as SVOC. Second, the lower loss rate of $C_9H_{11}O_3^+$ in the absence of oxidants compared to other fragments including $C_2H_4O_2^+$ (Fig. A.44) suggests its lower volatility in that its concentration is mainly affected by particle-phase wall losses compared to particle- and gas-phase wall losses for levoglucosan. Third, in two separate WB experiments in dry and humid conditions, ammonium sulfate seeds (approximately 40 $\mu g\, m^{-3}$) were injected into the chamber. The increase in the ammonium sulfate concentration in both dry and humid conditions resulted in an increase in the total OA concentration. This increase, however, was not uniform across all organic fragments, suggesting different partitioning behavior of different species: The levoglucosan-marker fragment ($C_2H_4O_2^+$) showed a significant increase and the lignin-related fragment, $C_9H_{11}O_3^+$ showed no dependence on the injection (Fig. A.45). The latter also suggests that the lignin-like

compounds producing $C_9H_{11}O_3^+$ are predominantly present in the particle phase under the conditions of the experiments as a significant increase in absorbing mass or surface area has no effect on them. These results highlight the role heterogeneous reactions for these compounds.

5.3.2 Characterization of oxidized biomass burning aerosols

In this section, the residual FTIR and AMS spectra are used to characterize the oxidized OA formed during the course of aging, including SOA and aged POA. Fig. 5.3 shows the average residual spectra for each aging scenario, calculated for approximately 3–4 hours after the initiation of aging concurrent with the sampling on the second PTFE filters. As expected, in the absence of UV lights or oxidants, no significant change in the wall-loss corrected AMS and FTIR spectra is observed from the beginning to the end of experiments. Thus, fragment concentrations or absorbances in the residual AMS and FTIR spectra, respectively, are noisy and close-to-zero (Figs. 5.2a and 5.3a), implying negligible emergence of new species.

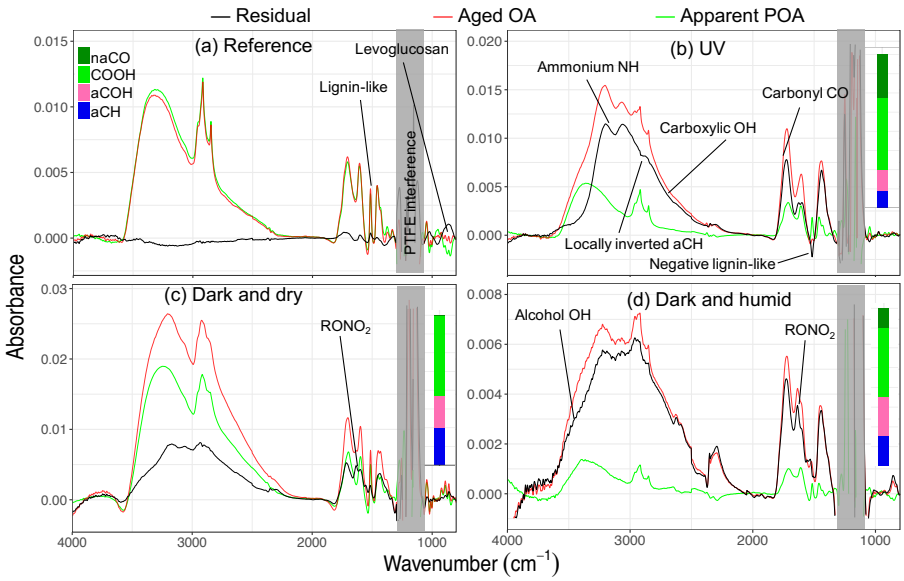

Figure 5.2 – FTIR spectra of aged aerosols (red), their apparent primary fraction (green), and the residual OA (black) for different aging experiments. Panels a–d represent Exp. 1, 4,7, and 8, respectively. Normalized functional group composition from peak fitting is shown for the residual spectra. Ammonium subtracted spectra for panels b and c are shown in Fig. A.47.

When emissions are aged with UV lights, the OA concentration and composition change significantly, allowing us to estimate meaningful residual spectra by subtracting wall loss-corrected POA spectra from those of aged OA. In this procedure, we obtained positive values for the majority of FTIR absorbances and AMS fragment concentrations, suggesting a significant formation of new oxidized species. However, we also calculated negative values for some elements (e.g., levoglucosan, hydrocarbons, and lignin-like compounds) related to the decay

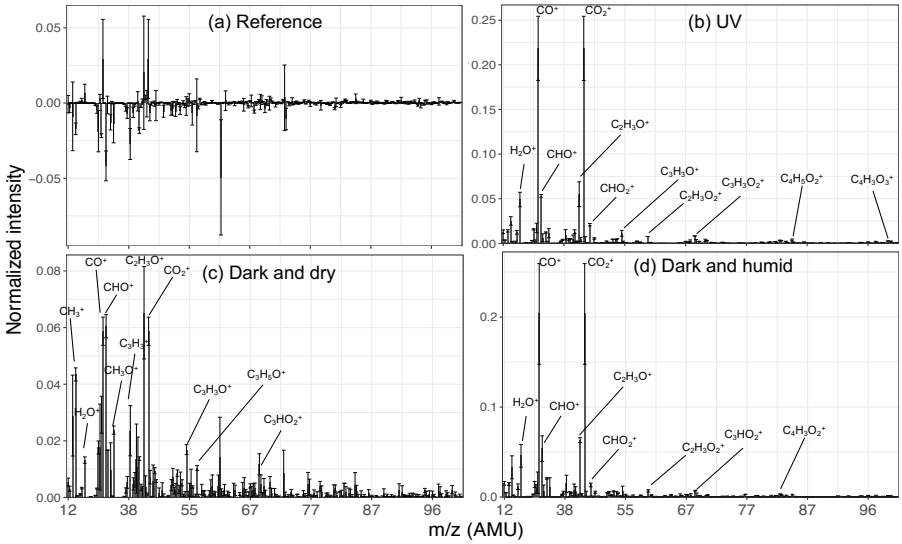

Figure 5.3 – Normalized AMS residual spectra of for biomass burning emissions aged with UV lights and the nitrate radical in dry and humid conditions. Spectra are averaged over experiments of the same aging category and error bars show the range across the experiments.

of fresh POA species. Since POA transformation was discussed in the previous section, we only consider the positive elements of the residual spectra here. As can be seen from Fig. 5.3b, the normalized residual spectrum in the UV experiments is mainly composed of CO_2^+ (and the fragments directly estimated from it: CO^+, H_2O^+, and OH^+), suggesting the abundance of carboxylic acids (Aiken et al., 2007). Other major oxygenated fragments that are prominent in primary bbOA (CHO^+ and $C_2H_3O^+$) appear to be weaker in this spectrum. A few heavier mass fragments with two or more oxygen atoms (e.g., CHO_2^+, $C_2H_3O_2^+$, $C_4H_5O_2^+$, and $C_4H_3O_3^+$) have non-zero values and can also be indicators of acids (Lambe et al., 2012) and polyfunctional organics. Although the concentrations of these fragments are considerably lower compared to that of CO_2^+, they are still statistically greater than the uncertainties across the experiments (Fig. 5.3b). The sharp low-frequency carbonyl peak in the residual FTIR spectrum (1700 cm^{-1}) besides the broad OH peak of dimerized acids (2400–3400 cm^{-1}) are other indicators of carboxylic acids in the residual OA (Fig. 5.2b). The latter, which is partially masked by the ammonium NH stretching peaks, becomes more prominent upon ammonium peaks subtraction (Fig. A.47a). Although the relative contributions of SOA and oxidized POA to the residual spectra are not clear, the formation of carboxylic acids via the oxidation of major wood burning VOCs and the abundance of carboxylic acids in the secondary bbOA have already been reported (Yazdani et al., 2021c; George et al., 2015; Chhabra et al., 2011a).

The difference in the residual spectra of WB and PB when these emissions are aged with the nitrate radical is not substantial and the average is reported here (Fig. 5.3c–d). CHO^+ and $C_2H_3O^+$ are among the most prominent fragments in the residual spectra in dark (NO_3) and dry conditions (Fig. 5.3c). This observation suggests the abundance of non-acid oxygenated species in the SOA formed by the nitrate radical. Some heavier mass fragments with single

oxygen atom (e.g., $C_3H_3O^+$ and $C_3H_5O^+$) are also observed in the residual spectrum and support this hypothesis (Fig. 5.3c). Apart from the oxygenated mass fragments, high abundances of light $C_xH_y^+$ fragments (CH_3^+, and $C_3H_3^+$) are observed in the residual AMS spectrum. In dark (NO_3) and humid conditions, the residual AMS mass spectra are fairly similar to those of UV experiments and they are mainly composed CO_2^+ (Fig. 5.2d). In the residual FTIR spectra of dark aging experiments, higher abundance of the aCOH group relative to UV experiments is observed (Fig. 5.2c–d), suggesting a different FG formation in reactions with the nitrate radical. The formation of organic nitrates is also confirmed in the residual FTIR spectra (Fig. 5.2c–d) as well as by AMS.

The composition of OA changes with aging due to the condensation of SOA onto the existing primary aerosols in addition to the change in the composition of POA and SOA over time due to different aging mechanisms. The latter causes the positive residual AMS spectra (negative elements omitted) to evolve during the course of aging. First, we examine the variation of the residual OA (represented by positive residual spectra) composition during the aging process simply by observing its oxidation trajectory in the f_{44}-f_{43} ($f_{CO_2^+}$-$f_{C_2H_3O^+}$) plot. As can be seen from Fig. 5.4, the trajectory for total OA (the sum of apparent POA and residual OA) spans a much smaller range of the plot and is located in the lower section of the triangle. For the residual OAs formed using UV lights, however, the trajectories start in the lower section of the triangle (Ng et al., 2011a) and continue upward until they end close to the upper vertex. The end points are also close to the location of the OOA factor , suggesting the extent of aging in the chamber is relevant to what observed in the atmosphere. Taking Exp. 4 as an example (Fig. A.48a), we observe that f_{44}-f_{43} ratios for the total and residual OA increase gradually until the end of experiment, while the wall-loss corrected OA mass concentration reaches a constant level about 1 h after the start of aging and the wall loss corrected organic carbon (OC) decreases gradually. Since, there is a net loss of carbon from the particle phase and OA becomes more oxidized without an increase in its mass, other aging mechanisms such as heterogeneous OH reactions and fragmentation appear to be significant. The f_{44}-f_{43} for the residual OA is reminiscent of the extent of POA transformation in this case (Fig. A.48a).

The residual OA from WB and PB emissions aged with the nitrate radical in dry conditions occupy the lower part of the triangle plot with no significant change in f_{44} and f_{43} with aging (Fig. 5.4). In humid conditions with the nitrate radical, the trajectories resemble those of UV experiments except for the elevated f_{43} (Fig. 5.4), suggesting the importance of aqueous nitrate radical reactions for oxidizing bbOA.

PCA extends the comparison to other fragments rather than CO_2^+ and $C_2H_3O^+$. We applied PCA to chamber AMS spectra before the separation of apparent POA and the residual OA. Thereafter, the residual chamber spectra and atmospheric PMF factors from Aiken et al. (2009) (Mexico City) and Finokalia (unpublished data) were projected onto the PC space (Fig. 5.5). The closeness of points in the PC space implies their spectral similarity. The first three principal components (PCs), whose loadings are shown in Fig. A.49, describe 87 % of the total variance in AMS spectra. These PCs have with high loadings only for few fragments, making

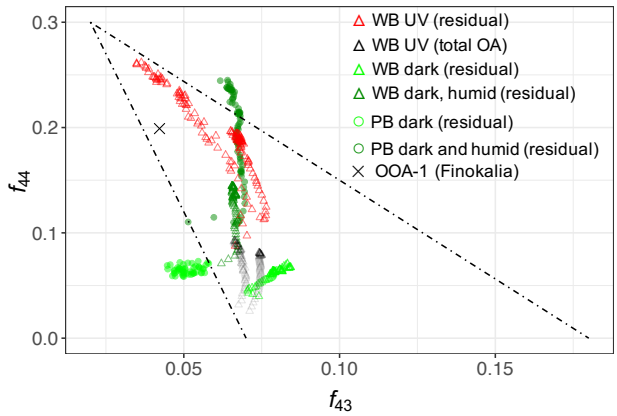

Figure 5.4 – Oxidation trajectories of residual spectra of wood burning (WB) and pellet burning (PB) emissions using UV lights and nitrate radical in dry and humid conditions.

their interpretation straightforward. PC1 mainly indicates the general direction of aging, PC2 indicates the abundance of the CHO^+ fragment (usually considered as a surrogate for alcohols), and PC3 captures the degradation of biomass burning markers (e.g., levoglucosan) with aging. As shown in Fig. 5.5, the atmospheric bbOA factor from Aiken et al. (2009) is located close to the primary wood burning aerosols in the PC space, suggesting their similar composition. The residual OA formed by the nitrate radical in dry conditions is the least aged among all (e.g., higher PC1 loading) located closer to aged chamber OA (mixture of POA and and residual OA). The residual OA aged with the nitrate radical in humid conditions is, however, more aged than that in dry conditions and it is located closer to the atmospheric semi-volatile OA factor (OOA-2). The residual OA in the UV experiment has the most aged spectrum and is located the closest to the atmospheric low-volatility oxygenated OA (OOA-1) in the PC1-PC2 and PC1-PC3 biplots (Fig. 5.5).

5.3.3 Implications for atmospheric aerosols

We found that the AMS spectra of chamber primary and residual bbOA are similar to the atmospheric bbOA and OOA factors, respectively, suggesting the compositional similarity between the two. We also expect that a similar evolution of primary POA due to evaporation and/or oxidation to occur in the atmosphere even at relatively short timescales on the order of a day. The net effect appears to be a more oxidized OA (higher f_{44}:f_{43}) compared to what expected from the mere condensation SOA on the existing POA. This can subsequently affect the hygroscopicity, cloud condensation nuclei (CCN) activity, radiative forcing, and residence time of organic aerosols in the atmosphere (Kanakidou et al., 2005).

The signatures of biomass burning markers, levoglucosan and lignin-like compounds, in spite of being observed in atmospheric aerosols, are among the fastest degrading ones in the FTIR and AMS spectra, making the identification of aged bbOA in the atmosphere challenging. We found that heavier lignin-related fragments (e.g., $C_{10}H_{13}O_3{}^+$) are more stable regarding

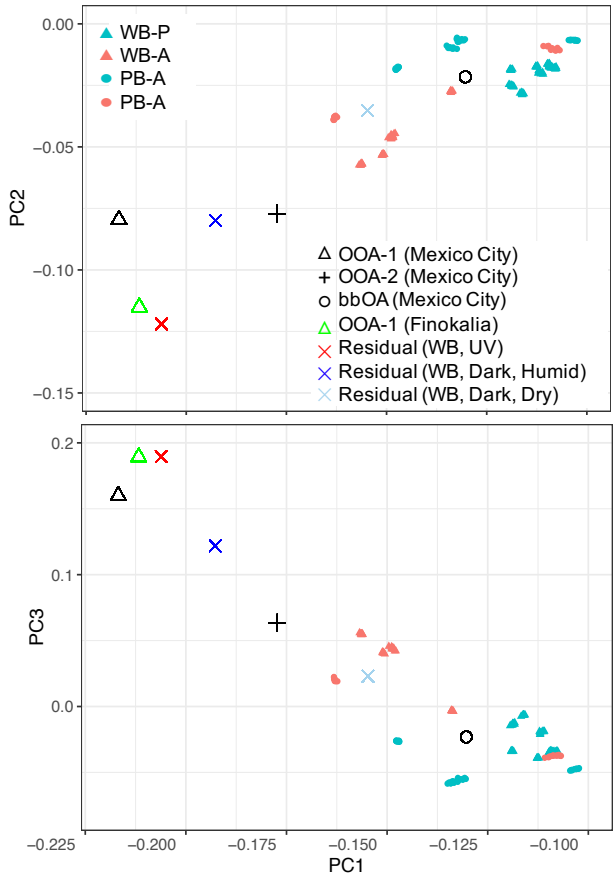

Figure 5.5 – PC1-PC2 and PC1-PC3 biplots of primary and aged chamber OA, residual OA, and ambient PMF factors. Only primary and aged chamber during filter sampling have been shown.

aging than lighter ones and their ratio to lighter, less stable marker fragments (e.g., $C_9H_{11}O_3^+$ and $C_8H_9O_2^+$) can be informative about the extent of aging. Heavy, oxygenated fragments that appear to be specific to secondary bbOA can be an alternative for the identification of aged bbOA. Yazdani et al. (2021b), for example, showed that the $C_7H_5O_4^+$ fragment is highly correlated with acids from FTIR in WB SOA but it is not detected in coal combustion SOA.

Yazdani et al. (2021c) used lignin- and levoglucosan- related mid-infrared signatures to identify smoke-impacted atmospheric $PM_{2.5}$ samples collected on PTFE filters and showed a reasonable agreement of this method with other smoke identification approaches. However, these signatures were not observed or were very weak in the FTIR spectra of certain samples collected in known fire periods (Yazdani et al., 2021e). For these samples, however, a unique spectral profile with very high abundance of carboxylic acids was observed (Fig. A.50) resembling the aged bbOA (with UV) of this work (Fig. A.47) except for the more prominent methylene peaks in the atmospheric samples caused by leaf wax (Hawkins and Russell, 2010). These chamber studies (including Yazdani et al., 2021c) support the emergence of distinct

spectral profiles that can be used to identify aged atmospheric bbOA when fresh markers are not detected in the FTIR spectra.

5.4 Conclusion

In this work, we developed a procedure to quantify the evolution of primary bbOA with aging. We detected some form of transformation (oxidation and/or evaporation) for up to 15 % of the POA mass at short timescales. The estimates of POA transformation with aging, which are more intense during the day-time conditions, are believed to be a lower bound estimate due to the extensive overlap of mass fragments related to primary and secondary bbOA in the AMS spectra.

AMS and FTIR measurements suggest that biomass burning markers such as lignin-like compounds and anhydrosugars in addition to hydrocarbons are among the POA compounds that degrade the most. Unlike hydrocarbons and lignin-like compounds that degrade during both day- and night-time oxidations, anhydrosugars were observed to degrade effectively only during the day-time oxidation. Since the degradation occurs for semi-volatile (levoglucosan) and less volatile (larger lignin-like) compounds, both gas-particle partitioning and heterogeneous reactions are believed to play a role.

We found that the aged fraction of bbOA (including bbSOA) resembles that of the oxidized atmospheric OA especially when aged with UV or in the dark and humid conditions . This fraction becomes more oxidized with continued aging and in certain cases this oxidation happens without a significant increase in the bbOA mass, highlighting the role of heterogeneous reactions.

6 Identification of smoke-impacted $PM_{2.5}$ samples with mid-infrared spectroscopy in a monitoring network

Manuscript in preparation

Authors : Yazdani, Amir; Takahama, Satoshi; Nenes, Athanasios; Sean, Raffuse; Dillner, Ann M.; Sullivan, Amy P.

Contributions: AMD and ST and AY conceived of the project and manuscript. APS performed IC measurements. SR performed satellite smoke detection. AY performed FTIR smoke detection and collected FTIR spectra of laboratory standards. AY performed the FTIR analysis, the data analysis, and wrote the manuscript. ST and AMD edited the manuscript and provided regular feedback on the analysis. APS and SR provided input on the analysis and the further editing of the manuscript. ST provided overall supervision of the project.

Reference: Yazdani et al. (2021e)

Abstract Biomass burning, including residential wood burning, wildfire, and prescribed burning, is an important source of primary and secondary organic aerosols (POA and SOA, respectively) with an increasing effect on climate and air pollution due to soaring wildfire activities around the world. We introduce a fast and non-destructive method for identification of smoke-impacted $PM_{2.5}$ aerosols collected on PTFE filters. This method uses the fingerprint signatures of biomass burning markers (i.e., anhydrosaccharides and lignin-like compounds) in the mid-infrared spectra of PTFE filter samples. We evaluate the method for samples collected at selected sites in the Inter-agency Monitoring of PROtected Visual Environments (IMPROVE) network during wildfire and prescribed burning events in 2013 against satellite observations and find reasonable agreement. The new method allows us to quantify the concentration of biomass burning tracers and to estimate a lower bound for the contribution of biomass burning organic aerosols (bbOA) to total atmospheric fine particulate matter. Our estimates of levoglucosan concentration agree very well with those measured by the high-performance anion exchange chromatography (HPAEC) for more than 300 atmospheric samples ($R^2 = 0.92$). A smoke classifier model is developed using the biomass tracer signatures and the carbonyl group peak in the mid-infrared spectra. This classifier provides results that are comparable to the smoke identification using water-soluble organic carbon (WSOC) and HPAEC anhydrosaccharide measurements. We use the new method, which is one of the few scalable to large air pollution monitoring networks, to study bbOA in around 20,000 filter samples collected in the US IMPROVE network in 2015 and find that 20 % of samples are impacted by wood smoke. In addition, indications of secondary bbOA condensation are observed in several of these samples. It is believed that the bbOA concentration is underestimated on average by 3.6 times using the levoglucosan signatures for these samples due to the marker degradation and SOA condensation.

6.1 Introduction

Natural (wildfire) and anthropogenically-driven (residential biomass combustion, prescribed burnings, agricultural fires) biomass burnings are becoming increasingly important sources of atmospheric organic matter (OM) (Westerling, 2016; Ford et al., 2018; Puxbaum et al., 2007). These aerosols have been the subject of recent studies as they can impact air quality, visibility, and climate directly and indirectly (Smith et al., 2013; IPCC, 2013). The oxidative stress caused by reactive oxygen species (ROS) is believed to be one of the contributors to the negative health effects of PM pollution (Tao et al., 2003). Using the dithiothreitol (DTT) assay, Verma et al. (2015) found that biomass burning organic aerosols (bbOAs) are one of the most important OM components regarding the ROS generation capability (Verma et al., 2015). Biomass burning is a major source of primary and secondary organic aerosols (POA and SOA, respectively) and light absorbing carbon (brown and black carbon) (Puxbaum et al., 2007; Qi et al., 2019; Wong et al., 2019). Large wildfires have been increasing in frequency in recent years due to various reasons including global warming and a similar trend is expected to continue (O'Neill et al., 2021). For example, emissions from wildfires in California are predicted to increase up to 100

% by 2100 above the baseline period (1961–1990) (Hurteau et al., 2014).

In order to understand the effect of biomass burning in the atmosphere, both primary and secondary bbOA should be quantified properly. The quantification of primary bbOA is usually done by relying on specific molecular markers and statistical models (e.g., positive matrix factorization and the chemical mass balance model) (e.g., Srivastava et al., 2021; Sullivan et al., 2008; Zhang et al., 2010; Ramadan et al., 2000; Lee et al., 2008). Water-soluble potassium (K^+) has been used as an stable, inorganic marker to estimate the contribution of primary biomass burning to atmospheric $PM_{2.5}$ (Ramadan et al., 2000; Lee et al., 2008) but has interference from other sources such as dust (Duvall et al., 2008). Anhydromonosaccharides (levoglucosan, mannosan, and galactosan), which are produced from pyrolysis of cellulose and hemicellulose, are more specific to organic biomass burning markers and are identifiable by different analytical methods such as aerosol mass spectrometry (AMS; Schneider et al., 2006), nuclear magnetic resonance (^{1}H-NMR; Paglione et al., 2014), ion chromatography (Sullivan et al., 2008), and Fourier transform infrared spectroscopy (FTIR; Yazdani et al., 2021c). However, the majority of the mentioned methods are destructive to the samples and cannot be scaled to large datasets (e.g., large monitoring networks with thousands of samples) due to high labor costs, expensive equipments, and being time consuming. Another problem with relying on the mentioned markers is that they are only found in primary bbOA, are semi-volatile (Bertrand et al., 2018a), and tend to degrade with aging (Hennigan et al., 2010). As a result, the contribution of primary bbOA is underestimated using the mentioned methods (Hennigan et al., 2010). Environmental chamber experiments suggest up to seven times enhancement of the OM mass concentration after aging with the hydroxyl radical due to SOA condensation at atmospherically relevant OH exposures (e.g., Yazdani et al., 2021c; Bertrand et al., 2017; Tiitta et al., 2016). Recent studies have tried to identify secondary bbOA burning markers but still little is known about these molecules (Nozière et al., 2015; Li et al., 2020a) and thus the contribution of secondary bbOA is often not estimated properly in the atmosphere via measurements. Moreover, simulations of bbOA using chemical transport models (CTMs) have often difficulties reproducing its measured concentrations (Fountoukis et al., 2014; Theodoritsi et al., 2020).

Despite advanced and expensive methods for biomass burning characterization, what can be accomplished to understand the impact of biomass burning over a wide spatial and temporal scale (permitted by existing monitoring network infrastructure) is still very limited. FTIR on PTFE filters is a fast, nondestructive, and cost effective method for quantification of the majority of functional groups (FGs) in OM and is shown to be able to characterize around 70–80 % OM mass (Boris et al., 2019; Russell et al., 2009a; Liu et al., 2012a). Specific mid-infrared spectral profiles have also been reported for bbOA in the atmosphere and environmental chambers (Russell et al., 2011; Takahama et al., 2011; Bürki et al., 2020; Yazdani et al., 2021c). In addition, Yazdani et al. (2021c) showed recently that anhydrosaccharides and lignin-like compounds that are ubiquitous in bbOA have rather unique signatures in the mid-infrared spectra. These characteristics make FTIR one of the only methods suitable for the analysis of large number of samples collected in networks such as the US Inter-agency Monitoring of PROtected Visual Environments (IMPROVE) network.

In this work, the identification and quantification of atmospheric bbOA by FTIR is evaluated against remote sensing and high-performance anion exchange chromatography (HPAEC). A smoke classifier is developed by combining FTIR and HPAEC measurements. Thereafter, the new bbOA identification and quantification method is applied to around 20000 samples collected in the IMPROVE network in 2015 to better understand the impact of biomass burning on OM concentrations at the scale of this monitoring network. Finally, indications of biomass burning SOA condensation in several samples of this dataset is demonstrated and discussed.

6.2 Methods

The sampling procedure for the filter samples selected from the IMPROVE network are discussed in Sect. 6.2.1. The basis of smoke identification using remote sensing for the periods and locations of the filter samples are discussed in Sect. 6.2.2. Quantification of biomass burning markers using ion chromatography is explained in Sect. 6.2.3. FTIR spectral analysis and smoke detection method is discussed in Sect. 6.2.4. Section 6.2.5 describes the development of smoke classifier using FTIR spectra.

6.2.1 Sampling $PM_{2.5}$ aerosols in the IMPROVE network

25 mm PTFE filters from Pall Corporation were used to collect particulate matter with an aerodynamic diameter smaller than 2.5 µm every third day at a flow rate of 22.8 L min^{-1} from midnight to midnight at selected sites of the Interagency Monitoring of PROtected Visual Environments (IMPROVE) network (http://vista.cira.colostate.edu/improve/) in years 2008, 2011, 2013, and 2015 (Fig. 6.1). From samples collected in 2013, 122 filters were selected in the period of forest fire (Rim Fire, from June until October) at Yosemite National Park, Hoover (northeast of Yosemite NP), and Bliss State Park (Lake Tahoe) monitoring sites and 87 filters were selected in the period of prescribed burns (from January until April) in the southeastern US in 2013 at Cape Romain National Wildlife Refuge (South Carolina), Saint Marks (Florida), and Okefenokee National Wildlife Refuge (Georgia) monitoring sites. In total, 303 PTFE samples were selected from the IMPROVE network from 2011 and 2008 in spring, summer, and fall, for which potential influence of fire was anticipated. In 2015, the first year for which FTIR spectra are available for the entire network, more than 20000 filters from 162 monitoring sites across the US were analyzed.

In parallel to PTFE filters, quartz-fiber filters (25 mm diameter, Pallflex Tissuquartz, Pall Corporation, Port Washington, NY) were collected and analyzed using a thermal optical reflectance (TOR) instrument for organic and elemental carbon (OC and EC) using the IMPROVE A protocol (Chow et al., 2007). TOR-OC concentrations have been used to support the results where needed.

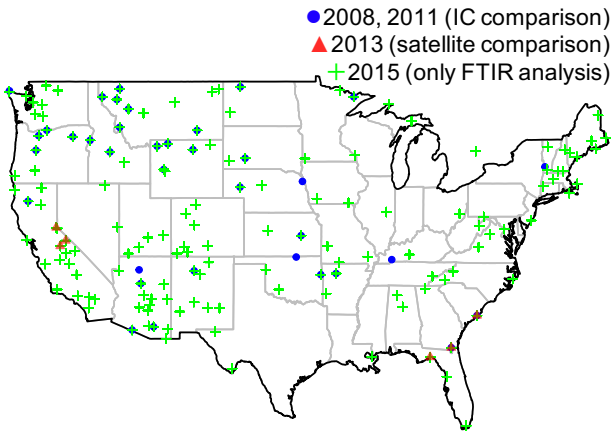

Figure 6.1 – The IMPROVE network sites where $PM_{2.5}$ samples of this study were collected.

6.2.2 Fire and smoke plume detection using remote sensing

We characterized sampling days as “probably" or “maybe" fire impacted using a qualitative weigh-of-evidence approach. Fire detects and smoke plume extents from the National Oceanic and Atmospheric Administration’s Hazard Mapping System (NOAA HMS) served as the primary sources of fire information. The NOAA HMS fire detect dataset compiles thermal anomaly (“hot spot”) data from a suite of polar-orbiting and geostationary satellites along with human quality control to minimize false detections, such as from persistent heat sources like metal smelters (McNamara et al., 2004). NOAA HMS smoke plume extents are digitized outlines of smoke as identified in visible satellite imagery by trained analysts (Brey et al., 2018). For each sampling event (site and day), four proximity scores were calculated: number of fire detects within 10–100 km of the site, the minimum distance to nearest fire detected, density of nearby hot spots using a kernel density function, distance from the smoke plume extent to site (highest score if the site is within a plume). Sampling events with low scores across all metrics were categorized as “no evidence of smoke impact". Those with high scores in any individual metric or high average scores were examined individually using one or more additional analyses, including assessment of visible imagery, back-trajectory analysis, review of nearby $PM_{2.5}$ monitors, and wildland fire smoke dispersion modeling using BlueSky Playground (Larkin et al., 2010). Using this weight-of-evidence approach and expert judgment, sampling events were characterized as “maybe” or “probably” smoke-impacted. This analysis was conducted independently and prior to the FTIR-based characterization for the corresponding periods and sampling locations of the 228 PTFE filter samples collected in 2013.

6.2.3 Smoke detection using HPAEC-PAD and WSOC

The method, which is similar to that described by Sullivan et al. (2008, 2011a,b), was used to analyze 303 PTFE filters collected in 2008 and 2011. Briefly, after obtaining the FTIR spectra, the whole filter was extracted in 15 mL of deionized water by sonication. A subsequent

filtration step with a PTFE syringe filter was performed on the extract to discard the non-water-soluble fraction of aerosols. The extracts were, thereafter, analyzed, in the same day, for levoglucosan (1,6-anhydro-β-D-glucopyranose) and its stereoisomers, mannosan (1,6-anhydro-β-D-mannopyranose) and galactosan (1,6-anhydro-β-D-galactopyranose), using high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD). A Dionex DX-500 series ion chromatograph equipped with a Dionex GP-50 pump and Dionex ED-50 electrochemical detector (amperometric mode, waveform A) was used. Separation was done with Dionex CarboPac PA-10 guard (4 × 50 mm) and analytical (4 × 250 mm) columns and each run took 59 min to complete. Each run included an isocratic elution for 10 min at 10 mM NaOH to detect anhydrosugars, 19 min of linear gradient from 10 ro 70 mM for the detection of sugars (e.g., galactose, mannose, and glucose) followed by the cleaning of the column at 180 mM for 19 min, and finally the reequilibration step for 16 min. Although this method is able to separate carbohydrates and anhydrosugars (important chemical markers of wood smoke), it has limitations regarding the separation of sugar alcohols (e.g., mannitol, and arabitol; detected in fungal spores; Bauer et al., 2008) and anhydrosugars. Based on the sampling time (24 h) and flow rate (22.8 $\mathrm{L\,min^{-1}}$) of this study the limit of detection (LOD) for the carbohydrates is estimated to be less than 0.1 $\mathrm{ng\,m^{-3}}$ and the detector response was found to be linear in the range concerning this study. The agreement of this method with GC-MS measurements for levoglucosan from PTFE filters have already been shown (Sullivan et al., 2011a).

A Sievers Turbo total organic carbon (TOC) Analyzer (Model 800) operating in Turbo mode was used to measure water-soluble organic carbon (WSOC). In this instrument, organic carbon (OC) is completely oxidized by the hydroxyl radical produced from the photolysis of water and persulfate and is turned in to CO_2. CO_2 is then measured using membrane based conductometric CO_2 sensors. The LOD was estimated to be approximately 0.06 $\mathrm{\mu g\,C\,m^{-3}}$.

Sullivan et al. (2006) found that bbOA and SOA are two main sources of WSOC. Thus, a substantial contribution of levoglucosan (as a bbOA marker) to WSOC indicates also a substantial contribution of fresh bbOA. However, the threshold above which a sample is considered smoke-impacted is rather subjective. Since, we were interested in identifying samples even with low contributions of bbOA and levoglucosan tends to degrade rapidly in the atmosphere (Hennigan et al., 2010), we chose a rather strict threshold: samples with the contribution of levoglucosan to WSOC over 1 % (levoglucosan C/WSOC > 0.01) were considered to be “smoke-impacted" and samples with the contribution below 1 % were considered uninfluenced by wood smoke. The levoglucosan C/WSOC > 0.01 threshold is roughly equivalent to fresh bbOA / OM > 0.07 (Appendix Sect. A.5.1). It should be noted that this criterion only considers the contribution of primary bbOA to WSOC via levoglucosan and might not be optimal when there is substantial degradation of bbOA markers or substantial formation of secondary bbOA.

6.2.4 FTIR analysis, post-processing, and smoke detection

FTIR spectra in the 400–4000 cm^{-1} range were obtained from PTFE filters with a Bruker-Tensor 27 FTIR instrument equipped with a liquid nitrogen-cooled, wide-band mercury-cadmium-telluride (MCT) detector at a resolution of 4 cm^{-1}. PTFE filters were kept in petri dishes (without refrigeration) before the analysis. FTIR spectra were baseline-corrected and blank-subtracted to eliminated the effect of light scattering, absorption by carbonaceous material and minimize the interference of PTFE absorbances with organic signatures. We used a smoothing spline to estimate the baseline for each spectrum (Kuzmiakova et al., 2016).

The fingerprint absorbances of anhydrosaccharides (three peaks in the 860–1050 cm^{-1} range) and lignin-like molecules (molecules with similar ring substitution to lignin such as monolignols and lignin pyrolysis products; single sharp peak at 1515 cm^{-1}) were used for identification of bbOA (Yazdani et al., 2021c). The concentration of anhydrosugars was quantified using the mentioned fingerprint absorbances and the absorption coefficient calculated from aerosolized levoglucosan standards. Aerosolized samples were generated following the approach of Ruthenburg et al. (2014) and calibration curves were produced using the weight of deposited aerosols and the height of levoglucosan peaks. The height of each peak was calculated from the absorbance at the peak location minus the average of absorbances at the location of the left and right shoulders. The absorption coefficient of lignin-like compounds was not estimated since these molecules refer to a class of substituted aromatics with various molecular structures.

The contributions of anhydrosaccharides and lignin-like compounds to bbOA depend on the burning conditions, the extent of exposure to oxidants, the fuel type, combustion efficiency, and OM concentrations (Fine et al., 2002; Robinson et al., 2006b; Bertrand et al., 2018a; Yazdani et al., 2021d). We considered measurements from Fine et al. (2002) as reference in that regard since they analyze bbOA from a variety of fuel types in atmospherically-relevant conditions.

6.2.5 Development of a smoke classifier

In order to develop a classifier that only needs FTIR spectra and is equivalent to that using WSOC and HPAEC-PAD levoglucosan, WSOC should be replaced by a relevant parameter obtained from the FTIR spectra. We used the carbonyl peak as a proxy to WSOC (Appendix A.5.2). Peak height of the FTIR signatures related to anhydrosaccharides (P_{LG1}, P_{LG2}, and P_{LG3}), lignin-like compounds (P_{Lig}), and the carbonyl group (P_{Carb}) in the FTIR spectra were used to develop a binary smoke classifier that agrees the best with the smoke identification using WSOC and HPAEC-PAD levoglucosan measurements. For this purpose, we considered five parameters: These parameters are the weights for the three levoglucosan peaks (a, b, c) that might interfere with different organic and inorganic peaks, the weights of the lignin-like compounds peak (d) and an offset value (e) taking into account inaccuracies in the peak height estimation caused by an imperfect baseline correction (Eq. 6.1). The carbonyl peak (P_{Carb}) was used as a proxy to WSOC in order to build a classifier that only requires FTIR

spectra to function.

$$A = \frac{aP_{\mathrm{LG1}} + bP_{\mathrm{LG2}} + cP_{\mathrm{LG2}} + dP_{\mathrm{Lig}} + e}{P_{\mathrm{Carb}}} > 0.01. \tag{6.1}$$

The dataset (303 samples from 2008 and 2011) was divided into a training set (202 samples) and a test set (101 samples). A floating point genetic algorithm (Mitchell, 1996) was used to find parameters that maximize the F_1 score, TP/(FP + 0.5(FP + FN)), on the test set. This parameter, which is a measure of accuracy of the classifier, is the harmonic mean of precision (the fraction of selected observations that are relevant, TP/(TP + FP)) and recall (the fraction of relevant observations that are selected, TP/(TP + FN)). TP (true positive) is the number of samples that FTIR classifies as smoke-impacted ($A > 0.01$) in agreement with the HPAEC-PAD criterion (considered as reference). FP (false positives) and FN (false negatives) are defined similarly by taking the HPAEC-PAD criterion as the reference. The F_1 score is the preferred performance measure for imbalanced datasets (Jeni et al., 2013) like our dataset with 25 % probability that a sample is smoke-impacted.

6.3 Results and discussions

In the following subsections, first the FTIR limit of detection and selectivity of bbOA marker signatures are evaluated in Sect. 6.3.1. FTIR smoke detection is compared to remote sensing detection of fires and smoke plumes in Sect. 6.3.2. FTIR quantification of bbOA markers and classification of samples based on the contribution fresh bbOA markers are compared to that from HPAEC in Sect. 6.3.3. In Sect. 6.3.4, we analyze the impact of biomass burning on the samples collected in the IMPROVE network across the US in 2015.

6.3.1 Selectivity and LOD of FTIR signatures of bbOA marker

Anhydrosaccharides

In this work, three peaks observed at 860, 890, and 920 cm^{-1} in the FTIR spectra (Fig. A.53; referred to as P_{LG1}, P_{LG2}, P_{LG3}, respectively) have been used for the identification of levoglucosan in atmospheric aerosols. The peaks in this range are believed to be related to the symmetric C–C–O vibrations in alcohols (Smith, 1998) and appear to have a unique pattern in anhydrosaccharides. As can be seen from the FTIR spectra of levoglucosan, mannosan, and galactosan in Fig. 5 of Kamińska et al. (2018), the mentioned three peaks appear to have similar heights in the FTIR spectrum of levoglucosan in ordered crystalline and plastic phases. Similar relative peak heights are observed in the aerosolized samples of levoglucosan in this work (Fig. A.53). The 890 cm^{-1} appears be the dominant peak for mannosan, while the 920 cm^{-1} peak appears to be slightly stronger for galactosan (Kamińska et al., 2018). Although

other factors such a the phase state of the molecules (e.g., liquid, ordered crystal, and plastic crystal) can affect the peak height ratios, the ratios can still be informative about the relative abundance of these three compounds. We considered an average of the three peaks to estimate the abundance of anhydrosaccharides. The mentioned peaks were not observed in the spectra of other alcohols such as straight-chain alcohols (e.g., docosanol; Yazdani et al., 2021c), saccharides (e.g., glucose, mannose, and galactose; Fig. A.54), and sugar alcohols (e.g., mannitol; Fig. A.54). In the complex atmospheric samples, emergence of peaks related to organics or inorganics in the 800–1000 cm^{-1} region can hinder the quantification of anhydrosaccharides. As an example, the bisulfate peak at around 870 cm^{-1} (Krost and McClenny, 1994) and the organonitrates peak at around 850 cm^{-1} (Yazdani et al., 2021c) might interfere with those of anhydrosaccharides.

The signal noise at the location of levoglucosan peaks was analyzed for 55 blank 25 mm PTFE filters after baseline correction and blank subtraction. Considering the sampling time and flow rates of this work, the LODs based on the $3\sigma_{\text{blank}}/b$, where b is the calibration slope (Fig. 6.2a), were estimated to be 7, 16, 27, and 12 $ng\,m^{-3}$ for P_{LG1}, P_{LG2}, P_{LG3}, and $P_{\text{LG, ave}}$ (average of the three levoglucosan peaks) as shown in Fig. 6.2b. The FTIR LODs are around two orders of magnitude higher than that of HPAEC-PAD for anhydrosaccharides but with a significant speed advantage (5 min for FTIR analysis compared to 100 min for HPAEC-PAD analysis) and without the need for any sample preparation.The LOD based on the standard error of estimate of the calibration curve, $3S_{yx}/b$, where $S_{yx} = \sqrt{\frac{\hat{y}_i - y_i}{n-2}}$, is 167 $ng\,m^{-3}$ (Fig. 6.2a). However, the latter is believed to be an overestimation of the LOD due to uncertainties associated with the weighing of the PTFE filter samples and the higher signal noise at higher concentrations (heteroscedacity). Based on the measurements of Fine et al. (2002) for OM from combustion of six different wood types, levoglucosan constitutes on average 8 % of OM mass (assumed to be 1.8 times OC; Malm and Hand, 2007) under atmospherically relevant dilutions and temperatures. This contribution is lower than that measured by Yazdani et al. (2021c) (15–29 %) in the environmental chamber experiments but is believed to represent atmospheric conditions more closely due to using atmospherically relevant dilutions and a variety of wood species found in the US. Using this ratio and the LOD of levoglucosan (based on $P_{\text{LG, ave}}$; 12 $ng\,m^{-3}$) the FTIR LOD for fresh bbOA was calculated to be approximately 0.15 $\mu g\,m^{-3}$.

Lignin pyrolysis products

A sharp 1515 cm^{-1} peak is observed in the FTIR spectra of aromatic compounds with specific ring substitutions (Yazdani et al., 2021c) due to the stretching vibrations of aromatic C=C (Feldman, 1985). These compounds include lignin, monolignols (conyferyl, coumaryl, and sinapyl alcohols) or molecules resulting from the pyrolysis of lignin such as substituted syringols and guaiacols (Bertrand et al., 2018a; Fine et al., 2001) and are referred to as lignin-like compounds here. The 1515 cm^{-1} peak has already been observed in the FTIR spectra of the water-soluble extracts of atmospheric samples likely affected by biomass burning (Duarte

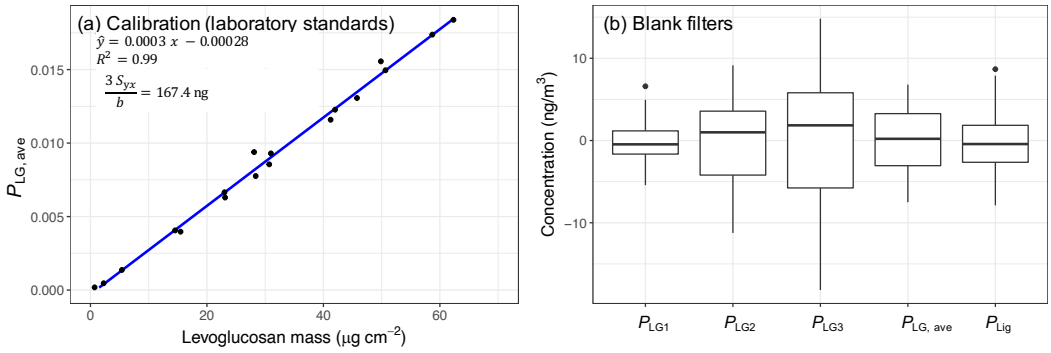

Figure 6.2 – The calibration curve used to derive the absorption coefficient of levoglucosan (a) and the box and whisker plot showing the variability of absorbances (converted to concentrations) at the location of the three levoglucosan peaks in blank filters after baseline correction and blank subtraction (the absorption coefficient of lignin-like compounds is assumed to be the same as that of levoglucosan in this plot).

et al., 2005). However, this peak is believed to have lower selectivity compared to the levoglucosan signatures at lower frequencies (fingerprint region). Although the methoxy group ($O-CH_3$) is observed in the structure of coumaryl and sinapyl alcohols, its absorption below 2850 cm^{-1} (Degen, 1968) is not distinguishable from normal aliphatic CH absorbances neither in the fresh wood burning spectra of Yazdani et al. (2021c) nor in atmospheric samples of this study. The 1515 cm^{-1} is prone to interference with PTFE absorbances, vibrorotational bands from water vapor in the FTIR chamber, and interference from the carbonyl peak (only in the case of high OM loading).

The FTIR LOD for fresh bbOA based on this peak in the chamber experiments of Yazdani et al. (2021c) was calculated to be approximately 16 ng m^{-3}. However, the relatively high OM concentrations in the chamber experiments of Yazdani et al. (2021c) (on average 30 µg m^{-3}) favors the partitioning of bbOA markers into the particle phase and thus tend to overestimate their contribution to fresh bbOA.

Hennigan et al. (2010) estimated the lifetime of levoglucosan in the atmosphere to be 0.7–2.2 days, assuming an average OH concentration of 10^6 molec. cm^{-1}. Given this lifetime, it is estimated that 5 µg m^{-3} of fresh bbOA becomes undetectable by FTIR (i.e., the levoglucosan concentration falls below the LOD of FTIR) after 2.5–7.7 days of exposure to typical summertime OH levels (without considering dilution effects). In addition, a significant underprediction of aged bbOA concentrations is also anticipated using this marker. Yazdani et al. (2021c) found that the FTIR signatures of lignin-like compounds diminish at an even faster rate than levoglucosan in the presence of OH. Unlike levoglucosan, these compounds also degrade substantially in the presence of the NO_3 radical. As a result, their lifetime is expected to be shorter especially when the precursors of the nitrate radical, NO_2 and O_3, are available during the nighttime.

6.3.2 Comparison of FTIR and satellite smoke detection

For the comparison between FTIR and satellite smoke detection methods, satellite measurements were taken as reference and confusion matrices were generated. Samples categorized either as “probably" or “maybe" impacted by smoke were all considered as smoke-impacted by satellite. Samples with either levoglucosan or lignin signatures three times above their LOD in the FTIR spectra were categorized as smoke-impacted by FTIR after the visual inspection of the corresponding peaks to avoid any misidentification due to interference from other compounds.

As can be seen from Table 6.1a, among 87 samples for the period of prescribed burns from January until April 2013 in the southeastern US, 20 samples are classified as smoke-impacted by both FTIR and satellite. The biomass burning signatures are visible in these samples with 24 h average OM concentrations (assumed to be 1.8 times TOR OC) ranging from 1 to 22.6 $\mu g\,m^{-3}$. The FTIR spectra of these samples usually have very prominent aliphatic CH and carboxylic acid signatures (Figs. 6.3 and A.55). For 51 samples, neither FTIR nor satellite observations detect any evidence of smoke. For 5 samples, only FTIR detects signatures of smoke, and for 10 samples only satellite confirms the existence fire or smoke. The 10 samples detected only by FTIR to be smoke-impacted have also spectral profiles (Fig A.55b) similar to those of smoke samples shown in previous studies (e.g., Russell et al., 2011; Bürki et al., 2020), supporting the hypothesis that they are smoke-impacted. The spectra of these samples usually have strong absorbances of organics, especially in the aliphatic CH (2800–3000 cm^{-1}) and carboxylic acid OH (broad band, 2400–3400 cm^{-1}) regions. Carboxylic acids are believed to be abundant in secondary bbOA (Yazdani et al., 2021c; Li et al., 2020a). The lack of detection by satellite for these samples may be attributed to the cloud coverage, or the low resolution of satellite images (approximately 1 km^2). For the 5 samples categorized only by satellite to be smoke-impacted (Fig A.55c), FTIR spectra are dominated by inorganics and the absorbances of organic bands are relatively weak. This observation might be due to the significant dilution and aging of smoke in the atmosphere that renders OM concentrations low and the marker signatures below their FTIR LODs. The low OM loading and lack of bbOA marker signatures can be also related to the direction of wind or high smoke plumes that do not considerably affect the samples collected at the surface level. The latter is, however, not expected for prescribed burns due to a lower intensity of these controlled burns. Nonetheless, the number of samples with conflicting FTIR and satellite smoke detections (i.e., false positives and negatives) is considerably lower (17 % of all observations) than those in agreement. In general, samples categorized as “probably" smoke-impacted by satellite have higher OM concentrations (from TOR OC) than those categorized as “maybe" smoke-impacted ($7.5 \pm 3.7\,\mu g\,m^{-3}$ against 2.9 ± 2.4 $\mu g\,m^{-3}$). A higher percentage of the former was detected to be smoke-impacted by FTIR (75% against 63 %).

As can be seen from Table 6.1b, from 122 samples collected in the Rim Fire period, both the FTIR and satellite methods confirm the impact of smoke for 31 samples. The OM concentration from TOR OC ranges from 2.0 to 54.0 $\mu g\,m^{-3}$ for these samples and the aliphatic CH peaks

Table 6.1 – Confusion matrices of FTIR and satellite smoke detection methods for prescribed burns (a) and RIM fire (b) in 2013. P and M indicate probably and maybe smoke-impacted, respectively.

(a) Prescribed burns 2013

Satellite	FTIR: Smoke	FTIR: No smoke
Smoke/fire	21 (9P, 12M)	5 (3P, 7M)
No smoke/fire	10	51

(b) Rim Fire 2013

Satellite	FTIR: Smoke	FTIR: No smoke
Smoke/fire	31 (14P, 17M)	26 (6P, 20M)
No smoke/fire	4	61

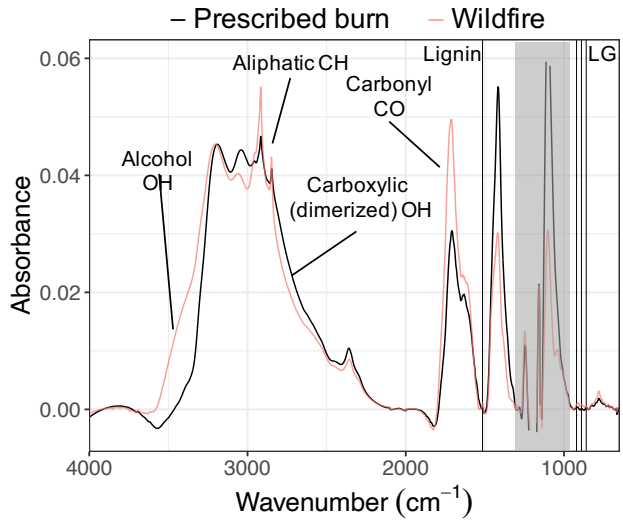

Figure 6.3 – Mean spectra of samples classified as smoke-impacted by both satellite observations and FTIR for the prescribed burn and wildfire datasets. The gray band shows the PTFE absorption region.Vertical lines show the signatures of levoglucosan (LG) and lignin-like compounds.

are prominent in their FTIR spectra (Figs. 6.3 and A.55d). The spectral profiles of 5 of these samples with prominent alcohol absorptions (Fig. A.55d) are reminiscent of those of fresh wood burning aerosols from the chamber experiments of Yazdani et al. (2021c). For 61 samples, FTIR and satellite detect no evidence of smoke. For 4 samples, only FTIR detects signatures of smoke. These samples have spectral profiles with prominent aliphatic CH peaks (Fig A.55e), which is one of the characteristics of smoke-impacted samples (Hawkins and Russell, 2010). For 26 samples, only the satellite method confirms the existence smoke. The number of misidentified samples in Rim Fire dataset is more than that for prescribed burns (21 % of all observations). The 24 h average of TOR OC for the misidentified samples ranges from 1 to 5.4 $\mu g\,m^{-3}$. Prominent aliphatic CH and carboxylic OH peaks are observed in the FTIR spectra of 10 and 2 of these samples, respectively (Fig A.55e). Although the latter suggests the impact of wood smoke, lignin and levoglucosan peaks are not identified for these samples either due to

the degradation of markers or interfering peaks. We believe that relatively high OM (especially acid) concentrations for the two mentioned samples are due to secondary bbOA formation (Yazdani et al., 2021c). In the wildfire dataset, two of the levoglucosan peaks (P_{LG2} and P_{LG3}) have significant interference with two peaks located at 875 and 910 cm^{-1}. These peaks, which have variable ratios, are believed to be related to inorganics because they are also observed in samples with OM concentrations as low as 1 $\mu g\,m^{-3}$ and high inorganic loadings (e.g., strong absorption of ammonium). It was necessary to subtract the interfering peaks for some of the samples to identify the levoglucosan signatures (e.g., Fig. A.56). For the samples with false negative error that have low organic loadings, the height of the smoke plume might cause the discrepancy between FTIR (ground-level) and satellite detections (ground-level and high altitude). The latter is more likely for wildfires as the intensity of fire and the smoke plum height are higher compared to prescribed burns. However, in general, smoke-impacted samples identified by satellite in the wildfire period have higher OM concentrations at the ground level ($7.3 \pm 9.1\ \mu g\,m^{-3}$, from TOR OC) than those collected in the prescribed burns period ($4.5 \pm 3.5\ \mu g\,m^{-3}$). A considerably higher percentage of samples categorized as “probably" smoke-impacted compared to those categorized as “maybe" smoke-impacted by satellite are detected to be smoke-impacted by FTIR (70 % against 46 %) in the wildfire dataset. The former category has also a considerably higher concentrations of OM ($13.1 \pm 13.1\ \mu g\,m^{-3}$ against $4.2 \pm 3.2\ \mu g\,m^{-3}$). The mean spectrum of smoke-impacted samples collected in the wildfire period has relatively stronger peaks related to organics and bbOA markers but lacks the broad acid peak compared to that of prescribed burns (Fig. 6.3). Both atmospheric processes and different fuel might cause the compositional difference in OM between these two categories.

6.3.3 Comparison of FTIR and HPAEC-PAD smoke marker measurements

In this section, a more quantitative approach has been taken by comparing smoke marker measurements from FTIR to those from HPAEC-PAD for 303 PTFE filters samples collected at selected sites of the IMPROVE network in 2011 and 2013. In order to compare FTIR levoglucosan measurements to those of levoglucosan, mannosan and galactosan from HPAEC-PAD, we chose samples for which concentrations of all three anhydrosaccharides were above their LOD in HPAEC-PAD. In total, 90 samples with this criterion were found in the dataset. The FTIR-based levoglucosan concentrations (from $P_{LG,\,ave}$) were found to be highly correlated with the concentrations of levoglucosan and the sum of anhydrosaccharides from HPAEC-PAD (R^2 = 0.93 and 0.88, respectively; Fig. 6.4). Although R^2 values were relatively high for the entire dataset, they decreased to 0.48 and 0.45, respectively, for a subset of samples with HPAEC-PAD levoglucosan concentrations below 0.5 $\mu g\,m^{-3}$, showing weaker agreement in low-concentration samples. This issue is believed to be mainly due to the interference with the 875 and 910 cm^{-1} peaks in samples with high loadings of inorganics and low loadings of levoglucosan. We also observed a higher R^2 for HPAEC-PAD levoglucosan against P_{LG1} (0.47) compared to P_{LG2} (0.39) and P_{LG3} (0.22) for the low-levoglucosan samples (below 0.5 $\mu g\,m^{-3}$), suggesting lower interference from other compounds for P_{LG3}. This result is while P_{LG2} and P_{LG3} have higher R^2 (0.94 and 0.93, respectively) compared to P_{LG1} (0.89) when the whole

dataset, including high-levoglucosan samples, is used.

In order to investigate the ability of FTIR to detect low concentrations of levoglucosan in atmospheric samples, the FTIR spectra of several filter samples with HPAEC-PAD levoglucosan concentrations below 10 ng m^{-3} were visually inspected. Among these spectra, for those the interfering peaks did not overwhelm the levoglucosan fingerprint region, two of the levoglucosan peaks were observed (P_{LG1} and P_{LG2}; Fig. A.57a). For samples with significant interference from inorganics, P_{LG1} was usually visible on the right shoulder of the interfering peak (Fig. A.57b) but since the other two levoglucosan peaks were not detected, the certainty of levoglucosan detection was lower. This analysis suggests that the main limiting factor for identification and quantification of levoglucosan using FTIR in atmospheric samples is the interference with inorganics and not the lower LOD compared to HPAEC-PAD. Peak fitting or equivalent statistical methods are required to properly quantify overlapping peaks in these cases. We also used partial least squares regression (PLSR) to predict HPAEC-PAD levoglucosan in low-levoglucosan samples using the levoglucosan fingerprint region in the FTIR spectra, which let a significant improvement in the correlation coefficient for these samples ($R^2 = 0.91$ compared to $R^2 = 0.48$ when using $P_{LG,\,ave}$), while maintaining reasonable fit quality for high-concentration samples (Appendix A.5.6).

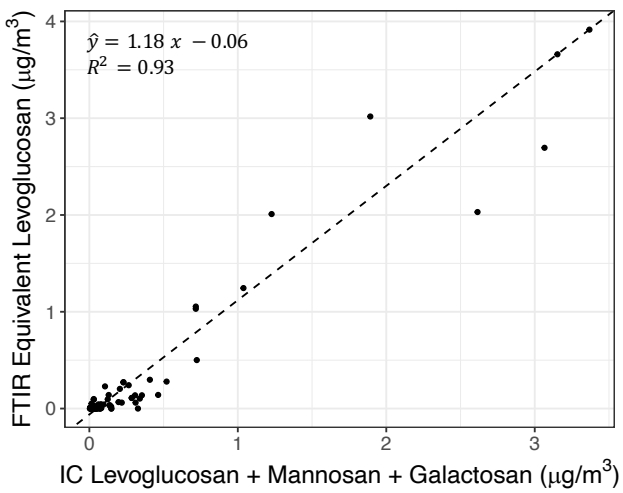

Figure 6.4 – Scatter plot of FTIR equivalent levoglucosan concentrations (from $P_{LG,\,ave}$) versus the sum of anhydrosaccharides measured by HPAEC-PAD.

The higher R^2 for FTIR-based levoglucosan concentration (from $P_{LG,\,ave}$) against the HPAEC-PAD sum of anhydrosaccharides compared to HPAEC-PAD levoglucosan suggests that $P_{LG,\,ave}$ is probably more representative of the sum of anhydrosaccharides rather than only levoglucosan. The slope of the calibration line was measured to be 1.18 (Fig. 6.4) for the sum of anhydrosaccharides and 1.50 when only considering HPAEC-PAD levoglucosan. Mannosan and Galactosan concentrations were measured to be 16.3 ± 10.9 % and 4.4 ± 4.0 % of that of levoglucosan from HPAEC-PAD for this dataset. The relatively higher concentrations of

anhydrosaccharides measured by FTIR (i.e, greater-than-one slopes of the calibration lines) might be related to the lower-than-unity extraction efficiency of anhydrosaccharides from PTFE filters before the HPAEC-PAD analysis. The slope for the sum of anhydrosaccharides (1.18) appears to be more consistent with the reported extraction efficiencies of levoglucosan from PTFE filters (Nozière et al., 2015).

Although FTIR absorbances related to lignin-like compounds and anhydrosaccharides are correlated for the majority of samples ($R^2 = 0.7$ for all samples), there are samples for which the ratio of lignin-like to levoglucosan absorbances deviate significantly from the average trend (Figs. 6.5 and A.57) probably due to variations in the fuel type and combustion conditions, different extent of oxidation of the markers in the atmosphere, and different interfering compounds. This deviation is especially prominent in samples with low lignin- and levoglucosan-related absorbances (lower than 0.01 and 0.002 respectively). As a result, FTIR signatures related to lignin-like compounds can add additional information about the contribution of bbOA. This feature can help compensate for the interference from inorganic compounds and the higher LOD of FTIR for anhydrosaccharides compared to HPAEC-PAD.

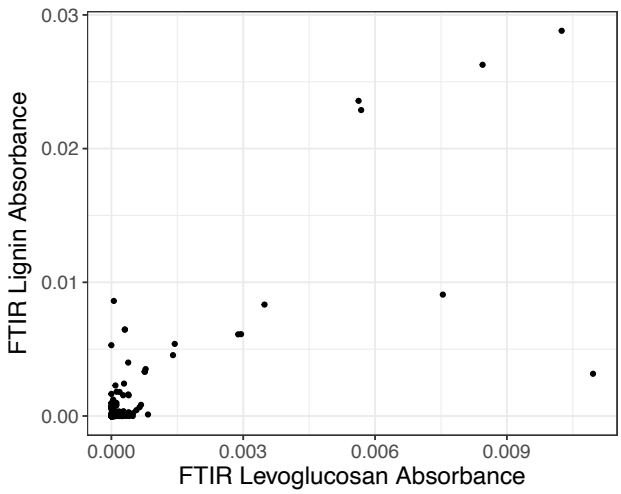

Figure 6.5 – Scatter plot of FTIR absorbances related to levoglucosan ($P_{\mathrm{LG,\,ave}}$) and lignin-like compounds (P_{Lig}).

FTIR smoke classifier

A smoke classifier was developed by optimizing the parameters in Eq. (6.1) using the genetic algorithm. The optimized parameters a–e for Eq. (6.1) are 1.9, 0.66, 0.14 , 0.74, and -9.3 $\times 10^{-5}$ respectively. Considering the fact that the peaks related to levoglucosan have often similar heights in the laboratory standards (Fig. A.53), these values suggest the relative importance of P_{LG1}, P_{LG2}, P_{Lig} for the smoke classifier. A slightly negative value for e probably takes into account an approximately constant error caused by baseline correction artifacts or interference from other compounds. P_{LG3} does not appear to be as informative as other levoglucosan peaks

probably because it is often suppressed by the interfering peak at 910 cm^{-1}. The classifier built using Eq. (6.1) performs reasonably well on the training set (F_1 = 0.90; Fig. A.52a) with 44 and 148 true positiver and negative results. There are eight false negatives, which are located below the HPAEC LG / WSOC < 0.025 threshold. Although the classifier fails to classify these samples as smoke-impacted, lignin and levoglucosan signatures are visible in the spectra of six of these samples. WSOC for false negative samples ranges from 0.28 to 1.30 $\mu g\,m^{-3}$. There are two false positives with HPAEC LG / WSOC > 0.005, which have clear lignin-like signatures in their FTIR spectra. WSOC is measured to be 1.20 and 1.95 $\mu g\,m^{-3}$ for these samples. The performance of the classifier degrades slightly on the test set (F_1 = 0.82; Fig. A.52b). The five false negatives in the test set have HPAEC LG / WSOC < 0.017 and WSOC ranges from 0.70 to 3.94 $\mu g\,m^{-3}$ for them. Lignin or levoglucosan signatures are visible for four of these samples. There are three false positives in the test set with HPAEC LG / WSOC > 0.005, for which WSOC ranges from 1.11 to 5.25 $\mu g\,m^{-3}$. For two of these samples the lignin signature is visible. Even in samples in the false negative error category, bbOA signatures are observed in the FTIR spectra but are not quantified correctly according to the threshold. In addition, the majority of false positives have prominent signatures of lignin-like compounds, which are not quantified by HPAEC, and thus are believed to be true smoke-impacted samples. This analysis shows that the FTIR classifier performs reasonably well. No significant change in the classifier performance was observed (i.e., similar F_1 score) by changing the threshold to 0.005 or 0.02.

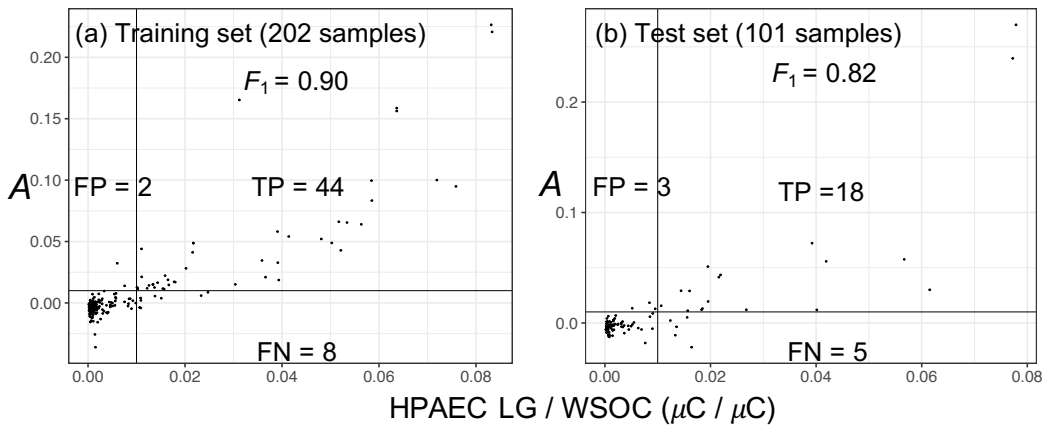

Figure 6.6 – Scatter plot comparing the smoke classifier that uses FTIR smoke signatures and the carbonyl peak to that using WSOC and HPAEC levoglucosan measurements. A is defined in Eq. (6.1). Vertical and horizontal solid black lines indicate the threshold (0.01) of smoke classification by each method.

Using the carbonyl peak as a proxy for WSOC does not lose any accuracy in classification (Appendix A.5.7), which suggests that FTIR might be used independently to detect smoke-impacted aerosols with comparable results to the detection with HPAEC and WSOC instruments combined. We also developed a classifier that uses levoglucosan concentration from a PLSR model instead of P_{LG1}, P_{LG2}, and P_{LG3} (Appendix A.5.7). PLSR is able to alleviate the interference from other (inorganic) peaks and improve the levoglucosan concentration

estimates for low-loading samples. The PLSR classifier performs slightly better over that reported here on both the training and test sets (F_1 = 0.97 and 0.85 for the training and test sets, respectively). The better performance is likely because the A_{PLS} parameter is better correlated to HPAEC LG / WSOC at lower values (close to the 0.01 threshold; Fig. A.61) than A (Fig. 6.6). In addition to working as a classifier, the value of A (or A_{PLS}) is informative about the extent of bbOA contribution to WSOC. The way this smoke classifier classifies samples based their FTIR levoglucosan and lignin signatures is visualized in Fig. A.62 for different thresholds.

6.3.4 Analysis of IMPROVE samples from 2015 using the FTIR smoke detection method

FTIR analysis on PTFE filters is one of the few methods capable of being scaled to an air quality monitoring network as extended as the IMPROVE network. We analyzed around 20000 filters collected in 2015 at 162 sites of the IMPROVE network across the US. For these filters, levoglucosan concentrations were estimated using FTIR ($P_{\mathrm{LG,\,ave}}$). The concentration of primary bbOA was estimated based on the contribution of levoglucosan to the fresh bbOA mass (8 % based on the data from Fine et al., 2002). Thereafter, the smoke classifier developed in the previous section was used to determine smoke-impacted samples.

Figure 6.7 shows monthly maximum concentrations of primary bbOA in the US estimated at the IMPROVE network sites and represented by contour lines. It should be noted that the spatial density of IMPROVE monitoring sites might not be optimal for a continuous representation everywhere across the US. However, the spatial density is relatively higher for the western and northwestern regions of the US where the majority of wildfire activities are observed and multiple sites are impacted. As can be seen from Fig. 6.7, the identification of smoke using FTIR on PTFE filters reveals the clear seasonal variation of bbOA concentrations and several smoke hot-spots in different months of 2015. The wildfire season in this year was reported to be the most destructive in the last ten years regarding the land area burned (National Interagency Fire Center, 2021). For August 2015, very high concentrations of bbOA (monthly maximum of 42 $\mu g\,m^{-3}$ and monthly average of 11 $\mu g\,m^{-3}$) are observed in the northwestern US (Fig. 6.7). The average contribution of fresh bbOA to OM for the filters collected in the northwestern US (Washington, Oregon, Idaho, Montana, and Wyoming) in this month is estimated to be approximately 12 % with its maximum contribution reaching 80 %. This observation is concurrent with one of the largest wildfires in the history of Washington State that burned more than 400,000 hectares (National Centers for Environmental Information, 2021). Our analysis suggests that daily mean concentrations of primary bbOA frequently rose above 10 $\mu g\,m^{-3}$ in monitoring sites located in California from July to December (e.g., in Fresno, Kaiser, Yosemite NP, Lava Beds NM, Lassen Volcanic NP, and Trinity). This observation is also consistent with the wildfire data from CALFIRE (2021), showing that wildfires in certain areas of California were not contained until late December. In addition to the wildfire season, high concentrations of bbOA are observed in several monitoring sites in colder months. For examples, the monitoring site in Phoenix, AZ, where residential wood burning for heating is

common (Pope et al., 2017), experienced daily mean bbOA concentrations up to 11 $\mu g\,m^{-3}$. The average wintertime contribution of primary bbOA to OM in this site is estimated to be approximately 15 % with its maximum contribution reaching 70 %, consistent with estimates of (Ramadan et al., 2000). The 1515 cm^{-1} peak related to lignin-like compounds also provides information about the distribution of bbOA that is consistent with that from levoglucosan with only minor differences (Fig. A.64).

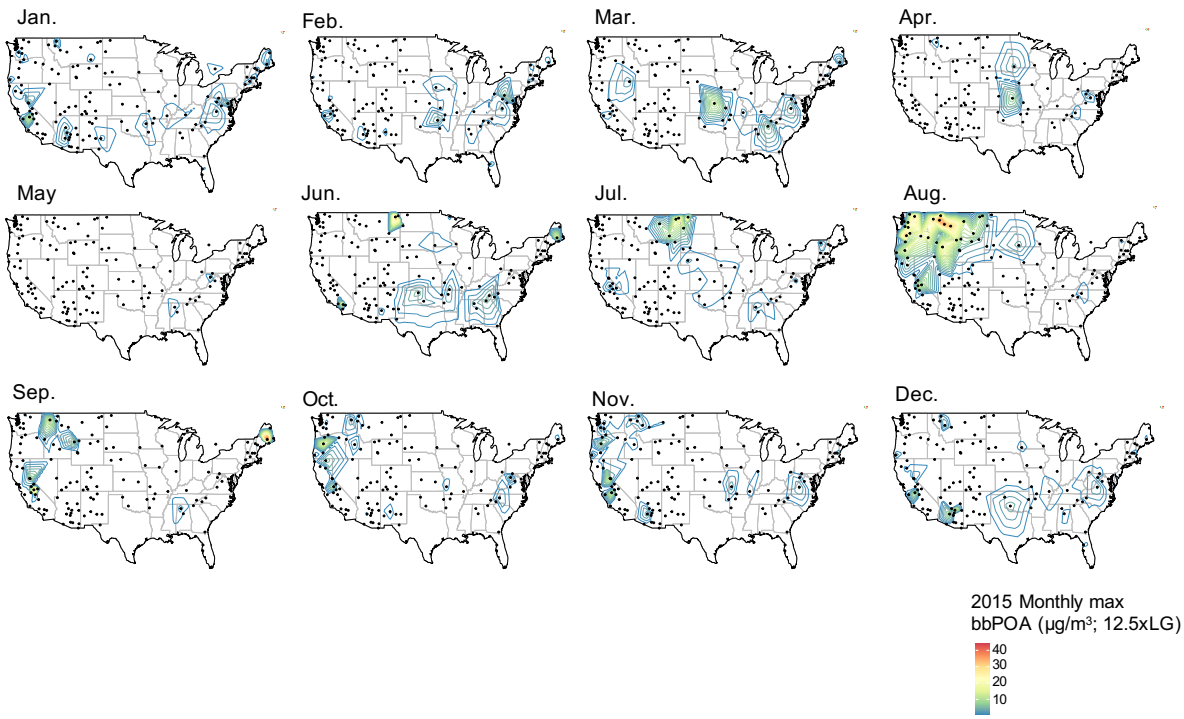

Figure 6.7 – Maximum monthly concentrations of primary bbOA in the US in 2015 estimated from FTIR levoglucosan signatures. These values represent lower bound estimates of primary bbOA concentrations assuming levoglucosan constitute 8 % of fresh bbOA mass.

The contribution of primary bbOA to OM and $PM_{2.5}$ across the US in 2015 is estimated to be on average 5 % and 3 %, respectively. The annual mean of levoglucosan concentration form all sites is estimated to be approximately 19 $ng\,m^{-3}$ in this year. This value is close to the measurements of Robinson et al. (2006b) for Pittsburgh from July 2001 to June 2002. The annual mean of levoglucosan for the southeastern US (sites located in Alabama, Georgia, and South Carolina) is estimated to be 28 $ng\,m^{-3}$ for 2015, which is considerably lower than that measured by Zhang et al. (2010) for this region in 2007 (107 $ng\,m^{-3}$). Different locations of monitoring sites and different periods of analysis between the two studies are believed to be the main reasons for this discrepancy.

The results of the smoke classifier developed based on the FTIR spectra are illustrated in Fig. 6.8 for this dataset. The fraction of smoke-impacted samples is the lowest in May and June but considerably increases in cold months and also during the wildfire season. In total, around 20

% of samples from the IMPROVE network are classified as smoke-impacted for the year 2015. This number approaches 100 % in certain sites located in the northwestern US during the wildfire season. Except for the northwestern US, which is impacted by wildfires, the fraction of smoke-impacted samples is the highest in winter months (Fig. A.65). The smoke-impacted samples appear to have similar distribution of OM concentration to normal samples but their OM / $PM_{2.5}$ is generally higher than other sample (Fig. A.66).

Figure 6.8 – Fraction of smoke impacted-samples in each month in 2015 determined by the FTIR smoke classifier.

6.3.5 Indications of aged bbOA

Atmospheric samples with relatively high OM concentrations (e.g., > 25 $\mu g\,m^{-3}$) and high contributions of levoglucosan to OM in the confirmed periods of biomass burning (e.g., wildfire) are believed to be the least aged smoke-dominated atmospheric samples. We compared the spectra of samples with these characteristics to that of fresh bbOA from chamber experiments of Yazdani et al. (2021c) with similar OM concentrations. This comparison suggests considerably weaker marker signatures but similar or more aged spectral profiles (e.g., prominent carbonyl peak) for the atmospheric samples (Fig. A.67). The lower contribution of primary bbOA markers even in the least aged smoke-dominated atmospheric samples compared to the chamber experiments is likely because these molecules tend to degrade and evaporate with dilution and chemical aging in the atmosphere (Hennigan et al., 2010; Yazdani et al., 2021d). As a results, the bbOA contributions predicted based on the levoglucosan signatures should be

considered as lower bound estimates especially when marker degradation and SOA formation are anticipated (e.g., for monitoring sites distant from the smoke source in summertime).

As an example, TOR measurements at the IMPROVE sites show high OM concentrations (estimated as 1.8 times OC; up to 35 $\mu g\,m^{-3}$) all over the northwestern US on August 19 (Fig. A.68a). This is concurrent with the widespread wildfire activities reported in that region and is also supported by the high levoglucosan concentrations for the samples collected at monitoring sites in the region (Fig. A.68b; and also elevated potassium and EC concentrations). These observations suggest that biomass burning is the main source of OM in those areas. However, the OM concentrations are considerably higher than those of primary bbOA estimated from FTIR levoglucosan (Fig. A.68b;). In addition, the high OM persists over a larger area compared to FTIR-based levoglucosan measurements. The latter can be justified by the chemical aging of biomass burning emissions and dilution that reduces the concentrations of fresh bbOA markers. Figure A.68c, shows the FTIR spectra of two samples collected in the Pasayten Wilderness and Columbia River Gorge sites on that date. While both samples have comparable OM concentrations (around 25 $\mu g\,m^{-3}$) and visible biomass burning markers in their FTIR spectra, the sample collected in the Columbia River Gorge site has substantially weaker signatures of levoglucosan and lignin. This sample is also substantially more oxidized with prominent peaks related to carboxylic acids, a major constituent of secondary bbOA (Yazdani et al., 2021c), suggesting the prominent contribution of secondary bbOA. If we assume that the OM in high-OM samples (e.g., > 20 $\mu g\,m^{-3}$) collected in confirmed wildfire locations and periods (in the northwestern US in August) can be completely attributed to biomass burning due to lack of any other dominant source, bbOA (bbPOA and bbSOA) concentrations are 3.6 ± 2.6 times more than those estimated using levoglucosan The latter is believed to be due a combined effect of SOA condensation and degradation of levoglucosan.. This value is also reminiscent of the OM concentration enhancement in the chamber experiments due to aging (Yazdani et al., 2021c; Bertrand et al., 2017; Tiitta et al., 2016).

We performed cluster analysis on smoke-impacted samples identified by the smoke classifier using k-means method (MacQueen, 1967). In this analysis, samples were clustered based on similarity in their 2-norm normalized FTIR spectra in the 1300–4000 cm^{-1} region. As can be seen from Fig. 6.9, biomass burning samples appear to be characterized not by a single distinct profile but several recurring profiles that vary in inorganic fraction, but also organic functional group contributions as observed through peak ratios. This analysis supports the existence of variations found in BB factor profile from FTIR positive matrix factorization (PMF) modeling in past studies (Hawkins and Russell, 2010; Russell et al., 2011; Takahama et al., 2011; Bürki et al., 2020). Cluster 2 and 4 show smoke dominated samples due to having dominant organic peaks. Cluster 2 has prominent acid and carbonyl peaks and weak signatures of fresh biomass markers. Even cluster 4, which appears to be more similar to the fresh bbOA spectra of environmental chamber experiments (e.g., Yazdani et al., 2021c) has considerably stronger carbonyl peaks compared fresh bbOA in the chamber experiments (e.g., Fig. A.68). The latter suggests that smoke-impacted aerosols in the atmosphere are relatively more oxidized.

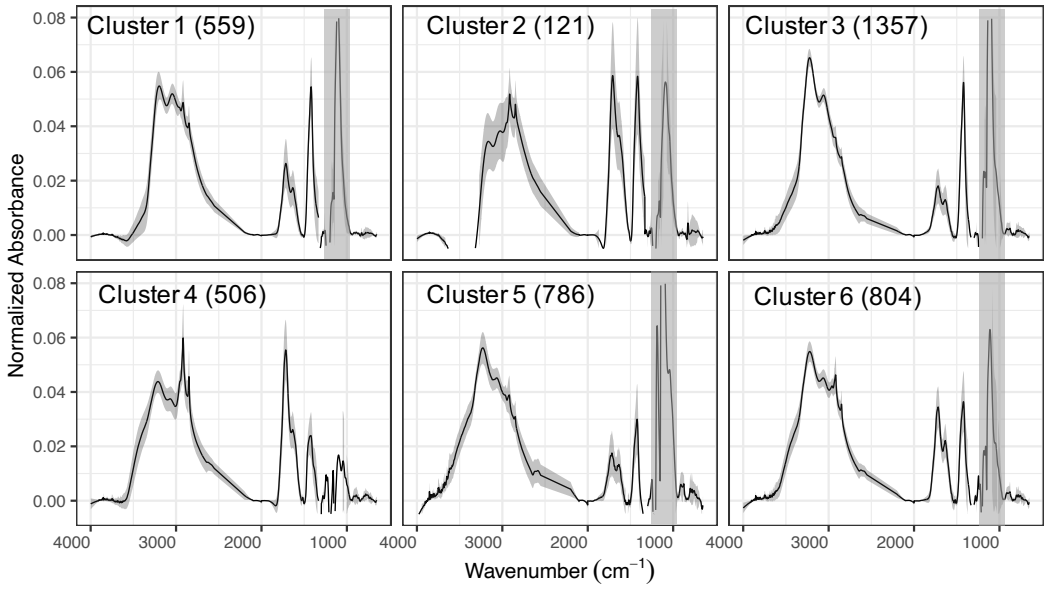

Figure 6.9 – The averaged normalized spectrum of six clusters of smoke-impacted samples. Gray band show the mean ± one standard deviation. The shaded vertical rectangles show the PTFE interference region. Number of samples in each cluster is indicated.

6.4 Concluding remarks

We found that FTIR is able to identify and quantify the fresh bbOA markers on PTFE filters with a LOD of a few ng m^{-3}. The good agreement of FTIR smoke detection using bbOA markers was found against remote sensing for wildfires and prescribed burning events in the US in 2013. Levoglucosan concentrations estimated by FTIR were highly correlated ($R^2 = 0.93$) to those from a more sensitive analytical method, HPAEC-PAD, for a series of PTFE filter samples collected in the IMPROVE network. A smoke classifier was developed that used only molecular markers and carbonyl absorbances in the FTIR spectra and performed comparably (F_1 score = 0.90) to the smoke identification with a combination of HPAEC-PAD anhydrosaccharides and WSOC measurements.

The new FTIR smoke detection method was used to study bbOA in the US at the scale of the IMPROVE network in 2015 and showed good agreement with the recorded wildfires and known residential wood burning events. We found that for around 20 % of samples collected across the US in that year, levoglucosan carbon constituted more than 1 % of WSOC. Samples affected by biomass burning were found to be characterized by several recurring FTIR profiles that varied in organic and inorganic compositions, consistent with different smoke FTIR PMF profiles identified in previous studies.

Atmospheric smoke-impacted samples were found to have weaker marker signatures compared to fresh bbOA of previous chambers studies (e.g., Yazdani et al., 2021c). Prominent carbonyl and carboxylic acid peaks were observed in atmospheric samples anticipated to be

aged bbOA. We found the concentration of atmospheric bbOA to be, on average, 3.6 times of that measured using levoglucosan concentrations due to the SOA condensation and the degradation of levoglucosan.

7 Quantified Functional Group Compositions in Household Fuel Burn Emissions using FTIR

Manuscript in preparation

Authors : Li, Emily Y.; Yazdani, Amir; Takahama, Satoshi; Champion, Wyatt; Shen, Guofeng; Jetter, James J.; Preston, William T.; Hays, Michael D.; Dillner, Ann M.

Contributions: WC wrote up the WBT method and clarified all the test phase differences. GS provided the artifact-correct OC/EC calculations. JJJ collected the filter samples for GC-MS and OC/EC analyses. MDH and WTP performed GC-MS analysis and provided PAH data. AMD collaborated on design of the filter holder for the transmission-mode FTIR. EYL collected the FTIR measurements on the cookstove burn emission filters and conceived the project. AY performed the FTIR and data analysis and the comparison with other techniques. EL and AY wrote the manuscript. ST and MTH provided regular supervision and feedback for the manuscript.

Reference: Li et al. (2021)

Abstract Globally, billions of people burn fuels indoors for cooking and heating, which contributes to millions of premature deaths and chronic illnesses annually. Additionally, residential burning contributes significantly to black carbon emissions, which are estimated to have the highest global warming impact second to carbon dioxide. In this study, we use Fourier transform infrared spectroscopy (FTIR) to analyze $PM_{2.5}$ emissions collected on Teflon membrane filters from fifteen cookstove types and five fuel types. We compare the FTIR measurements to the thermal optical method EC-OC, gravimetric $PM_{2.5}$, and polycyclic aromatic hydrocarbons (PAHs) measurements from GC-MS. Only emissions from three fuel types (charcoal, kerosene, and red oak wood) were above the minimum detection limit for the functional group (FG) analysis using FTIR and OC-EC analysis using the thermal optical transmittance (TOT) method. We find distinct spectral profiles for particulate emissions. Kerosene emissions are dominated by aromatics and show the most distinct profiles compared to the unburned fuel in the FTIR spectra. Charcoal particulate emissions are enriched in hydrocarbons likely from the volatile fraction of charcoal and red oak emissions possess the most similar spectra to the unburned fuel. We highlight the FGs associated the most with organic carbon (OC) using a multivariate statistical method and show that organic carbon (OC) estimates from collocated FTIR and TOT are highly correlated. We use the out-of-plane aromatic CH vibrational bands in the FTIR spectra for the first time to quantify the aromatic fraction organic matter (OM) and find that the abundance of aromatic compounds from FTIR is highly correlated with the sum of PAHs from GC-MS measurements. The baseline in the FTIR spectra of particulate emissions is also found to be informative about the abundance of elemental carbon (EC). Since FTIR analysis is fast and non-destructive and provides complementary functional group information, it can significantly reduce the need for thermal-optical measurements.

7.1 Introduction

Residential burning is a major source of organic carbon (OC), and contributes approximately 30% of global emissions of black carbon, which is estimated to have the second highest global warming impact after carbon dioxide (Ramanathan et al., 2008; Bond et al., 2013). The World Health Organization (WHO) estimates that 3–4 million premature deaths per year are associated with exposure to household air pollution, mainly from solid-fuel burning (WHO, 2014). Recent studies have typically categorized fine particulate matter ($PM_{2.5}$) emissions from household fuel burn emissions by mass, organic carbon (OC), and elemental carbon (EC) emission factors using gravimetric and thermal-optical methods (e.g., IMPROVE and NIOSH protocols) (Roden et al., 2006; Sharma and Jain, 2019; Jetter et al., 2012). However, thermal-optical EC/OC measurements take more than 45 minutes per sample and are completely destructive, while providing limited information (OC and EC alone are only broad categories of carbonaceous particles).

In contrast to thermal-optical methods, Fourier transform infrared spectroscopy (FTIR) measurements take few minutes per sample, are non-destructive and highly reproducible, and provide more cheminformatics. FTIR measurements have been shown to be in good agree-

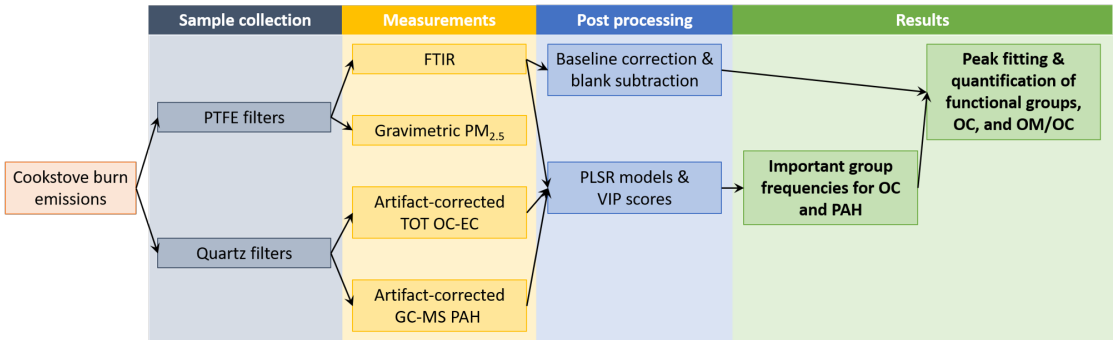

Figure 7.1 – Flow chart from sample collection to measurements, post-processing, and results.

ment with common methods for chemical characterization of organic matter (OM) such as aerosols mass spectrometry (AMS) and and thermal-optical analysis (Takahama et al., 2013; Dillner and Takahama, 2015; Reggente et al., 2016; Takahama et al., 2016; Yazdani et al., 2021a,c; Boris et al., 2019). However, FTIR has not been extensively used to characterize primary particulate emissions from residential burning.

In the current study, we used FTIR to quantify the organic functional group (FG) composition of $PM_{2.5}$ emissions from cookstove collected on polytetrafluoroethylene (PTFE) filters. We found that OC estimates from FTIR measurements are in good agreement with those from thermal optical transmittance (TOT). Our analysis highlighted spectral differences between unburned fuels (charcoal, kerosene, red oak wood, alcohol, and liquefied petroleum gas) and their particulate emissions during combustion. FTIR measurements also suggested high abundances of aromatics and polycyclic aromatic hydrocarbons (PAHs) in particulate emissions from charcoal, kerosene, and red oak wood, in agreement with gas chromatography-mass spectrometry (GS-MS) measurements.

7.2 Materials & Methods

Details of cookstove testing, emissions sampling, OC-EC measurements, GC-MS PAH measurements, FTIR measurements, and post-processing are illustrated via Figure 7.1's flow chart and described or cited to previous publications in subsections below.

In short, a range of stoves and fuels were tested for burn emissions, from which parallel PTFE and quartz fiber (Qf) air membrane filter samples were collected. PTFE filters were used to measure gravimetric $PM_{2.5}$ mass, and then scanned using transmission-mode FTIR. The quartz filters were used to measure TOT OC-EC as well as GC-MS PAHs. From these sets of measurements, partial least squares regression (PLSR) models were built for the identification of important FTIR group frequencies for OC and PAH, and peak fitting and quantification of FGs, OC, and OM/OC were performed.

Figure 7.2 – Photo showing 15 stoves used in this work. Charcoal stoves: (A) EcoZoom Jet, (B) Prakti Leo, (C) CH4400, (D) Jiko Koa (inset), and (E) Éclair (inset). Kerosene stoves: (F) Model 2668 / Butterfly Wick and (G) Model 2412 / Butterfly Pressure. Red oak wood stoves: (H) Jiko Poa, (I) Envirofit G3300, (J) Philips, (K) Home Stove (Biolite), (L) EcoChula XXL (inset), and (M) 3-Stone Fire (inset). Alcohol stove: (N) CleanCook (inset). LPG stove: (O) Solgas.

7.2.1 Cookstove emissions testing & sampling

Cookstove emission tests were conducted at the U.S. EPA cookstove test facility in Research Triangle Park, NC. Details of the facility and protocols are described in previous publications (Jetter and Kariher, 2009; Jetter et al., 2012; Shen et al., 2017a,b; Xie et al., 2018). We analyzed 152 samples from 15 different stoves and five fuel types, pictured in Figure 7.2. Fuels tested covered charcoal (56 samples), kerosene (21 samples), red oak wood (69 samples), alcohol (3 samples), and liquefied petroleum gas (LPG, 3 samples).

In this study, particulate emissions from stoves during the water boiling test (WBT) were sampled. The WBT protocol (version 4, Global Alliance for Clean Cookstoves, 2014) includes three test phases: (1) a high-power cold-start (CS) phase in which the stove, pot, and water are at ambient temperature; (2) a high-power hot-start (HS) in which the stove is hot; and (3) a low-power simmer phase (SIM) in which the water temperature is kept 3 °C below the boiling point (Fig. A.69) (Jetter et al., 2012). A modified protocol described in Jetter et al. (2012) was used for the charcoal stove.

$PM_{2.5}$ was sampled isokinetically on Qf (47-mm diameter Tissuquartz™ pure quartz no binder, Pall Corporation) filters and PTFE membrane filters (47-mm diameter Teflo® membrane, Pall Corporation) positioned in parallel and downstream of $PM_{2.5}$ cyclones (URG; Chapel Hill, NC) at a flow rate of 16.7 liters per minute. Aerosols were also sampled downstream of the PTFE filter using a Qf back filter for artifact correction. The filter-based PM sampling was conducted in a primary dilution tunnel for low-emission fuel-cookstove combinations (e.g.,

forced-draft biomass stoves), and in a secondary dilution tunnel for those with high emissions (e.g., natural-draft biomass stoves) to avoid overloading filters. After sampling, filters were stored in filter petri dishes at 253 K to minimize volatilization and chemical reactions.

7.2.2 Measurements

From the PTFE and Qf filter samples collected, three quantitative measurements were obtained on all samples for use in FTIR post-processing: gravimetric $PM_{2.5}$ mass, TOT OC-EC measurements, and GC-MS PAH measurements. FTIR spectra were gathered in transmission mode on the collected PTFE filters.

$PM_{2.5}$ mass was determined by measuring the PTFE filters gravimetrically with a microbalance (MC5, Sartorius, Germany) before and after sampling. These masses are used for artifact-correction on TOT OC-EC measurements and GC-MS PAH measurements, as well as for comparison with functional group quantifications from FTIR post-processing.

OC and EC on Qf samples were measured using a thermal-optical transmittance instrument (TOT; Sunset Laboratory, Portland, OR) and a revised NIOSH method 5040 (NIOSH, 1999). Blank PTFE and Qf filters were also collected at the cookstove test facility and the associated $PM_{2.5}$ mass, OC, and EC backgrounds were measured and subtracted from test samples. In order to make measurements on Qf (EC, OC, GC-MS) and PTFE (FTIR, gravimetric) filters more comparable (i.e., to minimize the differences due to volatilization and adsorption artifacts for semivolatile aerosols) artifact correction was performed. For this purpose, the values measured on the Qf back filters were subtracted from those on the Qf parallel samples in parallel to the PTFE filter.

Thermal extraction GC-MS (TEx-GC-MS) was utilized to identify and quantify individual PAHs in 31 sampled cookstove emissions as described by Shen *et al.* 2017(Shen et al., 2017b). Briefly, a 0.4 cm^2 Qf punch (or three punches for artifact filters and low-emission samples from LPG) was carefully placed inside a glass tube pre-baked to 550 °C and spiked with 1 µL of an internal standard mixture containing deuterated PAH compounds. Following automated tube insertion into the oven unit, the sample was heated to 325 °C at a rate of 50 °C/min and held for 11 min; 50 ml/min of He flowed over the sample. Desorbate was directed to a cryogenically cooled programmable temperature vaporization inlet at -85 °C operating in splitless mode. The GC inlet was rapidly heated at 720 °C/min to 325 °C. The sample was chromatographed on an ultralow-bleed capillary column (DB-5, 30 m × 0.25 mm I.D. × 0.25 µm film thickness) with helium as the carrier gas (1 ml/min). The GC oven was temperature-programmed as follows: 65 °C for 10 min ramped to 325 °C at a rate of 10 °C/min and held for 15 min. Chromatographed compounds were detected using an Agilent 5973 MS detector operating in selected ion monitoring mode. The internal standard method and a multilevel, linear ($R^2 > 0.9$) calibration was used (0.1–1 ng) for PAH quantification. A mid-level calibration check was performed daily prior to sample analysis and was used as a continuing calibration in cases where measured and fixed concentrations were not within 20%. Daily checks were

within 20% for most targets even when replicated over months. Mid-level check recoveries for PAHs were within $105 \pm 33\%$ on average. PAHs below their detection limit were reported as "ND". Retention time shifts were negligible (< 1%) throughout the analysis. Target analyte validation was achieved using the retention times and isotopic fragment ratios exceeding a 3:1 signal-noise ratio. Carryover tests were performed by reheating the TEx sample/apparatus immediately following the initial extraction and then examining the GC-MS chromatogram for the presence of target compounds, which indicated an extraction efficiency of > 98%. Naphthalene was the only PAH detected (0.16 ± 0.11 ng) above its method detection limit (0.02 ng) in the blank filters. Similar to EC and OC, artifact correction was performed for PAHs by subtracting value of the quartz back filters from those of the parallel filters.

For FTIR analysis, the mid-infrared spectra in the 400–4000 cm^{-1} range were obtained using a Bruker Tensor 27 FTIR instrument equipped with a α deuterated lanthanum alanine doped triglycine sulfate (DLaTGS) detector, at a resolution of 4 cm^{-1}, in transmission mode, averaged over 64 scans.

7.2.3 Post-processing

In the following subsections, the post-processing of FTIR spectra to identify and quantify FGs is described or cited to previous publications.

Baseline correction and blank subtraction

The FTIR spectra were baseline-corrected using smoothing splines to exclude the contribution of light scattering by the PTFE filter membrane and particles, and absorption by carbonaceous material (Kuzmiakova et al., 2016; Russo et al., 2014; Parks et al., 2019). After baseline correction, a scaled version of the baseline-corrected blank filter spectrum was subtracted from the baseline-corrected sample spectra (Yazdani et al., 2021c). This procedure allowed us to identify and quantify some organic FG absorptions such as the out-of-plane (OOP) aromatic CH bands.

Determination of influential functional groups for TOT OC

In order to identify the FGs that contribute the most to TOT OC, we statistically combined collocated TOT OC (on Qf filters) with FTIR (on PTFE filter) measurements. For this purpose, artifact-corrected TOT OC concentrations regressed against the baseline-corrected, blank-subtracted FTIR spectra using partial least squares regression (PLSR). Thereafter, influential FGs for OC were determined based on their VIP scores. By regressing TOT OC concentrations against FTIR spectra, we seek the solution of the following equation:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}, \tag{7.1}$$

where **X** (n×p) is the FTIR spectra matrix (n samples and p wavenumbers), **y** (n×1) is the vector TOT OC concentration, and **e** is the vector of residuals. We chose the univariate PLSR to solve the equation above (Wold et al., 1983). Univariate PLSR projects **X** onto basis **P** with orthogonal scores **T** and residual matrix **E** (Eq. 7.2) such that the covariance between each score column and **y** is maximized. In Eq. (7.3), **b** is the regression coefficient of **y** as a function of scores (**T**) and **f** is the vector of residuals.

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E}, \tag{7.2}$$

$$\mathbf{y} = \mathbf{T}\mathbf{b} + \mathbf{f}. \tag{7.3}$$

After solving the regression equation using PLSR, a repeated 10-fold cross validation was applied to indicate the optimal number of latent variables. We used variable importance in projection (VIP) (Wold et al., 1993; Chong and Jun, 2005; Takahama et al., 2016), to identify the important FGs. The VIP score of the jth wavenumber is calculated by considering all h latent variables in the model as shown in Eq. (7.4):

$$VIP_j = \sqrt{p\frac{\sum_{k=1}^{h} SS(b_k\mathbf{t}_k)(w_{jk}/\|\mathbf{w}_k\|)^2}{\sum_{k=1}^{h} SS(b_k\mathbf{t}_k)}}, \tag{7.4}$$

where $SS(b_k\mathbf{t}_k) = b_k^2\mathbf{t}_k^t\mathbf{t}_k$. Since the average of squared VIP scores is equal to one, wavenumbers with VIP scores greater than one are considered to have higher-than-average contribution to the response variable (i.e., are influential).

Quantification of functional groups

After determining the important functional groups (by VIP scores method), their abundances were quantified using a multiple peak-fitting algorithm (Takahama et al., 2013; Reggente et al., 2019b). Aliphatic CH (aCH), aromatic CH (rCH), alcohol COH (aCOH), carboxylic acid (COOH), and non-acid carbonyl (naCO) were quantified in this work. OC concentrations and OM:OC ratios were then calculated using FG abundances with few assumptions about the number of carbon atoms attached to each FG (Chhabra et al., 2011b; Russell, 2003; Maria et al., 2002).

7.3 Results & Discussion

In the following sub sections, first, we compare OC, EC, and $PM_{2.5}$ measurements among different fuel types, test phases, and stoves. Thereafter, FTIR spectra of particulate emissions are analyzed qualitatively and compared to those of unburned fuels. TOT OC measurements are then combined statistically with the FTIR spectra to identify FGs the highest association with OC. Finally, influential FGs are quantified to estimate OC, OM, and OM/OC, as well as to

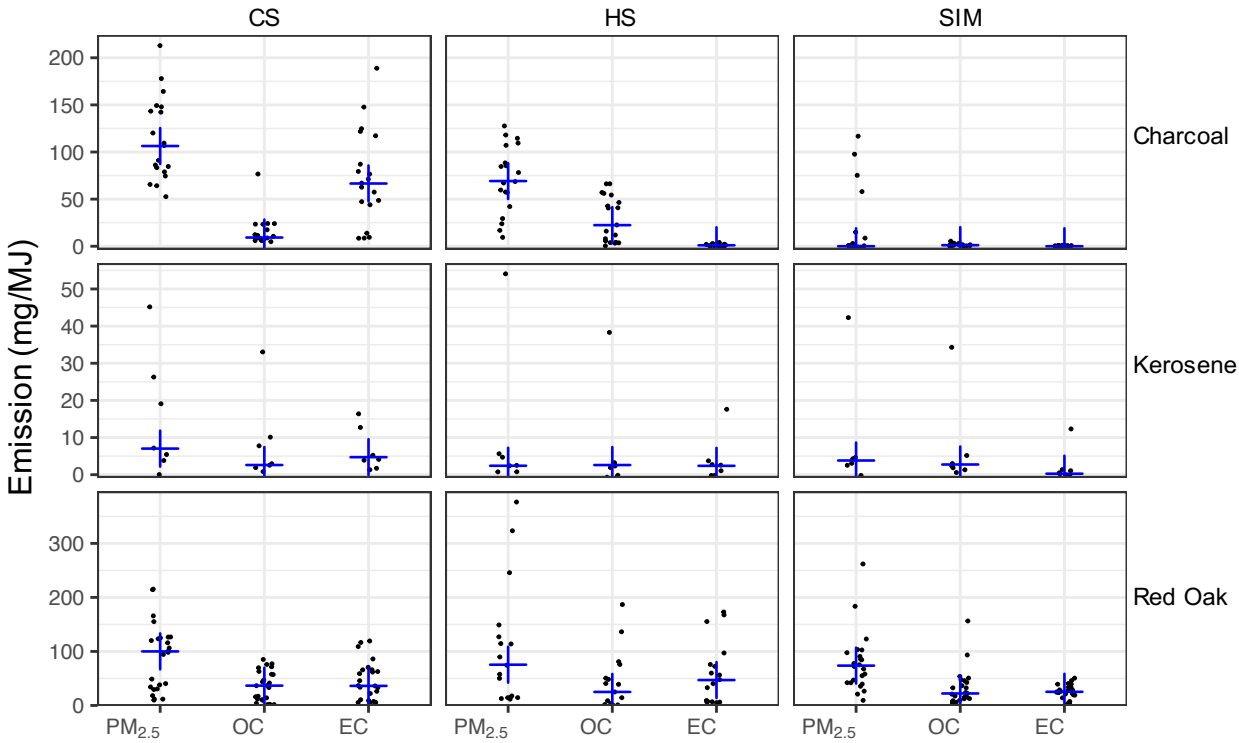

Figure 7.3 – Artifact-corrected emission factors (mg/MJ) of $PM_{2.5}$ (gravimetric), TOT OC, and EC separated by fuel type and test phase (CS: cold start, HS: hot start, SIM: simmering). Blue crosses show the median for each category.

compare with TOT OC and GC-MS PAH measurements.

7.3.1 OC, EC, and $PM_{2.5}$ emission factors

OC, EC, and $PM_{2.5}$ emission factors, expressed in grams of emission per MJ of energy delivered to the pot, are shown in Fig. 7.3. We observe a significant variation across red oak stoves in terms emissions factors. In particular, the 3-Stone fire and the Envirofit G3300 stove (natural draft) have relatively high emission factors of EC and OC accompanied by lower modified combustion efficiencies (MCEs; Fig. A.70). On the other hand, gasifier stoves such as EcoChula XXL have higher MCEs and around 10 times lower EC and OC emissions. Lower absorbances are also noticeable in the FTIR spectra of EcoChula XXL stove especially in the CS and HS phases. EC and OC emission factors appear to be slightly lower in the SIM phase compared to CS and HS phases. Nevertheless, the variations observed in the PM composition form red oak combustion across phases (characterized by elemental-carbon-to-total-carbon ratio, EC/TC, where TC = EC + OC) is not substantial when compared to the variability in each phase (Fig. A.70 and). The average EC/TC for red oak combustion aerosols is 0.58 ± 0.19. This value is on the higher end of those reported in previous studies (0.07–0.64) (Roden et al., 2006; Novakov

et al., 2000).

Kerosene combustion has on average 2 to 10 times lower EC, OC, and $PM_{2.5}$ emission factors compared to red oak and charcoal. The Butterfly model 2412 pressure-style cookstove that uses kerosene as fuel shows relatively high emission across all phases in three experiments. The high OC emissions are corroborated by high absorbances in the FTIR spectra (Fig. A.70) and are believed to be because of its orifice being clogged. Omitting the mentioned outlier samples, the variation of emission factors across phases are small. The average EC/TC ratio for kerosene is 0.44 ± 0.31 and appears to be in general lower in the SIM phase (Fig. A.70).

The variation in emission factors across phases are more significant for charcoal combustion emissions compared to other fuels. The HS phase has the highest OC emission factor (Fig. 7.3). The EC emission factor and the EC/TC ratio increase significantly in the CS phase. The average EC/TC ratios of charcoal combustion aerosols for the CS phase and all phases combined are 0.78 ± 0.23 and 0.30 ± 0.40, respectively.

Inorganic salts (e.g., ammonium sulfate and ammonium nitrate) are believed to contribute negligibly to the emissions for the majority of samples as EC and OM (estimated as OM/OC×OC, where OM/OC is estimated from FTIR measurements) constitute the majority of the gravimetrically-measured $PM_{2.5}$ mass (Fig. A.71). However, for a few wood and charcoal combustion samples $PM_{2.5}$ is considerably higher than the sum of EC and OM. For example, the EcoZoom and Jiko Koa stoves (both ceramic-insulated) have higher $PM_{2.5}$ and low EC during the SIM phase (Fig. A.70) accompanied by a relatively lower MCE. These are among samples with relatively high inorganic nitrate absorbances in their FTIR spectra (Fig. A.31).

7.3.2 FTIR spectra

Figure 7.4 exhibits distinct spectral profiles for $PM_{2.5}$ emitted from combustion of different fuel types. The FTIR spectra of alcohol and LPG particulate emissions were omitted as they were noisy and had organic absorptions comparable to those of blank filters. This observation is corroborated by the negative or close-to-zero TOT OC measurements for alcohol and LPG emissions.

As can be seen from Fig. 7.4, the average spectra of red oak combustion aerosols for different phases are very similar to the spectrum of wood (Pandey, 1999). The strong and broad band at 3500 cm^{-1} related to aCOH is observed in the spectra of all phases. This group is abundant in wood constituents (lignin, cellulose, and hemicellulose). In addition, medium aCH absorbances at 2800–3000 cm^{-1}, medium carbonyl CO absorbances at around 1700 cm^{-1}, and strong aromatic C=C absorbances at 1600 cm^{-1} are observed in the spectra of red oak combustion aerosols. The absorption band attributed to the lignin aromatic ring stretching vibration at 1515 cm^{-1} is also prominent in several spectra (Boeriu et al., 2004), suggesting the existence of aromatic compounds with similar structure to monolignols. Aromatic CH OOP bands at around 750 cm^{-1} (Centrone et al., 2005) and occasionally aromatic stretching

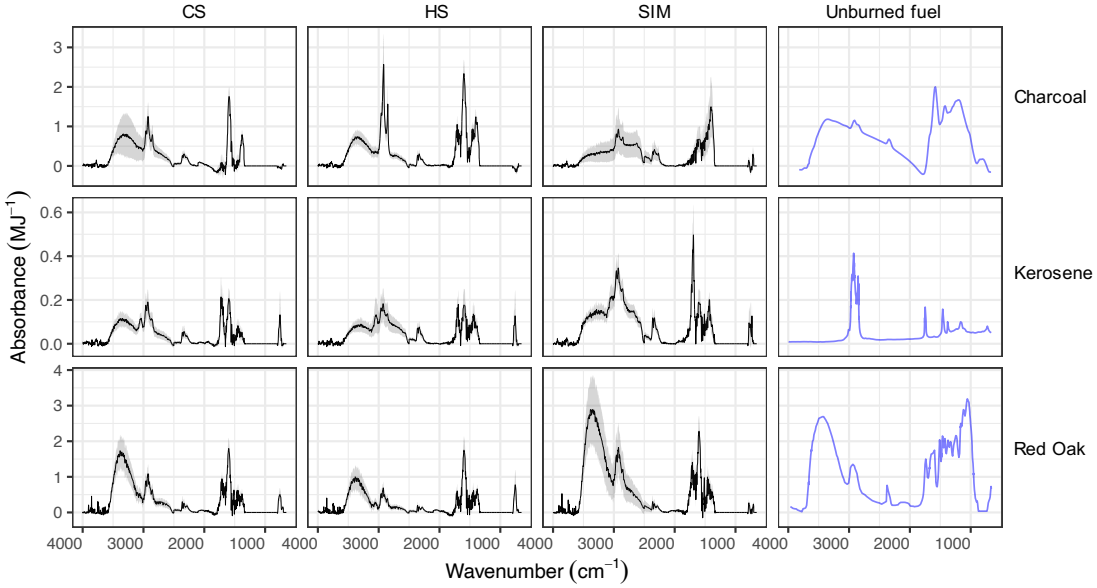

Figure 7.4 – Average mid-infrared spectra of unburned fuels and their particulate emissions separated by source and phase. Emission spectra are shown in terms of emission factors (absorbance divided by MJ energy delivered to the pot). Shaded bands show the mean spectrum ± 0.5 standard deviation. CS = cold start, HS = hot start, SIM = simmering. Blue spectra are taken from unburned fuels (arbitrary scale). The 1000–1300 cm^{-1}, which interferes with PTFE filters has been omitted.

aCH bands at around 3050 cm^{-1} are also visible in the FTIR spectra. The mean spectrum of the SIM phase has the weakest aromatic OOP band among all phases. Levoglucosan signatures described by Yazdani et al. (2021c) are observed in several spectra for all three phases. Inorganic nitrate absorbances at 1400 and 830 cm^{-1} are visible in some red oak samples especially in CS and HS phases. Among all fuel types, the spectra of particulate emissions from red oak present the greatest similarity to the unburned fuel except for some samples that are dominated by aromatic signatures. No systematic difference between the spectral profiles of emissions from different stove types (e.g., 3-stone fire, natural draft, and forced draft) was observed (Fig. A.73).

Kerosene is predominately composed of hydrocarbons with 6–16 carbons. The majority of these hydrocarbons are straight-chain and cyclic and the rest are aromatic (Collins, 2007). The mid-infrared spectrum of kerosene from Xia et al. (2017) has relatively strong aCH peaks and weak rCH stretching (3050 cm^{-1}) absorbances. The elemental analysis presented in Fig. A.74 also shows the insignificant contribution of elements other than H and C and a molar H/C ratio about 2.1. The relative aCH absorbances, however, are much weaker in the particulate emissions of kerosene combustion in all three phases with high CH_3/CH_2 peak ratios that indicate branched or small hydrocarbons. By contrast, rCH absorptions are more prominent in the particulate emissions (Fig. 7.4), suggesting a significant formation of aromatic and PAHs during the combustion process. In the majority of spectra, moderate aCOH absorbances are observed. In several samples taken from the emissions of the Butterfly model 2668 wick-style cookstove, a sharp CO carbonyl peak at 1695 cm^{-1} is observed (Fig. A.73). This peak is accompanied by a broad peak at 2400–3400 cm^{-1} (with doublets at 2400 and 2600 cm^{-1} attributed to the dimerized carboxylic acid OH), suggesting high abundances of carboxylic acids (Pavia et al., 2008). The carbonyl peak with a relatively low frequency and the high abundance of rCH indicate the existence of aromatic acids (Pavia et al., 2008). Aromatic signatures appear to be more prominent in the CS and HS phases. The combustion mechanism used (wick or orifice) appears to have a dominant effect on the composition of the emitted OM (Fig. A.73).

The charcoal spectrum taken from (Guo and Bustin, 1998) has weak aCH and carbonyl CO and strong aromatic C=C and aCOH absorbances. The weak aCH peak are consistent with the low hydrogen content of the fuel (Fig. A.74). However, the mid-infrared spectra of particulate emissions from charcoal combustion show strong aCH peaks (Fig. 7.4). The prominent aCH are believed to be associated with volatile compounds emitted from charcoal. The same observation has been reported for coal combustion particulate emissions (Yazdani et al., 2021c). The aCH peaks are especially prominent in the HS phase for which coal is added to the hot bed (Fig. A.69). In addition, the carbonyl CO peak at 1700 cm^{-1} emerges in the charcoal combustion spectra in the HS phase, while being absent in the charcoal spectrum itself. This observation suggests the formation of carbonyls during charcoal pyrolysis and burning. Charcoal combustion in the cold start phase has the highest EC emission factor (Fig. 7.3), concurrent with prominent C=C absorbances at around 1600 cm^{-1} in the FTIR spectra. The ratio of the aromatic C=C peak (1600 cm^{-1}) to aromatic CH (750 cm^{-1}) is higher

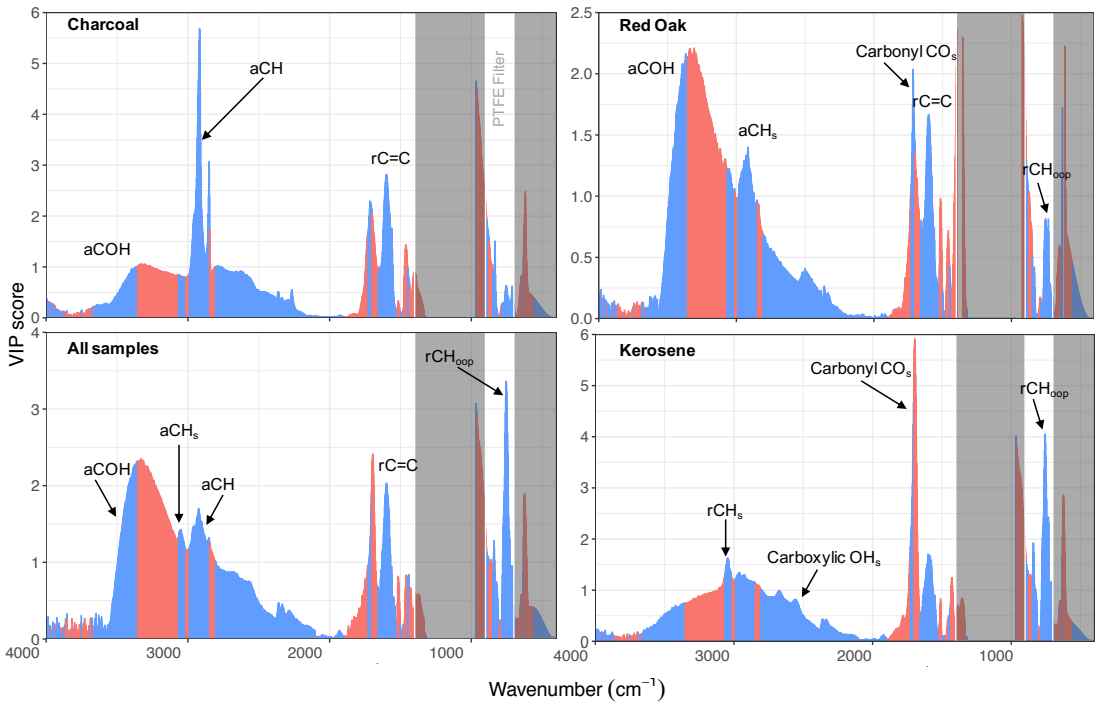

Figure 7.5 – VIP scores of TOT OC regressed against baseline-corrected mid-infrared absorbances for each fuel type. Important group frequencies for each fuel are indicated. The regions with negative (positive) regression coefficients are shown in red (blue).

for charcoal burning aerosols compared to other fuels, suggesting that aromatic compounds are poor in hydrogen (=C–H groups).

7.3.3 Influential group frequencies for OC

As seen in Fig. 7.5, the absorption region of the aCOH group in the 3400–3500 cm^{-1} range has the highest VIP scores for red oak combustion, suggesting this group is the the most influential FG for the OC emitted from red oak combustion. After aCOH, the carbonyl CO, aromatic ring C=C, and out-of-plane aromatic CH, and aCH regions have the highest VIP scores, respectively. As expected, although the inorganic nitrate peak is prominent in spectra of the oak burning aerosols at 1400 cm^{-1}, this group does not have either high VIP scores or positive regression coefficients of TOT OC. The greater-than-one VIP scores in the 2400–3400 cm^{-1} range are believed to indicate the alcohol group in sugar moieties, which are abundant in wood burning aerosols. The 3100–3400 cm^{-1} region, in which ammonium absorbs strongly, has high VIP scores with negative regression coefficients for all fuel. The PLSR models use the information from this region to correct the ammonium interference with organics.

For aerosols emitted from kerosene combustion, both the out-of-plane and the stretching bands of rCH have the highest high VIP scores, indicating the importance of aromatic compounds for OC emitted from kerosene combustion. The carbonyl peak along with the region

indicating the dimerized acid OH have also high VIP scores, suggesting that carboxylic acids are also important constituents of OC. Although aCH is the most important group in unburned kerosene, it does appear to be as important in OC emitted from kerosene combustion (i.e., no high VIP scores).

For aerosols emitted from charcoal combustion, the aCH region (2800–3000 cm^{-1}) has considerably higher VIP scores than other FGs, suggesting that the majority of OC is due to this functional group. Regions attributed to the aromatic ring C=C, carbonyl CO, and aCOH have, in order, the highest VIP scores after aCH.

When the PLSR model is developed using all samples in stead of each fuel, the resulting VIP scores are similar to the average of previously calculated VIP scores for each fuel. However, the regression is pulled more toward those of red oak aerosols due to their higher OC emissions.

Based on the VIP scores, aromatic CH (rCH), alcohol COH (aCOH), aliphatic CH (aCH), aromatic C=C, non-acid carbonyl (naCO), and carboxylic acid COOH are the best predictors for OC emitted from the combustion of red oak, kerosene, and charcoal. Since FGs such as amine, organonitrate, and carboxylate absorb in the same region as the aromatic C=C, the quantification of C=C group is uncertain. Moreover, the absorption coefficient of C=C band is variable and is enhanced by the irregularity of aromatic molecules caused by different ring substitutions (Russo et al., 2014), adding to this uncertainty. We used the rCH FG as an alternative means to quantify carbon and hydrogen atoms of aromatic rings.

7.3.4 Quantifying functional groups from FTIR spectra

aCH, rCH, naCO, COOH, and COH groups were quantified by peak fitting to mid-infrared. The results of peak fitting are shown in Fig. 7.6 in terms of OM emission factors and OM/OC ratios. Red oak aerosols have the highest mean OM emissions factors for all phases (Fig. 7.6), consistent with TOT OC measurements (Fig. 7.3). Among different phases, the simmering phase has a slightly lowest OM emission. For all three phases, aCOH, aCH, and rCH groups constitute the majority OM. The rCH group abundance is significantly lower in the simmering phase. The lower concentration of this group in the SIM phase is also observed for other fuel types. This is also reflected in the higher OM/OC ratio of this phase compared to other phases. The average OM/OC ratio of red oak combustion aerosols is estimated to be around 1.7.

Kerosene combustion OM is estimated to have the lowest average OM emissions factors compared to other fuels in all phases (Fig. 7.6; around 4 times lower than red oak). No clear difference is observed between the mean OM emission factor among different phases. The rCH is estimated to be the most prominent FG for the kerosene burning aerosols followed in order by aCOH, aCH, and COOH. The average OM/OC ratio is 1.8 for the kerosene aerosols. Like red oak, the simmering phase has the highest OM/OC ratio due to a higher abundance of abundance COOH (more oxygenated) and lower abundance of rCH (likely due to a more complete combustion).

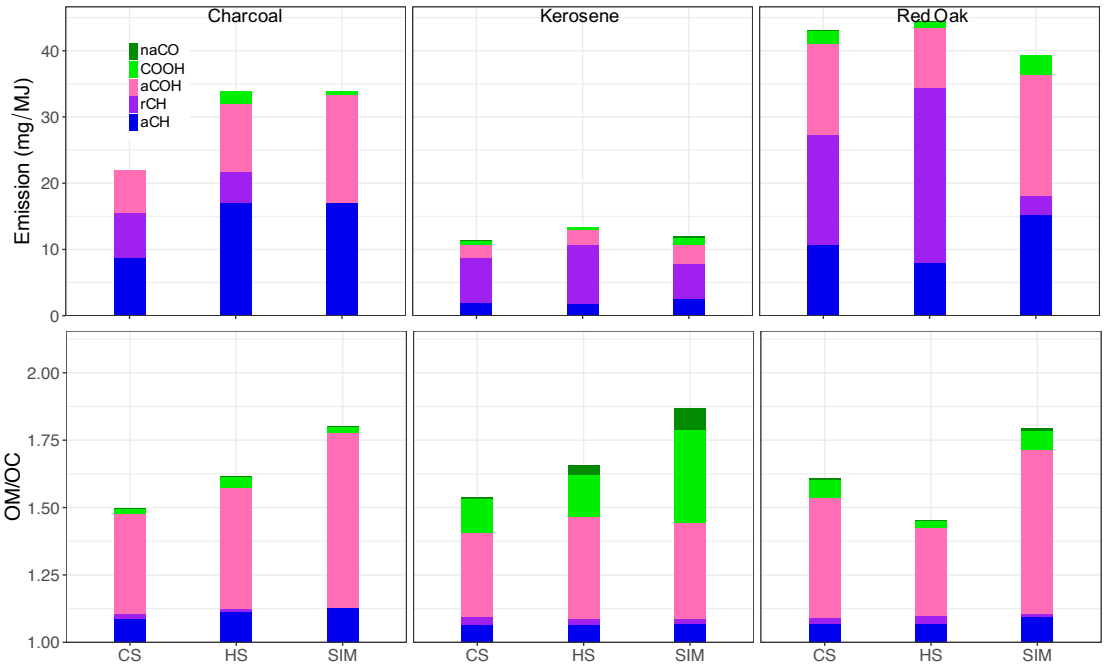

Figure 7.6 – OM emission factors calculated from mid-infrared spectroscopy, separated by functional group contribution, and averaged over each phase (upper row). OM/OC ratios calculated from mid-infrared spectroscopy, separated by functional group contribution (lower row). CS = cold start, HS = hot start, SIM = simmering.

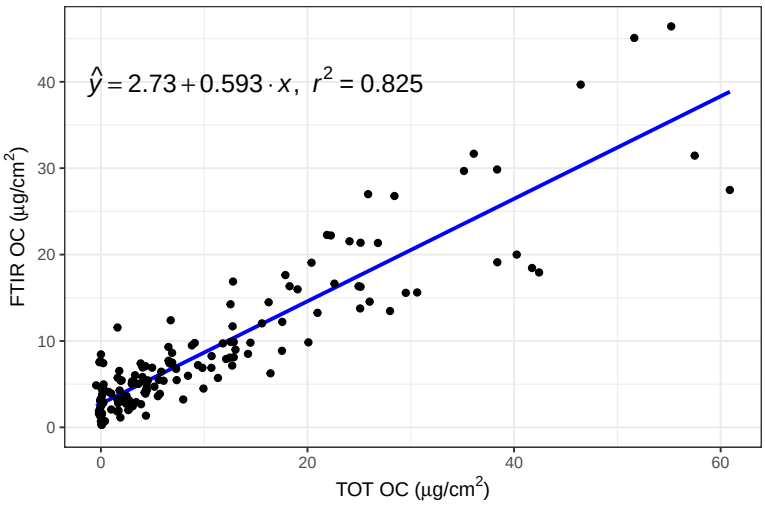

Figure 7.7 – Scatter plot comparing OC (areal density on filters) estimated from TOT and FTIR.

Charcoal burning aerosols have average OM emission factors slightly lower than those of red oak combustion (Fig. 7.6). The aCH group is the most important FG for charcoal burning aerosols constituting up to 50 % of OM. The aCOH and rCH are the most abundant groups after aCH. Other FGs exist in negligible amounts in charcoal burning aerosols. The average OM/OC ratio for charcoal burning aerosols is 1.6, slightly lower than that of other fuels. Like other fuel types, the simmering phase has the highest OM/OC ratio and the lowest abundance of aromatics.

FTIR OC estimated from the abundances of functional groups is highly correlated with TOT OC measurements (R^2 = 0.83; Fig. 7.7). However, FTIR OC underpredicts TOT OC by around 40% although artifact correction has been performed for TOT OC using Qf back filters. This underprediction can be attributed to quartz adsorption artifacts, PTFE volatilization artifacts (Subramanian et al., 2004), and the undermentioned of the fractional carbon associated with each FG (possibly due to presence of unmeasured functional groups) (Bürki et al., 2020). The operationally-defined EC and OC separation threshold for TOT might also be another cause of this discrepancy

The concentration of rCH was estimated in this work for the first time using the peak at 750 cm^{-1} calibrated to laboratory standards of aerosolized anthracene. These estimates are highly correlated (R^2 = 0.84; slope = 19.6) with the total PAH concentration by CG/MS (Fig. A.77). The VIP scores of GC/MS total PAH concentration regressed against mid-infrared absorbances also indicate the importance of the 750 cm^{-1} band for PAHs (Fig. A.78). The magnitude of PAHs and aromatic estimated using this band is almost an order of magnitude higher than that measured by GC/MS probably due to accounting for other unmeasured aromatic and PAH compounds. Previous studies correlate aromatics and PAHs with mutagenicity and carcinogenicity (Riedel et al., 2018; Shiraiwa et al., 2017b; Gibbs-Flournoy et al., 2018; Mutlu Esra et al., 2016). Our analysis shows that these compounds are abundant even in relatively

clean fuels (i.e., with low emission factor) like kerosene, suggesting their harmfulness. This observation along with the highly oxygenated FGs present in kerosene OM is consistent with previous studies providing some evidence that kerosene emissions may impair lung function and increase infectious illness, asthma, and cancer risks (Lam et al., 2012).

Graphitic carbon contributes to the mid-infrared spectra through electronic transition absorptions (Russo et al., 2014; Parks et al., 2021). After eliminating the variability of FTIR baseline due to PTFE filter membrane, we regressed the baseline value at 4000 cm^{-1} against EC and OC (major components of $PM_{2.5}$ in this study). The results of this regression suggest a relatively large and statistically significant coefficient for EC (2.2 times the coefficient of OC, $R^2 = 0.72$; Fig. A.80). This observation suggests that EC contributes significantly to the FTIR baseline of combustion aerosols. Thus, this baseline might be used to extract information about EC.

8 Conclusions and outlook

8.1 Summary and conclusion

Chemical characterization of OM is crucial for understanding its sources and its impacts on the climate and human health. Atmospheric OM is a complex mixture of thousands of species from different sources that undergo various processes during their lifetime in the atmosphere. These features make the characterization of OM difficult. Mid-infrared spectroscopy has a great potential as an informative analytical tool for aerosol chemical analysis. Nevertheless, its signal complexity and lack of systematic methods for spectrum interpretation has impeded its widespread use by the community.

In this work mid-infrared spectroscopy was used as the main technique to characterize fine OM in samples collected on PTFE filters from the atmosphere, smog chambers, and cookstoves. We showed that FTIR spectroscopy provides measurements that are consistent with other analytical techniques (e.g., aerosol mass spectrometry, ion chromatography, gas chromatography, and thermal optical techniques) in terms of the OM mass concentration (Chapters 3 and 7), functional group (FG) composition (Chapters 3 –4), elemental ratios (Chapter 4), and the concentrations certain marker molecules (Chapters 3–7) either qualitatively or quantitatively. The statistical combination of FTIR with AMS measurements helped us better understand the molecular structure of species associated with different small and large mass fragments in the AMS spectra (Chapter 4). Since AMS is widely adopted by the aerosol community, a better interpretation of its complex spectra can be beneficial. This combination also allowed us to take advantage of the high-temporal resolution of AMS measurements and the direct FG measurements by FTIR in order to estimate the high-temporal-resolution FG composition of OM (not readily available from either technique individually). The latter was used to study the dynamic evolution of OM during the course of oxidation in terms of formation of different oxygenated FGs based on emission sources, oxidants used, and exposure time.

The ability of FTIR to provide information about specific marker molecules and the molecular structure of OM (apart from its FG composition) did not received enough attention in previous studies. These two features of FTIR spectroscopy were emphasized more in this work as they

provide very insightful information about OM sources and processes in the atmosphere (e.g., fragmentation). We found that in addition to the functional group composition, FTIR can be used to provide additional information about the average molecular structure of OM (e.g., mean molecular weight and carbon number; Chapter 1) through the variations in the mid-infrared spectral profiles. Certain bbOA marker molecules (anhydrosaccharides and lignin-like compounds) were also identified and quantified using their FTIR fingerprints, showing good agreement with other analytical techniques (Chapter 3-6). These signatures were used to quantify the evolution of fresh bbOA with aging (Chapter 5) and to identify smoke-impacted atmospheric samples (Chapter 6). The identification of marker molecules using FTIR on PTFE filters does not require any prior sample preparation, is fast and non-destructive, and can be used at the scale of large air pollution monitoring networks. These characteristics make FTIR advantageous over time consuming, destructive, and more expensive separation and mass spectroscopic techniques for the study of longterm aerosol-related trends at large scales.

In this thesis, several aspects of data post-processing on PTFE filter samples were developed and improved. For example, often absorptions below 1500 cm^{-1} were omitted in previous FTIR studies on PTFE filter samples. This omission was mainly due to the significant interference from C−F bands of the PTFE membrane in this region. In Chapter 3, the baseline correction algorithm was extended to wavenumbers as low as 400 cm^{-1}. It was shown that in spite of the significant interference, quantitative information about certain organic and inorganic species can be obtained from the 400–1500 cm^{-1} region (e.g., levoglucosan, PAHs, sulfate, inorganic and organic nitrate). In Chapter 7, after manipulating the baseline of FTIR spectra to account for the variations in scattering by PTFE filters, quantitative information about elemental carbon (EC) was obtained from the baseline. The latter can be especially important for the analysis of EC in monitoring networks, which is currently carried out using time-consuming thermal optical methods. As another example, a Fourier self-deconvolution code was developed to analyze overlapping peaks in the chamber studies (Chapter 3). This algorithm proved useful for the analysis of the 1500–1800 cm^{-1} region in the FTIR spectra of particulate matter, which contains several overlapping peaks (e.g., carbonyl, amine, amide, aromatic and alkene C=C, carboxylate C=O, and organic nitrate).

8.2 Outlook and improvements

We found that FTIR spectral profiles can be used to identify the chemical environment in which bonds are vibrating. This can especially be important in regard to the identification of the phase state of OM, which is a determining factor regarding its impact on climate change and air pollution (Shiraiwa et al., 2017a). Although the feasibility of determining the phase state of OM was discussed in this work, it was treated rather briefly and merits further attention given its importance. Using a wider set of atmospherically-relevant laboratory standards can be beneficial for a robust estimation of the phase state.

Although the common belief has been that the identification individual molecules with FTIR

is not feasible in the condensed phase mixtures due to the broad absorption bands (Nozière et al., 2015), we showed that FTIR can be used to identify and quantify bbOA marker molecules with concentrations in the low ng m^{-3} range in complex atmospheric mixtures. An effort to differentiate between signatures and general spectral profiles of different types of biomass burning (e.g., hardwood, softwood, grass, and leaves) can be useful for a more detailed source apportionment of atmospheric bbOA. In addition, the identification of marker signatures in the FTIR spectra for other important atmospheric POA and SOA sources such as fossil fuel combustion and biogenic emissions can considerably improve our understanding of the contribution these sources to atmospheric aerosols at large scales.

The wood smoke identification method developed in Chapter 6 shows the promising potential of FTIR to analyze bbOA at the scale of large monitoring networks. This method might be used to study bbOA concentrations, and their temporal and spatial distributions over multiple years to better characterize the trend and impact of biomass burning on the air pollution. In addition, this method can be used to validate the chemical transport models with regard to bbOA concentrations, and to help policymakers with respect to emission-reduction policies.

All FTIR analyses in this work were performed on PTFE filters and showed good agreement with other offline (e.g., TOR-OC, GS-MS, IC) and online (AMS) analytical techniques. However, an efforts to better characterize sampling artifacts such as the loss of semi-volatile compounds on these filters can be beneficial to understand the biases. Contrary to the US, the majority of monitoring networks in Europe employ quartz fiber filters, which are considerably more opaque in the mid-infrared range compared to PTFE filters. Our primary analysis of atmospheric $PM_{2.5}$ collected on quartz filters using a single-bounce ATR instrument (Fig. A.81) suggested that FTIR spectra obtained from quartz filters are suitable at least for qualitative analyses. The applicability of qualitative and quantitative FTIR analyses on quartz filter samples directly or indirectly (e.g., via electrospray extraction; Arangio et al., 2019) merits further investigation. In addition, the collection of particles on transparent substrates coupled with a semi-online FTIR analysis can alleviate certain shortcomings of FTIR analysis on PTFE filters related to its temporal resolution and interference of organic and inorganic peaks with those of the PTFE membrane. The latter has been the focus of a recent study by Dudani (2021).

In addition to hardware improvements suggested in the previous paragraph, the post-processing algorithms developed or improved in this work might be incorporated as a standard suite of analyses for future FTIR studies on PTFE filter samples.

A Appendix

A.1 Appendix to chapter 2

A.1.1 Correlation matrix heat map

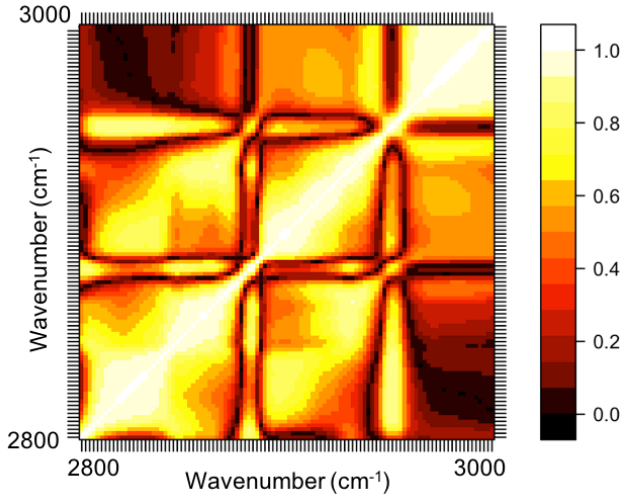

Figure A.1 – Correlation matrix heat map (absolute values) of mid-infrared spectra of the laboratory standards in aliphatic C−H region. In this heat map, absolute values of correlation coefficient of absorbances at each wavenumber with absorbances at other wavenumbers are demonstrated (ranging between zero to one).

A.1.2 Relating mixture property to pure compound property

Laboratory standards which have been used for model development are aerosols of single organic compounds, while atmospheric organic aerosols are generally complex mixtures of multitude of species (Hallquist et al., 2009). This fundamental difference highlights the importance of investigating the validity of the models for mixtures. Herein, the validity of the models developed on pure compounds is rationalized mathematically for estimating mean molecular properties of a non-interacting mixture.

In the aliphatic C−H region, a particular absorbance profile is observed due to different absorbance at each wavenumber. The absorbance profile is dependent on areal molar density n (mole per area of the filter) and the absorption coefficient $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\tilde{\boldsymbol{\nu}})$ of the compound, which is a function of wavenumber ($\tilde{\boldsymbol{\nu}}$). Thus, the absorbance profile $\boldsymbol{A}$ can be written as

$$\boldsymbol{A} = n\boldsymbol{\varepsilon}, \tag{A.1}$$

In this work, spectra are normalized before applying the models. This normalization step is done by a function denoted as g. The function g scales the profile between 0 and 1 regardless of the molar abundance, thus is scale invariant, meaning that

$$g(\boldsymbol{x}) = g(s\boldsymbol{x}), \tag{A.2}$$

where s is a an arbitrary scalar. After the normalization step, the model (function) f is applied to the spectra for estimating a molecular property (carbon number or molecular weight) of the laboratory standards or atmospheric samples. f is linear if

$$f(\sum_i \boldsymbol{x_i}) = \sum_i f(\boldsymbol{x_i}), \tag{A.3}$$

which is true for the linear calibration models used in this work. A pure compound i with the absorption coefficient $\boldsymbol{\varepsilon_i}$ is estimated to have the property Φ_i calculated by a scale-invariant model $f(g(.))$ (combining the model with the normalization step),

$$\Phi_i = f(g(\boldsymbol{A_i})) = f(g(\boldsymbol{\varepsilon}_i)). \tag{A.4}$$

For a mixture, the true mean property $\bar{\Phi}_{true}$ can be written as an molar average of the model estimates for pure compounds assuming no strong interaction between them in the mixture,

$$\bar{\Phi}_{true} = \frac{\sum_i n_i \Phi_i}{\sum_i n_i} = \frac{\sum_i n_i f(g(\boldsymbol{\varepsilon_i}))}{\sum_i n_i} \tag{A.5}$$

for which if the model is linear,

$$\frac{\sum_i n_i f(g(\boldsymbol{\varepsilon_i}))}{\sum_i n_i} = f\left(\frac{\sum_i n_i g(\boldsymbol{\varepsilon_i})}{\sum_i n_i}\right) = \bar{\Phi}_{lin}. \tag{A.6}$$

However, when applying the models to a mixture spectrum, the actual value of $\bar{\Phi}$ is estimated from the measured mixture absorbance profile, which is the sum of pure compound spectra, $\sum_i \boldsymbol{A_i}$ as

$$\bar{\Phi}_{mix} = f\left(g\left(\sum_i \boldsymbol{A_i}\right)\right). \tag{A.7}$$

Since the normalization function g scales the profile between 0 and 1, i.e. $g(\boldsymbol{x}) = \boldsymbol{x}/\max(\boldsymbol{x})$, the true mixture mean assuming a linear model will be:

$$\bar{\Phi}_{lin} = f\left(\frac{\sum_i n_i g(\boldsymbol{\varepsilon_i})}{\sum_i n_i}\right) = f\left(\sum_i \xi_i g(\boldsymbol{\varepsilon_i})\right) = f\left(\sum_i \frac{\xi_i \boldsymbol{\varepsilon_i}}{\max(\boldsymbol{\varepsilon_i})}\right), \tag{A.8}$$

where $\xi_i = n_i / \sum_i n_i$ is the mole fraction of the ith component in the mixture. However, the estimated molecular property for a mixture based on the mixture spectrum ($\bar{\Phi}_{mix}$) is

$$\bar{\Phi}_{mix} = f\left(\sum_i \boldsymbol{A_i}\right) = f\left(\frac{\sum_i n_i \boldsymbol{\varepsilon_i}}{\max(\sum_i n_i \boldsymbol{\varepsilon_i})}\right) = f\left(\frac{\sum_i \xi_i \boldsymbol{\varepsilon_i}}{\max(\sum_i \xi_i \boldsymbol{\varepsilon_i})}\right) = f\left(\sum_i \frac{\xi_i \boldsymbol{\varepsilon_i}}{\max(\sum_i \xi_i \boldsymbol{\varepsilon_i})}\right). \tag{A.9}$$

As a result, $\bar{\Phi}_{mix}$ and $\bar{\Phi}_{lin}$ are different because of their different denominators ($\max(\sum_i \xi_i \boldsymbol{\varepsilon_i})$ and $\max(\boldsymbol{\varepsilon_i})$). This means that the true mean property of a mixture is not necessarily the property estimated by applying the model to the mixture spectrum. The difference is, however, negligible as long as the models are linear and the compounds in the mixture have relatively similar absorption coefficients. These two conditions are valid for majority of compounds

considered in the laboratory standards.

A.1.3 Elemental carbon and carbon number

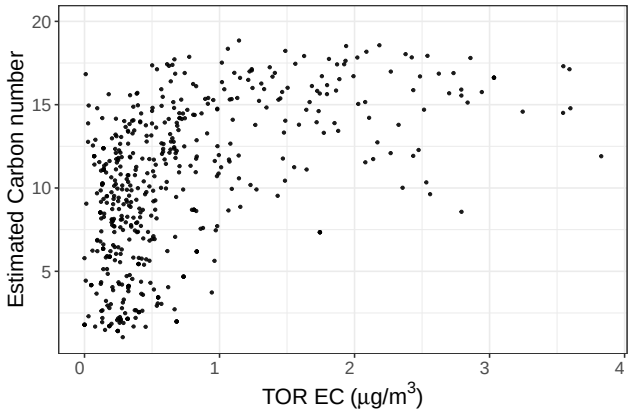

Figure A.2 – Scatter plot showing the relationship between collocated measurements of EC concentration and carbon number estimates by PLSR models in the IMPROVE network in 2011 and 2013.

A.1.4 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 ($\boldsymbol{y} = \frac{1}{1+((\boldsymbol{x}-\mu)/\sigma)^2}$) 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 consid-

ered ($\mu = 0 + N(0, m)$ 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 ($\boldsymbol{y} = \frac{1}{1+(\boldsymbol{x})^2}$; Fig. A.3). 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. A.3, 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. A.3, 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.

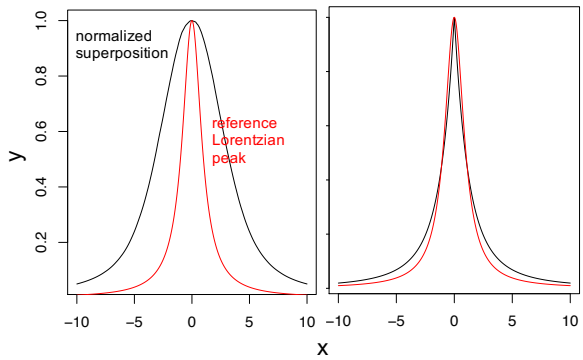

Figure A.3 – Normalized profile resulted from superposition of 10000 individual Lorentzian profiles ($\boldsymbol{y} = \frac{1}{1+((\boldsymbol{x}-\mu)/\sigma)^2}$ where $\mu = 0 + N(0, m)$ and $\sigma = 1 + N(0, n)$; black profile) and reference Lorentzian profile ($\boldsymbol{y} = \frac{1}{1+(\boldsymbol{x})^2}$; 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.

A.1.5 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. A.4 (liquid and solid laboratory standards are completely separated based on the PC3 scores).

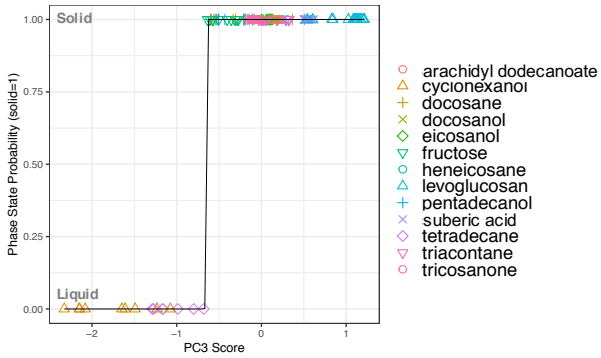

Figure A.4 – Phase state probability of the laboratory standards (liquid = 0, solid = 1) against their PC3 scores. Black curve indicates the phase state probability estimated by the logistic regression.

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. A.5). The glass transition temperature (T_g) (parametrized by Shiraiwa et al., 2017a) 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. A.5, 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.

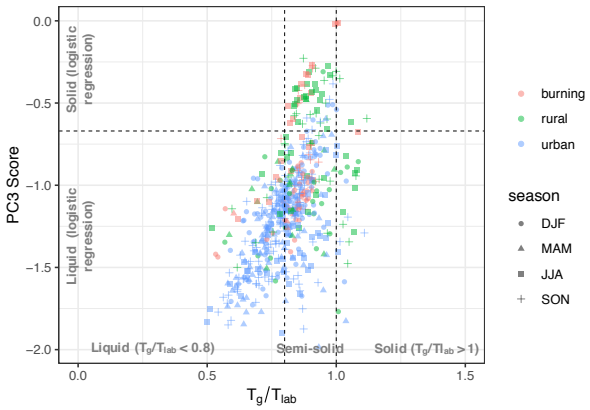

Figure A.5 – PC3 score (used in logistic regression for modeling phase state) regressed against glass transition temperature (T_g) calculated using the estimated mean molecular weight and the O/C ratio of this study.

A.1.6 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. A.6) 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.

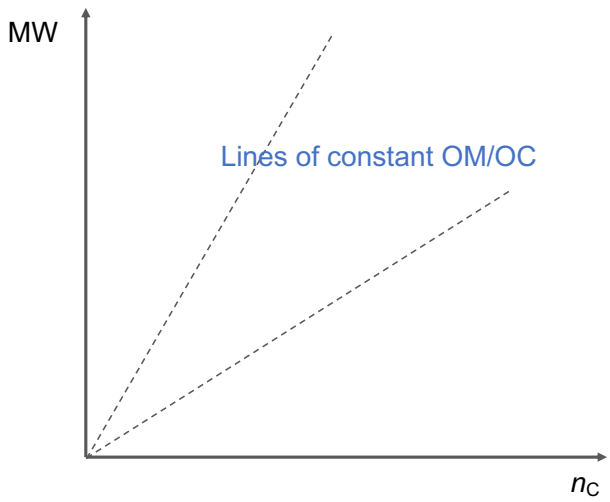

Figure A.6 – MW-n_C space considering the molecular weight (MW) and carbon number (n_C) of organic aerosols as independent variables. Dashed lines show constant OM/OC ratio lines.

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 A.1). 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. A.7), 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.

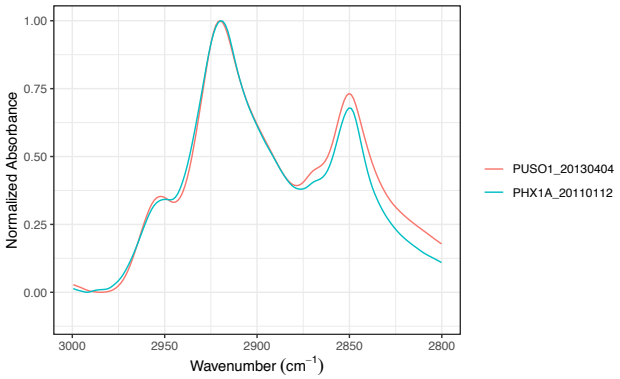

Figure A.7 – Normalized baseline-corrected spectra of two ambient samples which have similar estimated OM/OC ratios, but different estimated molecular weight and carbon number.

Table A.1 – Predicted molecular weight, carbon number and OM/OC ratio for individual samples having similar OM/OC ratio.

Site name/date	**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

Similar carbon number (n_C):

Two samples with similar estimated carbon number and fairly different molecular weight and OM/OC ratio are considered (Table A.2). The sample with higher molecular weight has slightly lower peak frequency (Fig. A.8) 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.

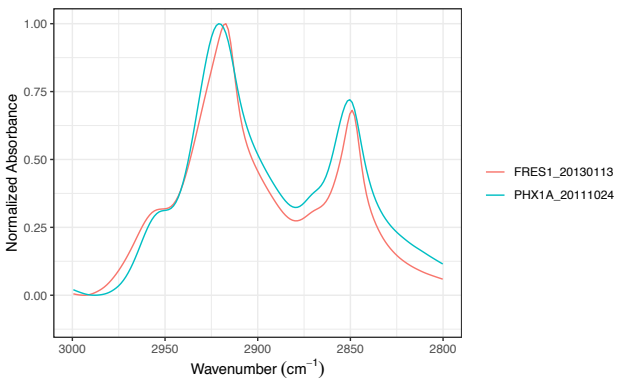

Figure A.8 – Normalized baseline-corrected spectra of two ambient samples which have similar carbon number but different molecular weight and OM/OC ratio.

Table A.2 – 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.

A.2 Appendix to chapter 3

A.2.1 Experimental set-up

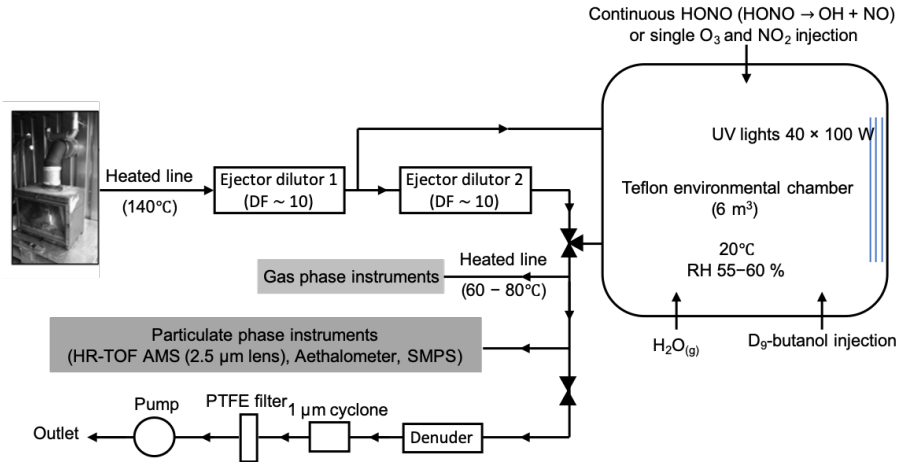

Figure A.9 – Schematic of the experimental set-up used for photo-oxidation and dark oxidation of biomass burning and coal combustion emissions.

A.2.2 Blank subtraction

As can be seen from Fig. A.10, the blank subtraction algorithm used in this work permits better identification of organic FG absorbances that are either overlapping with PTFE absorbances (e.g. the carbonyl peak) or are completely masked by them (e.g. the levoglucosan and organonitrate peaks). In addition, water vapor and CO_2 bands are eliminated with this method.

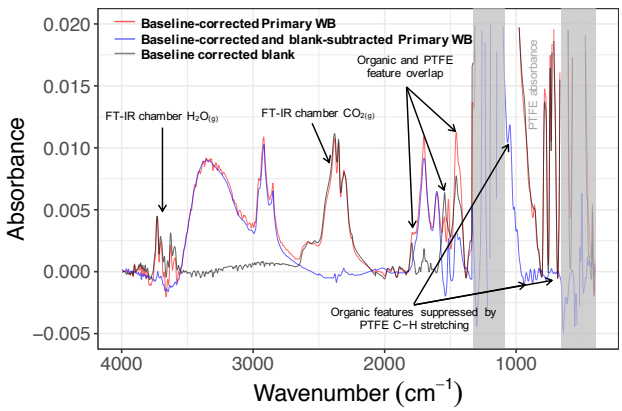

Figure A.10 – Comparison between baseline-corrected (red) and blank-subtracted baseline-corrected spectra (blue) of WB aerosols.

A.2.3 Dimensionality reduction

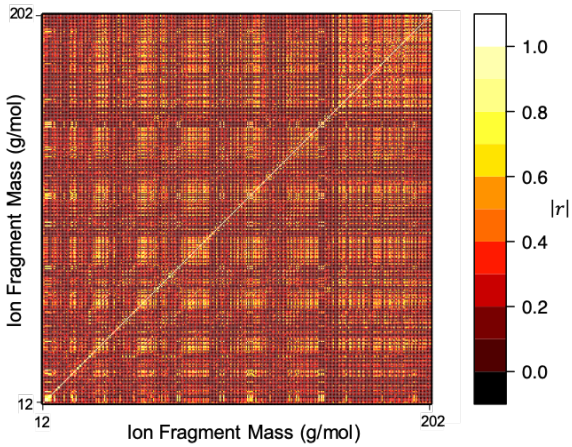

Figure A.11 – Heat map of correlation matrix of mass fragments.

A.2.4 Fourier self-deconvolution

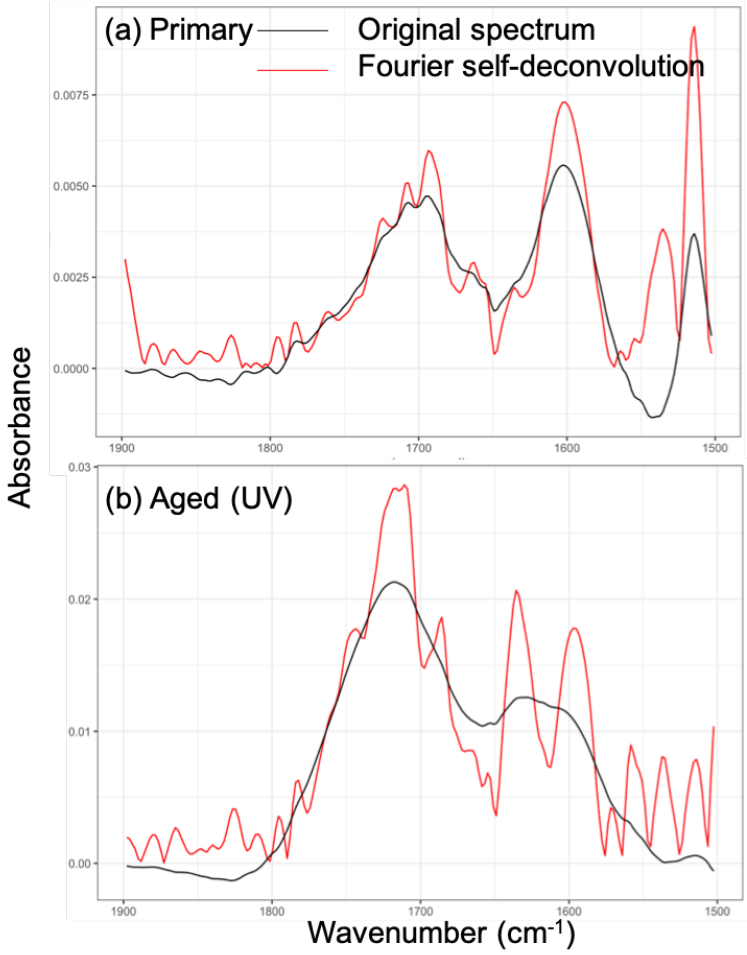

Figure A.12 – Example of Fourier self-deconvolution applied to the 1500–1800 cm^{-1} region of FTIR spectra of primary (a) and aged (b) wood burning aerosols using UV lights.

A.2.5 The MIR and AMS spectra and chemical composition of the chamber experiments

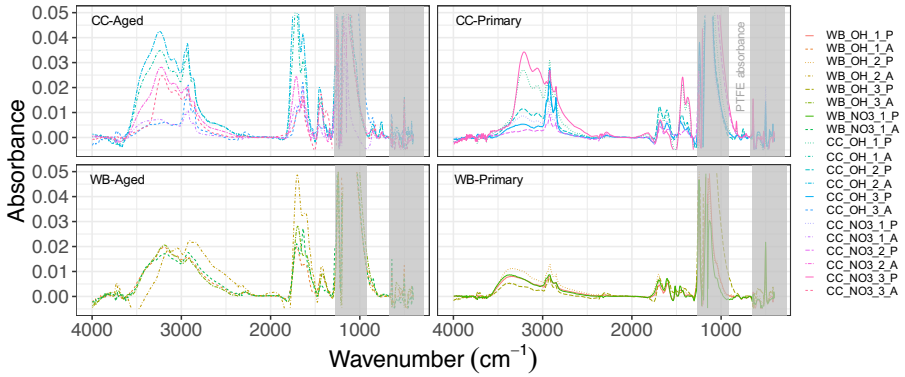

Figure A.13 – Baseline-corrected spectra of wood burning (WB) and coal combustion (CC) emissions separated by emission source and aerosol type.

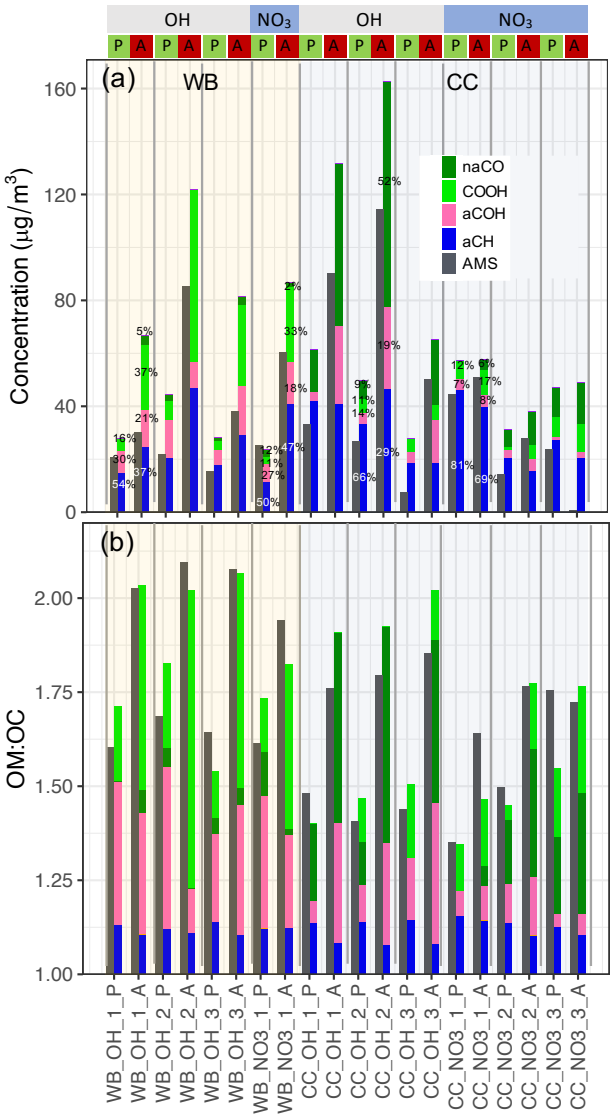

Figure A.14 – (a) Bar plot comparing AMS and MIR OA concentration estimates separated by functional group. The emission pairs (primary and aged emissions of the same experiment) are grouped by vertical lines. The type of aerosol (P: primary, A: aged), emission source (WB and CC), and oxidant used for aging (OH: hydroxyl radical, NO_3: nitrate radical) are indicated for each experiment. The percentage of each FG to total OA mass before and after aging is indicated for some experiments. (b) Bar plot comparing AMS OM:OC estimates with that of MIR spectroscopy separated by contribution of each functional group.

A.2.6 Aromatics and Polycyclic aromatic hydrocarbons (PAHs)

In this work, the following AMS fragment ions attributed to aromatics are quantified (Bruns et al., 2015; Pavia et al., 2008): $C_6H_5^+$, $C_7H_7^+$, $C_7H_5O^+$, $C_8H_5O_2^+$, $C_6H_6^+$, $C_6H_6O^+$, $C_8H_9^+$, $C_8H_{10}^+$, $C_8H_8^+$, $C_7H_7O^+$, and $C_7H_8O^+$. The summed concentration of these fragments is

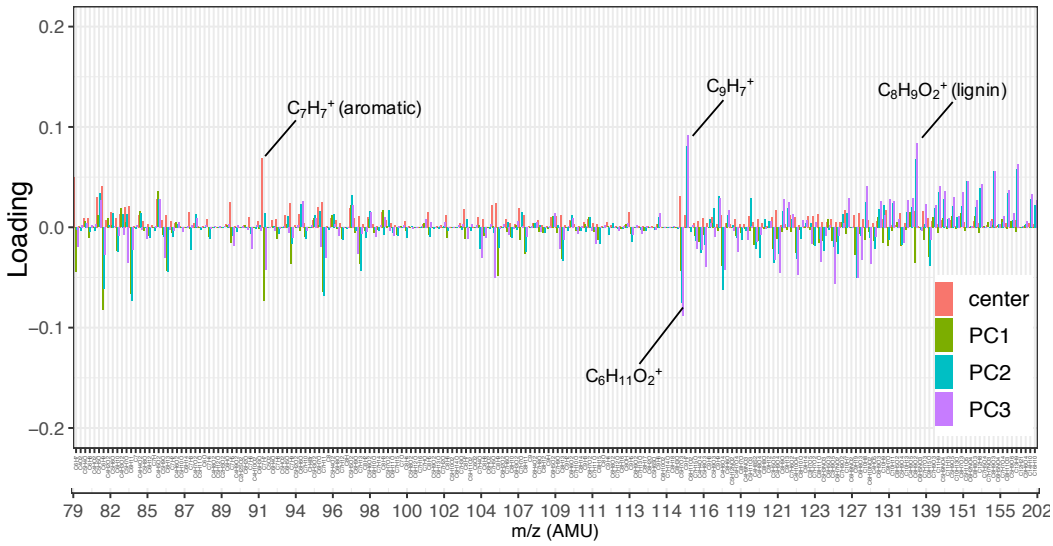

Figure A.15 – Loadings of the first three principal components and the normalized mean AMS mass spectrum (from m/z 80 to m/z 202). Fragment ions with high positive/negative loadings are indicated by their formula.

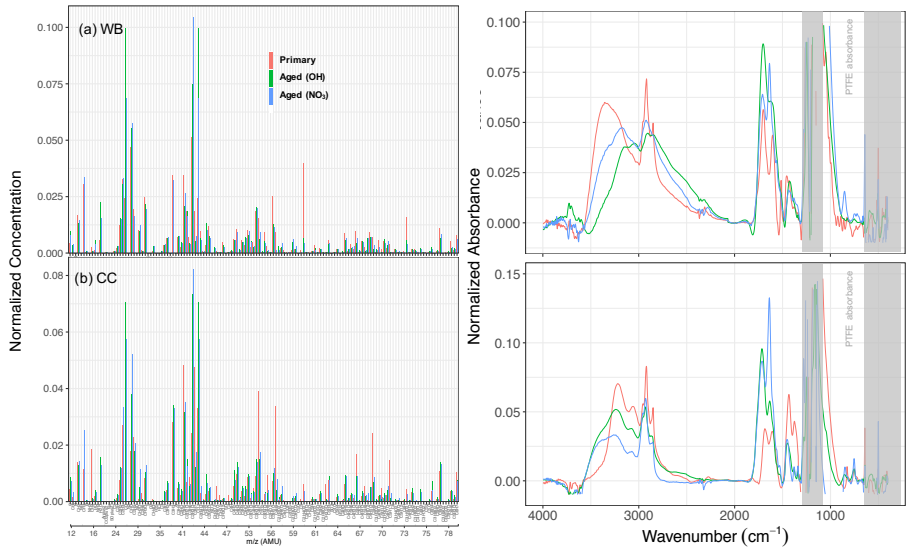

Figure A.16 – AMS normalized mass spectra (shown up to m/z 80) averaged during filter sampling for 20 minutes (left) and normalized mid-infrared spectra (Euclidean norm) of WB (a) and CC (b) on filter.

correlated with the MIR out-of-plane aromatic CH peak absorbance at 754 cm^{-1} ($R^2 = 0.70$; Fig. A.18). This correlation is considerably stronger than that of AMS total OA mass concentration with the MIR out-of-plane aromatic CH peak absorbance ($R^2 = 0.49$), suggesting a physical connection between the AMS fragments and the MIR absorbance at 754 cm^{-1}. The mentioned aromatic fragments contribute up to 6 % to the total OA mass. It is believed that oxygenated

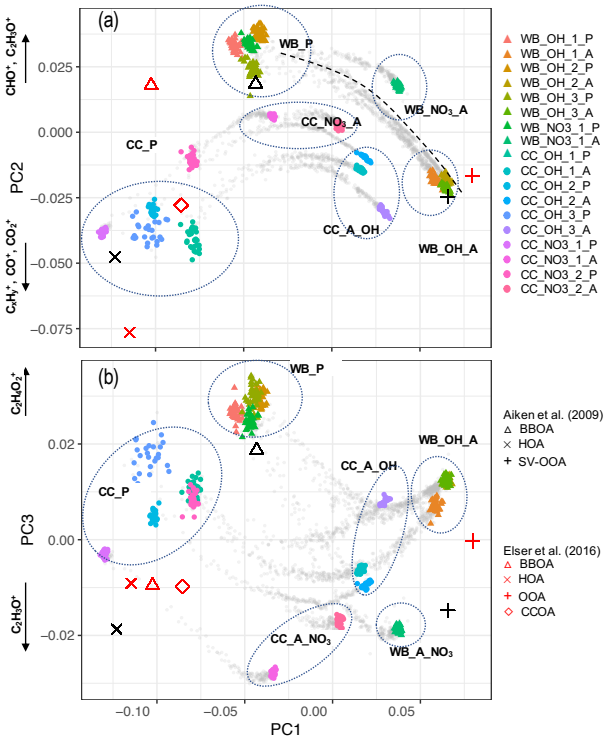

Figure A.17 – Biplots of PC2–PC1 (a), PC3–PC1 (b). The AMS measurements corresponding to filter sampling periods are color coded. The oxidation trajectories out of filter sampling periods are illustrated by gray dots. The clusters of samples based on their fuel and oxidant are indicated by dashed circles.

aromatics undergo fragmentation during electron impact ionization; thus, their contribution is possibly underestimated using AMS fragments.

From the PAH fragment ions suggested by Bruns et al. (2015) and Elser et al. (2016), $C_{10}H_7^+$, $C_{10}H_8^+$, $C_{10}H_9^+$, $C_{12}H_8^+$, $C_{13}H_9^+$, $C_{14}H_{10}^+$, $C_{16}H_{10}^+$ fragment ions were quantified in this work. The mentioned aromatic fragments contribute up to 3 % to the total OA mass. However, the fragments quantified in this work are not complete and oxygenated PAHs are also believed to undergo fragmentation during electron impact ionization (McLafferty and Tureček, 1993); thus, their contribution is possibly underestimated.

Mid-infrared aromatic CH absorbances are, however, not significantly different between oxygenated and non-oxygenated aromatics and PAHs. The mid-infrared OOP peak absorbance normalized by total OA mass can be an indicator of relative contribution of aromatic CH. This parameter is on average four times higher for CC aerosol compared to WB.

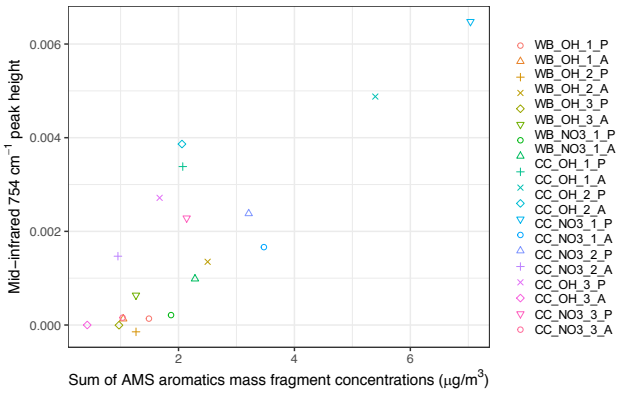

Figure A.18 – Scatter plot comparing MIR absorbance at 754 cm^{-1} (attributed the aromatic CH out-of-plane vibration) and sum of AMS fragment ions attributed to aromatics over the filter sampling periods for each experiment ($R^2 = 0.70$). The corresponding fragments were taken from Bruns et al. (2015) and Pavia et al. (2008).

A.2.7 Levoglucosan detection and quantification using MIR spectroscopy

In previous studies (e.g. Russell et al., 2009a; Faber et al., 2017), the alcohol COH abundances (calculated using MIR spectra) were compared to the abundance of AMS fragment ions attributed to levoglucosan fragmentation ($C_2H_4O_2^+$ and $C_3H_5O_2^+$), often resulting in low correlation coefficients. Levoglucosan, however, absorbs at several specific frequencies in mid-infrared range (860, 890, 920, 950, 990, 1012, and 1050 cm^{-1}) which are not observed for other compounds containing the aCOH group (Interference with other anhydrous sugars, monnasan and galactosan is possible; Fig. A.19a). In addition to lack of these specific absorption other compounds usually have high abundance of non-alcohol FGs that dominate the spectra. As a results, the specific levoglucosan peaks, which become identifiable after blank subtraction, have been used in this study to quantify levoglucosan (an important tracer of biomass burning). By scaling the mentioned peaks in a pure levoglucosan MIR spectrum to that of WB samples, it can be seen that approximately 40 % of the aCOH functional group in WB_OH_2_P (a primary WB sample having high concentration of levoglucosan) is due to levoglucosan (Fig. A.19b). By using the same method, we estimate that levoglucosan is responsible for, on average, 20 % of the aCOH group in primary WB aerosols and less than 10 % in aged WB aerosols, suggesting levoglucosan degradation with aging (Hennigan et al., 2010). The specific levoglucosan absorbances in primary and aged CC aerosols are generally below detection limit of FT-IR (negative absorptions at peak in 860-1050 cm^{-1} region; Fig. A.20), suggesting negligible amounts of levoglucosan in these samples.

The $C_2H_4O_2^+$ fragment ion (m/z 60 in unit mass resolution) in AMS spectrum is shown to be a tracer of levoglucosan (Aiken et al., 2009; Faber et al., 2017; Schneider et al., 2006) in spite of having interference from other molecules, such as long-chain carboxylic acids (Schneider et al., 2006). In order to calculate levoglucosan concentrations from AMS measurements, 0.3 % of OA was subtracted from the $C_2H_4O_2^+$ ion fragment signal (accounting for other molecules producing the same signal) and the remaining signal was called “levoglucosan-equivalent” concentration. The averaged “levoglucosan-equivalent” signal over the filter sampling periods is highly correlated to levoglucosan peak intensities in MIR spectra (Fig. A.20).

The decrease in levoglucosan-equivalent AMS signal (on average 1.11 h^{-1}) and FT-IR levoglucosan absorbances (on average 0.75 h^{-1}) with aging are much more that what can be attributed solely to particle-phase wall losses with a rate constant on the order 0.2–0.3 h^{-1}.

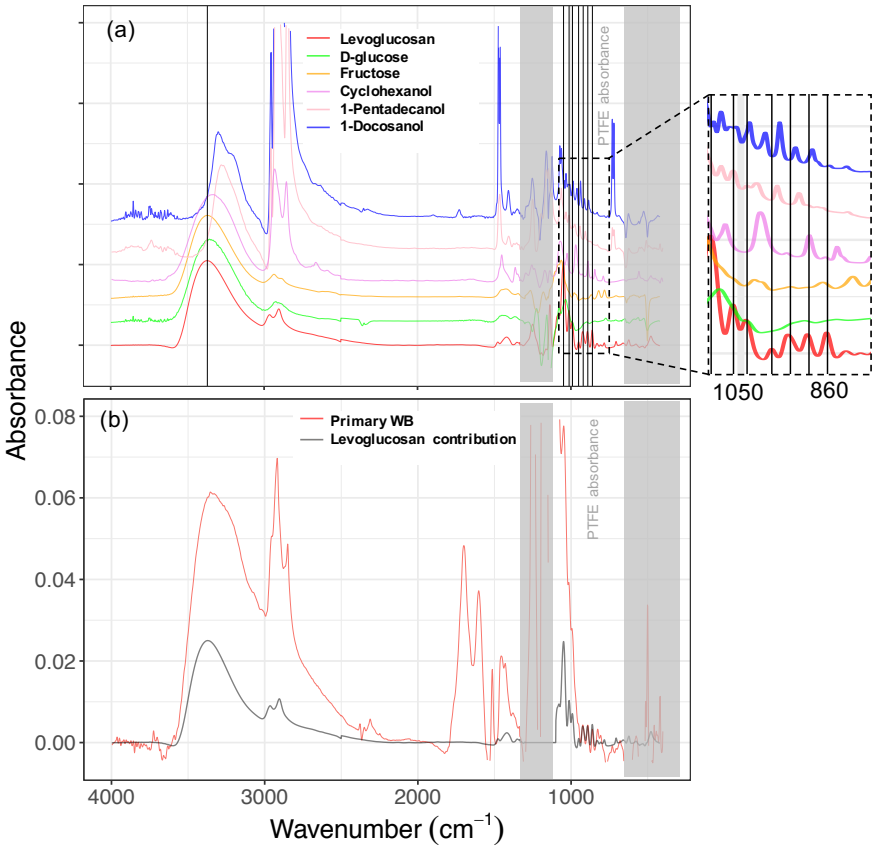

Figure A.19 – (a) Baseline corrected spectra of six different organic compounds containing the aCOH group (e.g. sugars, sugar derivatives, cyclic and straight-chain alcohols). Peak frequencies of levoglucosan are indicated by vertical lines (860, 890, 920, 950, 990, 1012, and 1050 cm^{-1}). (b) Baseline-corrected spectra of fresh wood burning (WB) aerosols and levoglucosan scaled based of its signature at 860–1050 cm^{-1}. The contribution of levoglucosan to the fresh WB aCOH is estimated to be approximately 40 % in this case (WB_OH_2_P). The spectra of d-glucose and fructose are taken from Ruthenburg et al. (2014).

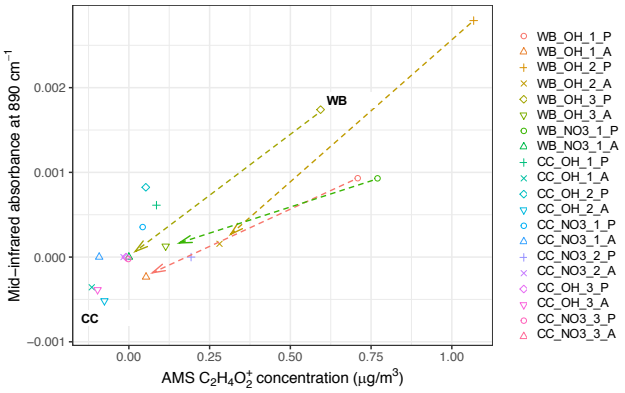

Figure A.20 – Scatter plot comparing the MIR absorbances attributed to levoglucosan and AMS $C_2H_4O_2{}^+$ fragment ion concentration averaged over the filter sampling periods ($R^2 = 0.76$). 0.03 % of OA concentration is subtracted from $C_2H_4O_2{}^+$ mass concentration to account for the non-levoglucosan sources (Aiken et al., 2009). Arrows show the change in levoglucosan with aging.

A.2.8 Lignin signatures

The peak at 1515 cm^{-1} is observed in lignin, WB aerosols, and coniferyl alcohol (a monolignol) due to aromatic ring stretching. A similar peak is observed in para-substituted compounds with a few cm^{-1} shift in the peak frequency from that of lignin. No similar peak is observed compounds without para substitutions. The methoxy peak at 2820 cm^{-1} is also relatively weak for lignin, coniferyl alcohol, and WB aerosols.

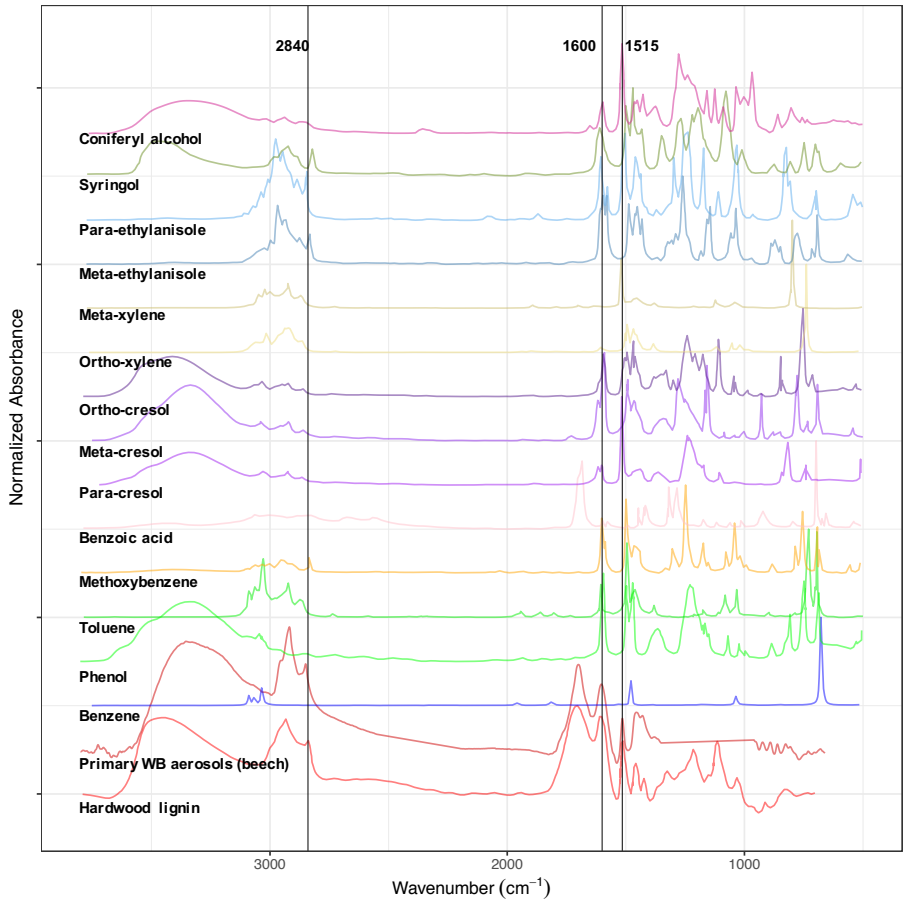

Figure A.21 – MIR spectra of hardwood lignin (KBr Pellet) from Boeriu et al. (2004), fresh beech wood burning aerosols (this work), coniferyl alcohol (film) from Boeriu et al. (2004), and some other substituted aromatic compounds (neat) from SpectraBase database (https://spectrabase.com/).

The peak at 1515 cm^{-1} is correlated with the AMS $C_8H_9O_2{}^+$ fragment ($R^2 = 0.68$), which is attributed G lignin fragmentation (Tolbert and Ragauskas, 2017; Saito et al., 2005).

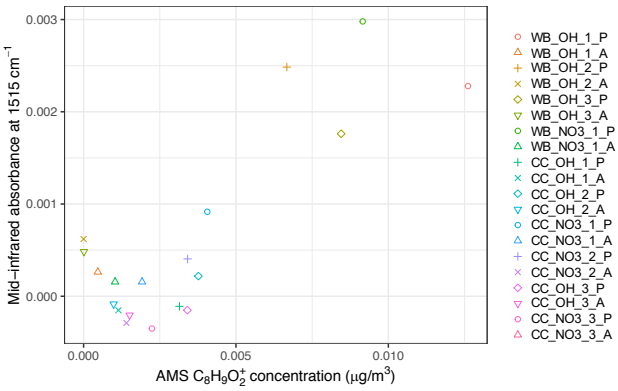

Figure A.22 – Scatter plot comparing the MIR 1520 cm^{-1} peak absorbance, attributed to lignin, and the AMS $C_8H_9O_2{}^+$ fragment ion concentration averaged over the filter sampling periods ($R^2 = 0.68$).

A.2.9 Atmospheric smoke impacted samples

Figure A.23 – Scatter plot comparing total $PM_{2.5}$ and OM mass in atmospheric samples. Orange and red circles indicate residential wood burning and wildfire samples identified by cluster analysis (Bürki et al., 2020), respectively. Blue triangles and green crosses shows burning samples identified based on lignin and levoglucosan signatures, respectively. Black circles indicate the existing atmospheric samples in the IMPROVE network (2011 and 2013; approximately 3050 samples). The dotted dashed lines delineate the range of OM mass fraction for samples designated as burning-influenced (the slope of 0.5 is arbitrarily chosen to guide the eye). OM was estimated by multiplying OC by 1.8 (assuming an average OM:OC ratio of 1.8). The Venn diagram at the top left compares the burning samples identified by each method.

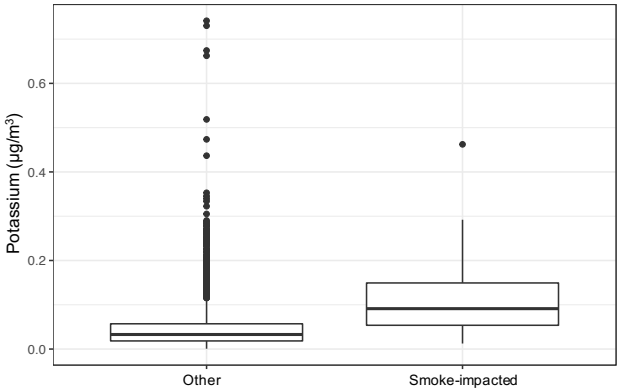

Figure A.24 – Box plot comparing potassium concentration of smoke-impacted and non-smoke-impacted samples in the IMPROVE network from X-ray fluorescence measurements.

A.3 Appendix to chapter 4

A.3.1 Methodology diagrams

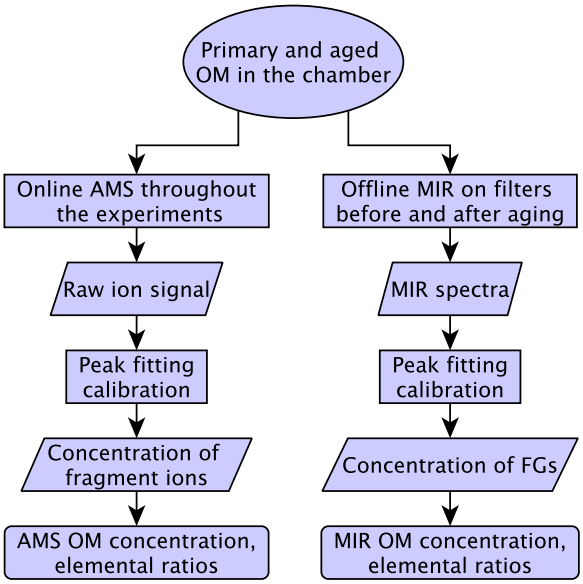

Figure A.25 – Diagram showing the MIR and AMS measurements of bulk OM and the initial postprocessing. Parallelograms, rectangles, and rounded rectangles show data, processes, and final postprocessed data, respectively.

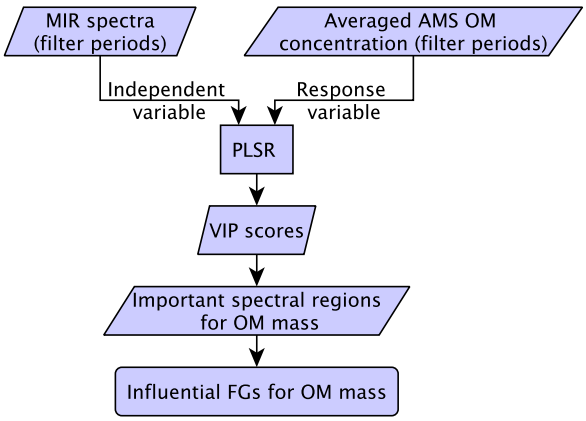

Figure A.26 – Diagram showing how the MIR spectra and AMS OM estimates are combined to determine the influential FGs.

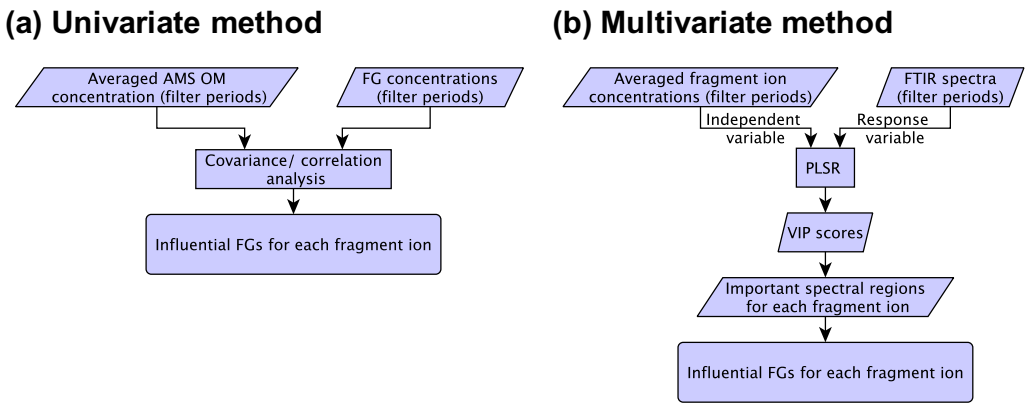

Figure A.27 – Diagrams showing how (a) MIR FG abundances or (b) MIR spectra are combined with the concentration of AMS fragment ions to determine which FGs produce certain fragment.

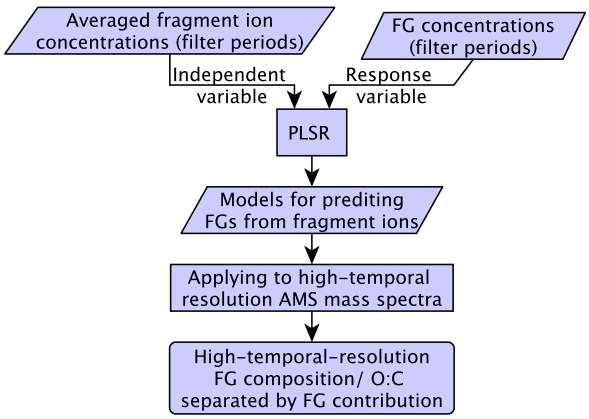

Figure A.28 – Diagram showing how high-temporal-resolution FG compositions are calculated by combining AMS and MIR measurements.

A.3.2 AMS fragment ion-FG connections

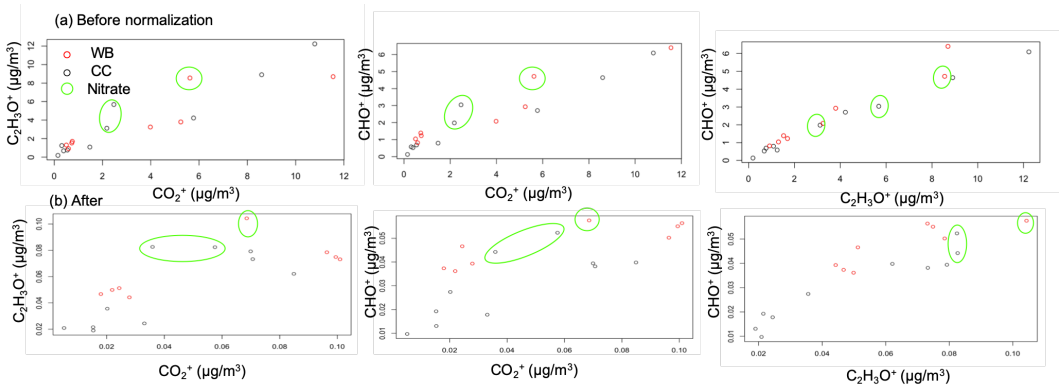

Figure A.29 – Scatter plot highlighting the correlation between major AMS mass fragments before and after normalization by total AMS OM. Green circles show the experiments for which the nitrate radical was used as oxidant.

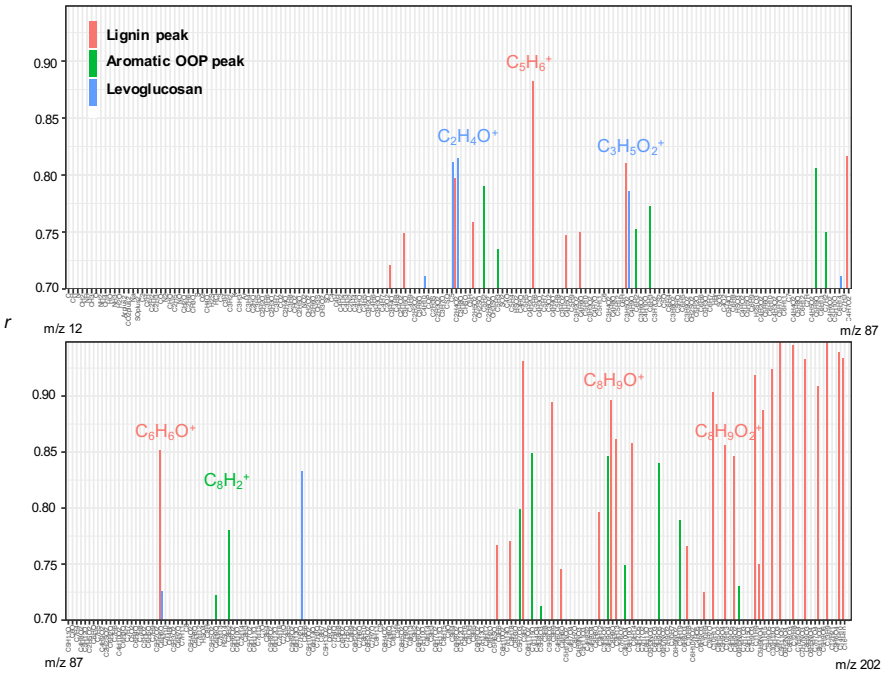

Figure A.30 – Barplot showing the Pearson correlation coefficients of lignin, levoglucosan and out-of-plane aromatic peaks in the MIR spectra and the concentration of AMS fragment ions. Only fragments with $r > 0.7$ have been shown.

Fig A.32 shows the results of multivariate linear regressions that regress standardized (unit standard deviation and zero mean) FG abundances against standardized concentrations of three major fragments (and vice versa). The parameter selection is done using an stepwise algorithm working based on the Akaike information criterion (AIC). Only experiments with the hydroxyl radical are considered as aging with the nitrate radical causes the mentioned

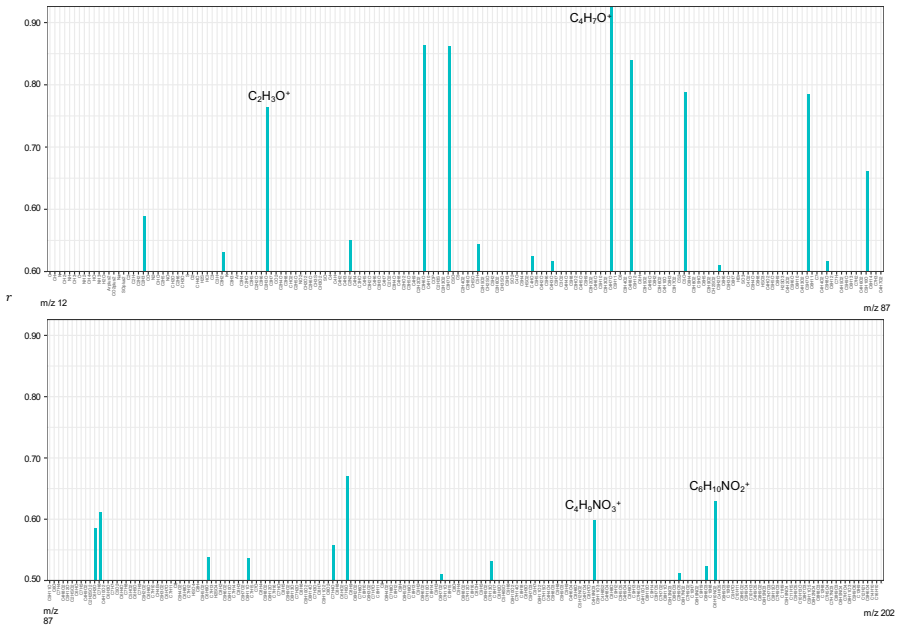

Figure A.31 – Barplot showing the Pearson correlation coefficient of $RONO_2$ peak in MIR spectra and the concentration of AMS organic fragment ions. Only fragments with $r > 0.5$ have been shown.

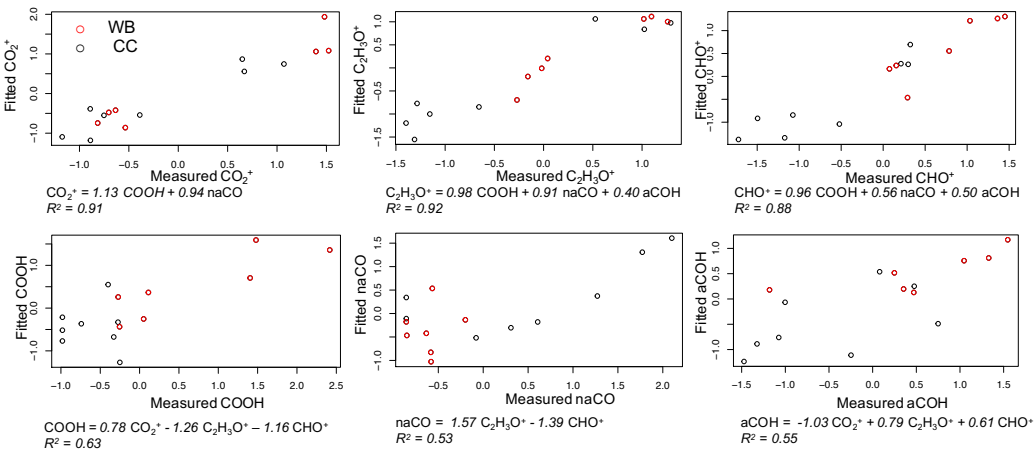

Figure A.32 – Scatter plots showing the measured and calculated AMS fragment concentrations using MIR FG concentrations and vice versa. The FG and fragment concentrations are standardized (zero mean and unit standard deviation) to make regression coefficients comparable. Regression equations are shown below each plot. The experiments for which the nitrate radical was been used have been omitted due to causing poor fit results.

fragments to vary significantly without affecting the abundance of oxygenated FGs except organonitrates.

A.3.3 PLS models

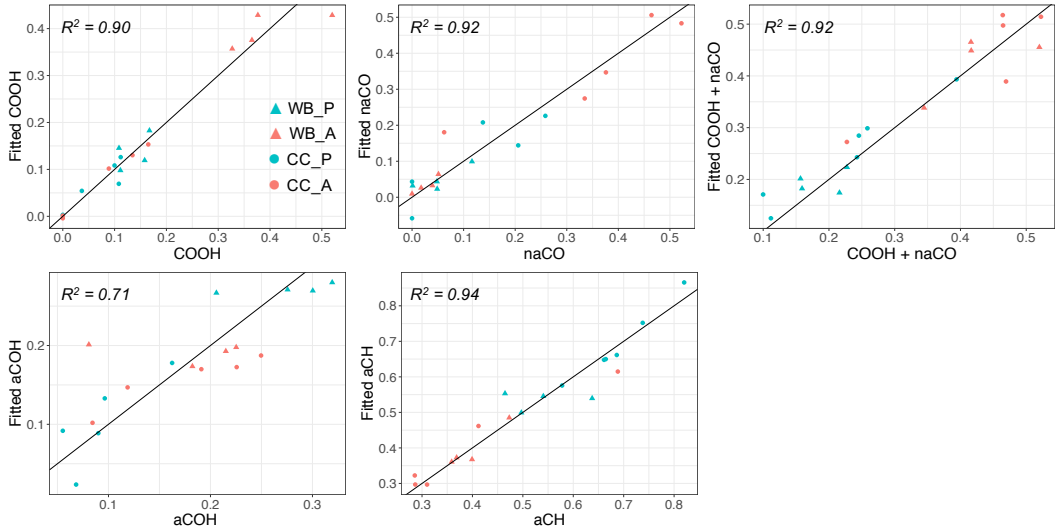

Figure A.33 – Scatter plots showing fitted (PLSR models) normalized functional group composition using AMS mass spectra against their values from the MIR peak fitting.

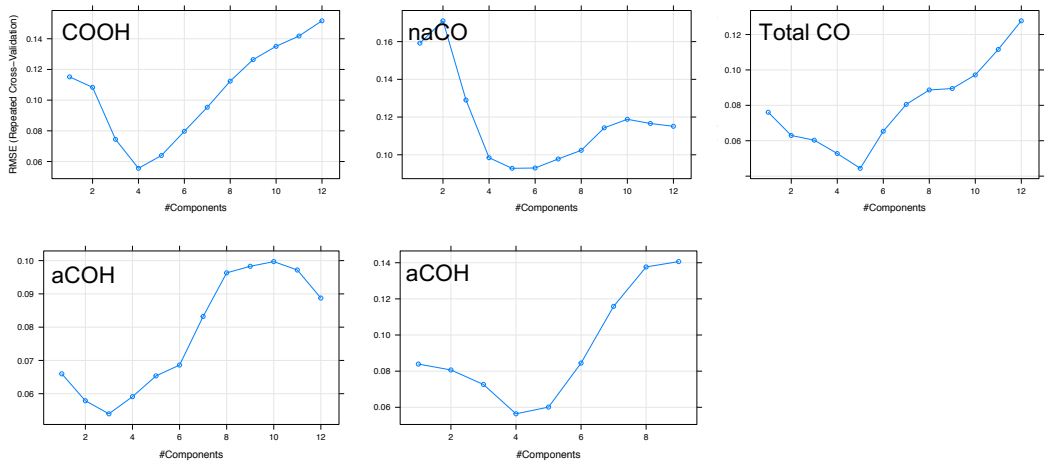

Figure A.34 – Plots showing the root mean squared error (RMSE) of predictions in the 10-fold cross validation for different number of latent variables in the PLSR models. The resulting PLSR models are used to predict normalized FG abundances using the AMS spectra.

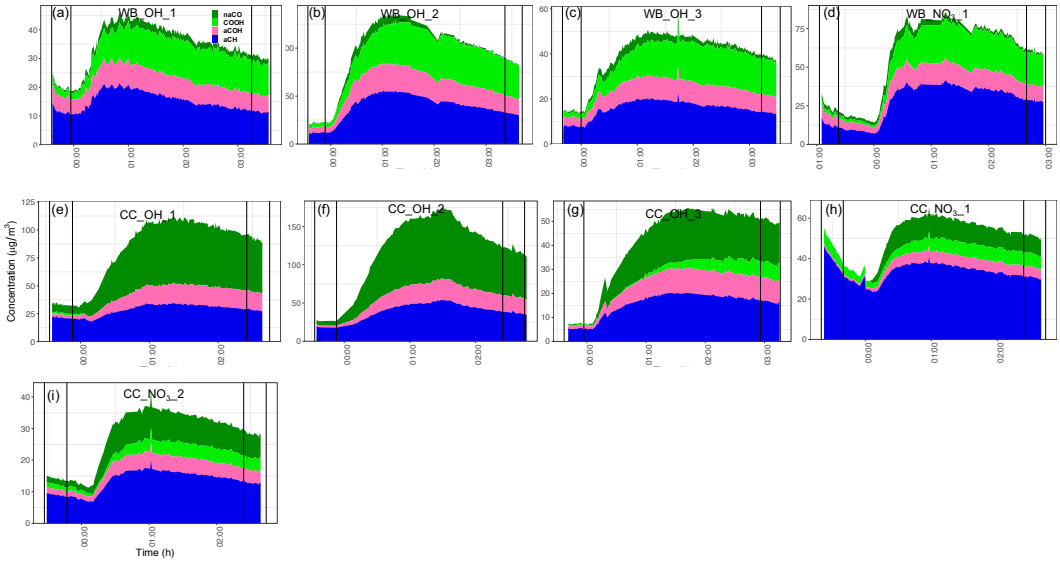

Figure A.35 – Time series of high-resolution stacked absolute functional group concentrations calculated using the AMS mass spectra. Total CO refers to the sum of COOH and naCO. The naming is the same as Fig. 4.4. The absolute concentrations are calculated by multiplying normalized concentrations by the total AMS OM and are not corrected for chamber wall losses.

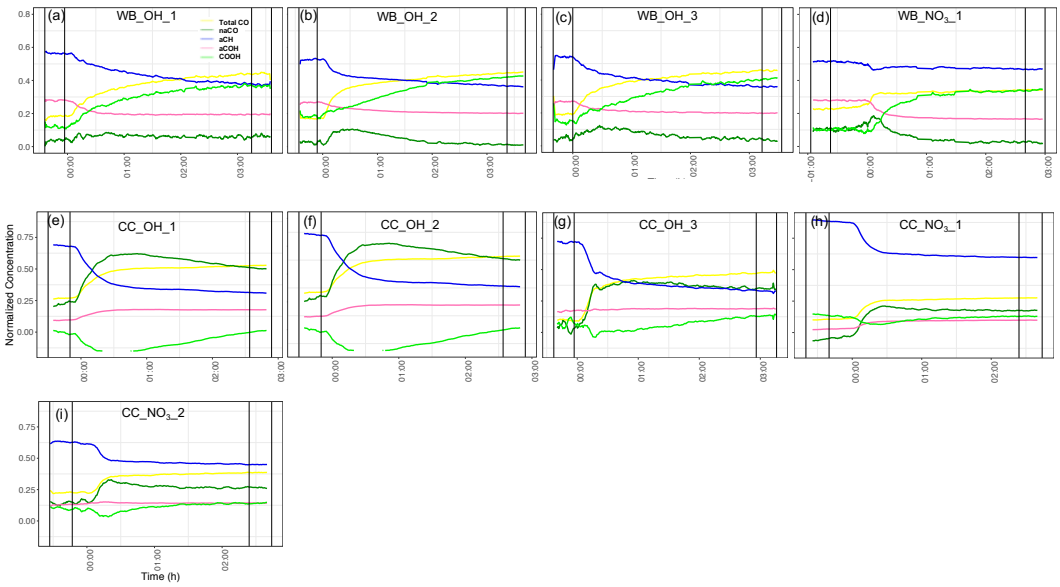

Figure A.36 – Time series of high-resolution normalized functional group concentrations calculated using the AMS mass spectra. Total CO refers to the sum of COOH and naCO. Vertical black lines show the filter sampling periods before and after aging. The naming is the same as Fig. 4.4 of the main manuscript. The time zero indicates the start of chemical aging.

As can be seen from Fig. A.36, the predicted COOH concentrations are slightly negative for two CC experiments (CC_OH_1 and CC_OH_2) during the course of aging. These unphysical values, which are not substantial compared to total OM concentration (up tp 15 % of total

OM), might be attributed to: uncertainties of each instrument especially uncertainties in MIR COOH estimation for primary CC samples; model uncertainties (especially as models were developed using very different experiments), and errors introduced due to the use of models to predict compositions fairly different from those of filter sampling periods. However, the dynamic trend of COOH is still captured in these cases.

Figure A.37 compares the oxidation trajectories from AMS and the reconstructed trajectories from the elemental ratios calculated from the interpolated FGs. The constructed trajectories are able to capture the curvatures observed in the original trajectories although some discrepancies are observed in the absolute values.

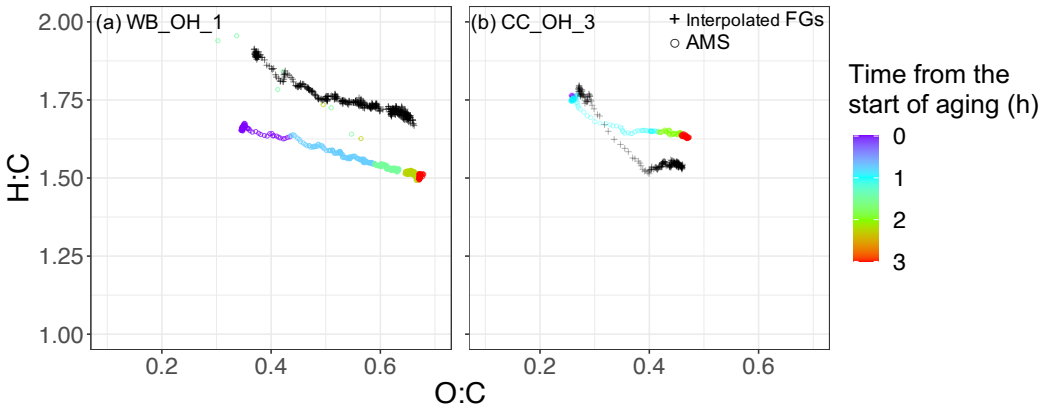

Figure A.37 – Comparison between van Krevelen trajectories from the AMS elemental ratios (circle) and elemental ratios reconstructed from the interpolated FGs (cross) for a WB and a CC experiment.

A.4 Appendix to chapter 5

A.4.1 Residual spectrum calculation

The procedure for obtaining FTIR residual spectra is explained an illustrated in an example (Exp. 5). First, a first-order loss is fitted to the AMS OA before the start of aging (in the −1.5–0 h range). The fitted curve is shown by a dashed curve in Fig. A.38a). The concentration of the apparent POA during the sampling of the second filter (F2) is estimated based on the fitted curve to be 0.68 that of primary aerosols (F1), assuming that particle wall loss is the dominant process for POA. The spectrum of apparent POA during the sampling of the second filter (F2) is estimated by scaling by 0.68 the spectrum of the first filter (F1) based on the first-order loss curve. The residual spectrum for the second filter (black, Fig. A.38b) is calculated by subtracting the apparent POA spectrum from that of the second filter (F2). Negative peaks observed in the residual spectrum suggest processes other than particle wall loss are also important for the POA during the course of aging.

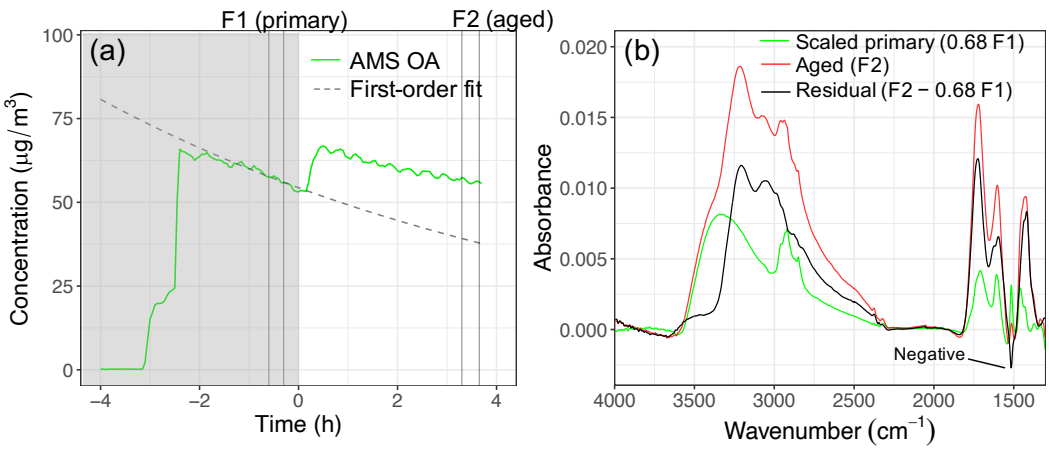

Figure A.38 – AMS OA concentration and apparent POA concentration using a first-order wall loss for Exp. 5 (a). Residual FTIR spectrum based on the wall loss rate estimated from AMS OA (b). Vertical lines indicate the periods of sampling for primary and aged aerosols.

A.4.2 AMS-FTIR comparison

OA concentrations estimated by FTIR and AMS are correlated ($r^2 = 0.75$; Fig. A.40). Aging with UV increases the OM:OC ratio more than the nitrate radical. We also observe that the general increase in OM:OC with aging is captured by both methods (Fig. A.40c). However, the absolute values are different. PTFE filters that belong to each experiment are as follows: Exp. 1 (filters F1, F2), Exp. 2 (F9, 10), Exp. 3 (F16, 17), Exp. 4 (F18, 19), Exp. 5 (F20, 21), Exp. 6 (F4, 5), Exp. 7 (F11, 12), Exp. 8 (F6, 7), and Exp. (F13, 14). F3, 8, and 18 are blank filters put in the chamber for few minutes. AMS OA is not corrected for CE.

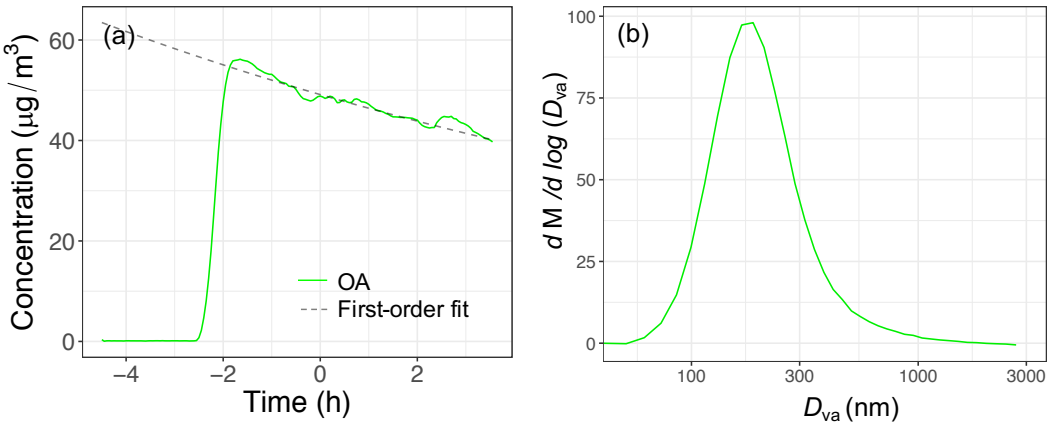

Figure A.39 – AMS OA concentration and the first-order fit for Exp. 1 (reference (a). Injection of biomass burning emissions happens at -2 h. AMS size-resolved organic mass. The results are based on particle time-of-flight data averaged from -2 to 2 h (b).

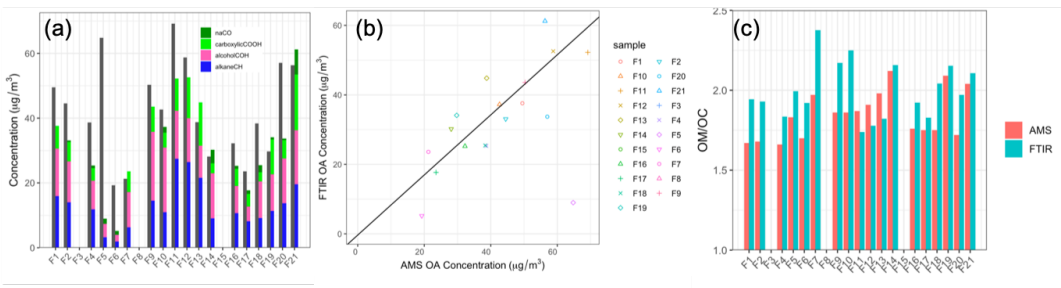

Figure A.40 – (a) OA mass concentration measured by FTIR (functional groups are color codded) and AMS (averaged over filter sampling periods). (b) Scatter plot comparing OA concentrations measured by AMS and FTIR. (c) OM:OC ratios measured by AMS and FTIR.

A.4.3 FTIR spectra of primary biomass burning aerosols

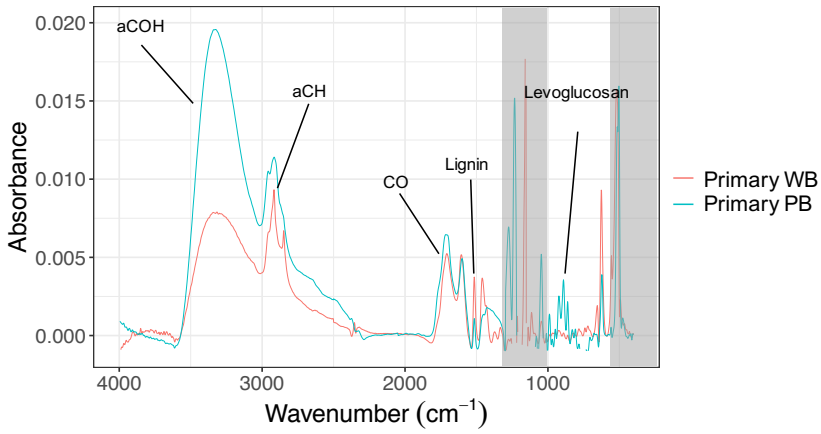

Figure A.41 – FTIR spectra of primary wood burning (WB) and pellet burning (PB) aerosols. Important functional groups or biomass burning markers are indicated.

A.4.4 Investigating lignin-related fragments

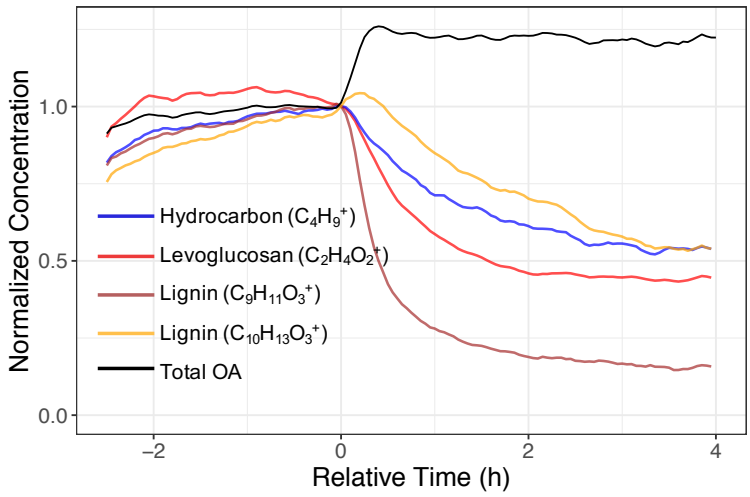

Figure A.42 – Smoothed time series of AMS OA concentration, and different tracer fragments in a UV experiment (Exp. 4).

A.4.5 POA transformation

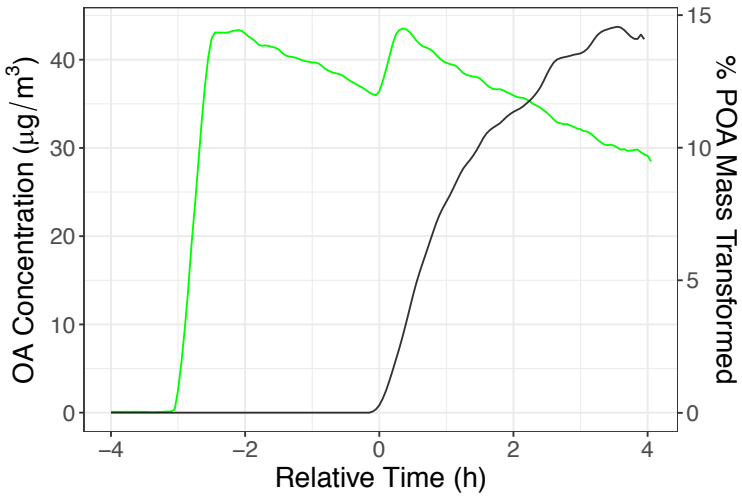

Figure A.43 – Smoothed time series of AMS OA concentration (green) and the percentage of POA transformed (black) in a UV experiment (Exp. 4).

A.4.6 Investigating the volatility of lignin-like compounds

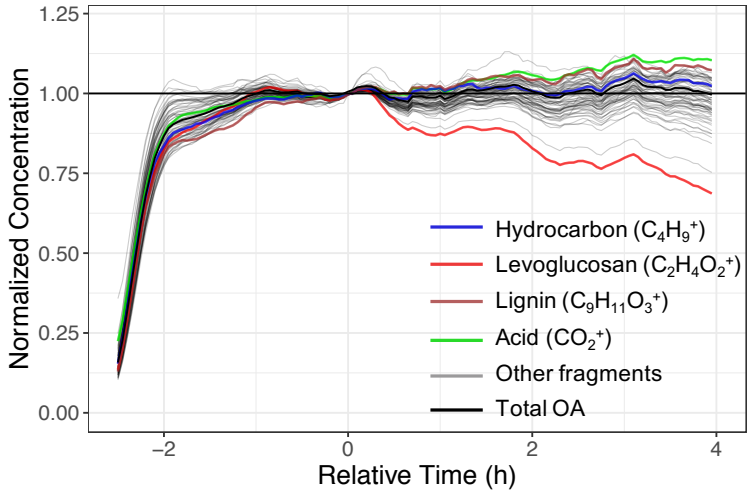

Figure A.44 – Normalized concentrations of AMS fragments in a reference experiment (Exp. 3).

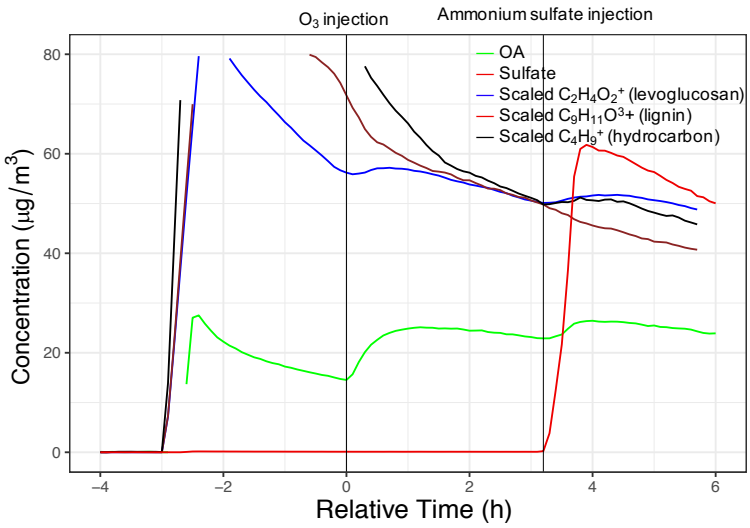

Figure A.45 – Smoothed time series of normalized AMS OA and sulfate concentration, and mass fragments attributed to certain tracers for a dark aging (nitrate) experiment in dry conditions.

In order to estimate the volatility of different species responsible for marker fragments, we assumed that the equilibrium state in gas-particle partitioning is reached at all temperatures and used the equation below:

$$A_i(T) = \frac{G_i(T) + A_i(T)}{1 + C_i^\star(T)/M(T)}, \tag{A.10}$$

where $A_i(T)$ and $G_i(T)$ are aerosol-phase and gas-phase concentrations of the fragment i at temperature T and $M(T)$ denotes the total organic mass. The sum of $A_i(T)$ and $G_i(T)$ are assumed to be the same between the bypass line and the thermodenuder line. The

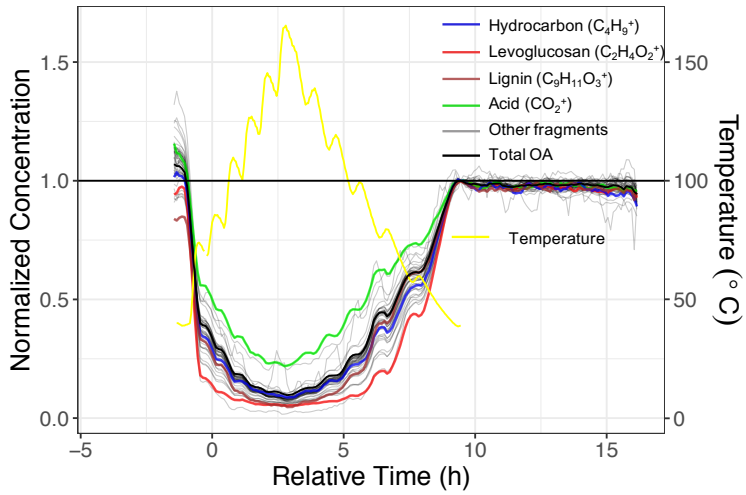

Figure A.46 – Normalized concentrations of individual AMS fragments after passing chamber OA through a thermodenuder.

Clausius-Clapeyron equation was used to relate saturation concentrations, $C_i^\star$ at different temperatures:

$$C_i^\star(T) = \frac{298}{T}\, C_i^\star(298)\, \exp(\frac{\Delta H_v}{R}(\frac{1}{T} - \frac{1}{298})). \tag{A.11}$$

We neglected particle losses due to thermophoresis. The total organic concentration has been assumed to be the value estimated by AMS considering a collection efficiency of unity. ΔH_v is assumed to be equal to 40 kJ mol^{-1}. The Levenberg-Marquardt nonlinear least-Squares algorithm was used to find the optimum $C^\star$ for different marker fragments. For $C_2H_4O_2^+$, $C_9H_{11}O_3^+$, $C_{10}H_{13}O_3^+$, CO_2^+, and $C_4H_9^+$, the $C^\star(298)$ was estimated to be 5.42, 1.83, 1.32, 0.76, and 2.94 µg m^{-3}, respectively. The $C_2H_4O_2^+$ is produced by the most volatile species (levoglucosan, anhydrosugars) and the CO_2^+ by the least volatile ones. However the differences are within one order of magnitude for all the fragments above. Due to simplifying assumptions (e.g., omission of thermophoresis or equilibrium state) these values should be considered as order of magnitude estimates. The value of $C^\star$ estimated for the levoglucosan is slightly lower than previous studies (Bertrand et al., 2018b) but still reasonably close considering the simplifying assumptions made here.

A.4.7 Ammonium subtracted spectra

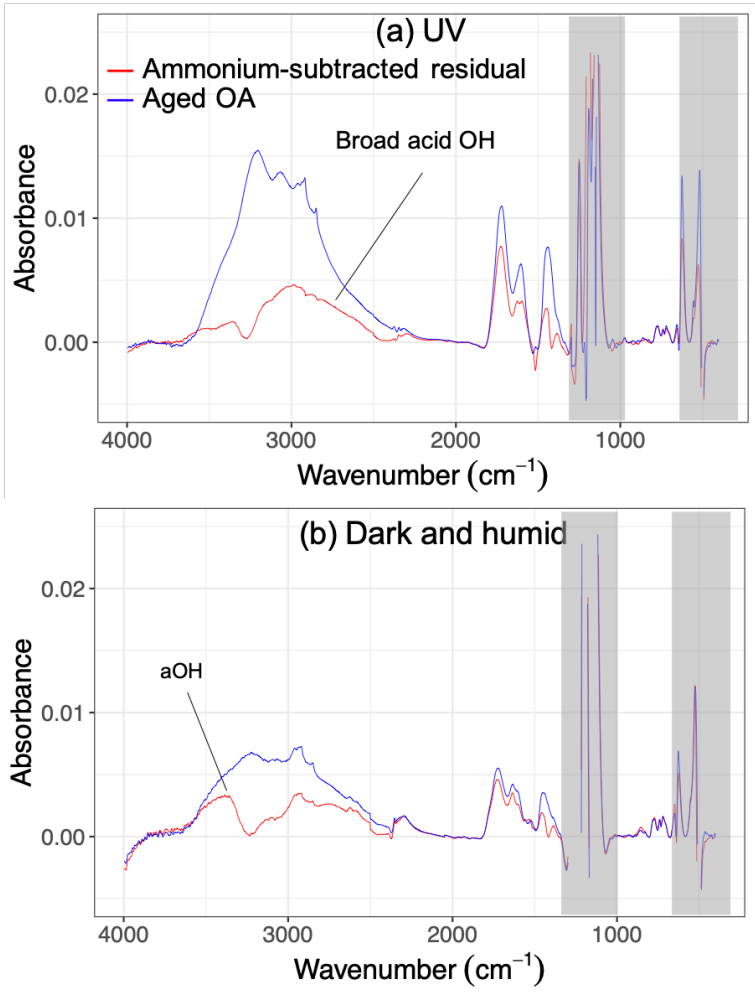

Figure A.47 – Ammonium-subtracted residual FTIR spectra of wood burning OA aged with UV and nitrate radical (Exp. 4, and 8, respectively).

A.4.8 f_{44}:f_{43} ratios in residual spectra

In Fig. A.48 the wall-loss corrected OA was calculated base on

$$C_{\mathrm{OA}}^{\mathrm{cor}}(t) = C_{\mathrm{OA}}^{\mathrm{obs}}(t) + k_{\mathrm{OA}} \int_{t_0}^{t} C_{\mathrm{OA}}^{\mathrm{obs}}(t)\, dt, \tag{A.12}$$

where $C_{\mathrm{OA}}^{\mathrm{obs}}(t)$ is the observed (measured) OA concentration at time t, $C_{\mathrm{OA}}^{\mathrm{cor}}(t)$ is the wall-loss-corrected OA at the t, k_{OA} is the first-order wall loss rate based on the AMS OA. OC was calculated from the sum of C concentration for all fragments. Dividing OA time series by OM:OC time series gives a similar trend with different absolute values.

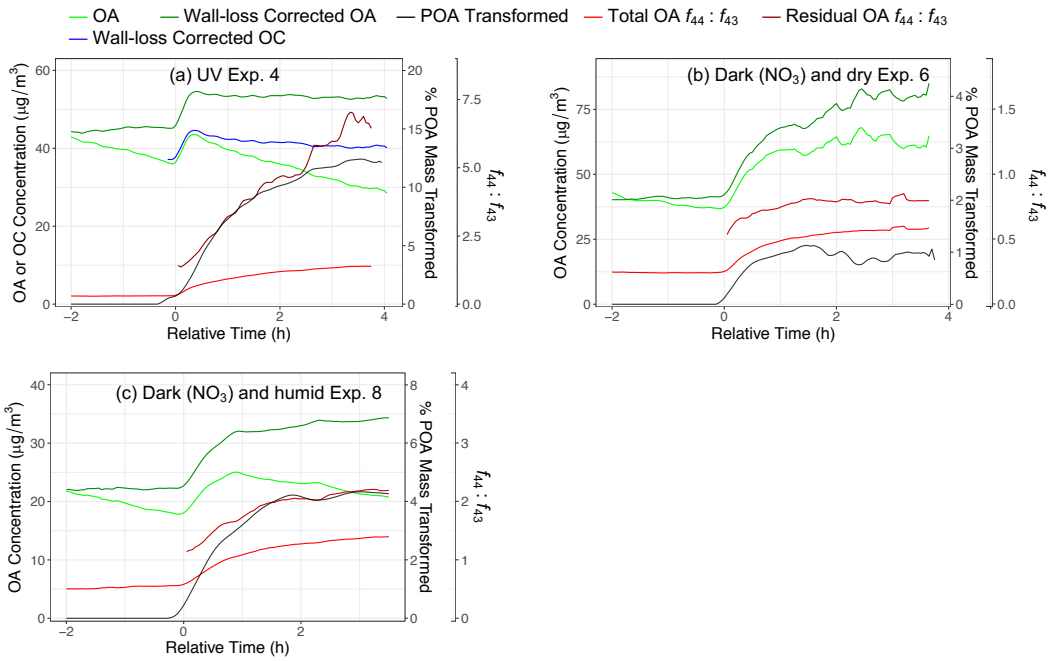

Figure A.48 – Time series of OA, OC concentrations, percentage of POA aged, and f_{44}:f_{43} for wood burning emissions in different aging scenarios.

A.4.9 Dimensionality reduction of AMS mass spectra

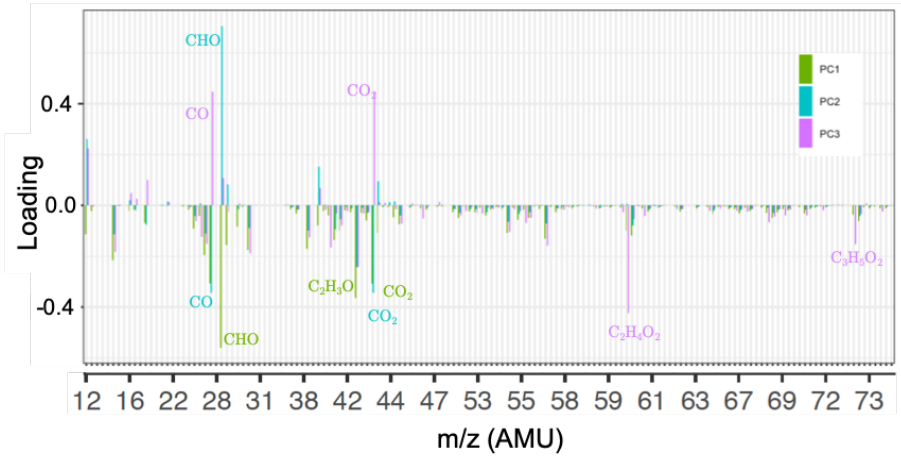

Figure A.49 – Loadings of the first three PCs.

A.4.10 Atmospheric smoke impacted FTIR spectrum

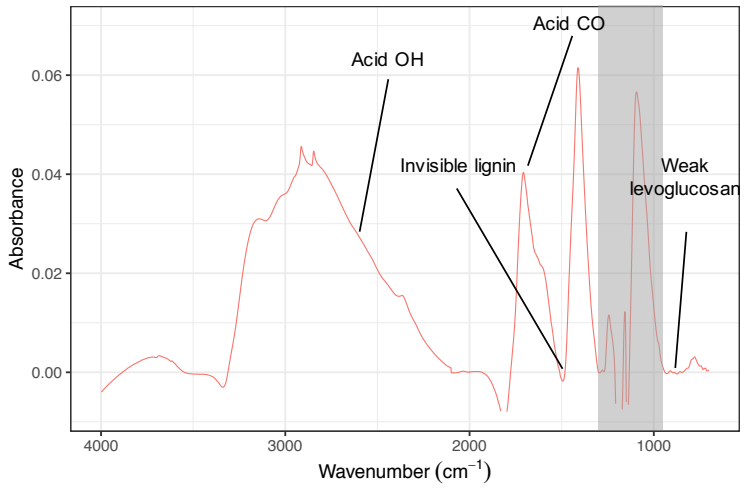

Figure A.50 – An example of atmospheric smoke impacted $PM_{2.5}$ samples (prescribed burning in the eastern US 2013, validated by satellite observations; Yazdani et al., 2021e) with high organic loading and high acid signatures and very weak levoglucosan and invisible lignin-like signatures.

A.5 Appendix to chapter 6

A.5.1 Contribution of fresh bbOA to OM

Assuming that levoglucosan constitutes on average 8 % of fresh bbOA (Fine et al., 2002), WSOC/OC is 0.42 (regression line slope in Fig. A.51), and the average OM:OC ratio is 1.8, the levoglucosan C/WSOC > 0.01 threshold is roughly equivalent to bbOA ($\mu g\,m^{-3}$)/OM ($\mu g\,m^{-3}$) > 0.07.

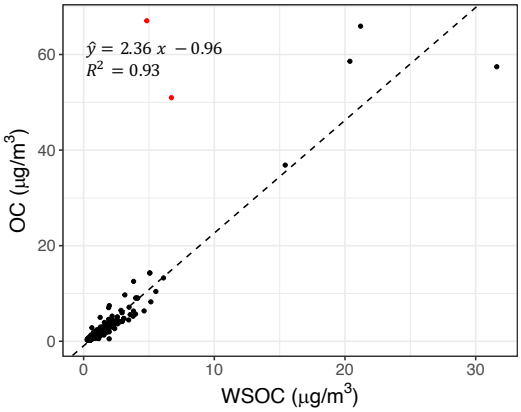

Figure A.51 – Scatter plot of TOR OC and WSOC. The dashed line shows the regression line. Red points are not considered in the regression.

A.5.2 The carbonyl peak as proxy to WSPC

This carbonyl peak represents non-acid carbonyls (e.g., ketones, aldehydes, esters) and carboxylic acids, which are polar and hydrophilic functional groups. The hydroxyl group was not considered here as its quantification is more involved and uncertain due to the significant overlap of the OH stretching band with ammonium NH stretching peaks in the 3200–3400 range cm^{-1}. In addition, the estimate of carbonyl abundance has been shown to be robust with respect to different estimation methods due to its well-defined peak Reggente et al. (2019a). The high correlation ($R^2 = 0.94$ omitting one point shown with the red circle) of WSOC measured by the carbon analyzer and the carbonyl peak in the FTIR spectra can be seen from Fig. A.52. The correlation holds reasonably well ($R^2 = 0.77$) even for samples with low WSOC concentrations (< 2 $\mu g\,m^{-3}$), suggesting that using the carbonyl peak as a proxy to WSOC does not introduce large errors for low-concentration samples, for which the identification of smoke is already challenging.

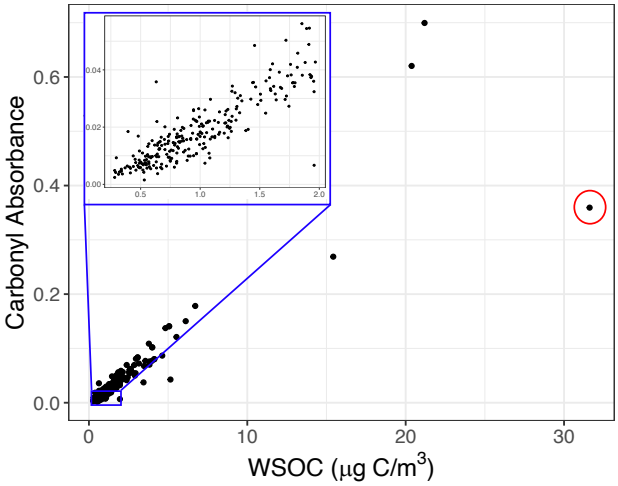

Figure A.52 – Scatter plot of the carbonyl peak absorbance in the FTIR spectra against WSOC from the carbon analyzer instrument.

A.5.3 Levoglucosan fingerprint in FTIR spectra of laboratory standards

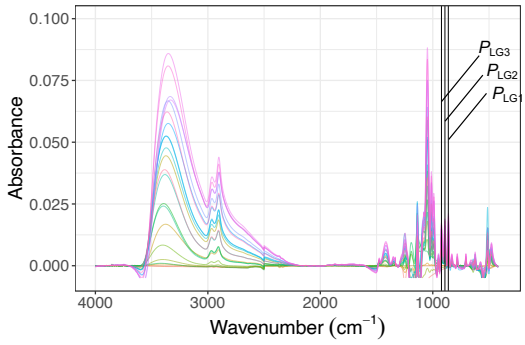

Figure A.53 – Baseline-corrected and blank-subtracted FTIR spectra of laboratory standards of aerosolized (atomized) levoglucosan collected on PTFE filters. The three peaks at 860, 890, and 920 cm^{-1} in the fingerprint region that have been used for the identification levuglucosan have been shown with vertical lines.

A.5.4 Selectivity of levoglucosan fingerprint signatures.

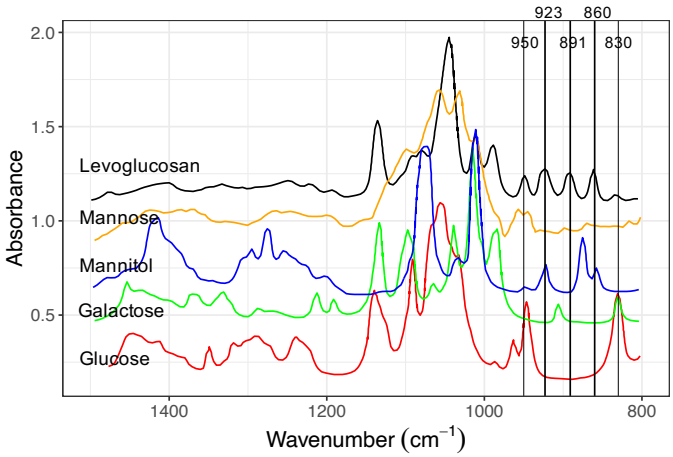

Figure A.54 – Analysis of interference of the fingerprint signatures of levoglucosan. The mid-infrared spectra of levoglucosan (transmission, KBr wafer), mannose (transmission, KBr wafer), galactose (transmission, KBr wafer), glucose (transmission, KBr wafer), and mannitol (transmission, KBr wafer) are taken from the SpectraBase database (https://spectrabase.com/).

A.5.5 Satellite vs. FTIR smoke detection

Table A.3 – Table comparing satellite and FTIR smoke detection methods for the period of prescribed burns and RIM Fire in 2013. X and O indicate the detection of levoglucosan and lignin-like compound signatures in the FTIR spectra respectively. M and P refer to “maybe" smoke-impacted and “probably " smoke-impacted categories by satellite.

(a) Prescribed burns 2013

Date	Okefenokee		St. Marks		Cape Romain	
	Satellite	FTIR	Satellite	FTIR	Satellite	FTIR
2013-01-13	M	O			M	
2013-01-16	P	X			M	
2013-01-19	M	XO		XO		
2013-01-22	M	X		XO	M	XO
2013-01-25	M	X	M	X	M	X
2013-01-31			M	XO		
2013-02-03		XO		XO		XO
2013-02-06					P	X
2013-02-18			M	X		O
2013-03-02		X				X
2013-03-05		XO			M	X
2013-03-08	M	X	P	XO	P	XO
2013-03-14	M	X	P	XO	M	X
2013-03-17	M					
2013-03-23						X
2013-03-29	P	X	P	X	P	XO
2013-04-10			M		M	

(b) RIM fire 2013

Date	Yosemite		Bliss S. P.		Hoover	
	Satellite	FTIR	Satellite	FTIR	Satellite	FTIR
2013-06-18	M	XO				
2013-06-21				O		
2013-06-30	M		M	XO	M	
2013-07-03	M		M		M	
2013-07-09			M			
2013-07-18	P		P		P	
2013-07-24	M	X			M	
2013-07-27	M	X			M	XO
2013-07-30	P	XO	P	XO	P	XO
2013-08-02	M	XO			P	XO
2013-08-05	M	XO	M		P	XO
2013-08-08	M		M		M	
2013-08-11	M	X			M	
2013-08-17		O	M	XO		
2013-08-20			M	XO		
2013-08-23	M				P	XO
2013-08-26	M		P	XO	P	XO
2013-08-29	P		P	XO	P	
2013-09-01	P	XO	M	XO	P	XO
2013-09-04	M		M	X	M	
2013-09-07	P		P	XO	M	XO
2013-09-10	P	XO	M	XO	M	O
2013-09-13	P	XO			M	
2013-09-16	M		M	O		
2013-09-19	M		M	O		
2013-09-28				XO		
2013-10-22	M					
2013-10-31				O		

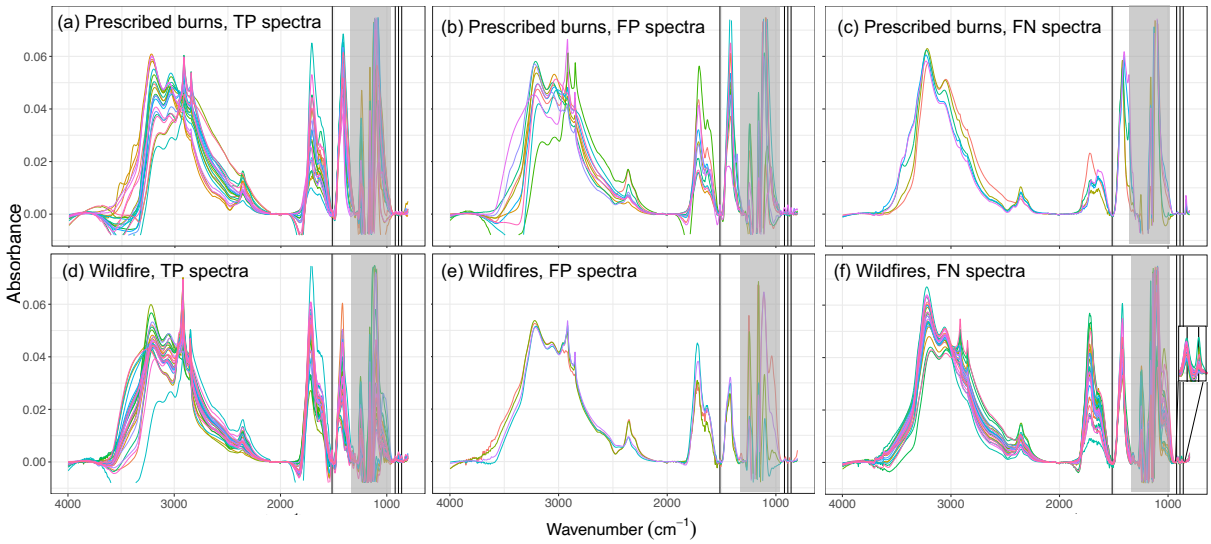

Figure A.55 – Normalized spectra of samples with true positive, false positive, and false negative results from FTIR smoke detection compared to satellite observations.

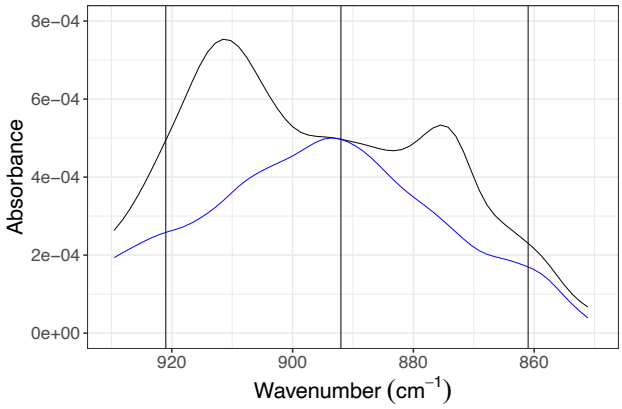

Figure A.56 – Levoglucosan region before and after subtraction of inorganic interference in a smoke-impacted sample. Vertical lines show the expected location of levoglucosan peaks.

A.5.6 HPAEC-PAD vs. FTIR smoke detection

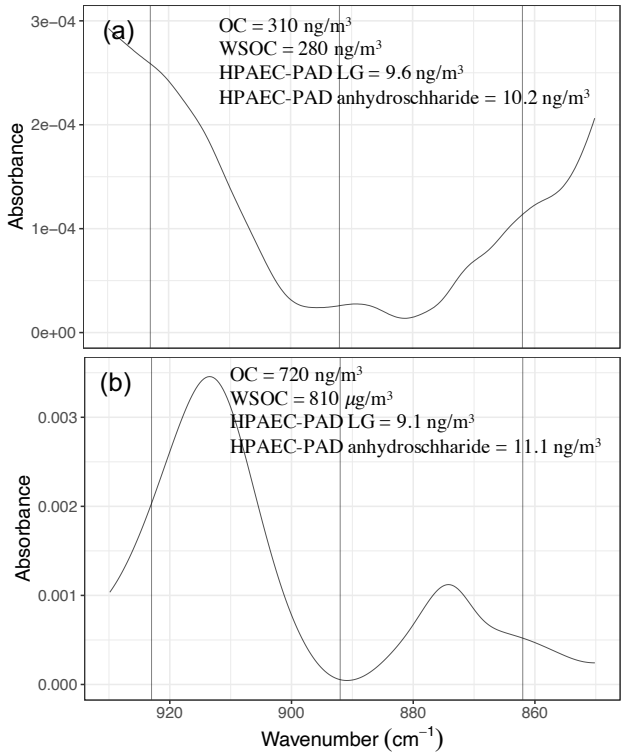

Figure A.57 – Levoglucosan region in two samples with low levoglucosan loadings. Vertical lines show the expected location of levoglucosan peaks. TOR-OC, WSOC, and concentrations of anhydrosaccharides from HPAEC-PAD are shown for each sample.

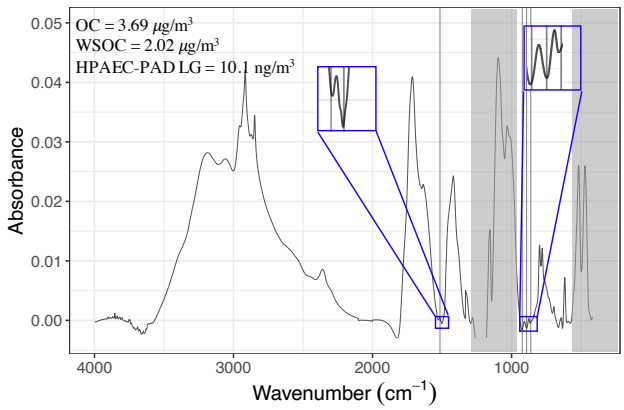

Figure A.58 – FTIR spectrum of an atmospheric sample with visible lignin-like compounds signature, while levoglucosan signatures are not detected due to the low concentration and/or existence of interfering peaks. TOR OC, WSOC, and concentration of levoglucosan from HPAEC-PAD are shown.

Estimating HPAEC-PAD-equivalent anhydrosaccharide concentrations using FTIR and PLSR

As discussed in the main text, in the levoglucosan fingerprint region (830–950 cm^{-1}) in the FTIR spectra of ambient samples, often two interfering peaks from inorganics are observed. This interference complicates the quantification of levoglucosan especially for samples with low levoglucosan loadings. Generally, HPAEC-PAD and FTIR levoglucosan measurements are highly correlated. However, the correlation becomes significantly weaker for the subset of samples with HPAEC-PAD levoglucosan below 0.5 $\mu g\,m^{-3}$ when FTIR levoglucosan is estimated simply from the average height of its three peaks. This phenomenon is probably due to the significant the interference from inorganic peaks and an imperfect baseline correction. In order to alleviate this issue, the absorbances in the 830–950 cm^{-1} region were regressed against HPAEC-PAD sum of anhydrosaccharides using partial least squares regression (PLSR; Wold et al., 1983). The regression was performed for the subset of samples with levoglucosan, mannosan, and galactosan concentrations above the detection limit of HPAEC-PAD (90 samples).

PLSR models have been shown to be able to quantify organic functional groups in atmospheric samples without the need for prior baseline correction or peak fitting (e.g., Reggente et al., 2019a) in the presence of interfering inorganic peaks (e.g., ammonium). As a result, PLSR is expected to perform better than a simple peak height measurement for the quantification of levoglucosan in the presence of interfering peaks. We used PLSR to find the coefficients of the following linear equation, $\boldsymbol{a}$:

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{a} + \boldsymbol{e}, \tag{A.13}$$

where $\mathbf{X}$ $(n \times p)$ is the FTIR spectra matrix with n samples and p wavenumbers in the mentioned region, $\boldsymbol{y}$ $(n \times 1)$ is the vector of HPAEC-PAD sum of anhydrosaccharides, and $\boldsymbol{e}$ is the vector of residuals. After solving the regression equation using PLSR, a 10-fold cross validation was used to choose the optimum number of latent variables.

Applying PLSR to all samples results in a fit with $R^2 = 0.92$, which is very similar to that obtained when regressing HPAEC-PAD sum of anhydrosaccharides against $P_{\mathrm{LG,\,ave}}$. This is because in this case high-concentration samples dominate the regression and for these samples, $P_{\mathrm{LG,\,ave}}$ is minimally affected by interference or an imperfect baseline correction. If we apply PLSR on the subset of samples with HPAEC-PAD levoglucosan < 0.5 $\mu g\,m^{-3}$, we obtain $R^2 = 0.91$ for the subset. This value suggests a significant improvement over the regression coefficient of the HPAEC-PAD sum of anhydrosaccharides against $P_{\mathrm{LG,\,ave}}$ ($R^2 = 0.48$). The PLSR model developed on the samples with low levoglucosan concentrations also performs reasonably well for the whole dataset, including high-concentration samples ($R^2 = 0.85$; Fig. A.59a).

As can be seen from Fig. A.59b, the regression coefficients (from PLSR) in the (830–950 cm^{-1}) range are maximum and positive around the expected locations for the levoglucosan peaks and negative of close to zero in the vicinity of interfering peaks.

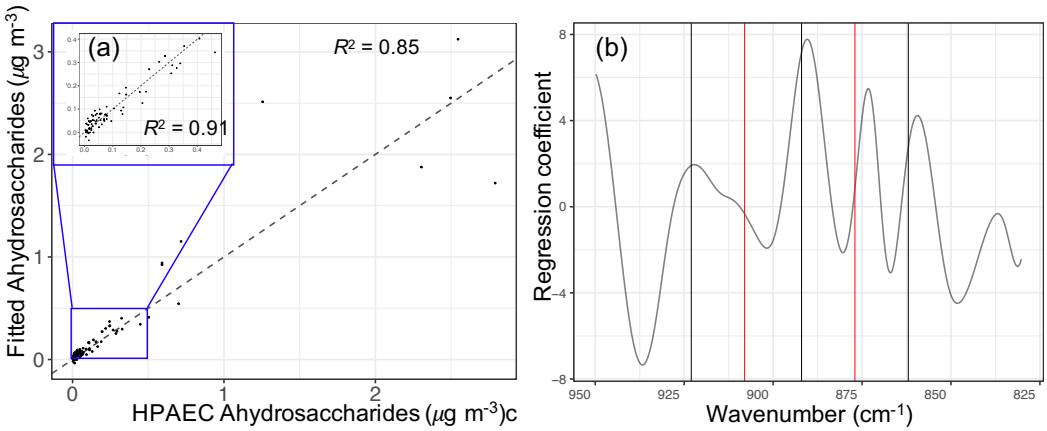

Figure A.59 – HPAEC-PAD sum of anhydrosaccharides versus fitted values using PLSR models on FTIR spectra (a) and regression coefficients of the wavenumbers in the levoglucosan fingerprint region calculated using PLSR. Vertical black black show the location levoglucosan peaks and red lines show the location interfering peaks.

A.5.7 Alternative smoke classifiers

In this section, alternative statistical models for building classifiers are discussed. The summary of these models has been show in Table A.4.

Table A.4 – Summary of alternative classifiers.

Classifier	Training set F_1 score	Test set F_1 score	Advantages
Simple (A)	0.90	0.82	Simple and physically interpretable
Simple (+PLSR, A_{PLS})	0.97	0.85	More accurate and robust with regard to the inorganic interference in low concentrations
Logistic (+PLSR)	0.88	0.88	Provides probabilities
SVM (+PLSR)	0.92	0.89	Accepts P_{Carb} as direct input. Has hyperparameters that can be tuned using cross validation. Thus, performance on the test set does not degrade significantly

Smoke classifier using WSOC and FTIR signatures

A smoke classifier was developed by optimizing the parameters in Eq. (A.14) using the genetic algorithm. In Eq. (A.14) WSOC is not substituted by the carbonyl peak. The results of this classifier (Fig. A.60) are similar to those of Eq. (6.1) in the main manuscript.

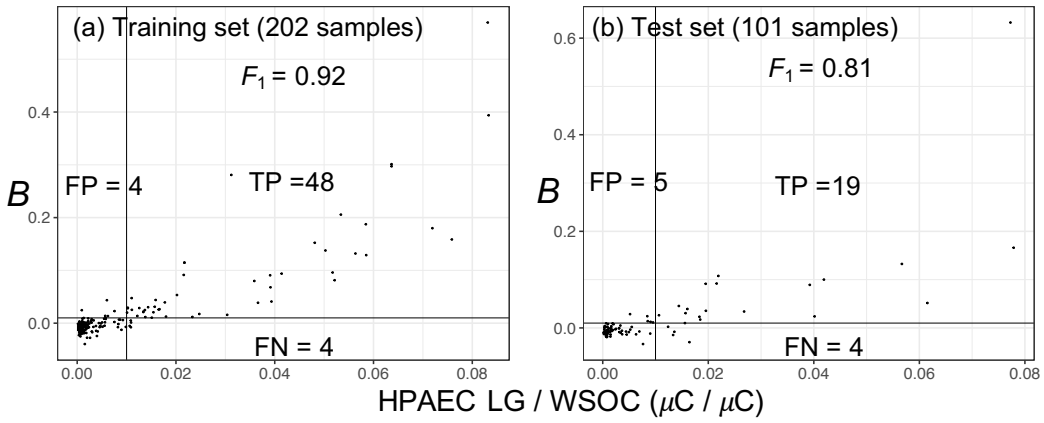

Figure A.60 – Comparison of the performance of the smoke classifier that uses FTIR smoke signatures and WSOC on training and test sets. B is defined in Eq. (A.14). Vertical and horizontal lines indicate the threshold (0.01) of smoke classification by each method.

$$B = \frac{aP_{\mathrm{LG1}} + bP_{\mathrm{LG2}} + cP_{\mathrm{LG2}} + dP_{\mathrm{Lig}} + e}{WSOC} > 0.01. \tag{A.14}$$

Building classifiers using PLSR results

The classifiers developed on the PLSR models are similar to those developed using P_{LG1}, P_{LG2}, and P_{LG3} peaks except that these three peaks are substituted by $P_{\mathrm{LG,\,PLS}}$ as in Eqs. (A.16) and (A.15). $P_{\mathrm{LG,\,PLS}}$ represents the response of the PLSR models developed by regressing the levoglucosan fingerprint region (830–950 cm^{-1}) in the FTIR spectra against HPAEC-PAD levoglcusoan measurements. The dataset (303 samples from 2008 and 2011) was divided into a training set (approximately two thirds) and a test set (approximately one third). PLSR models were developed on the training set and the optimization for a, b, and c was carried out using the genetic algorithm. The optimization was performed to maximize the F_1 scores for FTIR and HPAEC-PAD (assumed as reference) smoke classifications for two cases: using the carbonyl peak as a proxy to WSOC (Eq. A.15) and using FTIR biomass burning markers and WSOC data (Eq. A.16).

$$A_{\mathrm{PLS}} = \frac{aP_{\mathrm{LG,\,PLS}} + bP_{\mathrm{Lig}} + c}{P_{\mathrm{Carb}}} > 0.01 \text{ or} \tag{A.15}$$

$$B_{\mathrm{PLS}} = \frac{aP_{\mathrm{LG,\,PLS}} + bP_{\mathrm{Lig}} + c}{\mathrm{WSOC}} > 0.01. \tag{A.16}$$

Comparing Figs. A.61, A.60, and 6.6 , one can see that the performance of classifiers using PLSR models output is improved slightly compared to those using P_{LG1}, P_{LG2}, and P_{LG3} on the

training and test sets.

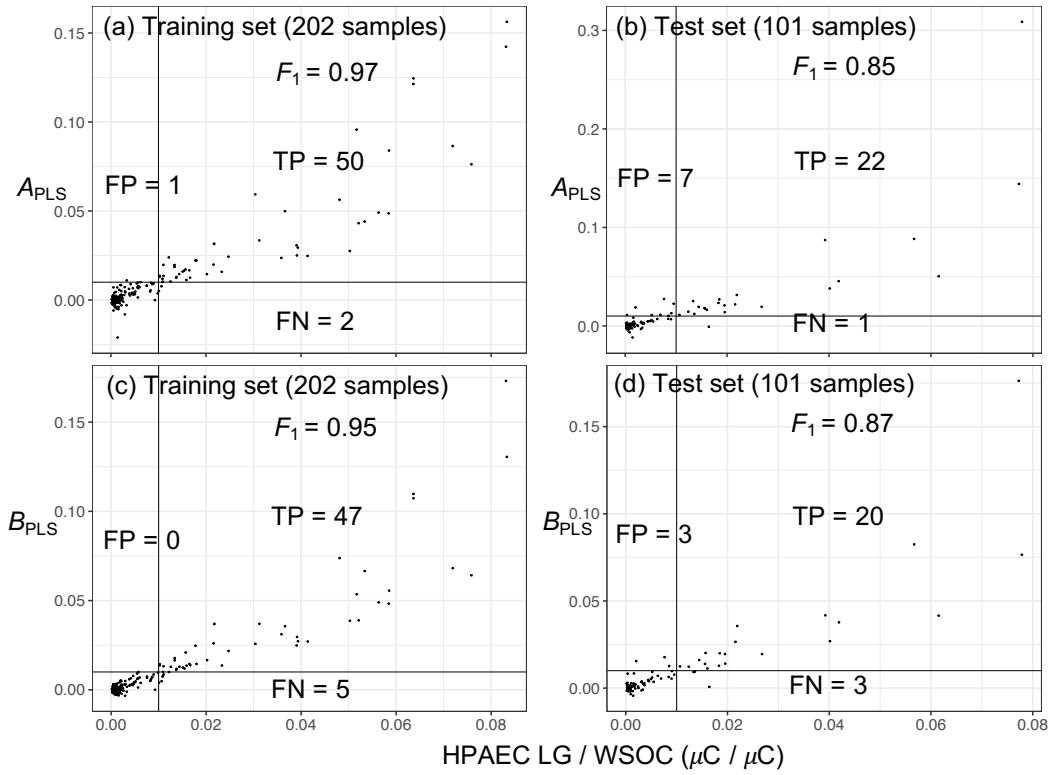

Figure A.61 – Comparison of the performances of smoke classifiers that use FTIR smoke signatures and WSOC (c–d) or the carbonyl peak (a–b) on training and test sets. A_{PLS} and B_{PLS} are defined in Eqs. (A.16) and (A.15). Vertical and horizontal black lines indicate the threshold (0.01) of smoke classification by each method.

The way the smoke classifier classifies samples based on their levoglucosan and lignin signatures is shown in Fig. A.61.

Building classifiers using PLSR results and logistic regression

A smoke classifier using logistic regression has been shown in Fig. A.63. In this classifier, two categories based on HPAEC-PAD measurements have been considered (no smoke, smoke-impacted). The advantage of this classifier is that it estimates a probability for the smoke-impacted samples. The logistic regression coefficients are solved by maximizing the likelihood of the equation below for the dataset. The $P_{\mathrm{LG,\,PLS}}$ appears to be the determining factor in this classifier.

$$\mathrm{logit}(\Pr(Y=1)) = \beta_0 + \beta_1 P_{\mathrm{LG,\,PLS}}/P_{\mathrm{Carb}} + \beta_2 P_{\mathrm{Lig}}/P_{\mathrm{Carb}}. \tag{A.17}$$

Adjusting the threshold of smoke-impacted to 0.28 gives the highest F_1 score for the training

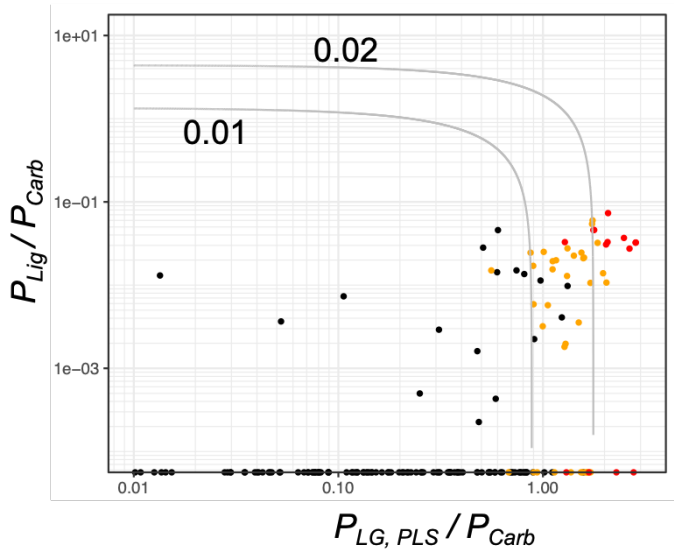

Figure A.62 – Scatter plot showing how the smoke classifier classifies samples based on their levoglucosan and lignin signatures and the selected threshold. Smoke-impacted samples having HPAEC-PAD LG / WSOC > 0.02 and maybe 0.01 < HPAEC-PAD LG / WSOC < 0.02 are shown in red and orange, respectively. One of the coordinates of the points located on the log axes is zero.

set (0.88). This threshold gives a reasonably high F_1 score on the test set as well (0.88).

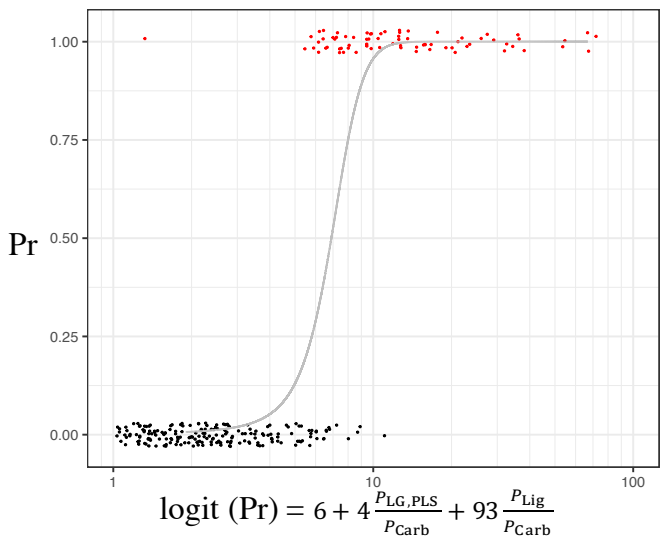

Figure A.63 – Logistic regression for smoke classification. Smoke-impacted samples (HPAEC-PAD LG / WSOC > 0.01) are shown in red. The solid curve shows the probably of a sample being smoke-impacted or maybe smoke-impacted.

Building classifiers using support vector machine

A soft-margin support vector machine classifier (Steinwart and Christmann, 2008) was used as another flexible method for building smoke classifiers . The dataset was divided into a training (202 samples) and a test set (101). The features used were either P_{LG1}, P_{LG2}, P_{LG3} , P_{Lig}, and P_{Carb} or $P_{\mathrm{LG,\,PLS}}$, P_{Lig}, and P_{carb}. A radial basis function (RBF) kernel was used,

$$K(x, x') = \exp(-\gamma ||x - x'||^2), \tag{A.18}$$

where x and x' are feature vectors of observation pairs. A 10-fold cross validation was used to find the optimum kernel parameter, γ, and the optimum regularization parameter in the Lagrange formulation. The F_1 scores obtained when using levoglucosan peaks were 0.83 and 0.77 for the training and test sets, respectively. The F_1 scores obtained when using the PLSR model output were 0.92 and 0.89 for the training and test sets, respectively. These results are comparable to those of the simple classifier used in the main text.

A.5.8 Smoke analysis of IMPROVE 2015 data

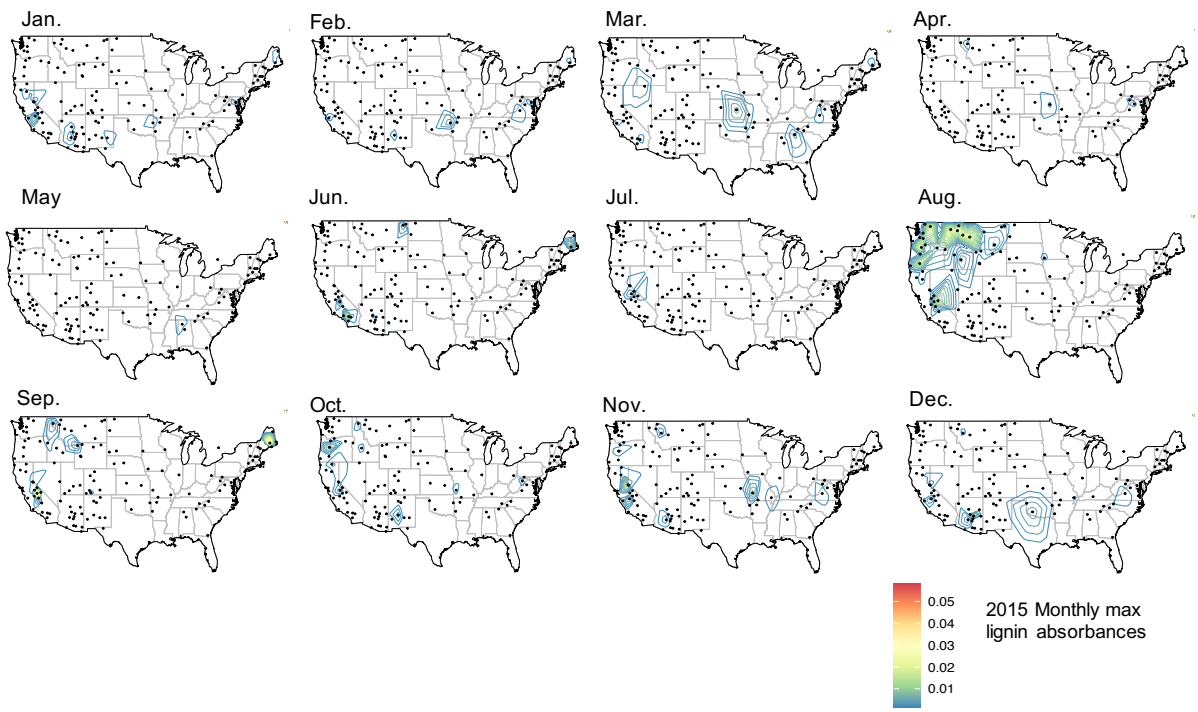

Figure A.64 – Monthly maximum absorbance of lignin-like compounds in the FTIR spectra from IMPROVE filter samples in 2015.

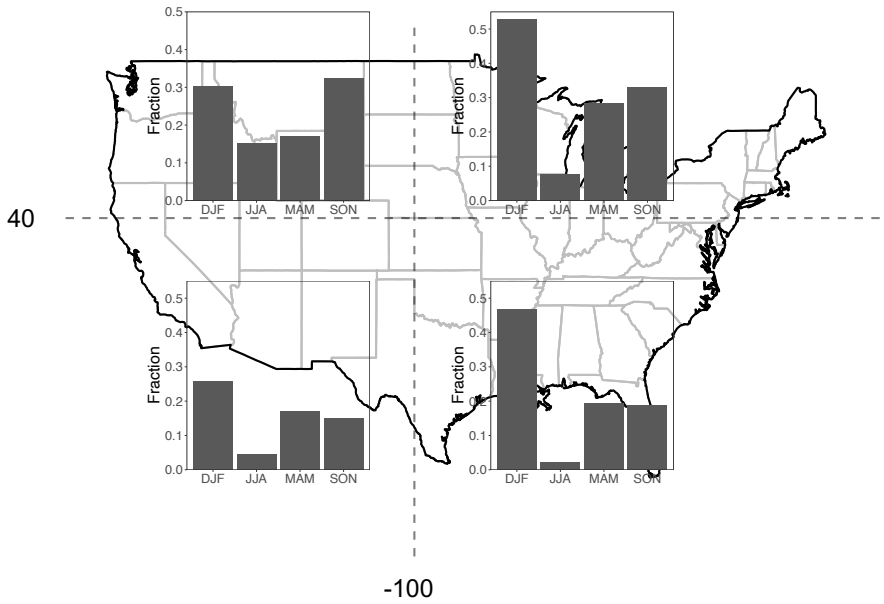

Figure A.65 – The fraction of smoke-impacted samples separated by season. The contiguous US is divided into four quadrants (vertical and horizontal lines indicating 40° N and −100° W).

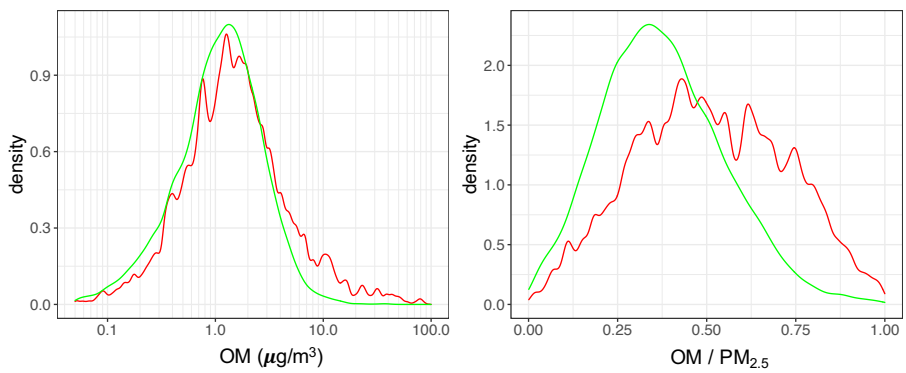

Figure A.66 – Kernel density estimates of OM concentrations (a) and OM / $PM_{2.5}$ (b) for smoke-impacted (red) and normal samples (green).

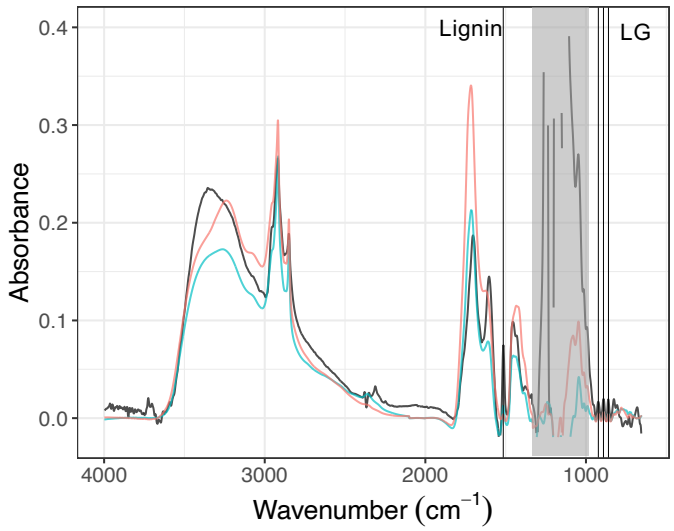

Figure A.67 – FTIR spectra of fresh wood burning fine aerosols (high LG/OM ratio) from the IMPROVE network (color) and environmental chamber experiments of Yazdani et al. (2021c) (black). Vertical lines indicate the location of bbOA marker signatures. The gray rectangle indicates the PTFE interference.

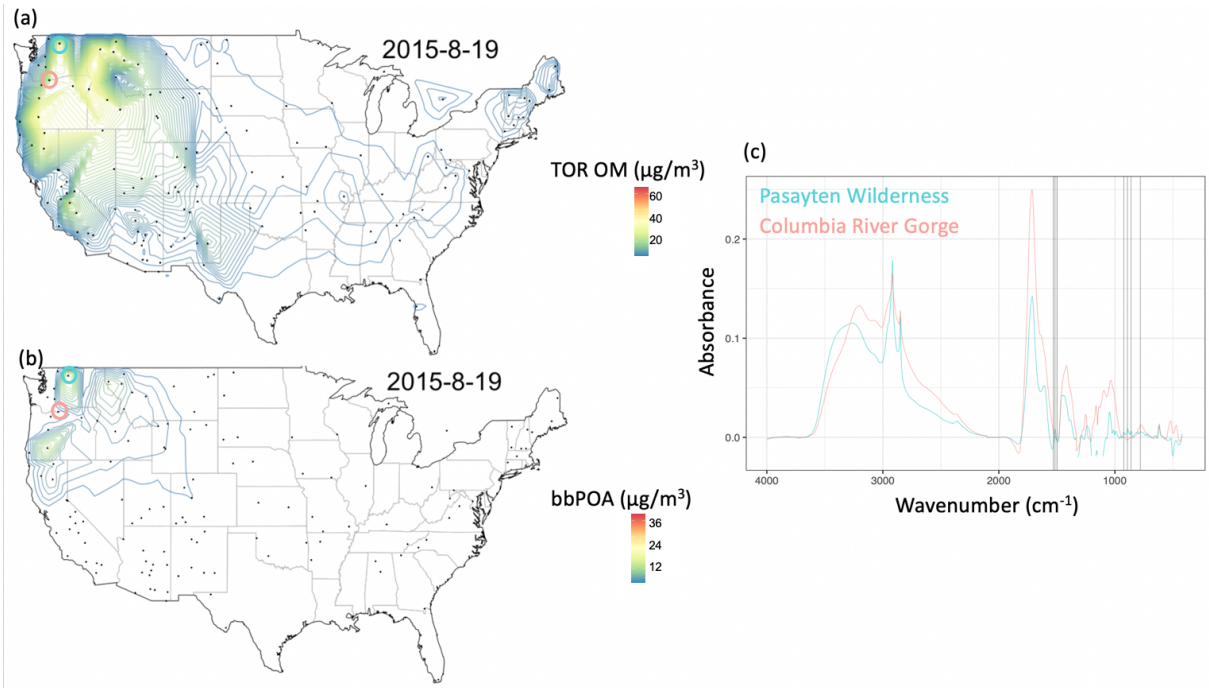

Figure A.68 – Comparison of OM (a) and bbPOA concentrations (b) for two sites impacted by smoke and their FTIR spectra (c).

A.6 Appendix to chapter 7

A.6.1 WBT protocol

The WBT protocol used in this study is illustrated as a schematic in Fig. A.69

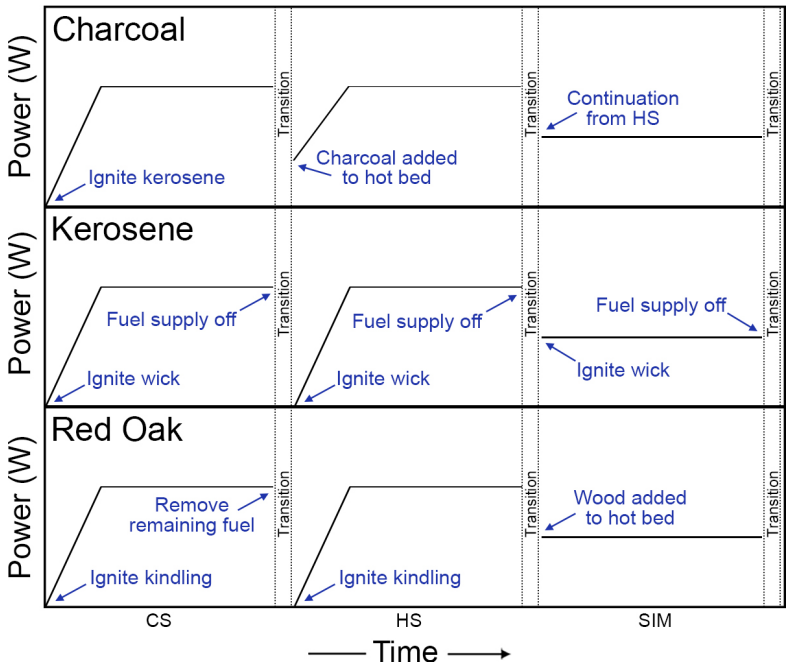

Figure A.69 – Schematic of WBT protocol for each fuel.

A.6.2 EC, OC, and $PM_{2.5}$ emissions

As can be seen from Fig. A.71, for the majority samples, the sum of EC and OM (OM/OC × OC, OM/OC is calculated from FTIR measurements) constitutes the majority of $PM_{2.5}$ mass. This suggests that $PM_{2.5}$ is mostly composed of EC and OM and the contribution of inorganics is negligible. There are three samples from alcohol burning that have non-zero EC and OC but zero $PM_{2.5}$. There wood and charcoal combustion samples in the $0.4 < (OM+EC)/PM_{2.5} < 0.7$ range a clear inorganic nitrate peak is observed (Fig. A.72).

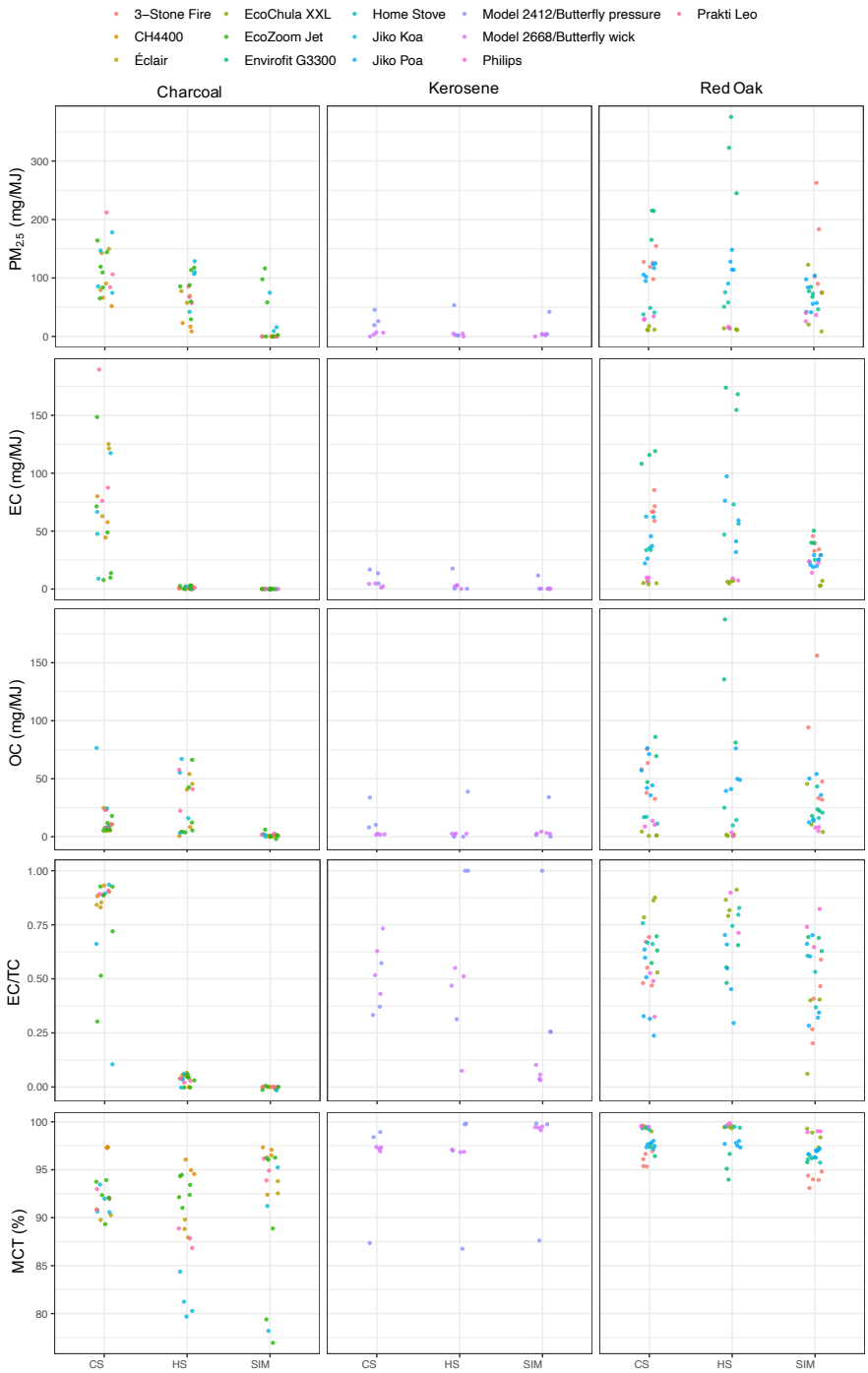

Figure A.70 – The emissions factors, EC/TC, and modified combustion efficiency (MCE; $(\Delta CO_2+\Delta CO)/\Delta CO_2$) separated by fuel type, test phase, and cookstove.

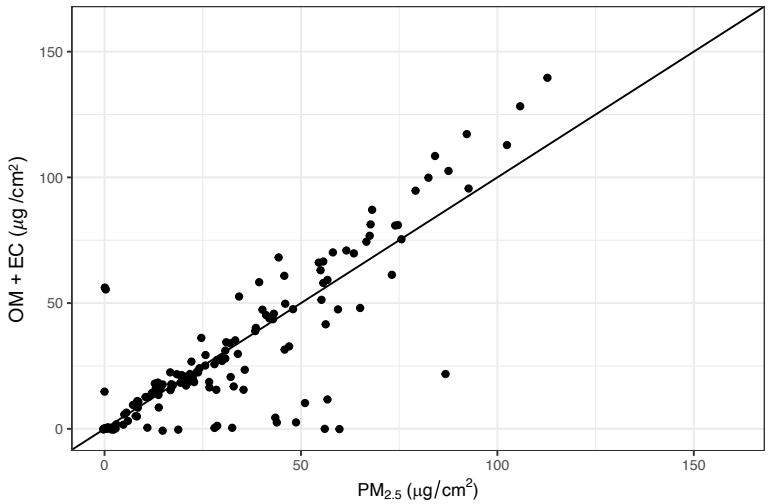

Figure A.71 – Scatter plot comparing $PM_{2.5}$ (gravimetric) measurements with the sum of artifact-corrected TOT OM (OM/OC × OC, where OM/OC is calculated from FTIR) and EC on quartz fiber filters.

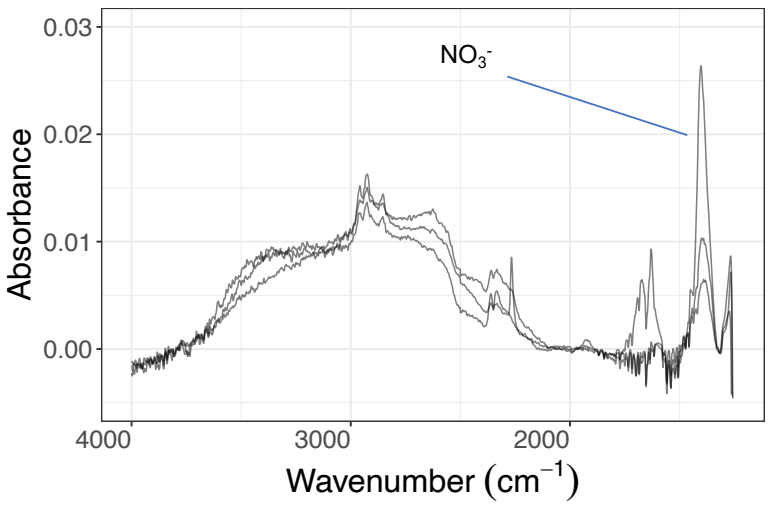

Figure A.72 – Spectra of three samples from charcoal particulate emissions with significant contributions of inorganic nitrate.

A.6.3 Mid-infrared spectra of $PM_{2.5}$ separated by stove type

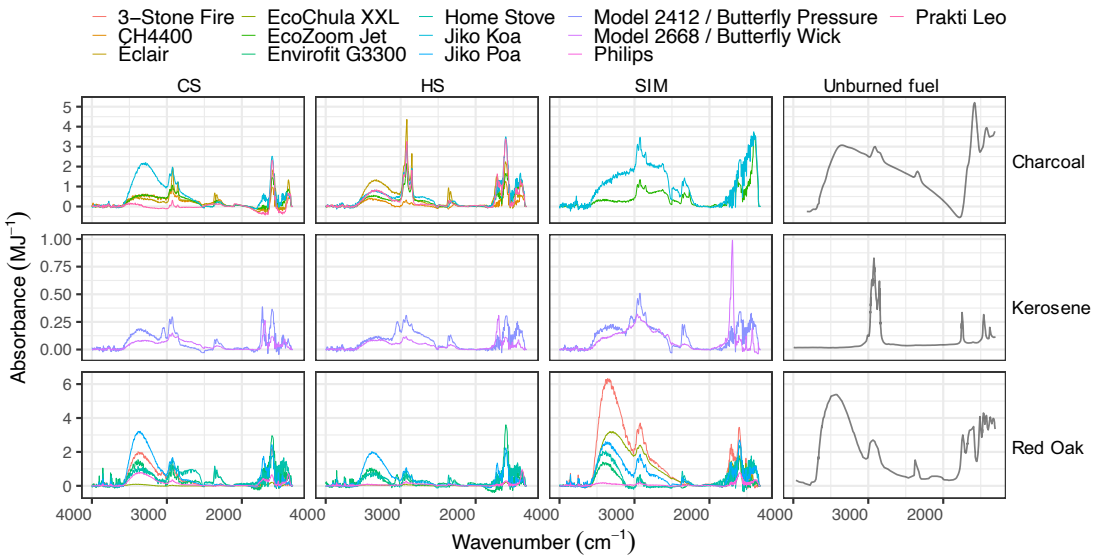

Figure A.73 – Average mid-infrared spectra of particulate emissions separated by source and phase and stove (normalized by energy delivered to the pot). CS = cold start, HS = hot start, SIM = simmering. Black spectra are for unburned fuels (wood, kerosene, and charcoal).

A.6.4 Fuels Analysis, Original Fuel Compositions

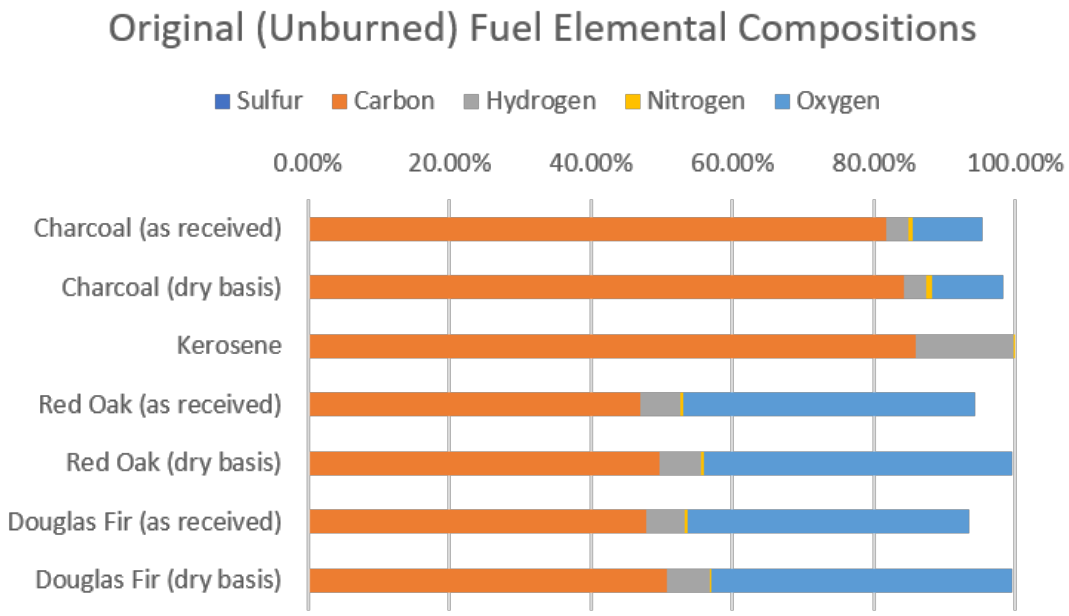

Figure A.74 – Elemental composition of unburned fuels. Data shown in Supplemental Table A.5.

Table A.5 – Original (pre-burn) fuel analysis.

Fuel	moisture	volatile	fixed carbon	ash	Sulfur	Carbon	Hydrogen	Nitrogen	Oxygen
Charcoal (as received)	2.95%	17.97%	77.32%	1.76%	0.02%	81.72%	3.04%	0.81%	9.70%
Charcoal (dry basis)		18.51%	79.67%	1.82%	0.02%	84.20%	3.13%	0.83%	10.00%
Kerosene					0.03%	85.93%	13.82%	0.01%	
Red Oak (as received)	5.28%	82.58%	11.78%	0.36%	0.08%	46.92%	5.74%	0.30%	41.33%
Red Oak (dry basis)		87.19%	12.44%	0.38%	0.08%	49.54%	6.05%	0.32%	43.64%
Douglas Fir (as received)	5.99%	80.51%	13.03%	0.47%	0.17%	47.58%	5.58%	0.32%	39.89%
Douglas Fir (dry basis)		85.64%	13.85%	0.51%	0.18%	50.61%	5.93%	0.34%	42.43%
Alcohol					0.03%	36.34%	11.68%	0.03%	
LPG						82.09%	17.91%		

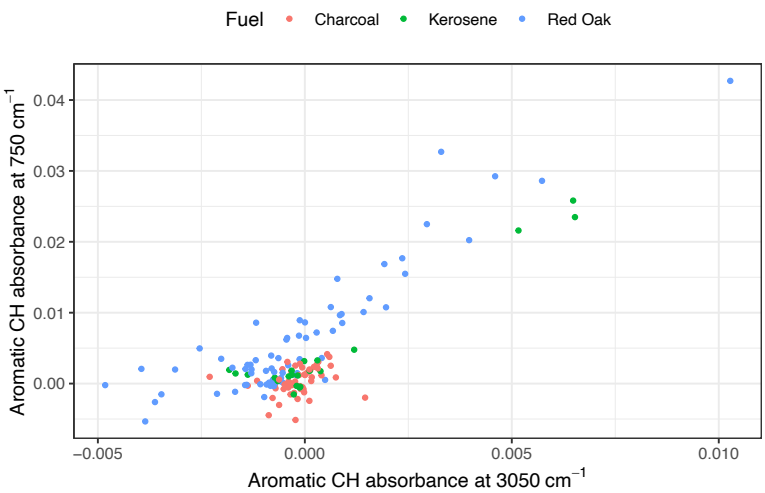

Figure A.75 – Scatter plot of aromatic CH stretching at 3050 cm^{-1} and out-of plane absorbances at 750 cm^{-1}.

A.6.5 Using out-of-plane vibrations for quantifying aromatic CH

The aromatic CH stretching absorbances usually have low absorption coefficients and and overlap with the ammonium NH stretching, making them hard to quantify. In this work, out-of-plane band, which has a relatively higher absorption coefficient, and do not overlap significantly with other group frequencies are used for the first time to quantify the aromatic CH group in the organic aerosols. The latter is highly correlated to that at 3050 cm^{-1} but is expected to have higher sensitivity (Fig. A.75). Contrary to the peak at 3050 cm^{-1}, the peak at 1600 cm^{-1} attributed to aromatic C=C is not highly correlated with the out-of-plane peak, especially for red oak burning aerosols likely due high abundance of lignin pyrolysis products substituted rings (Fig. A.76).

The aromatic CH concentration estimated using the peak at 750 cm^{-1} is also correlated (r = 0.84; slope = 19.6) with estimated total PAH concentration estimated by GC-MS (Fig. A.77. VIP scores of GC/MS total PAH concentration regressed against mid-infrared absorbances also shows high scores at 750 cm^{-1}, highlighting the importance of this region for PAHs (Fig. A.78). On the other hand, there is a non-linear relationship between EC concentration and the aromatic CH concentration (Fig. A.79).

A.6.6 EC information in the FTIR spectra

The absorbance at each wavenumber is the sum of functional group absorbances at that wavenumber, PTFE filter membrane light scattering, organic material light scattering, inorganic material light scattering and absorption of elemental carbon due to electronic transitions. In the current data set, inorganic salts are not present thus, except the filter light scattering, only EC and OC change the absorbance. There are some variations in filter light scattering

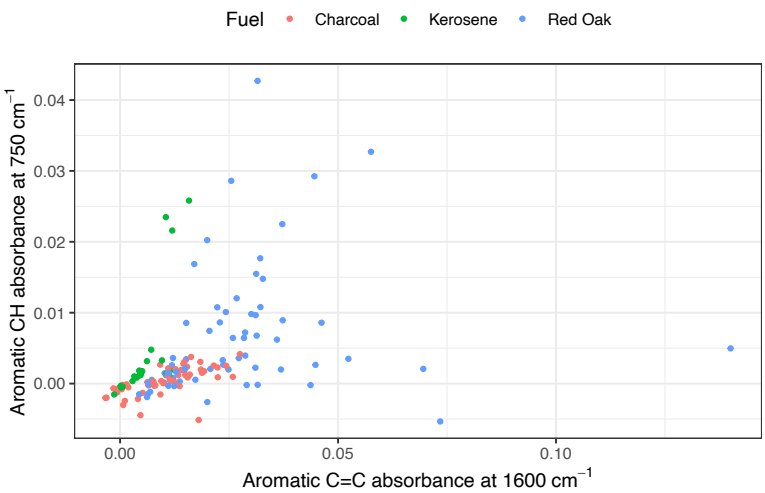

Figure A.76 – Scatter plot of aromatic C=C stretching at 1600 cm^{-1} and out-of plane absorbances at 750 cm^{-1}.

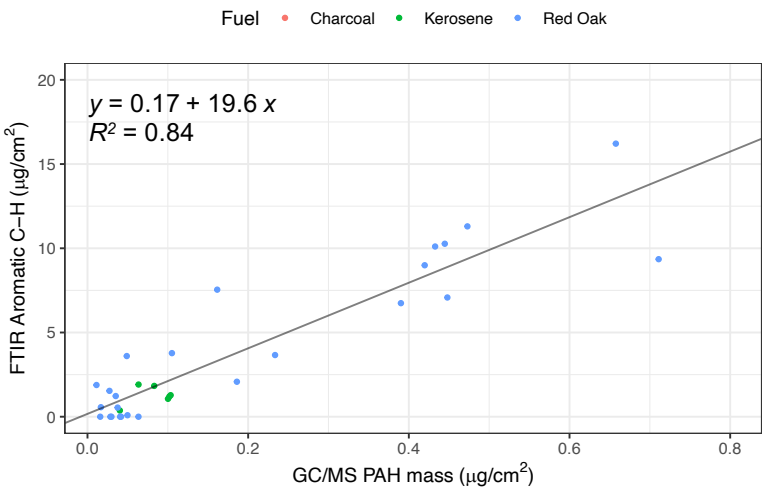

Figure A.77 – Scatter plot of aromatic CH concentration estimated using the peak at 750 cm^{-1} and GC/MS total PAH concentration.

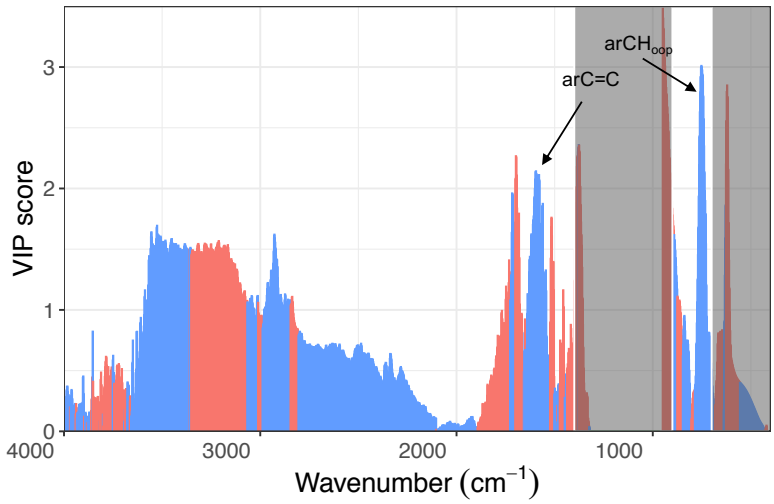

Figure A.78 – VIP scores of GC-MS sum of PAHs regressed against FTIR spectra.

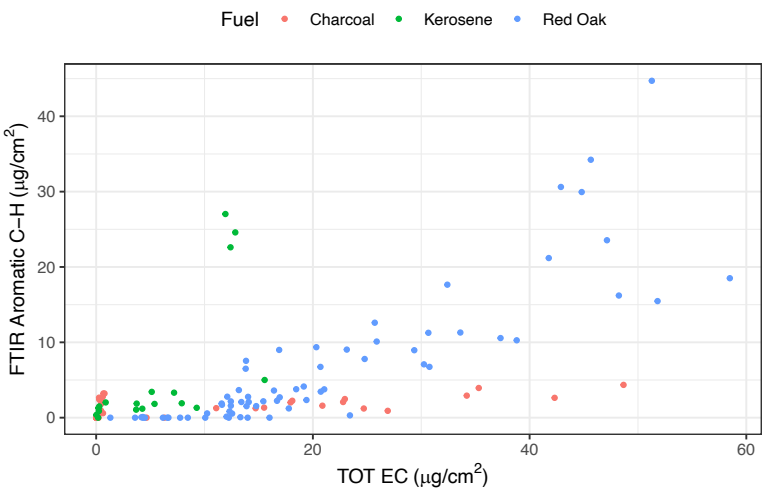

Figure A.79 – Scatter plot of aromatic CH concentration estimated using the peak at 750 cm^{-1} and TOT EC concentration.

from sample to sample due to factors such as thickness variation, fiber orientation and filter porosity. As the filter scattering is believed to be the major absorbing component due to its substantial mass and thickness compared to particulate matter, initially, its variation should be eliminated. For this purpose, the highest C–F absorption peak was chosen. Each spectrum was baseline-corrected, which separated the spectrum in two parts. The first part which is called the baseline-corrected spectrum, contains variations due to functional group absorptions, which appear usually as sharp to semi-broad peaks. The second part contains low level gradual variations, which usually arise from light scattering of filter and particulate matter, and EC absorption. In order to eliminate filter contribution to background, a scaled version of an empty filter background spectrum was subtracted from the background part of each spectrum. During the scaling process, the absorbance at each wavenumber for empty spectrum was multiplied by a factor which made the highest peak equal to that of the desired spectrum. By regressing the slope of corrected background absorption to EC and OC we get a relatively high correlation coefficient $r = 0.85$ (Fig. A.80).

The EC coefficient from ordinary least squares solution is approximately 2.2 times that of OC or approximately 4 times that of OM (assuming OM is 1.8 times OC).

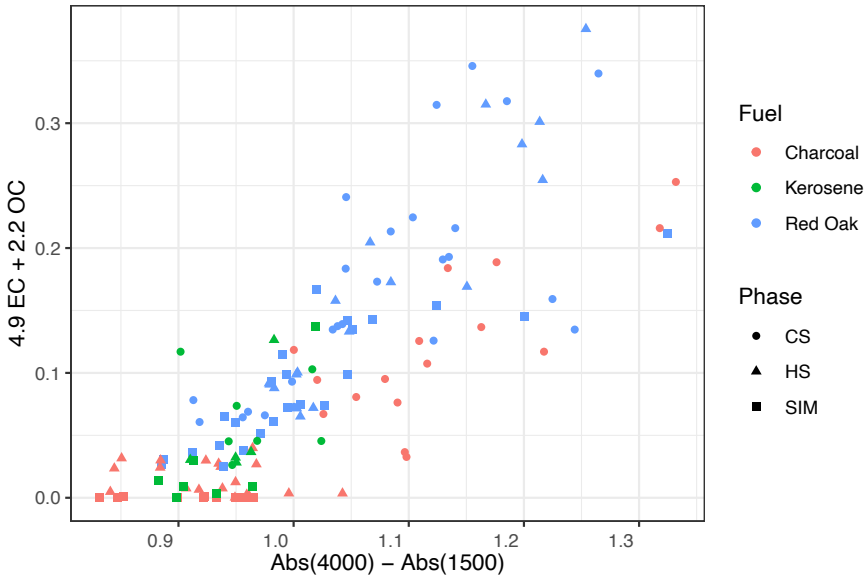

Figure A.80 – Baseline at 4000 cm^{-1} in the FTIR spectra regressed against EC and OC measured by TOT. The coefficients of the regression are shown on the vertical axis.

A.7 Appendix to conclusion

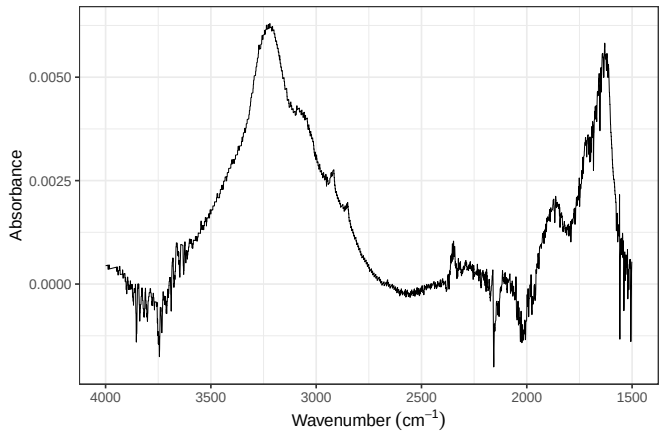

Figure A.81 – The single-bounce ATR-FTIR spectrum of atmospheric fine PM collected on a quartz fiber filter.

Bibliography

Abdi, H. and Williams, L. J. (2010). Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.*, 2(4):433–459.

Aiken, A. C., DeCarlo, P. F., and Jimenez, J. L. (2007). Elemental Analysis of Organic Species with Electron Ionization High-Resolution Mass Spectrometry. *Anal. Chem.*, 79(21):8350–8358.

Aiken, A. C., DeCarlo, P. F., Kroll, J. H., Worsnop, D. R., Huffman, J. A., Docherty, K. S., Ulbrich, I. M., Mohr, C., Kimmel, J. R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M., Ziemann, P. J., Canagaratna, M. R., Onasch, T. B., Alfarra, M. R., Prevot, A. S. H., Dommen, J., Duplissy, J., Metzger, A., Baltensperger, U., and Jimenez, J. L. (2008). O/C and OM/OC Ratios of Primary, Secondary, and Ambient Organic Aerosols with High-Resolution Time-of-Flight Aerosol Mass Spectrometry. *Environ. Sci. Technol.*, 42(12):4478–4485.

Aiken, A. C., Salcedo, D., Cubison, M. J., Huffman, J. A., DeCarlo, P. F., Ulbrich, I. M., Docherty, K. S., Sueper, D., Kimmel, J. R., Worsnop, D. R., Trimborn, A., Northway, M., Stone, E. A., Schauer, J. J., Volkamer, R. M., Fortner, E., de Foy, B., Wang, J., Laskin, A., Shutthanandan, V., Zheng, J., Zhang, R., Gaffney, J., Marley, N. A., Paredes-Miranda, G., Arnott, W. P., Molina, L. T., Sosa, G., and Jimenez, J. L. (2009). Mexico City aerosol analysis during MILAGRO using high resolution aerosol mass spectrometry at the urban supersite (T0) – Part 1: Fine particle composition and organic source apportionment. *Atmos. Chem. Phys.*, page 21.

Allen, D. T. and Palen, E. (1989). Recent advances in aerosol analysis by infrared spectroscopy. *Journal of Aerosol Science*, 20(4):441–455.

Allen, D. T., Palen, E. J., Haimov, M. I., Hering, S. V., and Young, J. R. (1994). Fourier Transform Infrared Spectroscopy of Aerosol Collected in a Low Pressure Impactor (LPI/FTIR): Method Development and Field Calibration. *Aerosol Science and Technology*, 21(4):325–342.

Alves, C., Gonçalves, C., Fernandes, A. P., Tarelho, L., and Pio, C. (2011). Fireplace and woodstove fine particle emissions from combustion of western Mediterranean wood types. *Atmos. Res.*, 101(3):692–700.

Andreae, M. O. and Gelencsér, A. (2006). Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols. *Atmospheric Chemistry and Physics*, 6(10):3131–3148.

Arangio, A., Delval, C., Ruggeri, G., Dudani, N., Yazdani, A., and Takahama, S. (2019). Electrospray Film Deposition for Solvent-Elimination Infrared Spectroscopy. *Appl. Spectrosc.*, page 000370281882133.

Arnold, S., Neuman, M., and Pluchino, A. B. (1984). Molecular spectroscopy of a single aerosol particle. *Optics Letters*, 9(1):4–6.

Arnold, S. and Pluchino, A. B. (1982). Infrared spectrum of a single aerosol particle by photothermal modulation of structure resonances. *Applied Optics*, 21(23):4194_1–4196.

Atkins, P., de Paula, J., and Keeler, J. (2017). *Atkins' Physical Chemistry*. Oxford University Press, Oxford, New York, eleventh edition.

Ayres, B. R., Allen, H. M., Draper, D. C., Brown, S. S., Wild, R. J., Jimenez, J. L., Day, D. A., Campuzano-Jost, P., Hu, W., de Gouw, J., Koss, A., Cohen, R. C., Duffey, K. C., Romer, P., Baumann, K., Edgerton, E., Takahama, S., Thornton, J. A., Lee, B. H., Lopez-Hilfiker, F. D., Mohr, C., Wennberg, P. O., Nguyen, T. B., Teng, A., Goldstein, A. H., Olson, K., and Fry, J. L. (2015). Organic nitrate aerosol formation via NO_3 + biogenic volatile organic compounds in the southeastern United States. *Atmos. Chem. Phys.*, 15(23):13377–13392.

Bäfver, L. S., Leckner, B., Tullin, C., and Berntsen, M. (2011). Particle emissions from pellets stoves and modern and old-type wood stoves. *Biomass Bioenergy*, 35(8):3648–3655.

Barmet, P., Dommen, J., DeCarlo, P. F., Tritscher, T., Praplan, A. P., Platt, S. M., Prévôt, A. S. H., Donahue, N. M., and Baltensperger, U. (2012). OH clock determination by proton transfer reaction mass spectrometry at an environmental chamber. *Atmos. Meas. Tech.*, 5(3):647–656.

Barnes, I., Bastian, V., Becker, K. H., and Tong, Z. (1990). Kinetics and products of the reactions of nitrate radical with monoalkenes, dialkenes, and monoterpenes. *J. Phys. Chem.*, 94(6):2413–2419.

Bateman, A. P., Nizkorodov, S. A., Laskin, J., and Laskin, A. (2011). Photolytic processing of secondary organic aerosols dissolved in cloud droplets. *Physical Chemistry Chemical Physics*, 13(26):12199–12212.

Bauer, H., Claeys, M., Vermeylen, R., Schueller, E., Weinke, G., Berger, A., and Puxbaum, H. (2008). Arabitol and mannitol as tracers for the quantification of airborne fungal spores. *Atmospheric Environment*, 42(3):588–593.

Beekmann, M., Prévôt, A. S. H., Drewnick, F., Sciare, J., Pandis, S. N., Denier van der Gon, H. a. C., Crippa, M., Freutel, F., Poulain, L., Ghersi, V., Rodriguez, E., Beirle, S., Zotter, P., von der Weiden-Reinmüller, S.-L., Bressi, M., Fountoukis, C., Petetin, H., Szidat, S., Schneider, J., Rosso, A., El Haddad, I., Megaritis, A., Zhang, Q. J., Michoud, V., Slowik, J. G., Moukhtar, S., Kolmonen, P., Stohl, A., Eckhardt, S., Borbon, A., Gros, V., Marchand, N., Jaffrezo, J. L., Schwarzenboeck, A., Colomb, A., Wiedensohler, A., Borrmann, S., Lawrence, M., Baklanov, A., and Baltensperger, U. (2015). In situ, satellite measurement and model evidence on

the dominant regional contribution to fine particulate matter levels in the Paris megacity. *Atmos. Chem. Phys.*, 15(16):9577–9591.

Bertrand, A., Stefenelli, G., Bruns, E. A., Pieber, S. M., Temime-Roussel, B., Slowik, J. G., Prévôt, A. S. H., Wortham, H., El Haddad, I., and Marchand, N. (2017). Primary emissions and secondary aerosol production potential from woodstoves for residential heating: Influence of the stove technology and combustion efficiency. *Atmos. Environ.*, 169:65–79.

Bertrand, A., Stefenelli, G., Jen, C. N., Pieber, S. M., Bruns, E. A., Ni, H., Temime-Roussel, B., Slowik, J. G., Goldstein, A. H., Haddad, I. E., Baltensperger, U., Prévôt, A. S. H., Wortham, H., and Marchand, N. (2018a). Evolution of the chemical fingerprint of biomass burning organic aerosol during aging. *Atmos. Chem. Phys.*, 18(10):7607–7624.

Bertrand, A., Stefenelli, G., Pieber, S. M., Bruns, E. A., Temime-Roussel, B., Slowik, J. G., Wortham, H., Prévôt, A. S. H., Haddad, I. E., and Marchand, N. (2018b). Influence of the vapor wall loss on the degradation rate constants in chamber experiments of levoglucosan and other biomass burning markers. *Atmos. Chem. Phys.*, 18(15):10915–10930.

Bock, P. and Gierlinger, N. (2019). Infrared and Raman spectra of lignin substructures: Coniferyl alcohol, abietin, and coniferyl aldehyde. *J. Raman Spectrosc.*, 50(6):778–792.

Boeriu, C. G., Bravo, D., Gosselink, R. J. A., and van Dam, J. E. G. (2004). Characterisation of structure-dependent functional properties of lignin with infrared spectroscopy. *Ind. Crops Prod.*, 20(2):205–218.

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P. K., Sarofim, M. C., Schultz, M. G., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S. K., Hopke, P. K., Jacobson, M. Z., Kaiser, J. W., Klimont, Z., Lohmann, U., Schwarz, J. P., Shindell, D., Storelvmo, T., Warren, S. G., and Zender, C. S. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552.

Boris, A. J., Takahama, S., Weakley, A. T., Debus, B. M., Fredrickson, C. D., Esparza-Sanchez, M., Burki, C., Reggente, M., Shaw, S. L., Edgerton, E. S., and Dillner, A. M. (2019). Quantifying organic matter and functional groups in particulate matter filter samples from the southeastern United States – Part 1: Methods. *Atmos. Meas. Tech.*, 12(10):5391–5415.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1983). Classification and Regression Trees. volume 40, page 874. Biometrics.

Brey, S. J., Ruminski, M., Atwood, S. A., and Fischer, E. V. (2018). Connecting smoke plumes to sources using Hazard Mapping System (HMS) smoke and fire location data over North America. *Atmospheric Chemistry and Physics*, 18(3):1745–1761.

Bruns, E., Krapf, M., Orasche, J., Huang, Y., Zimmermann, R., Drinovec, L., Močnik, G., El-Haddad, I., G. Slowik, J., Dommen, J., Baltensperger, U., and Prevot, A. (2015). Characterization of primary and secondary wood combustion products generated under different burner loads. *Atmos. Chem. Phys.*, 15:2825–2841.

Bruns, E. A., El Haddad, I., Slowik, J. G., Kilic, D., Klein, F., Baltensperger, U., and Prévôt, A. S. H. (2016). Identification of significant precursor gases of secondary organic aerosols from residential wood combustion. *Scientific Reports*, 6(1):1–9.

Bruns, E. A., Slowik, J. G., Haddad, I. E., Kilic, D., Klein, F., Dommen, J., Temime-Roussel, B., Marchand, N., Baltensperger, U., and Prévôt, A. S. H. (2017). Characterization of gas-phase organics using proton transfer reaction time-of-flight mass spectrometry: Fresh and aged residential wood combustion emissions. *Atmos. Chem. Phys.*, 17(1):705–720.

Bürki, C., Reggente, M., Dillner, A. M., Hand, J. L., Shaw, S. L., and Takahama, S. (2020). 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(3):1517–1538.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C. A., Apte, J. S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., Di, Q., Brunekreef, B., Frostad, J., Lim, S. S., Kan, H., Walker, K. D., Thurston, G. D., Hayes, R. B., Lim, C. C., Turner, M. C., Jerrett, M., Krewski, D., Gapstur, S. M., Diver, W. R., Ostro, B., Goldberg, D., Crouse, D. L., Martin, R. V., Peters, P., Pinault, L., Tjepkema, M., van Donkelaar, A., Villeneuve, P. J., Miller, A. B., Yin, P., Zhou, M., Wang, L., Janssen, N. A. H., Marra, M., Atkinson, R. W., Tsang, H., Thach, T. Q., Cannon, J. B., Allen, R. T., Hart, J. E., Laden, F., Cesaroni, G., Forastiere, F., Weinmayr, G., Jaensch, A., Nagel, G., Concin, H., and Spadaro, J. V. (2018). Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *PNAS*, 115(38):9592–9597.

CALFIRE (2021). 2015 Incident Archive. https://www.fire.ca.gov/.

Canagaratna, M. R., Jayne, J. T., Jimenez, J. L., Allan, J. D., Alfarra, M. R., Zhang, Q., Onasch, T. B., Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L. R., Trimborn, A. M., Northway, M. J., DeCarlo, P. F., Kolb, C. E., Davidovits, P., and Worsnop, D. R. (2007). Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer. *Mass Spectrom. Rev.*, 26(2):185–222.

Canagaratna, M. R., Jimenez, J. L., Kroll, J. H., Chen, Q., Kessler, S. H., Massoli, P., Hildebrandt Ruiz, L., Fortner, E., Williams, L. R., Wilson, K. R., Surratt, J. D., Donahue, N. M., Jayne, J. T., and Worsnop, D. R. (2015). Elemental ratio measurements of organic compounds using aerosol mass spectrometry: Characterization, improved calibration, and implications. *Atmos. Chem. Phys.*, 15(1):253–272.

Centrone, A., Brambilla, L., Renouard, T., Gherghel, L., Mathis, C., Müllen, K., and Zerbi, G. (2005). Structure of new carbonaceous materials: The role of vibrational spectroscopy. *Carbon*, 43(8):1593–1609.

Chen, Q., Ikemori, F., Higo, H., Asakawa, D., and Mochida, M. (2016). Chemical Structural Characteristics of HULIS and Other Fractionated Organic Matter in Urban Aerosols: Results from Mass Spectral and FT-IR Analysis. *Environ. Sci. Technol.*, 50(4):1721–1730.

Chhabra, P. S., Ng, N. L., Canagaratna, M. R., Corrigan, A. L., Russell, L. M., Worsnop, D. R., Flagan, R. C., and Seinfeld, J. H. (2011a). Elemental composition and oxidation of chamber organic aerosol. *Atmos. Chem. Phys.*, 11(17):8827–8845.

Chhabra, P. S., Ng, N. L., Canagaratna, M. R., Corrigan, A. L., Russell, L. M., Worsnop, D. R., Flagan, R. C., and Seinfeld, J. H. (2011b). Supplemental Material for Elemental Composition and Oxidation of Chamber Organic Aerosol. *Atmos. Chem. Phys.*, page 16.

Chong, I.-G. and Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.*, 78(1-2):103–112.

Chow, J. C., Watson, J. G., Chen, L.-W. A., Chang, M. C. O., Robinson, N. F., Trimble, D., and Kohl, S. (2007). The IMPROVE_A Temperature Protocol for Thermal/Optical Carbon Analysis: Maintaining Consistency with a Long-Term Database. *J Air Waste Manage.*, 57(9):1014–1023.

Chow, J. C., Watson, J. G., Lowenthal, D. H., and Magliano, K. L. (2005). Loss of PM2.5 Nitrate from Filter Samples in Central California. *J Air Waste Manage.*, 55(8):1158–1168.

Cocker III, D. R., Mader, B. T., Kalberer, M., Flagan, R. C., and Seinfeld, J. H. (2001). The effect of water on gas–particle partitioning of secondary organic aerosol: II. m-xylene and 1,3,5-trimethylbenzene photooxidation systems. *Atmos. Environ.*, 35(35):6073–6085.

Collins, C. D. (2007). Implementing Phytoremediation of Petroleum Hydrocarbons. In Willey, N., editor, *Phytoremediation: Methods and Reviews*, Methods in Biotechnology, pages 99–108. Humana Press, Totowa, NJ.

Corrigan, A. L., Russell, L. M., Takahama, S., Äijälä, M., Ehn, M., Junninen, H., Rinne, J., Petäjä, T., Kulmala, M., Vogel, A. L., Hoffmann, T., Ebben, C. J., Geiger, F. M., Chhabra, P., Seinfeld, J. H., Worsnop, D. R., Song, W., Auld, J., and Williams, J. (2013). Biogenic and biomass burning organic aerosol in a boreal forest at Hyytiälä, Finland, during HUMPPA-COPEC 2010. *Atmos. Chem. Phys.*, 13(24):12233–12256.

Corsetti, S., Rabl, T., McGloin, D., and Kiefer, J. (2017). Intermediate phases during solid to liquid transitions in long-chain n-alkanes. *Phys. Chem. Chem. Phys.*, 19(21):13941–13950.

Coury, C. and Dillner, A. M. (2008). 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(23):5923–5932.

Day, D. A., Liu, S., Russell, L. M., and Ziemann, P. J. (2010). Organonitrate group concentrations in submicron particles with high nitrate and organic fractions in coastal southern California. *Atmos. Environ.*, 44(16):1970–1979.

Debus, B., Takahama, S., Weakley, A. T., Seibert, K., and Dillner, A. M. (2019). 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. *Appl. Spectrosc.*, 73(3):271–283.

DeCarlo, P. F., Dunlea, E. J., Kimmel, J. R., Aiken, A. C., Sueper, D., Crounse, J., Wennberg, P. O., Emmons, L., Shinozuka, Y., Clarke, A., Zhou, J., Tomlinson, J., Collins, D. R., Knapp, D., Weinheimer, A. J., Montzka, D. D., Campos, T., and Jimenez, J. L. (2008). Fast airborne aerosol size and chemistry measurements above Mexico City and Central Mexico during the MILAGRO campaign. *Atmos. Chem. Phys.*, 8(14):4027–4048.

DeCarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K. S., Worsnop, D. R., and Jimenez, J. L. (2006). Field-Deployable, High-Resolution, Time-of-Flight Aerosol Mass Spectrometer. *Anal. Chem.*, 78(24):8281–8289.

Decesari, S., Facchini, M. C., Fuzzi, S., and Tagliavini, E. (2000). Characterization of water-soluble organic compounds in atmospheric aerosol: A new approach. *J. Geophys. Res.-Atmos.*, 105(D1):1481–1489.

Decesari, S., Mircea, M., Cavalli, F., Fuzzi, S., Moretti, F., Tagliavini, E., and Facchini, M. C. (2007). Source Attribution of Water-Soluble Organic Aerosol by Nuclear Magnetic Resonance Spectroscopy. *Environmental Science & Technology*, 41(7):2479–2484.

Degen, I. A. (1968). Detection of the Methoxyl Group by Infrared Spectroscopy. *Applied Spectroscopy*, 22(3):164–166.

DeRieux, W.-S. W., Li, Y., Lin, P., Laskin, J., Laskin, A., Bertram, A. K., Nizkorodov, S. A., and Shiraiwa, M. (2018). Predicting the glass transition temperature and viscosity of secondary organic material using molecular composition. *Atmos. Chem. Phys.*, 18(9):6331–6351.

Derkacheva, O. and Sukhov, D. (2008). Investigation of Lignins by FTIR Spectroscopy. *Macromol. Symp.*, 265(1):61–68.

Desiraju, G. R. and Steiner, T. (2001). *The Weak Hydrogen Bond: In Structural Chemistry and Biology*. Oxford University Press.

Dillner, A. M. and Takahama, S. (2015). Predicting ambient aerosol thermal-optical reflectance (TOR) measurements from infrared spectra: Organic carbon. *Atmos. Meas. Tech.*, 8(3):1097–1109.

Donahue, N. M., Epstein, S. A., Pandis, S. N., and Robinson, A. L. (2011). A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics. *Atmos. Chem. Phys.*, 11(7):3303–3318.

Donahue, N. M., Henry, K. M., Mentel, T. F., Kiendler-Scharr, A., Spindler, C., Bohn, B., Brauers, T., Dorn, H. P., Fuchs, H., Tillmann, R., Wahner, A., Saathoff, H., Naumann, K.-H., Möhler, O.,

Leisner, T., Müller, L., Reinnig, M.-C., Hoffmann, T., Salo, K., Hallquist, M., Frosch, M., Bilde, M., Tritscher, T., Barmet, P., Praplan, A. P., DeCarlo, P. F., Dommen, J., Prévôt, A. S. H., and Baltensperger, U. (2012). Aging of biogenic secondary organic aerosol via gas-phase OH radical reactions. *PNAS*, 109(34):13503–13508.

Donahue, N. M., Robinson, A. L., Stanier, C. O., and Pandis, S. N. (2006). Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics. *Environ. Sci. Technol.*, 40(8):2635–2643.

Duarte, R. M. B. O., Pio, C. A., and Duarte, A. C. (2005). Spectroscopic study of the water-soluble organic matter isolated from atmospheric aerosols collected under different atmospheric conditions. *Analytica Chimica Acta*, 530(1):7–14.

Duarte, R. M. B. O., Santos, E. B. H., Pio, C. A., and Duarte, A. C. (2007). Comparison of structural features of water-soluble organic matter from atmospheric aerosols with those of aquatic humic substances. *Atmos. Environ.*, 41(37):8100–8113.

Dudani, N. (2021). *Design and Fabrication of an Aerosol Collector Device for IR-spectroscopy Analysis*. PhD thesis, Ecole Polytechnique Federale de Lausanne Lausanne (EPFL), Lausanne, Switzerland.

Duplissy, J., DeCarlo, P. F., Dommen, J., Alfarra, M. R., Metzger, A., Barmpadimos, I., Prevot, A. S. H., Weingartner, E., Tritscher, T., Gysel, M., Aiken, A. C., Jimenez, J. L., Canagaratna, M. R., Worsnop, D. R., Collins, D. R., Tomlinson, J., and Baltensperger, U. (2011). Relating hygroscopicity and composition of organic aerosol particulate matter. *Atmos. Chem. Phys.*, 11(3):1155–1165.

Duvall, R. M., Majestic, B. J., Shafer, M. M., Chuang, P. Y., Simoneit, B. R. T., and Schauer, J. J. (2008). The water-soluble fraction of carbon, sulfur, and crustal elements in Asian aerosols and Asian soils. *Atmospheric Environment*, 42(23):5872–5884.

Eadon, G., Djerassi, C., Beynon, J. H., and Caprioli, R. M. (1971). The fragmentation of aliphatic ketones in the mass spectrometer: A detailed study of nonan-4-one using ion kinetic energy spectroscopy. *Int. J. Mass Spectrom.*, 5(8):917–933.

Elser, M., Huang, R.-J., Wolf, R., Slowik, J. G., Wang, Q., Canonaco, F., Li, G., Bozzetti, C., Daellenbach, K. R., Huang, Y., Zhang, R., Li, Z., Cao, J., Baltensperger, U., El-Haddad, I., and Prévôt, A. S. H. (2016). New insights into $PM_{2.5}$ chemical composition and sources in two major cities in China during extreme haze events using aerosol mass spectrometry. *Atmos. Chem. Phys.*, 16(5):3207–3225.

Faber, P., Drewnick, F., Bierl, R., and Borrmann, S. (2017). Complementary online aerosol mass spectrometry and offline FT-IR spectroscopy measurements: Prospects and challenges for the analysis of anthropogenic aerosol particle emissions. *Atmos. Environ.*, 166:92–98.

Farmer, D. K., Matsunaga, A., Docherty, K. S., Surratt, J. D., Seinfeld, J. H., Ziemann, P. J., and Jimenez, J. L. (2010). Response of an aerosol mass spectrometer to organonitrates and organosulfates and implications for atmospheric chemistry. *PNAS*, 107(15):6670–6675.

Feldman, D. (1985). *Wood—Chemistry, Ultrastructure, Reactions, by D. Fengel and G. Wegener, Walter de Gruyter, Berlin and New York, 1984, 613 Pp. Price: 245 DM*, volume 23. Walter de Gruyter, Berlin and New York.

Fine, P. M., Cass, G. R., and Simoneit, B. R. T. (2001). Chemical Characterization of Fine Particle Emissions from Fireplace Combustion of Woods Grown in the Northeastern United States. *Environmental Science & Technology*, 35(13):2665–2675.

Fine, P. M., Cass, G. R., and Simoneit, B. R. T. (2002). Chemical Characterization of Fine Particle Emissions from the Fireplace Combustion of Woods Grown in the Southern United States. *Environ. Sci. Technol.*, 36(7):1442–1451.

Ford, B., Martin, M. V., Zelasky, S. E., Fischer, E. V., Anenberg, S. C., Heald, C. L., and Pierce, J. R. (2018). Future Fire Impacts on Smoke Concentrations, Visibility, and Health in the Contiguous United States. *GeoHealth*, 2(8):229–247.

Fornaro, T., Burini, D., Biczysko, M., and Barone, V. (2015). Hydrogen-Bonding Effects on Infrared Spectra from Anharmonic Computations: Uracil–Water Complexes and Uracil Dimers. *J. Phys. Chem. A*, 119(18):4224–4236.

Fountoukis, C., Butler, T., Lawrence, M. G., Denier van der Gon, H. A. C., Visschedijk, A. J. H., Charalampidis, P., Pilinis, C., and Pandis, S. N. (2014). Impacts of controlling biomass burning emissions on wintertime carbonaceous aerosol in Europe. *Atmos. Environ.*, 87:175–182.

Frossard, A. A., Russell, L. M., Massoli, P., Bates, T. S., and Quinn, P. K. (2014). Side-by-Side Comparison of Four Techniques Explains the Apparent Differences in the Organic Composition of Generated and Ambient Marine Aerosol Particles. *Aerosol Sci. Tech.*, 48(3):v–x.

Frossard, A. A., Shaw, P. M., Russell, L. M., Kroll, J. H., Canagaratna, M. R., Worsnop, D. R., Quinn, P. K., and Bates, T. S. (2011). Springtime Arctic haze contributions of submicron organic particles from European and Asian combustion sources. *J. Geophys. Res.-Atmos.*, 116(D5):D05205.

Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W. H., Dallmann, T. R., Davis, L., Liu, S., Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R. A., and Goldstein, A. H. (2012). Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions. *PNAS*, 109(45):18318–18323.

George, I. J., Slowik, J., and Abbatt, J. P. D. (2008). Chemical aging of ambient organic aerosol from heterogeneous reaction with hydroxyl radicals. *Geophys. Res. Lett.*, 35(13):L13811.

George, K. M., Ruthenburg, T. C., Smith, J., Yu, L., Zhang, Q., Anastasio, C., and Dillner, A. M. (2015). FT-IR quantification of the carbonyl functional group in aqueous-phase secondary organic aerosol from phenols. *Atmos. Environ.*, 100:230–237.

Gibbs-Flournoy, E. A., Gilmour, M. I., Higuchi, M., Jetter, J., George, I., Copeland, L., Harrison, R., Moser, V. C., and Dye, J. A. (2018). Differential exposure and acute health impacts of inhaled solid-fuel emissions from rudimentary and advanced cookstoves in female CD-1 mice. *Environmental Research*, 161:35–48.

Gilardoni, S., Liu, S., Takahama, S., Russell, L. M., Allan, J. D., Steinbrecher, R., Jimenez, J. L., De Carlo, P. F., Dunlea, E. J., and Baumgardner, D. (2009). Characterization of organic ambient aerosol during MIRAGE 2006 on three platforms. *Atmos. Chem. Phys.*, 9(15):5417–5432.

Graber, E. R. and Rudich, Y. (2006). Atmospheric HULIS: How humic-like are they? A comprehensive and critical review. *Atmos. Chem. Phys.*, 6(3):729–753.

Grieshop, A. P., Donahue, N. M., and Robinson, A. L. (2009). Laboratory investigation of photochemical oxidation of organic aerosol from wood fires 2: Analysis of aerosol mass spectrometer data. *Atmos. Chem. Phys.*, 9(6):2227–2240.

Griffiths, P. R. and Haseth, J. A. D. (2007). *Fourier Transform Infrared Spectrometry*. John Wiley & Sons, Hoboken, New Jersey.

Guo, Y. and Bustin, R. M. (1998). FTIR spectroscopy and reflectance of modern charcoals and fungal decayed woods: Implications for studies of inertinite in coals. *International Journal of Coal Geology*, 37(1):29–53.

Haaland, D. M. and Thomas, E. V. (1988). Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry*, 60(11):1193–1202.

Haddad, I. E., D'Anna, B., Temime-Roussel, B., Nicolas, M., Boreave, A., Favez, O., Voisin, D., Sciare, J., George, C., Jaffrezo, J.-L., Wortham, H., and Marchand, N. (2013). Towards a better understanding of the origins, chemical composition and aging of oxygenated organic aerosols: Case study of a Mediterranean industrialized environment, Marseille. *Atmos. Chem. Phys.*, 13(15):7875–7894.

Hähner, G., Zwahlen, M., and Caseri, W. (2005). Chain-Length Dependence of the Conformational Order in Self-Assembled Dialkylammonium Monolayers on Mica Studied with Soft X-ray Absorption. *Langmuir*, 21(4):1424–1427.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R.,

and Wildt, J. (2009). The formation, properties and impact of secondary organic aerosol: Current and emerging issues. *Atmos. Chem. Phys.*, 9(14):5155–5236.

Hand, J. L., Prenni, A. J., Schichtel, B. A., Malm, W. C., and Chow, J. C. (2019). Trends in remote PM2.5 residual mass across the United States: Implications for aerosol mass reconstruction in the IMPROVE network. *Atmos. Environ.*, 203:141–152.

Haque, M. M., Kawamura, K., Deshmukh, D. K., Fang, C., Song, W., Mengying, B., and Zhang, Y.-L. (2019). Characterization of organic aerosols from a Chinese megacity during winter: Predominance of fossil fuel combustion. *Atmos. Chem. Phys.*, 19(7):5147–5164.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition.* Springer Series in Statistics. Springer-Verlag, New York, second edition.

Hastings, S. H., Watson, A. T., Williams, R. B., and Anderson, J. A. (1952). Determination of Hydrocarbon Functional Groups by Infrared Spectroscopy. *Anal. Chem.*, 24(4):612–618.

Hawkins, L. N. and Russell, L. M. (2010). Oxidation of ketone groups in transported biomass burning aerosol from the 2008 Northern California Lightning Series fires. *Atmos. Environ.*, 44(34):4142–4154.

Hawkins, L. N., Russell, L. M., Covert, D. S., Quinn, P. K., and Bates, T. S. (2010). Carboxylic acids, sulfates, and organosulfates in processed continental organic aerosol over the southeast Pacific Ocean during VOCALS-REx 2008. *Journal of Geophysical Research: Atmospheres*, 115(D13).

Heald, C. L., Kroll, J. H., Jimenez, J. L., Docherty, K. S., DeCarlo, P. F., Aiken, A. C., Chen, Q., Martin, S. T., Farmer, D. K., and Artaxo, P. (2010). A simplified description of the evolution of organic aerosol composition in the atmosphere. *Geophys. Res. Lett.*, 37(8).

Hearn, J. D., Lovett, A. J., and Smith, G. D. (2005). Ozonolysis of oleic acid particles: Evidence for a surface reaction and secondary reactions involving Criegee intermediates. *Phys. Chem. Chem. Phys.*, 7(3):501–511.

Helland, I. S. (1988). On the structure of partial least squares regression. *Communications in Statistics - Simulation and Computation*, 17(2):581–607.

Hennigan, C. J., Miracolo, M. A., Engelhart, G. J., May, A. A., Presto, A. A., Lee, T., Sullivan, A. P., McMeeking, G. R., Coe, H., Wold, C. E., Hao, W.-M., Gilman, J. B., Kuster, W. C., de Gouw, J., Schichtel, B. A., Collett, J. L. J., Kreidenweis, S. M., and Robinson, A. L. (2011). Chemical and physical transformations of organic aerosol from the photo-oxidation of open biomass burning emissions in an environmental chamber. *Atmospheric Chemistry and Physics*, 11(15):7669–7686.

Hennigan, C. J., Sullivan, A. P., Collett, J. L., and Robinson, A. L. (2010). Levoglucosan stability in biomass burning particles exposed to hydroxyl radicals. *Geophys. Res. Lett.*, 37(9):L09806.

Henry, K. M. and Donahue, N. M. (2012). Photochemical aging of alpha-pinene secondary organic aerosol: Effects of OH radical sources and photolysis. *J. Phys. Chem. A*, 116(24):5932–5940.

Hergert, H. L. (1960). Infrared Spectra of Lignin and Related Compounds. II. Conifer Lignin and Model Compounds. *J. Org. Chem.*, 25(3):405–413.

Heringa, M. F., DeCarlo, P. F., Chirico, R., Tritscher, T., Dommen, J., Weingartner, E., Richter, R., Wehrle, G., Prévôt, A. S. H., and Baltensperger, U. (2011). Investigations of primary and secondary particulate matter of different wood combustion appliances with a high-resolution time-of-flight aerosol mass spectrometer. *Atmos. Chem. Phys.*, 11(12):5945–5957.

Hermans, J., Ongay, S., Markov, V., and Bischoff, R. (2017). Physicochemical Parameters Affecting the Electrospray Ionization Efficiency of Amino Acids after Acylation. *Anal. Chem.*, 89(17):9159–9166.

Hinds, W. C. (1999). *Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles*. Wiley-Interscience, New York, 2 edition edition.

Holmes, B. J. and Petrucci, G. A. (2006). Water-Soluble Oligomer Formation from Acid-Catalyzed Reactions of Levoglucosan in Proxies of Atmospheric Aqueous Aerosols. *Environ. Sci. Technol.*, 40(16):4983–4989.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 24(6):417–441.

Hung, H.-M., Katrib, Y., and Martin, S. T. (2005). Products and Mechanisms of the Reaction of Oleic Acid with Ozone and Nitrate Radical. *J. Phys. Chem. A*, 109(20):4517–4530.

Hurteau, M. D., Westerling, A. L., Wiedinmyer, C., and Bryant, B. P. (2014). Projected Effects of Climate and Development on California Wildfire Emissions through 2100. *Environmental Science & Technology*, 48(4):2298–2304.

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Iyer, S., Lopez-Hilfiker, F., Lee, B. H., Thornton, J. A., and Kurtén, T. (2016). Modeling the Detection of Organic and Inorganic Compounds Using Iodide-Based Chemical Ionization. *J. Phys. Chem. A*, 120(4):576–587.

Jang, M. and Kamens, R. M. (2001a). Atmospheric Secondary Aerosol Formation by Heterogeneous Reactions of Aldehydes in the Presence of a Sulfuric Acid Aerosol Catalyst. *Environ. Sci. Technol.*, 35(24):4758–4766.

Jang, M. and Kamens, R. M. (2001b). Characterization of Secondary Aerosol from the Photooxidation of Toluene in the Presence of NOx and 1-Propene. *Environ. Sci. Technol.*, 35(18):3626–3639.

Jathar, S. H., Cappa, C. D., Wexler, A. S., Seinfeld, J. H., and Kleeman, M. J. (2015). Multi-generational oxidation model to simulate secondary organic aerosol in a 3-D air quality model. *Geosci. Model Dev.*, 8(8):2553–2567.

Jeni, L. A., Cohn, J. F., and De La Torre, F. (2013). Facing Imbalanced Data–Recommendations for the Use of Performance Metrics. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 245–251.

Jetter, J., Zhao, Y., Smith, K. R., Khan, B., Yelverton, T., DeCarlo, P., and Hays, M. D. (2012). Pollutant Emissions and Energy Efficiency under Controlled Conditions for Household Biomass Cookstoves and Implications for Metrics Useful in Setting International Test Standards. *Environ. Sci. Technol.*, 46(19):10827–10834.

Jetter, J. J. and Kariher, P. (2009). Solid-fuel household cook stoves: Characterization of performance and emissions. *Biomass and Bioenergy*, 33(2):294–305.

Jiang, J., El-Haddad, I., Aksoyoglu, S., Stefenelli, G., Bertrand, A., Marchand, N., Canonaco, F., Petit, J.-E., Favez, O., Gilardoni, S., Baltensperger, U., and Prévôt, A. S. H. (2020). Influence of biomass burning vapor wall loss correction on modeling organic aerosols in Europe by CAMx v6.50. *Geosci. Model Dev.*, pages 1–22.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., E, Dunlea, J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R. (2009). Evolution of Organic Aerosols in the Atmosphere. *Science*, 326(5959):1525–1529.

Johansson, L. S., Leckner, B., Gustavsson, L., Cooper, D., Tullin, C., and Potter, A. (2004). Emission characteristics of modern and old-type residential boilers fired with wood logs and wood pellets. *Atmos. Environ.*, 38(25):4183–4195.

Johnson, S. A., Kumar, R., and Cunningham, P. T. (1982). Airborne Detection of Acidic Sulfate Aerosol Using an ATR Impactor. *Aerosol Science and Technology*, 2(3):401–405.

Junninen, H., Mønster, J., Rey, M., Cancelinha, J., Douglas, K., Duane, M., Forcina, V., Müller, A., Lagler, F., Marelli, L., Borowiak, A., Niedzialek, J., Paradiz, B., Mira-Salama, D., Jimenez, J., Hansen, U., Astorga, C., Stanczyk, K., Viana, M., Querol, X., Duvall, R. M., Norris, G. A., Tsakovski, S., Wåhlin, P., Horák, J., and Larsen, B. R. (2009). Quantifying the Impact of Residential Heating on the Urban Air Quality in a Typical European Coal Combustion Region. *Environ. Sci. Technol.*, 43(20):7964–7970.

Kalberer, M. (2004). Identification of Polymers as Major Components of Atmospheric Organic Aerosols. *Science*, 303(5664):1659–1662.

Kalberer, M., Sax, M., and Samburova, V. (2006). Molecular Size Evolution of Oligomers in Organic Aerosols Collected in Urban Atmospheres and Generated in a Smog Chamber. *Environ. Sci. Technol.*, 40(19):5917–5922.

Kaltsonoudis, C., Kostenidou, E., Louvaris, E., Psichoudaki, M., Tsiligiannis, E., Florou, K., Liangou, A., and Pandis, S. N. (2017). Characterization of fresh and aged organic aerosol emissions from meat charbroiling. *Atmos. Chem. Phys.*, 17(11):7143–7155.

Kamińska, E., Madejczyk, O., Tarnacka, M., Jurkiewicz, K., Wolnica, K., Śmiszek-Lindert, W. E., Kamiński, K., and Paluch, M. (2018). Anhydrosaccharides—A new class of the fragile plastic crystals. *J. Chem. Phys.*, 148(7):074501.

Kanakidou, M., Seinfeld, J. H., Pandis, S. N., Barnes, I., Dentener, F. J., Facchini, M. C., Dingenen, R. V., Ervens, B., Nenes, A., Nielsen, C. J., Swietlicki, E., Putaud, J. P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G. K., Winterhalter, R., Myhre, C. E. L., Tsigaridis, K., Vignati, E., Stephanou, E. G., and Wilson, J. (2005). Organic aerosol and global climate modelling: A review. *Atmos. Chem. Phys.*, page 71.

Karl, T. G., Christian, T. J., Yokelson, R. J., Artaxo, P., Hao, W. M., and Guenther, A. (2007). The Tropical Forest and Fire Emissions Experiment: Method evaluation of volatile organic compound emissions measured by PTR-MS, FTIR, and GC from tropical biomass burning. *Atmospheric Chemistry and Physics*, 7(22):5883–5897.

Kauppinen, J. K., Moffatt, D. J., Mantsch, H. H., and Cameron, D. G. (1981). Fourier Self-Deconvolution: A Method for Resolving Intrinsically Overlapped Bands. *Applied Spectroscopy*, 35(3):271–276.

Kautzman, K. E., Surratt, J. D., Chan, M. N., Chan, A. W. H., Hersey, S. P., Chhabra, P. S., Dalleska, N. F., Wennberg, P. O., Flagan, R. C., and Seinfeld, J. H. (2010). Chemical Composition of Gas- and Aerosol-Phase Products from the Photooxidation of Naphthalene. *J. Phys. Chem. A*, 114(2):913–934.

Kellner, R. and Malissa, H. (1989). Fourier Transform Infrared Microscopy—a Tool for Speciation of Impactor-Sampled Single Particles or Particle Clusters. *Aerosol Science and Technology*, 10(2):397–407.

Kelly, A. M. (2013). *Condensed-Phase Molecular Spectroscopy and Photophysics.* John Wiley & Sons, Inc., Hoboken, NJ, 1st edition.

Klein, F., Pieber, S. M., Ni, H., Stefenelli, G., Bertrand, A., Kilic, D., Pospisilova, V., Temime-Roussel, B., Marchand, N., El Haddad, I., Slowik, J. G., Baltensperger, U., Cao, J., Huang, R.-j., and Prévôt, A. S. H. (2018). Characterization of Gas-Phase Organics Using Proton Transfer Reaction Time-of-Flight Mass Spectrometry: Residential Coal Combustion. *Environ. Sci. Technol.*, 52(5):2612–2617.

Kodros, J. K., Papanastasiou, D. K., Paglione, M., Masiol, M., Squizzato, S., Florou, K., Skyllakou, K., Kaltsonoudis, C., Nenes, A., and Pandis, S. N. (2020). Rapid dark aging of biomass burning as an overlooked source of oxidized organic aerosol. *PNAS*, 117(52):33028–33033.

Kroll, J. H., Donahue, N. M., Jimenez, J. L., Kessler, S. H., Canagaratna, M. R., Wilson, K. R., Altieri, K. E., Mazzoleni, L. R., Wozniak, A. S., Bluhm, H., Mysak, E. R., Smith, J. D., Kolb, C. E., and Worsnop, D. R. (2011). Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol. *Nat. Chem.*, 3(2):133–139.

Kroll, J. H. and Seinfeld, J. H. (2008). Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere. *Atmos. Environ.*, 42(16):3593–3624.

Krost, K. J. and McClenny, W. A. (1994). FT-IR Transmission Spectroscopy for Quantitation of Ammonium Bisulfate in Fine-Particulate Matter Collected on Teflon® Filters. *Applied Spectroscopy*, 48(6):702–705.

Kumar, N. K., Corbin, J. C., Bruns, E. A., Massabó, D., Slowik, J. G., Drinovec, L., Močnik, G., Prati, P., Vlachou, A., Baltensperger, U., Gysel, M., El-Haddad, I., and Prévôt, A. S. H. (2018). Production of particulate brown carbon during atmospheric aging of residential wood-burning emissions. *Atmos. Chem. Phys.*, 18(24):17843–17861.

Kuzmiakova, A., Dillner, A. M., and Takahama, S. (2016). An automated baseline correction protocol for infrared spectra of atmospheric aerosols collected on polytetrafluoroethylene (Teflon) filters. *Atmos. Meas. Tech.*, 9(6):2615–2631.

Lack, D. A., Langridge, J. M., Bahreini, R., Cappa, C. D., Middlebrook, A. M., and Schwarz, J. P. (2012). Brown carbon and internal mixing in biomass burning particles. *Proceedings of the National Academy of Sciences*, 109(37):14802–14807.

Lam, N. L., Smith, K. R., Gauthier, A., and Bates, M. N. (2012). Kerosene: A review of household uses and their hazards in low-and middle-income countries. *Journal of Toxicology and Environmental Health - Part B: Critical Reviews*, 15(6):396–432.

Lambe, A. T., Onasch, T. B., Croasdale, D. R., Wright, J. P., Martin, A. T., Franklin, J. P., Massoli, P., Kroll, J. H., Canagaratna, M. R., Brune, W. H., Worsnop, D. R., and Davidovits, P. (2012). Transitions from Functionalization to Fragmentation Reactions of Laboratory Secondary Organic Aerosol (SOA) Generated from the OH Oxidation of Alkane Precursors. *Environ. Sci. Technol.*, 46(10):5430–5437.

Lanz, V. A., Prévôt, A. S. H., Alfarra, M. R., Weimer, S., Mohr, C., DeCarlo, P. F., Gianini, M. F. D., Hueglin, C., Schneider, J., Favez, O., D'Anna, B., George, C., and Baltensperger, U. (2010). Characterization of aerosol chemical composition with aerosol mass spectrometry in Central Europe: An overview. *Atmos. Chem. Phys.*, 10(21):10453–10471.

Larkin, S., Brown, T., Lahm, P., and Zimmerman, T. (2010). Wildland fire decision support system air quality tools. *Fire Management Today*, 70(2):5.

Lee, S., Liu, W., Wang, Y., Russell, A. G., and Edgerton, E. S. (2008). Source apportionment of PM2.5: Comparing PMF and CMB results for four ambient monitoring sites in the southeastern United States. *Atmospheric Environment*, 42(18):4126–4137.

Li, E. Y., Yazdani, A., Takahama, S., Champion, W., Shen, G., Jetter, J. J., Hays, M. D., and Dillner, A. M. (2021). Quantified Functional Group Compositions in Household Fuel Burn Emissions using FTIR. *Manuscript in preparation.*

Li, J., Li, J., Wang, G., Zhang, T., Dai, W., Ho, K. F., Wang, Q., Shao, Y., Wu, C., and Li, L. (2020a). Molecular characteristics of organic compositions in fresh and aged biomass burning aerosols. *Science of The Total Environment*, 741:140247.

Li, X., Han, J., Hopke, P. K., Hu, J., Shu, Q., Chang, Q., and Ying, Q. (2019). Quantifying primary and secondary humic-like substances in urban aerosol based on emission source characterization and a source-oriented air quality model. *Atmos. Chem. Phys.*, 19(4):2327–2341.

Li, Y., Day, D. A., Stark, H., Jimenez, J., and Shiraiwa, M. (2020b). Predictions of the glass transition temperature and viscosity of organic aerosols from volatility distributions. *Atmos. Meas. Tech.*, 20(13):8103–8122.

Li, Y., Pöschl, U., and Shiraiwa, M. (2016). Molecular corridors and parameterizations of volatility in the chemical evolution of organic aerosols. *Atmos. Chem. Phys.*, 16(5):3327–3344.

Li, Y. J., Yeung, J. W. T., Leung, T. P. I., Lau, A. P. S., and Chan, C. K. (2012). Characterization of Organic Particles from Incense Burning Using an Aerodyne High-Resolution Time-of-Flight Aerosol Mass Spectrometer. *Aerosol Sci. Tech.*, 46(6):654–665.

Lii, J.-H., Chen, K.-H., and Allinger, N. L. (2004). Alcohols, Ethers, Carbohydrates, and Related Compounds Part V. The Bohlmann Torsional Effect. *The Journal of Physical Chemistry A*, 108(15):3006–3015.

Lim, Y. B. and Ziemann, P. J. (2009). Chemistry of Secondary Organic Aerosol Formation from OH Radical-Initiated Reactions of Linear, Branched, and Cyclic Alkanes in the Presence of NO_x. *Aerosol Sci. Tech.*, 43(6):604–619.

Lipp, E. D. (1986). Application of Fourier Self-Deconvolution to the FT-IR Spectra of Polydimethylsiloxane Oligomers for Determining Chain Length. *Appl. Spectrosc.*, 40(7):1009–1011.

Liu, C., Zhang, C., Mu, Y., Liu, J., and Zhang, Y. (2017). Emission of volatile organic compounds from domestic coal stove with the actual alternation of flaming and smoldering combustion processes. *Environmental Pollution*, 221:385–391.

Liu, S., Ahlm, L., Day, D. A., Russell, L. M., Zhao, Y., Gentner, D. R., Weber, R. J., Goldstein, A. H., Jaoui, M., Offenberg, J. H., Kleindienst, T. E., Rubitschun, C., Surratt, J. D., Sheesley, R. J., and

Scheller, S. (2012a). Secondary organic aerosol formation from fossil fuel sources contribute majority of summertime organic mass at Bakersfield. *J. Geophys. Res.-Atmos.*, 117(D24).

Liu, S., Day, D. A., Shields, J. E., and Russell, L. M. (2011). Ozone-driven daytime formation of secondary organic aerosol containing carboxylic acid groups and alkane groups. *Atmos. Chem. Phys.*, 11(16):8321–8341.

Liu, S., Shilling, J. E., Song, C., Hiranuma, N., Zaveri, R. A., and Russell, L. M. (2012b). Hydrolysis of Organonitrate Functional Groups in Aerosol Particles. *Aerosol Sci. Tech.*, 46(12):1359–1369.

Lopez-Hilfiker, F. D., Pospisilova, V., Huang, W., Kalberer, M., Mohr, C., Stefenelli, G., Thornton, J. A., Baltensperger, U., Prevot, A. S. H., and Slowik, J. G. (2019). An extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF) for online measurement of atmospheric aerosol particles. *Atmos. Meas. Tech.*, 12(9):4867–4886.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 5.1:281–298.

Mader, P. P., MacPhee, R. D., Lofberg, R. T., and Larson, G. P. (1952). Composition of Organic Portion of Atmospheric Aerosols in the Los Angeles Area. *Industrial & Engineering Chemistry*, 44(6):1352–1355.

Malm, W. C. and Hand, J. L. (2007). An examination of the physical and optical properties of aerosols collected in the IMPROVE program. *Atmospheric Environment*, 41(16):3407–3427.

Maria, S. F. and Russell, L. M. (2005). Organic and Inorganic Aerosol Below-Cloud Scavenging by Suburban New Jersey Precipitation. *Environmental Science & Technology*, 39(13):4793–4800.

Maria, S. F., Russell, L. M., Turpin, B. J., and Porcja, R. J. (2002). FTIR measurements of functional groups and organic mass in aerosol samples over the Caribbean. *Atmos. Environ.*, 36(33):5185–5196.

Maria, S. F., Russell, L. M., Turpin, B. J., Porcja, R. J., Campos, T. L., Weber, R. J., and Huebert, B. J. (2003). 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(D23):8637.

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

Mcclenny, W. A., Childers, J. W., Rōhl, R., and Palmer, R. A. (1985). FTIR transmission spectrometry for the nondestructive determination of ammonium and sulfate in ambient aerosols collected on teflon filters. *Atmos. Environ.*, 19(11):1891–1898.

McFiggans, G., Coe, H., Burgess, R., Allan, J., Cubison, M., Alfarra, M. R., Saunders, R., Saiz-Lopez, A., Plane, J. M. C., Wevill, D., Carpenter, L., Rickard, A. R., and Monks, P. S. (2004). Direct evidence for coastal iodine particles from Laminaria macroalgae – linkage to emissions of molecular iodine. *Atmos. Chem. Phys.*, 4(3):701–713.

McHale, J. L. (2017). *Molecular Spectroscopy*. CRC Press, Boca Raton, FL.

McKendry, P. (2002). Energy production from biomass (Part 1): Overview of biomass. *Bioresour. Technol.*, 83(1):37–46.

McLafferty, F. W. and Tureček, F. (1993). *Interpretation of Mass Spectra*, volume 8. University Science Books, Mill Valley, CA, USA, fourth edition.

McNamara, D., Stephens, G., Ruminski, M., and Kasheta, T. (2004). The Hazard Mapping System (HMS)—NOAA's Multi-Sensor Fire and Smoke Detection Program Using Environmental Satellites. In *13th Conference on Satellite Meteorology and Oceanography*, volume Vol. 22.

Mitchell, M. (1996). *An Introduction to Genetic Algorithms*. Complex Adaptive Systems. A Bradford Book, Cambridge, MA, USA.

Murphy, B. N., Donahue, N. M., Fountoukis, C., Dall'Osto, M., O'Dowd, C., Kiendler-Scharr, A., and Pandis, S. N. (2012). Functionalization and fragmentation during ambient organic aerosol aging: Application of the 2-D volatility basis set to field studies. *Atmos. Chem. Phys.*, 12(22):10797–10816.

Mutlu Esra, Warren Sarah H., Ebersviller Seth M., Kooter Ingeborg M., Schmid Judith E., Dye Janice A., Linak William P., Gilmour M. Ian, Jetter James J., Higuchi Mark, and DeMarini David M. (2016). Mutagenicity and Pollutant Emission Factors of Solid-Fuel Cookstoves: Comparison with Other Combustion Sources. *Environmental Health Perspectives*, 124(7):974–982.

Nah, T., Kessler, S. H., Daumit, K. E., Kroll, J. H., Leone, S. R., and Wilson, K. R. (2014). Influence of Molecular Structure and Chemical Functionality on the Heterogeneous OH-Initiated Oxidation of Unsaturated Organic Particles. *J. Phys. Chem. A*, 118(23):4106–4119.

National Centers for Environmental Information (2021). Wildfires - August 2015. https://www.ncdc.noaa.gov/sotc/fire/201508.

National Interagency Fire Center (2021). National Fire News. https://www.nifc.gov/fire-information/nfn.

Ng, N. L., Brown, S. S., Archibald, A. T., Atlas, E., Cohen, R. C., Crowley, J. N., Day, D. A., Donahue, N. M., Fry, J. L., Fuchs, H., Griffin, R. J., Guzman, M. I., Herrmann, H., Hodzic, A., Iinuma, Y., Jimenez, J. L., Kiendler-Scharr, A., Lee, B. H., Luecken, D. J., Mao, J., McLaren, R., Mutzel, A., Osthoff, H. D., Ouyang, B., Picquet-Varrault, B., Platt, U., Pye, H. O. T., Rudich, Y., Schwantes, R. H., Shiraiwa, M., Stutz, J., Thornton, J. A., Tilgner, A., Williams, B. J., and Zaveri, R. A. (2017). Nitrate radicals and biogenic volatile organic compounds: Oxidation, mechanisms, and organic aerosol. *Atmos. Chem. Phys.*, 17(3):2103–2162.

Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Chhabra, P. S., Seinfeld, J. H., and Worsnop, D. R. (2011a). Changes in organic aerosol composition with aging inferred from aerosol mass spectra. *Atmos. Chem. Phys.*, 11(13):6465–6474.

Ng, N. L., Canagaratna, M. R., Zhang, Q., Jimenez, J. L., Tian, J., Ulbrich, I. M., Kroll, J. H., Docherty, K. S., Chhabra, P. S., Bahreini, R., Murphy, S. M., Seinfeld, J. H., Hildebrandt, L., Donahue, N. M., DeCarlo, P. F., Lanz, V. A., Prévôt, A. S. H., Dinar, E., Rudich, Y., and Worsnop, D. R. (2010). Organic aerosol components observed in Northern Hemispheric datasets from Aerosol Mass Spectrometry. *Atmos. Chem. Phys.*, 10(10):4625–4641.

Ng, N. L., Herndon, S. C., Trimborn, A., Canagaratna, M. R., Croteau, P. L., Onasch, T. B., Sueper, D., Worsnop, D. R., Zhang, Q., Sun, Y. L., and Jayne, J. T. (2011b). An Aerosol Chemical Speciation Monitor (ACSM) for Routine Monitoring of the Composition and Mass Concentrations of Ambient Aerosol. *Aerosol Science and Technology*, 45(7):780–794.

NIOSH (1999). Elemental Carbon (Diesel Particulate): Method 5040. *Manual of Analytical Methods*, pages 1–9.

Nizkorodov, S. A., Laskin, J., and Laskin, A. (2011). Molecular chemistry of organic aerosols through the application of high resolution mass spectrometry. *Phys. Chem. Chem. Phys.*, 13(9):3612–3629.

Novakov, T., Andreae, M. O., Gabriel, R., Kirchstetter, T. W., Mayol-Bracero, O. L., and Ramanathan, V. (2000). Origin of carbonaceous aerosols over the tropical Indian Ocean: Biomass burning or fossil fuels? *Geophysical Research Letters*, 27(24):4061–4064.

Nozière, B., Kalberer, M., Claeys, M., Allan, J., D'Anna, B., Decesari, S., Finessi, E., Glasius, M., Grgić, I., Hamilton, J. F., Hoffmann, T., Iinuma, Y., Jaoui, M., Kahnt, A., Kampf, C. J., Kourtchev, I., Maenhaut, W., Marsden, N., Saarikoski, S., Schnelle-Kreis, J., Surratt, J. D., Szidat, S., Szmigielski, R., and Wisthaler, A. (2015). The Molecular Identification of Organic Compounds in the Atmosphere: State of the Art and Challenges. *Chem. Rev.*, 115(10):3919–3983.

O'Neill, S., Diao, M., Raffuse, S., Al-Hamdan, M., Barik, M., Jia, Y., Reid, S., Zou, Y., Tong, D., West, J., Wilkins, J., Marsha, A., Freedman, F., Vargo, J., Larkin, N., Alvarado, E., and Loesche, P. (2021). A Multi-Analysis Approach for Estimating Regional Health Impacts from the 2017 Northern California Wildfires. *Journal of the Air & Waste Management Association*, 0(ja):null.

Orendorff, C. J., Ducey, M. W., and Pemberton, J. E. (2002). Quantitative Correlation of Raman Spectral Indicators in Determining Conformational Order in Alkyl Chains. *J. Phys. Chem. A*, 106(30):6991–6998.

Orthous-Daunay, F. R., Quirico, E., Beck, P., Brissaud, O., Dartois, E., Pino, T., and Schmitt, B. (2013). Mid-infrared study of the molecular structure variability of insoluble organic matter from primitive chondrites. *Icarus*, 223(1):534–543.

Paglione, M., Gilardoni, S., Rinaldi, M., Decesari, S., Zanca, N., Sandrini, S., Giulianelli, L., Bacco, D., Ferrari, S., Poluzzi, V., Scotto, F., Trentini, A., Poulain, L., Herrmann, H., Wiedensohler, A., Canonaco, F., Prévôt, A. S. H., Massoli, P., Carbone, C., Facchini, M. C., and Fuzzi, S. (2020). The impact of biomass burning and aqueous-phase processing on air quality: A multi-year source apportionment study in the Po Valley, Italy. *Atmospheric Chemistry and Physics*, 20(3):1233–1254.

Paglione, M., Saarikoski, S., Carbone, S., Hillamo, R., Facchini, M. C., Finessi, E., Giulianelli, L., Carbone, C., Fuzzi, S., Moretti, F., Tagliavini, E., Swietlicki, E., Eriksson Stenström, K., Prévôt, A. S. H., Massoli, P., Canaragatna, M., Worsnop, D., and Decesari, S. (2014). Primary and secondary biomass burning aerosols determined by proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy during the 2008 EUCAARI campaign in the Po Valley (Italy). *Atmospheric Chemistry and Physics*, 14(10):5089–5110.

Palen, E. J., Allen, D. T., Pandis, S. N., Paulson, S., Seinfeld, J. H., and Flagan, R. C. (1993). Fourier transform infrared analysis of aerosol formed in the photooxidation of 1-octene. *Atmospheric Environment. Part A. General Topics*, 27(9):1471–1477.

Palen, E. J., Allen, D. T., Pandis, S. N., Paulson, S. E., Seinfeld, J. H., and Flagan, R. C. (1992). Fourier transform infrared analysis of aerosol formed in the photo-oxidation of isoprene and beta-pinene. *Atmospheric Environment. Part A. General Topics*, 26(7):1239–1251.

Pandey, K. K. (1999). A study of chemical structure of soft and hardwood and wood polymers by FTIR spectroscopy. *Journal of Applied Polymer Science*, 71(12):1969–1975.

Pankow, J. F. and Barsanti, K. C. (2009). The carbon number-polarity grid: A means to manage the complexity of the mix of organic compounds when modeling atmospheric organic particulate matter. *Atmos. Environ.*, 43(17):2829–2835.

Parks, D. A., Griffiths, P. R., Weakley, A. T., and Miller, A. L. (2021). Quantifying elemental and organic carbon in diesel particulate matter by mid-infrared spectrometry. *Aerosol Sci. Tech.*, 0(0):1–14.

Parks, D. A., Raj, K. V., Berry, C. A., Weakley, A. T., Griffiths, P. R., and Miller, A. L. (2019). Towards a Field-Portable Real-Time Organic and Elemental Carbon Monitor. *Mining, Metallurgy & Exploration*, 36(4):765–772.

Pathak, R. K., Stanier, C. O., Donahue, N. M., and Pandis, S. N. (2007). Ozonolysis of alpha-pinene at atmospherically relevant concentrations: Temperature dependence of aerosol mass fractions (yields). *J. Geophys. Res.-Atmos.*, 112(D3).

Pavia, D. L., Lampman, G. M., Kriz, G. S., and Vyvyan, J. A. (2008). *Introduction to Spectroscopy*. Brooks Cole, Belmont, CA, fourth edition.

Pereira, K. L., Rovelli, G., Song, Y. C., Mayhew, A. W., Reid, J. P., and Hamilton, J. F. (2019). A new aerosol flow reactor to study secondary organic aerosol. *Atmospheric Measurement Techniques*, 12(8):4519–4541.

Platt, S. M., El Haddad, I., Zardini, A. A., Clairotte, M., Astorga, C., Wolf, R., Slowik, J. G., Temime-Roussel, B., Marchand, N., Ježek, I., Drinovec, L., Močnik, G., Möhler, O., Richter, R., Barmet, P., Bianchi, F., Baltensperger, U., and Prévôt, A. S. H. (2013). Secondary organic aerosol formation from gasoline vehicle emissions in a new mobile environmental reaction chamber. *Atmos. Chem. Phys.*, 13(18):9141–9158.

Pollard, M. and Jaklevic, J. (1988). FOURIER TRANSFORM INFRARED SPECTROMETRY OF AMBIENT AEROSOLS. Technical report, Environmental Protection Agency, Washangton D.C.

Pope, C. A., Ezzati, M., and Dockery, D. W. (2009). Fine-Particulate Air Pollution and Life Expectancy in the United States. *N. Engl. J. Med.*, 360(4):376–386.

Pope, R., Stanley, K. M., Domsky, I., Yip, F., Nohre, L., and Mirabelli, M. C. (2017). The relationship of high PM2.5 days and subsequent asthma-related hospital encounters during the fireplace season in Phoenix, AZ, 2008–2012. *Air Qual. Atmos. Hlth.*, 10(2):161–169.

Price, D. J., Chen, C.-L., Russell, L. M., Lamjiri, M. A., Betha, R., Sanchez, K., Liu, J., Lee, A. K. Y., and Cocker, D. R. (2017). More unsaturated, cooking-type hydrocarbon-like organic aerosol particle emissions from renewable diesel compared to ultra low sulfur diesel in at-sea operations of a research vessel. *Aerosol Sci. Tech.*, 51(2):135–146.

Puxbaum, H., Caseiro, A., Sánchez-Ochoa, A., Kasper-Giebl, A., Claeys, M., Gelencsér, A., Legrand, M., Preunkert, S., and Pio, C. (2007). Levoglucosan levels at background sites in Europe for assessing the impact of biomass combustion on the European aerosol background. *J. Geophys. Res.-Atmos.*, 112(D23):D23S05.

Qi, L., Chen, M., Stefenelli, G., Pospisilova, V., Tong, Y., Bertrand, A., Hueglin, C., Ge, X., Baltensperger, U., Prévôt, A. S. H., and Slowik, J. G. (2019). Organic aerosol source apportionment in Zurich using an extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF-MS) – Part 2: Biomass burning influences in winter. *Atmos. Chem. Phys.*, 19(12):8037–8062.

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

Ramadan, Z., Song, X.-H., and Hopke, P. K. (2000). Identification of Sources of Phoenix Aerosol by Positive Matrix Factorization. *Journal of the Air & Waste Management Association*, 50(8):1308–1320.

Ramanathan, V., Rodhe, H., Agrawal, M., Akimoto, H., Auffhammer, M., Chopra, U. K., Emberson, L., Hasnain, S. I., Iyngararasan, M., Jayaraman, A., Lawrence, M., Nakajima, T., Ruchirawat, M., Singh, A. K., Vincent, J. R., and Zhang, Y. (2008). Atmospheric Brown Clouds: Regional Assessment Report with Focus on Asia Summary. Technical report, United Nations Environment Programme, Nairobi, Kenya.

Reggente, M., Dillner, A. M., and Takahama, S. (2016). Predicting ambient aerosol thermal–optical reflectance (TOR) measurements from infrared spectra: Extending the predictions to different years and different sites. *Atmos. Meas. Tech.*, 9(2):441–454.

Reggente, M., Dillner, A. M., and Takahama, S. (2019a). Analysis of functional groups in atmospheric aerosols by infrared spectroscopy: Systematic intercomparison of calibration methods for US measurement network samples. *Atmos. Meas. Tech.*, 12(4):2287–2312.

Reggente, M., Höhn, R., and Takahama, S. (2019b). An open platform for Aerosol InfraRed Spectroscopy analysis – AIRSpec. *Atmos. Meas. Tech.*, 12(4):2313–2329.

Riedel, T. P., DeMarini, D. M., Zavala, J., Warren, S. H., Corse, E. W., Offenberg, J. H., Kleindienst, T. E., and Lewandowski, M. (2018). Mutagenic atmospheres resulting from the photooxidation of aromatic hydrocarbon and NOx mixtures. *Atmospheric Environment*, 178:164–172.

Robinson, A. L., Donahue, N. M., and Rogge, W. F. (2006a). Photochemical oxidation and changes in molecular composition of organic aerosol in the regional context. *Journal of Geophysical Research: Atmospheres*, 111(D3).

Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop, A. P., Lane, T. E., Pierce, J. R., and Pandis, S. N. (2007). Rethinking Organic Aerosols: Semivolatile Emissions and Photochemical Aging. *Science*, 315(5816):1259–1262.

Robinson, A. L., Subramanian, R., Donahue, N. M., Bernardo-Bricker, A., and Rogge, W. F. (2006b). Source Apportionment of Molecular Markers and Organic Aerosol. 2. Biomass Smoke. *Environmental Science & Technology*, 40(24):7811–7819.

Roden, C. A., Bond, T. C., Conway, S., and Pinel, A. B. O. (2006). Emission Factors and Real-Time Optical Properties of Particles Emitted from Traditional Wood Burning Cookstoves. *Environmental Science & Technology*, 40(21):6750–6757.

Ruggeri, G. (2017). *On the Functional Group Composition of Organic Aerosol.* PhD thesis, Ecole Polytechnique Federale de Lausanne Lausanne (EPFL).

Russell, L. M. (2003). Aerosol Organic-Mass-to-Organic-Carbon Ratio Measurements. *Environ. Sci. Technol.*, 37(13):2982–2987.

Russell, L. M., Bahadur, R., Hawkins, L. N., Allan, J., Baumgardner, D., Quinn, P. K., and Bates, T. S. (2009a). Organic aerosol characterization by complementary measurements of chemical bonds and molecular fragments. *Atmos. Environ.*, 43(38):6100–6105.

Russell, L. M., Bahadur, R., and Ziemann, P. J. (2011). Identifying organic aerosol sources by comparing functional group composition in chamber and atmospheric particles. *PNAS*, 108(9):3516–3521.

Russell, L. M., Takahama, S., Liu, S., Hawkins, L. N., Covert, D. S., Quinn, P. K., and Bates, T. S. (2009b). Oxygenated fraction and mass of organic aerosol from direct emission and atmospheric processing measured on the R/V Ronald Brown during TEXAQS/GoMACCS 2006. *J. Geophys. Res.-Atmos.*, 114(D7):D00F05.

Russo, C., Stanzione, F., Tregrossi, A., and Ciajolo, A. (2014). Infrared spectroscopy of some carbon-based materials relevant in combustion: Qualitative and quantitative analysis of hydrogen. *Carbon*, 74:127–138.

Ruthenburg, T. C., Perlin, P. C., Liu, V., McDade, C. E., and Dillner, A. M. (2014). 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.

Saito, K., Kato, T., Takamori, H., Kishimoto, T., and Fukushima, K. (2005). A New Analysis of the Depolymerized Fragments of Lignin Polymer Using ToF-SIMS. *Biomacromolecules*, 6(5):2688–2696.

Sauvain, J.-J., Vu Duc, T., and Guillemin, M. (2003). Exposure to carcinogenic polycyclic aromatic compounds and health risk assessment for diesel-exhaust exposed workers. *Int. Arch. Occup. Environ. Health*, 76(6):443–455.

Schneider, J., Weimer, S., Drewnick, F., Borrmann, S., Helas, G., Gwaze, P., Schmid, O., Andreae, M. O., and Kirchner, U. (2006). Mass spectrometric analysis and aerodynamic properties of various types of combustion-related aerosol particles. *Int. J. Mass Spectrom.*, 258(1):37–49.

Seinfeld, J. H. and Pandis, S. N. (2016). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons, Hoboken, NJ.

Sharma, D. and Jain, S. (2019). Impact of intervention of biomass cookstove technologies and kitchen characteristics on indoor air quality and human exposure in rural settings of India. *Environment International*, 123:240–255.

Shen, D. K. and Gu, S. (2009). The mechanism for thermal decomposition of cellulose and its main products. *Bioresour. Technol.*, 100(24):6496–6504.

Shen, G., Gaddam, C. K., Ebersviller, S. M., Vander Wal, R. L., Williams, C., Faircloth, J. W., Jetter, J. J., and Hays, M. D. (2017a). A Laboratory Comparison of Emission Factors, Number Size Distributions, and Morphology of Ultrafine Particles from 11 Different Household Cookstove-Fuel Systems. *Environmental Science and Technology*, 51(11):6522–6532.

Shen, G., Preston, W., Ebersviller, S. M., Williams, C., Faircloth, J. W., Jetter, J. J., and Hays, M. D. (2017b). Polycyclic Aromatic Hydrocarbons in Fine Particulate Matter Emitted from Burning Kerosene, Liquid Petroleum Gas, and Wood Fuels in Household Cookstoves. *Energy and Fuels*, 31(3):3081–3090.

Shiraiwa, M., Berkemeier, T., Schilling-Fahnestock, K. A., Seinfeld, J. H., and Pöschl, U. (2014). Molecular corridors and kinetic regimes in the multiphase chemical evolution of secondary organic aerosol. *Atmos. Chem. Phys.*, 14(16):8323–8341.

Shiraiwa, M., Li, Y., Tsimpidi, A. P., Karydis, V. A., Berkemeier, T., Pandis, S. N., Lelieveld, J., Koop, T., and Pöschl, U. (2017a). Global distribution of particle phase state in atmospheric secondary organic aerosols. *Nat. Commun.*, 8:15002.

Shiraiwa, M., Ueda, K., Pozzer, A., Lammel, G., Kampf, C. J., Fushimi, A., Enami, S., Arangio, A. M., Fröhlich-Nowoisky, J., Fujitani, Y., Furuyama, A., Lakey, P. S. J., Lelieveld, J., Lucas, K., Morino, Y., Pöschl, U., Takahama, S., Takami, A., Tong, H., Weber, B., Yoshino, A., and Sato, K. (2017b). Aerosol Health Effects from Molecular to Global Scales. *Environ. Sci. Technol.*, 51(23):13545–13567.

Simon, H., Bhave, P. V., Swall, J. L., Frank, N. H., and Malm, W. C. (2011). Determining the spatial and seasonal variability in OM/OC ratios across the US using multiple regression. *Atmos. Chem. Phys.*, 11(6):2933–2949.

Simoneit, B. R. T., Rogge, W. F., Mazurek, M. A., Standley, L. J., Hildemann, L. M., and Cass, G. R. (1993). Lignin pyrolysis products, lignans, and resin acids as specific tracers of plant classes in emissions from biomass combustion. *Environ. Sci. Technol.*, 27(12):2533–2541.

Slade, J. H. and Knopf, D. A. (2013). Heterogeneous OH oxidation of biomass burning organic aerosol surrogate compounds: Assessment of volatilisation products and the role of OH concentration on the reactive uptake kinetics. *Physical Chemistry Chemical Physics*, 15(16):5898–5915.

Smith, B. C. (1998). *Infrared Spectral Interpretation: A Systematic Approach*. CRC Press, Florida,, 1st edition.

Smith, K. R., Frumkin, H., Balakrishnan, K., Butler, C. D., Chafe, Z. A., Fairlie, I., Kinney, P., Kjellstrom, T., Mauzerall, D. L., McKone, T. E., McMichael, A. J., and Schneider, M. (2013). Energy and Human Health. *Annual Review of Public Health*, 34(1):159–188.

Sobkowiak, M. and Painter, P. (1992). Determination of the aliphatic and aromatic CH contents of coals by FT-i.r.: Studies of coal extracts. *Fuel*, 71(10):1105–1125.

Sobkowiak, M. and Painter, P. (1995). A Comparison of DRIFT and KBr Pellet Methodologies for the Quantitative Analysis of Functional Groups in Coal by Infrared Spectroscopy. *Energy Fuels*, 9(2):359–363.

Srivastava, D., Xu, J., Vu, T. V., Liu, D., Li, L., Fu, P., Hou, S., Moreno Palmerola, N., Shi, Z., and Harrison, R. M. (2021). Insight into $PM_{2.5}$ sources by applying positive matrix factorization (PMF) at urban and rural sites of Beijing. *Atmospheric Chemistry and Physics*, 21(19):14703–14724.

Stefenelli, G., Jiang, J., Bertrand, A., Bruns, E. A., Pieber, S. M., Baltensperger, U., Marchand, N., Aksoyoglu, S., Prévôt, A. S. H., Slowik, J. G., and Haddad, I. E. (2019). Secondary organic aerosol formation from smoldering and flaming combustion of biomass: A box model parametrization based on volatility basis set. *Atmos. Chem. Phys.*, 19(17):11461–11484.

Steinwart, I. and Christmann, A. (2008). *Support Vector Machines*. Springer Science & Business Media.

Subramanian, R., Khlystov, A. Y., Cabada, J. C., and Robinson, A. L. (2004). Positive and Negative Artifacts in Particulate Organic Carbon Measurements with Denuded and Undenuded Sampler Configurations Special Issue of Aerosol Science and Technology on Findings from the Fine Particulate Matter Supersites Program. *Aerosol Sci. Tech.*, 38(sup1):27–48.

Sullivan, A. P., Frank, N., Kenski, D. M., and Collett, J. L. (2011a). Application of high-performance anion-exchange chromatography–pulsed amperometric detection for measuring carbohydrates in routine daily filter samples collected by a national network: 2. Examination of sugar alcohols/polyols, sugars, and anhydrosugars in the upper Midwest. *Journal of Geophysical Research: Atmospheres*, 116(D8).

Sullivan, A. P., Frank, N., Onstad, G., Simpson, C. D., and Collett, J. L. (2011b). Application of high-performance anion-exchange chromatography–pulsed amperometric detection for measuring carbohydrates in routine daily filter samples collected by a national network: 1. Determination of the impact of biomass burning in the upper Midwest. *J. Geophys. Res.-Atmos.*, 116(D8):D08302.

Sullivan, A. P., Holden, A. S., Patterson, L. A., McMeeking, G. R., Kreidenweis, S. M., Malm, W. C., Hao, W. M., Wold, C. E., and Collett, J. L. (2008). A method for smoke marker measurements and its potential application for determining the contribution of biomass burning from wildfires and prescribed fires to ambient PM2.5 organic carbon. *J. Geophys. Res.-Atmos.*, 113(D22):D22302.

Sullivan, A. P., Peltier, R. E., Brock, C. A., de Gouw, J. A., Holloway, J. S., Warneke, C., Wollny, A. G., and Weber, R. J. (2006). Airborne measurements of carbonaceous aerosol soluble in water over northeastern United States: Method development and an investigation into water-soluble organic carbon sources. *Journal of Geophysical Research: Atmospheres*, 111(D23).

Taira, M. and Kanda, Y. (1990). Continuous generation system for low-concentration gaseous nitrous acid. *Anal. Chem.*, 62(6):630–633.

Takahama, S., Dillner, A. M., Weakley, A. T., Reggente, M., Bürki, C., Lbadaoui-Darvas, M., Debus, B., Kuzmiakova, A., and Wexler, A. S. (2019). Atmospheric particulate matter characterization by Fourier transform infrared spectroscopy: A review of statistical calibration strategies for carbonaceous aerosol quantification in US measurement networks. *Atmospheric Measurement Techniques*, 12(1):525–567.

Takahama, S., Johnson, A., and Russell, L. M. (2013). Quantification of Carboxylic and Carbonyl Functional Groups in Organic Aerosol Infrared Absorbance Spectra. *Aerosol Sci. Tech.*, 47(3):310–325.

Takahama, S., Ruggeri, G., and Dillner, A. M. (2016). Analysis of functional groups in atmospheric aerosols by infrared spectroscopy: Sparse methods for statistical selection of relevant absorption bands. *Atmos. Meas. Tech.*, 9(7):3429–3454.

Takahama, S., Schwartz, R. E., Russell, L. M., Macdonald, A. M., Sharma, S., and Leaitch, W. R. (2011). Organic functional groups in aerosol particles from burning and non-burning forest emissions at a high-elevation mountain site. *Atmos. Chem. Phys.*, 11(13):6367–6386.

Tao, F., Gonzalez-Flecha, B., and Kobzik, L. (2003). Reactive oxygen species in pulmonary inflammation by ambient particulates. *Free Radical Biology & Medicine*, 35(4):327–340.

Theodoritsi, G. N., Ciarelli, G., and Pandis, S. N. (2020). Simulation of the evolution of biomass burning organic aerosol with different volatility basis set schemes in PMCAMx-SRv1.0. *Geoscientific Model Development Discussions*, pages 1–33.

Thomas, M. (2017). *Theoretical Modeling of Vibrational Spectra in the Liquid Phase*. PhD thesis, Springer International Publishing, Cham.

Thomas, M., Brehm, M., Fligg, R., Vöhringer, P., and Kirchner, B. (2013). Computing vibrational spectra from ab initio molecular dynamics. *Physical Chemistry Chemical Physics*, 15(18):6608.

Tiitta, P., Leskinen, A., Hao, L., Yli-Pirilä, P., Kortelainen, M., Grigonyte, J., Tissari, J., Lamberg, H., Hartikainen, A., Kuuspalo, K., Kortelainen, A.-M., Virtanen, A., Lehtinen, K. E. J., Komppula, M., Pieber, S., Prévôt, A. S. H., Onasch, T. B., Worsnop, D. R., Czech, H., Zimmermann, R., Jokiniemi, J., and Sippula, O. (2016). Transformation of logwood combustion emissions in a smog chamber: Formation of secondary organic aerosol and changes in the primary organic aerosol upon daytime and nighttime aging. *Atmos. Chem. Phys.*, 16(20):13251–13269.

Tolbert, A. and Ragauskas, A. J. (2017). Advances in understanding the surface chemistry of lignocellulosic biomass via time-of-flight secondary ion mass spectrometry. *Energy Sci. Eng.*, 5(1):5–20.

Tolocka, M. P., Jang, M., Ginter, J. M., Cox, F. J., Kamens, R. M., and Johnston, M. V. (2004). Formation of Oligomers in Secondary Organic Aerosol. *Environ. Sci. Technol.*, 38(5):1428–1434.

Trump, E. R. and Donahue, N. M. (2014). Oligomer formation within secondary organic aerosols: Equilibrium and dynamic considerations. *Atmos. Chem. Phys.*, 14(7):3691–3701.

Turpin, B. J. and Lim, H.-J. (2001). Species Contributions to PM2.5 Mass Concentrations: Revisiting Common Assumptions for Estimating Organic Mass. *Aerosol Sci. Tech.*, 35(1):602–610.

Turpin, B. J., Saxena, P., and Andrews, E. (2000). Measuring and simulating particulate organics in the atmosphere: Problems and prospects. *Atmos. Environ.*, 34(18):2983–3013.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L. (2009). Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data. *Atmos. Chem. Phys.*, 9(9):2891–2918.

Vasireddy, S., Morreale, B., Cugini, A., Song, C., and Spivey, J. J. (2011). Clean liquid fuels from direct coal liquefaction: Chemistry, catalysis, technological status and challenges. *Energy Environ. Sci.*, 4(2):311–345.

Verma, V., Fang, T., Xu, L., Peltier, R. E., Russell, A. G., Ng, N. L., and Weber, R. J. (2015). Organic Aerosols Associated with the Generation of Reactive Oxygen Species (ROS) by Water-Soluble $PM_{2.5}$. *Environ. Sci. Technol.*, 49(7):4646–4656.

Volkamer, R., Jimenez, J. L., Martini, F. S., Dzepina, K., Zhang, Q., Salcedo, D., Molina, L. T., Worsnop, D. R., and Molina, M. J. (2006). Secondary organic aerosol formation from anthropogenic air pollution: Rapid and higher than expected. *Geophys. Res. Lett.*, 33(17).

Wang, N., Jorga, S. D., Pierce, J. R., Donahue, N. M., and Pandis, S. N. (2018a). Particle wall-loss correction methods in smog chamber experiments. *Atmos. Meas. Tech.*, 11(12):6577–6588.

Wang, N., Kostenidou, E., Donahue, N. M., and Pandis, S. N. (2018b). Multi-generation chemical aging of alpha-pinene ozonolysis products by reactions with OH. *Atmos. Chem. Phys.*, 18(5):3589–3601.

Westerling, A. L. (2016). Increasing western US forest wildfire activity: Sensitivity to changes in the timing of spring. *Philos. Trans. R. Soc. B: Biological Sciences*, 371(1696):20150178.

WHO (2014). Indoor Air Quality Guidelines: Household Fuel Combustion. *World Health Organization*, pages 1–172.

Williams, B. J., Jayne, J. T., Lambe, A. T., Hohaus, T., Kimmel, J. R., Sueper, D., Brooks, W., Williams, L. R., Trimborn, A. M., Martinez, R. E., Hayes, P. L., Jimenez, J. L., Kreisberg, N. M., Hering, S. V., Worton, D. R., Goldstein, A. H., and Worsnop, D. R. (2014). The First Combined Thermal Desorption Aerosol Gas Chromatograph—Aerosol Mass Spectrometer (TAG-AMS). *Aerosol Science and Technology*, 48(4):358–370.

Wold, S., Johansson, E., and Cocchi, M. (1993). 3D QSAR in drug design: Theory, methods and applications. *ESCOM, Leiden, Holland*, pages 523–550.

Wold, S., Martens, H., and Wold, H. (1983). The multivariate calibration problem in chemistry solved by the PLS method. In Kågström, B. and Ruhe, A., editors, *Matrix Pencils*, Lect. Notes Math., pages 286–293. Springer Berlin Heidelberg.

Wong, J. P. S., Tsagkaraki, M., Tsiodra, I., Mihalopoulos, N., Violaki, K., Kanakidou, M., Sciare, J., Nenes, A., and Weber, R. J. (2019). Atmospheric evolution of molecular-weight-separated brown carbon from biomass burning. *Atmospheric Chemistry and Physics*, 19(11):7319–7334.

Xia, W., Zhou, C., and Peng, Y. (2017). Improving flotation performance of intruded coal using heavy oil as a collector. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 39(11):1124–1130.

Xie, M., Shen, G., Holder, A. L., Hays, M. D., and Jetter, J. J. (2018). Light absorption of organic carbon emitted from burning wood, charcoal, and kerosene in household cookstoves. *Environmental Pollution*, 240:60–67.

Xie, Q., Li, Y., Yue, S., Su, S., Cao, D., Xu, Y., Chen, J., Tong, H., Su, H., Cheng, Y., Zhao, W., Hu, W., Wang, Z., Yang, T., Pan, X., Sun, Y., Wang, Z., Liu, C.-Q., Kawamura, K., Jiang, G., Shiraiwa, M., and Fu, P. (2020). Increase of High Molecular Weight Organosulfate With Intensifying Urban Air Pollution in the Megacity Beijing. *J. Geophys. Res.-Atmos.*, 125(10):e2019JD032200.

Xu, W., Lambe, A., Silva, P., Hu, W., Onasch, T., Williams, L., Croteau, P., Zhang, X., Renbaum-Wolff, L., Fortner, E., Jimenez, J. L., Jayne, J., Worsnop, D., and Canagaratna, M. (2018). Laboratory evaluation of species-dependent relative ionization efficiencies in the Aerodyne Aerosol Mass Spectrometer. *Aerosol Sci. Tech.*, 52(6):626–641.

Yadav, I. C. and Devi, N. L. (2019). Biomass Burning, Regional Air Quality, and Climate Change. In Nriagu, J., editor, *Encyclopedia of Environmental Health (Second Edition)*, pages 386–391. Elsevier, Oxford.

Yang, H., Yan, R., Chen, H., Lee, D. H., and Zheng, C. (2007). Characteristics of hemicellulose, cellulose and lignin pyrolysis. *Fuel*, 86(12-13):1781–1788.

Yazdani, A., Dillner, A. M., and Takahama, S. (2021a). Estimating mean molecular weight, carbon number, and OM/OC with mid-infrared spectroscopy in organic particulate matter samples from a monitoring network. *Atmos. Meas. Tech.*, 14(7):4805–4827.

Yazdani, A., Dudani, N., El Haddad, I., Bertrand, A., Prévôt, A. S. H., and Dillner, A. M. (2021b). Fragment ion-functional group relationships in organic aerosols using aerosol mass spectrometry and mid-infrared spectroscopy. *Atmos. Meas. Tech. Discuss. [preprint]*.

Yazdani, A., Dudani, N., Takahama, S., Bertrand, A., Prévôt, A. S. H., El Haddad, I., and Dillner, A. M. (2021c). Characterization of primary and aged wood burning and coal combustion organic aerosols in environmental chamber and its implications for atmospheric aerosols. *Atmos. Chem. Phys.*, 21(13):10273–10293.

Yazdani, A., Takahama, S., Kodros John K., Paglione Marco, Masiol Mauro, Stefania, Squizzato, Florou, K., Pandis, S. N., and Nenes, A. (2021d). Chemical evolution of primary and formation of secondary biomass burning aerosols during daytime and nighttime. *Manuscript in preparation*.

Yazdani, A., Takahama, S., Raffuse, S., Dillner, A. M., and Sullivan, A. P. (2021e). Identification of smoke-impacted PM2.5 samples with mid-infrared spectroscopy in a monitoring network. *Manuscript in preparation*.

Yuan, Q., Lai, S., Song, J., Ding, X., Zheng, L., Wang, X., Zhao, Y., Zheng, J., Yue, D., Zhong, L., Niu, X., and Zhang, Y. (2018). Seasonal cycles of secondary organic aerosol tracers in rural Guangzhou, Southern China: The importance of atmospheric oxidants. *Environ. Pollut.*, 240:884–893.

Zahardis, J., Geddes, S., and Petrucci, G. A. (2011). Improved Understanding of Atmospheric Organic Aerosols via Innovations in Soft Ionization Aerosol Mass Spectrometry. *Anal. Chem.*, 83(7):2409–2415.

Zhang, F., Shang, X., Chen, H., Xie, G., Fu, Y., Wu, D., Sun, W., Liu, P., Zhang, C., Mu, Y., Zeng, L., Wan, M., Wang, Y., Xiao, H., Wang, G., and Chen, J. (2020). Significant impact of coal combustion on VOCs emissions in winter in a North China rural site. *Sci. Total Environ.*, 720:137617.

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. (2007). Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. *Geophys. Res. Lett.*, 34(13).

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R., and Sun, Y. (2011). Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: A review. *Anal. Bioanal. Chem.*, 401(10):3045–3067.

Zhang, Q., Worsnop, D. R., Canagaratna, M. R., and Jimenez, J. L. (2005). Hydrocarbon-like and oxygenated organic aerosols in Pittsburgh: Insights into sources and processes of organic aerosols. *Atmos. Chem. Phys.*, 5(12):3289–3311.

Zhang, W., Jiang, S., Wang, K., Wang, L., Xu, Y., Wu, Z., Shao, H., Wang, Y., and Miao, M. (2015). Thermogravimetric Dynamics and FTIR Analysis on Oxidation Properties of Low-Rank Coal at Low and Moderate Temperatures. *Int. J. Coal Prep. Util.*, 35(1):39–50.

Zhang, X., Hecobian, A., Zheng, M., Frank, N. H., and Weber, R. J. (2010). Biomass burning impact on $PM_{2.5}$ over the southeastern US during 2007: Integrating chemically speciated FRM filter measurements, MODIS fire counts and PMF analysis. *Atmospheric Chemistry and Physics*, 10(14):6839–6853.

Zhao, R., Mungall, E. L., Lee, A. K. Y., Aljawhary, D., and Abbatt, J. P. D. (2014). Aqueous-phase photooxidation of levoglucosan – a mechanistic study using Aerosol Time of Flight Chemical Ionization Mass Spectrometry (Aerosol-ToF-CIMS). *Atmos. Chem. Phys.*, 14(7):8819–8850.

Ziemann, P. J. and Atkinson, R. (2012). Kinetics, products, and mechanisms of secondary organic aerosol formation. *Chem. Soc. Rev.*, 41(19):6582–6605.

Amir YAZDANI

Rouges Terres 17, 2068 Hauterive (NE), Switzerland
+41 78 914 67 27
amir.yazdani69@yahoo.com
Skype: amiryz69
LinkedIn

Strengths

- **Data scientist, mechanical and environmental engineer** with extensive knowledge of aerosols, fluid dynamics, and machine learning
- **Project management:** coordinating interdisciplinary projects, multitasking
- **Innovative and problem solver**
- **Trilingual:** English, French, and Farsi

Education

Ph.D. in Environmental Engineering 2017-2021
Ecole Polytechnique fédérale de Lausanne (EPFL), Switzerland

M.Sc. in Mechanical Engineering (Energy Conversion) 2012-2015
Sharif University of Technology, Tehran, Iran

B.Sc. in Mechanical Engineering 2008-2012
Sharif University of Technology, Tehran, Iran

High school (Mathematics and Physics) 2004-2008
Salam High School, Tehran, Iran

Core Experience

Doctoral Assistant 2017-2021
Full-time, Laboratory of Atmospheric Processes and their Impacts (LAPI), EPFL

- Leading five projects in collaboration with international institutes for chemical characterization of aerosols using infrared spectroscopy (FTIR) and mass spectrometry
- Development of a fast, cost-efficient technique for quantification of biomass burning atmospheric aerosols at large air pollution monitoring networks
- Application of machine learning techniques for interpretation of high-dimensional data and building predictive models

Research and Development Engineer 2015-2017
Full-time, R&D Department, MAPNA Turbine Engineering & Manufacturing Co., Karaj, Iran

- Development of a computer code for the aerodynamic design of axial compressors
- Aerodynamic design of a high-pressure multi-stage compressor to upgrade a power gas turbine manufactured by Siemens
- Evaluation and optimization of gas turbine compressors using computational fluid dynamics (CFD)

Research and Development Engineer 2013-2015
Part-time, Advanced Technologies Incubator Center of Sharif University of Technology

- Design of a high-efficiency cascade impactor for the collection of aerosols based on size
- Design and fabrication of the optical and sheath flow modules of a cell counter

Amir YAZDANI
amir.yazdani69@yahoo.com | +41 78 914 67 27 | Rouges Terres 17, 2068 Hauterive (NE), Switzerland

Additional Experience

Internship
Aerosol Technology Laboratory, Sharif University of Technology, Tehran, Iran. 240 hrs 2011 and 2012

- Simulation of particle transport and deposition in pharmaceutical inhalers.

Certificates

PRINCE2 Foundation certificate (project management)

Skills

Programming
MATLAB (advanced, GUI),
R (advanced, statistical modeling, Shiny),
C++ (basic), Python (intermediate)

CAM/CAD/CAE, CFD and misc. software
Siemens NX, Sloid Works, Auto CAD,
ANSYS FLUENT, CFX, Gambit, TurboGrid,
Thermoflow, Tecplot, Latex, Microsoft Office

Technical

- Computational fluid dynamics (CFD)
- Turbulent flow
- Gas turbines and turbo machinery design
- Science communication and technical writing
- Statistics and machine learning
- Data visualization
- Infrared spectroscopy
- Mass spectrometry
- Aerosol physics and chemistry
- Light scattering
- Manufacturing processes

Languages

English: Fluent (C1-C2) **Farsi:** Native **German:** Elementary (A2)
French: Upper-intermediate (B2) **Arabic:** Elementary (A2)

Personal Information

Name: Amir **Surname:** Yazdani **Nationality:** Iranian **Swiss driver's license (B)**
Gender: Male **Birth Date:** 4/14/1990 **Swiss residence/work permit:** B (formation avec activité)

Achievements

- Nominated for EPFL Doctoral Program Thesis Distinction
- Ranked 35th and 12th among more than 300,000 and 450,000 participants in Iran's 2008 national university entrance exams for mathematics/physics and languages disciplines, respectively.
- Iran's National Elites Foundation grant for B.Sc. and M.Sc. studies.
- Merit-based admission to the Master's program (no entrance exam) at Sharif University of Technology (the best technical university in Iran) in 2012.

Extracurricular Activities

Playing santour (Persian instrument): First place in the santour solo performance contest in Tehran in 2007 and first place in the Shamseh festival ensemble contest in 2012.

Selected Publications

- Yazdani, A., Dudani, N., Takahama, S., Bertrand, A., Prévôt, A. S. H., El Haddad, I., and Dillner, A. M.: Characterization of primary and aged wood burning and coal combustion organic aerosols in an environmental chamber and its implications for atmospheric aerosols, Atmos. Chem. Phys., 21, 10273–10293, https://doi.org/10.5194/acp-21-10273-2021, 2021.

- Yazdani A., Normandie, M., Yousefi, M., Saidi, M.S., Ahmadi, G.: Transport and Deposition of Pharmaceutical Particles in Three Commercial Spacer-MDI Combinations, Comput. Biol. Med., 145-155, https://doi.org/10.1016/j.compbiomed.2014.08.001, 2014.

More on ResearchGate: