

1 **A molecular-level approach for characterizing water-insoluble components of ambient organic**
2 **aerosol particulates using ultra-high resolution mass spectrometry**

3
4 **A. S. Willoughby¹, A. S. Wozniak¹, and P. G. Hatcher¹**

5 [1] {Department of Chemistry & Biochemistry, Old Dominion University, Norfolk, VA 23529, USA}

6 Correspondence to: P. G. Hatcher (phatcher@odu.edu)

7
8 **Abstract**

9 The chemical composition of organic aerosols in the atmosphere is strongly influenced by human emissions. The effect these
10 have on the environment, human health, and climate change is determined by the molecular nature of these chemical species. The
11 complexity of organic aerosol samples limits the ability to study the chemical composition, and therefore, the associated properties
12 and the impacts they have. Many studies address the water-soluble fraction of organic aerosols, and have had much success in
13 identifying specific molecular formulas for thousands of compounds present. However, little attention is given to the water-insoluble
14 portion, which can contain most of the fossil material that is emitted through human activity. Here we compare the organic aerosols
15 present in water extracts and organic solvent extracts (pyridine and acetonitrile) of an ambient aerosol sample collected in a rural
16 location that is impacted by natural and anthropogenic emission sources. A semi-quantitative method was developed using proton
17 nuclear magnetic resonance spectroscopy to determine that the amount of organic matter extracted by pyridine is comparable to that of
18 water. Electrospray ionization Fourier transform ion cyclotron resonance mass spectra show that pyridine extracts a molecularly
19 unique fraction of organic matter compared to water or acetonitrile, which extract chemically similar organic matter components. The
20 molecular formulas unique to pyridine were less polar, more aliphatic, and reveal formulas containing sulfur to be an important
21 component of insoluble aerosol organic matter.

23 **1 Introduction**

24 The introduction and improvement of advanced spectroscopic methods has provided a wealth of new chemical information
25 regarding organic aerosols in the last few decades. Establishing the chemical identity of the individual components that make up
26 organic aerosols has remained an important goal in order to understand the relationships between their sources, transport, molecular
27 identities and transformations and their impacts and fates in the environment. The importance of organic aerosols to climate and
28 biogeochemical cycling has been well documented (Andreae and Crutzen, 1997; Andreae and Gelencsér, 2006; Bond et al., 2013;
29 Booth et al., 2012; Crutzen and Andreae, 1990; Jacobson, 2000; Ramanathan et al., 2001). The specific molecular nature of organic
30 aerosols directly influences the impact they have on the environment. For example, light absorption is caused by electronic transition
31 associated with molecular double bonds and the number of double bonds and aromatic rings present within the structure has been
32 linked to the ability of that molecule to absorb ultraviolet or visible light (Andreae and Gelencsér, 2006). Given the connections
33 between chemical characteristics and environmental impact, establishing the relationship between aerosol organic matter (OM) source
34 and associated chemical characteristics is important for modeling and predicting the net impact they have on environmental systems.
35 To date, uncertainties in the molecular makeup of organic aerosols limit our ability to make these linkages between aerosol OM
36 source, chemical characteristics, and environmental impact.

37 Between 10-90% of total aerosol mass is comprised of OM (Jimenez et al., 2009; Kanakidou et al., 2005; and references
38 within), and 10-70% of that OM is water-soluble (Decesari et al., 2007; Kleefeld et al., 2002; Sullivan et al., 2004; Zappoli et al.,
39 1999) depending on its physical and chemical composition. A suite of analytical techniques have provided bulk and specific chemical

40 information about carbonaceous species present within fractions of atmospheric OM samples (Hallquist et al., 2009), including
41 ambient aerosols, fog water and rainwater, and laboratory generated secondary organic aerosols (SOA). These efforts have shown
42 aerosol OM to be made up of a highly diverse suite of oxygenated compounds that include aliphatic and conjugated functional groups
43 (e.g., Decesari et al., 2007; Jimenez et al., 2009; Mayol-Bracero et al., 2002; Wozniak et al., 2008), which influence the water-
44 solubility as well as the light-absorbing capacity of the compounds (Andreae and Gelencsér, 2006; Robertson and O'Reilly, 1987).
45 Recently, ultra-high resolution mass spectrometry has shown an extraordinary capacity for characterizing aerosol OM, and is the only
46 current technique able to provide elemental formula information for the thousands of compounds present within individual ambient
47 aerosols (Lin et al., 2012; Mazzoleni et al., 2012; Reemtsma et al., 2006; Schmitt-Kopplin et al., 2010; Wozniak et al., 2008) and
48 laboratory generated aerosols (Bateman et al., 2010; Heaton et al., 2009; Laskin et al., 2010; Reinhardt et al., 2007) without prior
49 chromatographic separation. In addition to oxygenated compounds, many other functionalized species, including those containing
50 nitrogen and sulfur, have been identified in ambient aerosols, fog water (LeClair et al., 2012; Mazzoleni et al., 2010), and rainwater
51 (Altieri et al., 2012; Altieri et al., 2009; Mitra et al., 2013) from various sources.

52 Numerous publications characterize some fraction of aerosol OM using various solvent systems, ionization sources, and mass
53 analyzers. To date, much of the work has focused on the water-soluble fraction, leaving little information regarding the water-
54 insoluble fraction that comprises 30 - 90% of the OM (Decesari et al., 2007; Kleefeld et al., 2002; Sullivan et al., 2004; Zappoli et al.,
55 1999). Radiocarbon data for organic aerosols collected in rural environments suggests that much of the water-insoluble carbon is
56 fossil-derived and from anthropogenic sources, whereas the water-soluble carbon is contemporary and biogenically-derived (Szidat et

57 al., 2004; Wozniak et al., 2012a; Wozniak et al., 2012b). Limiting molecular analysis to the water-soluble OM (WSOM) means that a
58 quantitatively important component of organic aerosols is missing. Studies have examined the methanol-soluble (Heaton et al., 2009)
59 and acetonitrile-soluble (Bateman et al., 2010; Heaton et al., 2009) fractions of SOAs and other laboratory generated aerosols. Studies
60 of SOA, which are thought to be highly water-soluble, have shown that the water and acetonitrile extracted OM fractions are
61 extremely similar at the molecular level (Bateman et al., 2010; Heaton et al., 2009). However, Bateman et al. (2010) examined a
62 laboratory-generated biomass burning aerosol and determined that the acetonitrile-soluble component of aerosol OM had
63 characteristically lower oxygen to carbon (O/C) ratios than the water-soluble component, suggesting molecular differences between
64 water-soluble and solvent-soluble components of some organic aerosols. These differences imply that the water-insoluble materials in
65 aerosol OM have molecular characteristics distinct from water-soluble OM. However, the differences in water-insoluble and water-
66 soluble aerosol OM have still not been extensively explored.

67 In the present study we evaluate the specific molecular composition of the water-insoluble fraction of ambient aerosol OM.
68 We employ Fourier transform ion cyclotron resonance mass spectrometry with negative electrospray ionization (ESI-FTICR-MS) for
69 molecular characterization and NMR spectroscopy for additional structural characterization. We select pyridine and acetonitrile, two
70 solvents that exhibit different polarities than water, have different solvating characteristics, and are compatible with ESI-FTICR-MS
71 and NMR.

72 Three ambient aerosol particulate samples were collected from a coastal Virginia (USA) site influenced by a mixture of
73 anthropogenic and biogenic aerosol OM sources typical to the east coast of the United States. The suitability of pyridine and

74 acetonitrile as solvents for characterizing aerosol water-insoluble organic matter (WIOM) by ESI-FTICR-MS was established by
75 comparing the molecular formulas assigned to the pyridine-soluble organic matter (PSOM), acetonitrile-soluble organic matter
76 (ASOM), and WSOM extracts of these ambient aerosol particulate samples.

77 **2 Methods**

78 **2.1 Sample Collection**

79 Three ambient aerosol total suspended particulate (TSP) samples were collected during the summer (16-17 August 2011; 24-25
80 June 2013; and 25-26 June 2013) at the Virginia Institute of Marine Science in Gloucester Point, VA (37.2482 N, 76.5005 W). Air
81 was drawn through a pre-combusted (4 h, 475°C) quartz microfiber filter (Whatman QM/A, 20.3 x 25.4 cm, 419 cm² exposed area, 0.6
82 µm effective pore size) using a TSP high volume air sampler (Model GS2310, Thermo Andersen, Smyrna, GA) at an average flow
83 rate of 0.81 m³ min⁻¹. Air particles were collected for 24 hours with total air volumes ranging between 1124-1169 m³, and are
84 expected to have contributions from several nearby biogenic and anthropogenic sources (e.g., estuarine waters, natural vegetation,
85 light vehicle traffic, oil refinery, coal-fired power plant) as is typical of the eastern United States. The samples were transferred to a
86 pre-combusted foil pouch immediately after collection and stored at -8°C until analysis. A new and pre-combusted QM/A filter blank
87 was stored under identical conditions as the sample, and was analyzed as a storage filter blank for the 2011 sample. A new and pre-
88 combusted QM/A filter was attached to the air sampler immediately prior to the 2013 sampling (24 June 2013) and then stored under
89 identical conditions as the samples, and was analyzed as a field filter blank for the 2013 samples.

90 **2.2 Aerosol Mass and Carbon Measurements**

91 The QM/A filters were weighed before and after sampling to determine the total aerosol mass (24.3 – 29.7 mg TSP) and
92 concentration (20.8 - 26.2 $\mu\text{g}/\text{m}^3$). Triplicate aerosol core plug (2.84 cm^2 area) samples were acidified to remove inorganic carbon by
93 treating the plugs with 1M HCl followed by drying in an oven (4h, 80°C). Acidified filter plugs were evaluated for total carbon (C =
94 organic + elemental/black – inorganic carbonates) and nitrogen (N) content using a Flash EA 1112 elemental analyzer
95 (ThermoFinnigan). Quantification was achieved using an aspartic acid standard calibration curve. Acidified blank QM/A filters were
96 evaluated for blank subtraction; however, C and N quantities for all of the blank QM/A filter plugs were below the detection levels of
97 the instrument.

98 **2.3 Solvent Extractions**

99 Replicate solvent extracts of the aerosols and the respective filter blank were obtained by combining three aerosol plugs or
100 blank plugs (3 aerosol plugs contain between 0.116 and 0.147 mg C) with either 15 mL water (Millipore Synergy Ultrapure Water
101 System), 5 mL pyridine (Acros, 99+%), or 5 mL acetonitrile (Fisher Optima, 99.9%) and thoroughly mixed on an orbital shaker table
102 (150 RPM, 4h, 21°C). Soluble particles were isolated using a syringe filter with a pre-combusted glass fiber filter (0.7 μm pore size).
103 The water extracts were then desalted in order to remove salts that can limit the ionization of OM by ESI and concentrated using an
104 established procedure for PPL solid-phase extraction cartridges (Dittmar et al., 2008). PPL is expected to retain 60-75% of OM
105 (Dittmar et al., 2008; Stubbins et al., 2012). Low molecular weight and exceptionally hydrophilic compounds are expected to be
106 among the losses. Compounds with a low molecular weight (<200 Da) are not detected under the FTICR-MS conditions used here, so
107 the negative effects of PPL extraction are expected to be minimal. The desalted sample was eluted in 5 mL of methanol (WSOM,

108 Acros, 99.9%). The pyridine filtrate (PSOM) and acetonitrile filtrate (ASOM) did not require further manipulation, and all three
109 extracts were thus prepared so that the final solutions had the same filter plug to solvent volume ratio (3 plugs in 5 mL of solvent).
110 Samples were stored at -8°C until FTICR-MS analysis, typically within 24 hours of preparation.

111 **2.4 Extraction efficiency determinations**

112 The filtrates from the water-extracts were evaluated for non-purgeable organic carbon using a Shimadzu TOC-VCPH analyzer
113 to measure water soluble organic carbon (WSOC) content (Wozniak et al., 2008; Wozniak et al., 2012a). The WSOC content was then
114 compared with the total aerosol organic carbon content to find that 54 – 60% of the total aerosol organic carbon was WSOC.

115 An important aspect of choosing a suitable solvent for WIOM characterization is its extraction efficiency which could not be
116 accurately determined for acetonitrile and pyridine using standard TOC analysis. Upon evaporation of solvent extracts, these organic
117 solvents adhere to natural organic matter artificially inflating the carbon content in acetonitrile and pyridine extracts. Therefore, a
118 method was developed based on ¹H NMR using glucose (98%, Acros Organics) as a standard, which is 100% soluble in pyridine at
119 low concentrations (<1mg/mL). Acetonitrile interferes with our ¹H NMR quantification strategy due to a strong signal from
120 acetonitrile hydrogen (occurring at ~2 ppm) that overlap with the signal from OM (occurring between 0.1-4.4 ppm). This overlapping
121 signal impedes our ability to determine the amount of proton signal derived from the sample and precludes a reliable calculation of
122 extraction efficiency for acetonitrile. A known mass of glucose was dissolved into pyridine-D₅ (100% atom D, Acros Organics)
123 providing known carbon and hydrogen concentrations to compare to ¹H NMR spectra. Aerosol plugs of known OC masses were each
124 dissolved into pyridine-D₅ and water. The WSOM and PSOM samples were then subjected to ¹H NMR spectroscopy using a Bruker

125 Daltonics 400 MHz NMR with a BBI probe. WSOM samples were diluted using D₂O (100% atom D, Acros Organics) at a ratio of
126 90:10 WSOM:D₂O. PSOM samples were analyzed in a solvent system of 100% pyridine-D₅. All samples were analyzed for 4000
127 scans using a standard Bruker water-suppression pulse program, where the 90° pulse and the transmitter offset were optimized
128 individually for each sample.

129 The signals obtained from ¹H NMR spectra were integrated over the range of 0.1 to 4.4 ppm to get a total signal response, and
130 also integrated over three ranges of chemical shifts to determine contributions from the major proton types (Moretti et al., 2008;
131 Shakya et al., 2012), aliphatic, unsaturated alkyl, and oxygenated aliphatics. Aliphatic hydrogen (H-C, 0.6-1.8 ppm for WSOM and
132 0.7-1.95 ppm for PSOM), unsaturated alkyl or hydrogen α to unsaturated carbons (H-C-C=, 1.8-3.2 ppm for WSOM and 1.95-3.2 ppm
133 for PSOM), and oxygenated aliphatic hydrogen including alcohol, ether, and ester function groups (H-C-O, 3.2-4.4 ppm in WSOM
134 and PSOM). Though aromatic protons (6.5-8.2 ppm) are used in other studies, we omit them from analysis due to the interference of
135 exchanged protons in the pyridine-D₅ solvent. The region of 0.1-4.4 ppm is appropriate for this study because it is free of signals from
136 the solvent and contains the majority of signal for these types of samples. The signal response for each proton region was normalized
137 to the total signal response between 0.1-4.4 to establish the percent contribution for each proton type. These percentages were used to
138 calculate the average H/C ratio for each sample to be used to convert hydrogen to carbon content (see Supplementary Table 1). An
139 H/C ratio of 2 is used for the aliphatic hydrogen and unsaturated alkyl hydrogen, and an H/C ratio of 1.1 is used for the oxygenated
140 aliphatics (Decesari et al., 2007; Moretti et al., 2008; Shakya et al., 2012).

141 The signal obtained for glucose protons dissolved in pyridine-D₅ were used to establish a glucose response factor (area per
142 mass unit hydrogen). The total area between 0.1-4.4 ppm in each sample was converted into a mass of dissolved hydrogen using the
143 glucose response factor. The dissolved hydrogen mass was converted to dissolved carbon mass using the calculated average H/C ratio
144 (1.94-1.98). Comparison of the calculated dissolved carbon amount to the starting mass of carbon gives a relative percentage of
145 extractable organic matter. Details on these values are provided in Supplementary Table 2.

146

147 **2.5 ESI-FTICR-MS**

148 Immediately prior to analysis, the WSOM replicates (in methanol) and filter blank extract were diluted by 2 using LC/MS
149 grade water with a small amount (<0.1% total volume) of ammonium hydroxide to enhance ionization efficiency. The PSOM,
150 ASOM, and respective filter blank extracts were diluted by 2 using methanol with a small amount (<0.1% total volume) of ammonium
151 hydroxide. Prior to sample analysis, the instrument was externally calibrated using a polyethylene glycol standard. Each of the
152 samples was introduced to an Apollo II electrospray ionization source (negative ion mode) at a flow rate of 120 nL/min on a Bruker
153 Daltonics 12 Tesla Apex Qe FTICR-MS housed at the College of Sciences Major Instrumentation Cluster at Old Dominion
154 University. Spray voltages were optimized for each sample. Ions were accumulated in the hexapole for 0.5-2.0 seconds before
155 transfer into the ICR cell, where exactly 300 transients were co-added. Experimental duplicates were evaluated for each aerosol
156 sample and solvent mixture to ensure good experimental and instrumental reproducibility.

157 **2.6 Data Processing**

158 Each spectrum was calibrated internally using naturally occurring molecules (fatty acids, dicarboxylic acids, and other
159 homologous series with only carbon, hydrogen, and oxygen in the molecular formula) within the sample (Sleighter et al., 2008). Salt
160 peaks (mass defect 0.4-0.98 for $m/z < 400$, and mass defect 0.6-0.97 for $m/z < 400$), blank peaks (those present in the respective QM/A
161 filter blank), and ^{13}C isotopologue peaks were removed from each mass list prior to formula assignments. Additionally, each set of
162 duplicates was evaluated for common m/z , where only common m/z were used for molecular formula assignments. Each set of
163 duplicates was threshold-corrected for peaks that were below the S/N 3 threshold, but above S/N 2.5 (for example, if a peak was
164 present at S/N 3.1 in one sample, and S/N 2.9 in the second sample, then it was considered common; Sleighter et al., 2012). This
165 corrects for minor differences in peak magnitude that may cause a peak to go undetected, when it is present at a magnitude slightly
166 below the method detection limit. Each set of duplicates had more than 67% of the peaks (300-600 m/z) in common, indicating good
167 instrumental and experimental reproducibility (Sleighter et al., 2012).

168 Molecular formulas were assigned to m/z common to duplicates using an in-house MatLab (The Math Works, Inc., Natick,
169 MA) code according to the criteria $^{12}\text{C}_{5-50}$, $^1\text{H}_{5-100}$, $^{14}\text{N}_{0-5}$, $^{16}\text{O}_{1-30}$, $^{32}\text{S}_{0-2}$, and $^{31}\text{P}_{0-2}$, where the subscripts indicate the range of atoms
170 allowed in a formula. The assigned formulas were screened to eliminate any chemically unreasonable formulas for naturally
171 occurring organic compounds. The criteria for formula assignments are consistent with published procedures (Stubbins et al., 2010;
172 Wozniak et al., 2008). Most (79%-96%) of the common peaks between 200-800 m/z could be assigned a formula. A large majority
173 (>90%) of the formulas are within 0.5 ppm agreement of the measured m/z , and all formulas are within 1 ppm error.

174 **3 Results and Discussion**

175 **3.1 NMR spectroscopy**

176 One of the initial concerns for this comparison was to determine whether the organic solvents chosen were as efficacious as
177 water for removing OM from aerosol particulates. We resorted to ^1H NMR spectroscopy to evaluate the relative proportion of OM
178 extracted by water and pyridine. Pyridine has been demonstrated to extract significant quantities of macromolecular OM from natural
179 samples that have distinctly low water-solubility (e.g., coals, soils, kerogen, etc.; McKee and Hatcher, 2010; Salmon et al., 2011; Wu
180 et al., 2003). Extracts of pyridine lend themselves well to recovery estimates using ^1H NMR. By integrating the peaks in the main
181 resonance absorption region of the spectrum of aerosol PSOM, between 0.1 and 4.4 ppm, we determined the area response from the
182 sample and used the glucose response factor (area per mass unit of hydrogen) to determine the amount of carbon present in the PSOM.
183 The starting mass of carbon for each aerosol is known, and the percent of extractable carbon, and therefore extractable OM, could be
184 determined. We determined that 31 – 59% of the aerosol OM is soluble in pyridine, which is comparable to what is found for WSOM
185 (54 – 60%). We recognize that assumptions were made in order to determine these values, and that OM solubility in pyridine will vary
186 with sample type; however, we are confident in these values due to the matched signal response in both PSOM and WSOM samples.

187 Acetonitrile has been used in the characterization of laboratory-generated aerosol OM in previous studies (Bateman et al.,
188 2010; Bateman et al., 2008; Heaton et al., 2009; Laskin et al., 2010; Reinhardt et al., 2007) and is examined by FTICR-MS in this
189 study for this reason. We did not evaluate the acetonitrile extraction efficiency by NMR because this solvent displays its main signal
190 in the 0.1 to 4.4 ppm region and a comparison like the one made with pyridine was not feasible. However, we can speculate that the
191 efficiency is comparable to that of water and pyridine, considering its relative polarity as a solvent.

192 While the primary motivation for obtaining ^1H NMR spectra was to evaluate extraction efficiencies, the information contained
193 therein is valuable for bulk characterization. A more detailed structure characterization is beyond the scope of this manuscript but will
194 be the subject of future work. Figure 1 shows the ^1H NMR spectra for the WSOM and PSOM fractions for one of the aerosol samples
195 (collected 25-26 June 2013) and the table inset gives the chemical shifts and average relative intensities (for all three aerosol samples)
196 and standard deviations. Both spectra are dominated by aliphatic signals if one does not consider the strong resonances in the aromatic
197 region of the PSOM spectrum assigned to pyridine protons. The peak positions are different for PSOM giving the impression that
198 functionalized structures in WSOM are different from those of PSOM. For example, the methylenic peak (CH_2) in WSOM spectrum
199 falls between 1.1-1.2 ppm, and is between 1.2-1.5 in the PSOM spectrum. However, it is likely that we are observing peak shifts due
200 to solvent interactions (Sanders and Hunter, 1993), and some of the peaks in the PSOM spectrum are shifted downfield with respect to
201 the WSOM spectrum. A majority of the signal is due to aliphatic hydrogen in both spectra. However, this group is proportionally of
202 greater intensity (73.9% of spectral intensity) in the PSOM spectrum relative to all other resonances. Methyl protons (CH_3 at 0.7-0.8
203 ppm for WSOM and 0.8-0.9 ppm for PSOM) are nearly in the same relative proportions to methylenic (CH_2 at 1.1-1.2 ppm for
204 WSOM and 1.2-1.5 ppm for PSOM) protons in both extracts (a CH_2/CH_3 ratio of 4 for both WSOM and PSOM). This suggests
205 aliphatic structures of similar chain lengths.

206 Additional differences between the WSOM and PSOM spectra are found in the regions downfield of 1.7 ppm, regions typically
207 associated with protons near electron withdrawing functional groups (e.g., hydroxyl, carboxyl, carbonyl, and amino groups). The
208 WSOM extract shows higher relative peak areas in this region, evidence for a greater relative abundance of these types of resonances.

209 Protons found in this region make up 41.8% of the spectral area for WSOM and 25.4% of the spectral area for PSOM. The peaks
210 between 2.0 and 2.8 ppm in both spectra are characteristic of hydrogen attached to a carbon alpha to an unsaturated carbon (H-C-C=),
211 which represent alkenes and carbonyl groups. The relative abundance is higher in the WSOM than PSOM (33.5% and 21.6%,
212 respectively). This higher abundance suggests that WSOM is more selective for unsaturated compounds, which absorb light (Andreae
213 and Gelencsér, 2006), and WSOM may contribute more light absorption on a per carbon basis than PSOM. The broad peak between
214 3.2 and 4.0 ppm in the WSOM spectrum is attributed to hydrogen attached to carbon adjacent to oxygen or nitrogen could represent
215 polyols, carbohydrate-like materials, or peptide-like materials. Their relative abundance in the PSOM spectrum is less than in the
216 WSOM spectrum (3.8% versus 8.3% spectral intensity) demonstrating that the two solvents extract chemically distinct portions of
217 aerosol OM. The NMR spectral information thus shows a greater preponderance of signals from protons near functionalized structures
218 and unsaturated carbons in the WSOM and a higher relative signal for aliphatic protons in the PSOM.

219 **3.2 Mass spectra and molecular formula characteristics**

220 The ESI-FTICR mass spectra for ambient aerosol WSOM are comparable to previously published FTICR-MS spectra of
221 aerosol WSOM extracts (Lin et al., 2012; Mazzoleni et al., 2012; Schmitt-Kopplin et al., 2010; Wozniak et al., 2008) averaging
222 thousands of peaks across a broad range of 200-800 m/z; the PSOM and ASOM also contained thousands of peaks across the same
223 mass range. A representative full spectrum for each solvent can be found in the supplemental information (Supplemental Fig. 1).
224 There are clear spectral differences between the WSOM and the solvent extracts (i.e., PSOM and ASOM), including differences in the
225 presence of some peaks, as well as the relative magnitudes of common peaks (Fig. 2).

226 Figure 2 shows a WSOM, PSOM, and ASOM spectra for one of the aerosol samples (collected August 16-17, 2011) expanded
227 at a single nominal mass ($m/z=427$). This distribution of peak intensities is representative of odd nominal masses across the spectral
228 range. Each of the peaks with S/N of at least 3 is labeled with colored shapes to denote the elemental makeup of the assigned
229 molecular formula. There are a few peaks that were not assigned a molecular formula because no chemically reasonable formula was
230 possible using the given criteria suggesting these molecular formulas contain elements other than C, H, N, O, S, and/or P. A Kendrick
231 mass defect plot for formulas differing by a CH_2 group (Supplemental Fig. 2; Kendrick 1963) shows that the formulas identified in
232 Fig. 2 are involved in Kendrick mass defect series that span nearly the entire mass range (200-800 m/z). One of the striking
233 differences is the presence of more peaks at the low mass defect region (below 427.1 m/z) in the PSOM (Fig. 2b). PSOM has the most
234 peaks in this low mass defect region (masses having a decimal less than 0.1) throughout the entire spectrum when compared to
235 WSOM and ASOM (Supplemental Fig. 3). Peaks with low mass defect are either deficient in hydrogen or rich in oxygen, which
236 suggests that PSOM contains either more unsaturated formulas or more highly oxygenated species than does the WSOM. Those
237 specific peaks at 427.05294 m/z and 427.05518 m/z represent CHOS compounds with high O/C ratios (>0.4) and moderate H/C ratios
238 (1.0-1.5), indicating that the selectivity is likely due to the compounds being more highly oxygenated rather than hydrogen deficient.
239 In addition to peak presence and absence, there is a difference in the peak intensity distributions between the three solvents. The most
240 intense peak in the WSOM (Fig. 2a) is in the center of all peaks for that nominal mass at 427.19723 m/z , and has been assigned as a
241 formula containing only CHO ($\text{C}_{21}\text{H}_{32}\text{O}_9$). This peak is present in both ASOM and PSOM, but the relative intensity is considerably
242 lower. The most intense peak (427.09168 m/z) in the ASOM and PSOM is a CHOS formula ($\text{C}_{15}\text{H}_{24}\text{O}_{12}\text{S}$) located at a lower mass

243 defect than the most intense peak in the WSOM spectrum. This peak is also present in WSOM, but at a much lower relative intensity.
244 These differences in relative magnitude of peaks and the presence or absence of some peaks reflect the differences in the ability of
245 these three solvents to extract and detect certain compounds. Recently, the relative magnitude of peaks for compounds detected using
246 ESI-FTICR-MS have been shown to be quantitatively significant and reproducible in a consistent solvent system (Kamga et al., 2014),
247 and have also been used to highlight the differences in sample composition in ESI-FTICR-MS studies of aerosol OM (e.g., Mazzoleni
248 et al., 2012). However, differences in solvent composition have not been evaluated to determine how the solvent affects ESI
249 efficiency, so the remainder of the discussion of the data focuses on differences in presence and absence of peaks in water, pyridine
250 and acetonitrile extracts to provide a qualitative view of WIOM components detected by ESI-FTICR-MS.

251 A majority of the detected peaks in each extract were assigned molecular formulas within 0.5 ppm error. All of the assigned
252 formulas present in any of the three aerosol samples were combined for each of the three solvents to generate master lists containing
253 every molecular formula assigned to each WSOM, PSOM, and ASOM. Additionally, the master lists for the three solvents were
254 combined to evaluate the molecular properties of the aerosol sample as a whole. The master lists for each solvent were compared for
255 distinct molecular characteristics. The molecular characteristics for each solvent are arranged by their elemental makeup (e.g., number
256 and percentage of CHO formulas) and numerical averages for O/C, H/C and a modified aromaticity index (AI_{mod} , Eq. 1), shown in
257 Table 1. For simplicity and due to low number frequency, all formulas containing phosphorus (CHOP, CHONP, and CHOSP) have
258 been grouped together and reported as CHOP(N,S). AI_{mod} is a metric established by Koch and Dittmar (2006) that estimates the
259 degree of aromaticity of an organic compound using only its molecular formula. AI_{mod} is calculated as follows:

260

261
$$AI_{\text{mod}} = (1 + C - 0.5O - S - 0.5H)/(C - 0.5O - S - N - P), \quad (1)$$

262

263 for any molecular formula comprised of C, H, N, O, S, and P. AI_{mod} is similar to other aromaticity metrics, such as carbon normalized
264 double bond equivalent (Hockaday et al., 2006), and it includes heteroatoms as points of unsaturation and conservatively assumes that
265 half of all O participate in a double bond, such as that of a carbonyl group. A low AI_{mod} indicates a low degree of aromaticity where a
266 value of zero is an aliphatic compound, a value between 0-0.5 is representative of olefinic compounds (containing at least one double
267 bond) and includes alicyclic molecules. A high AI_{mod} indicates a higher degree of aromaticity where a compound having a value
268 above 0.5 is aromatic, and a value of 0.67 or higher indicates condensed aromatic compounds (fused aromatic rings; Koch and
269 Dittmar, 2006). Aromatic and condensed aromatic compounds play an important role in the light-absorbing ability of organic aerosols
270 (Andreae and Gelencsér, 2006).

271 Before analyzing the molecular properties of the individual solvents, every molecular formula (5106 formulas total) assigned
272 to any of the three solvents was combined to evaluate the aerosol samples as a whole. The Venn diagram shown in Fig. 3a shows the
273 percentage of the total formulas unique to each of the solvents, and the ones found in one or more of the solvents. Of all the formulas,
274 WSOM analyzes the largest fraction (3396, 67%). There are 3152 formulas (62%) identified in ASOM, and 2397 formulas (47%)
275 found in PSOM. There are many formulas common between WSOM and the organic solvent extracts, where 1077 formulas (21%) are
276 found in all three solvents, 1367 (27%) are common with PSOM, and 2072 (41%) are common with ASOM. There are 1710 formulas

277 (33%) found in the PSOM and/or ASOM spectra that are not present in WSOM spectra (Fig. 3a). The formulas present in ASOM
278 and/or PSOM, but not in WSOM, are representative of compounds that are either (1) not water-soluble or (2) outside the analytical
279 window of WSOM using ESI-FTICR-MS. Because of this, we refer to these compounds as “water-insoluble” organic matter (WIOM).
280 It is important to emphasize that the ability of a compound to be detected by ESI-FTICR-MS is highly dependent on its ability to be
281 ionized by negative electrospray ionization, thus any compound that does not contain a polar ionizable functional group (e.g.,
282 hydrocarbons) will not be analyzed in any of the solvents and some unknown portion of WIOM will go undetected. One thing to note
283 is that 67% of the formulas are found in WSOM suggesting that WSOM is more complex molecularly than WIOM despite extracting
284 similar amounts of material. This is likely due, in part, to the poor ionization efficiency of aliphatic material and compounds low in
285 oxygen. The ¹H NMR of the PSOM shows us that pyridine is selective for more aliphatic compounds and compounds that are not
286 extensively substituted with polar functional groups.

287 The 5106 individual formulas assigned to the three solvents are made up of 2051 CHO (40%), 1472 CHOS (29%), 857 CHON
288 (17%), 599 CHONS (12%), and 127 CHOP(N,S) (2.5%, Fig. 3b). Most (>95%) of the detected compounds are classified as either
289 aliphatic (40%, 2043 formulas) or olefinic (56%, 2837 formulas) based on their low AI_{mod} values (Fig. 3c). Both anthropogenic and
290 biogenic emission sources release OM that fall under the classification of aliphatic and olefinic such as alkanes, alkenes, alkanolic
291 acids, alkenolic acids, alkanals, alkenals, and terpenes (Rogge et al., 1993a, b, c), which can act as precursors to SOA. The polar
292 compounds are formed through atmospheric processing reactions (i.e., photooxidation or reaction with NO_x and SO_x) with these
293 biogenically and anthropogenically released precursor molecules (Alfarra et al., 2006; Andreae and Gelencsér, 2006; Jacobson, 2000).

294 Less than 5% of all formulas are classified as aromatic (3.6%, 183 formulas) or condensed aromatic (0.8%, 43 formulas) based on
295 their high AI_{mod} . The low abundance of these aromatic species identified in these samples suggests only small contributions from
296 combustion sources, which are known sources of carbonaceous aromatic compounds such as black carbon, oxygenated aromatic
297 compounds, and polycyclic aromatic hydrocarbons. The sampling site from which these aerosols were collected is influenced by local
298 combustion sources (i.e., coal-fired power plant and light vehicular traffic), but not to a large extent. This observation is supported by
299 a previous study of aerosol OM from a similar rural site located 36 km north of our sampling site. They show that aromatic
300 compounds made up a quantitatively small (less than 1% on average) component of the total organic carbon and that only half of that
301 aromatic OM was derived from fossil sources (Wozniak et al., 2012b). The small amount of aromatic material in the WSOM sample
302 detected by FTICR-MS is supported by the lack of intensity (<3% of total spectral intensity) of peaks that represent aromatic protons
303 (chemical shift 6.5 – 8.2 ppm) in the ^1H NMR spectra (Supplemental Fig. 4a). Unfortunately, aromatic protons from the aerosol OM
304 are not distinguishable from the exchanged protons in the pyridine- D_5 ^1H NMR spectra (Supplemental Fig. 4b); therefore, no
305 conclusion can be made about the amount of aromatic compounds in the PSOM based on ^1H NMR.

306 The WSOM mass spectra contain the greatest number of molecular formulas (3396), which are dominated by 1563 CHO
307 (46%) and 868 CHOS (26%) formulas followed by CHON (671 formulas, 20%), CHONS (214 formulas, 6.3%), and CHOP(N,S) (80
308 formulas, 2.4%), as displayed in Fig. 3b. This distribution of molecular formulas, specifically the dominance by CHO formulas, is
309 consistent with other ambient aerosol samples collected in the eastern United States (Wozniak et al., 2008). There are more CHOS

310 than CHON formulas, suggesting that sulfur species (e.g., SO_4^{2-}) were important to the atmospheric processes in this region at the time
311 of sampling.

312 The PSOM contains the fewest number of total formulas (2397); however, the molecular characteristics of PSOM are distinct
313 from that of WSOM or ASOM. PSOM contains mostly CHOS (976 formulas, 41%) and CHO (613 formulas, 26%) formulas,
314 followed by CHONS (489 formulas, 20%), CHON (294 formulas, 12%), and CHOP(N,S) (25 formulas, 1.0%) formulas (Fig. 3b).
315 Nearly half of the PSOM formulas (1030 formulas, 43%) are not found in WSOM (WIOM_{pyr}), indicating that they are either truly
316 water-insoluble compounds or do not ionize well in water due to being suppressed by other WSOM components with higher
317 ionization efficiencies. The WIOM_{pyr} is dominated by sulfur-containing species, 442 CHOS (43%) and 311 CHONS (30%) formulas,
318 with smaller contributions from CHO (155 formulas, 15%), CHON (101 formulas, 9.8%), and CHOP(N,S) (21 formulas, 2.0%)
319 compounds (Table 1). A large majority (88%) of the sulfur formulas in the WIOM_{pyr} have an O/S ratio of at least 4, suggesting that
320 many of these formulas represent organosulfates and nitrooxyorganosulfates. Additionally, many of the sulfur compounds are
321 aliphatic (65%) according to their AI_{mod} values. This sulfur predominance suggests that pyridine could either be 1) selective towards
322 organic compounds with sulfur as part of its molecular structure and/or 2) selective for aliphatic compounds, which could be more
323 susceptible than unsaturated compounds to reactions with sulfur species that are co-emitted with the carbonaceous aerosols in the
324 atmosphere. A study by Schmitt-Kopplin et al. (2010) showed that compounds with higher H/C ratios can react more efficiently with
325 sulfate aerosols to form organosulfates, providing support for the latter of the two possibilities.

326 The ASOM fraction contains 3152 formulas, and 1080 of those formulas were not assigned to peaks in the WSOM spectra.
327 Like WSOM, the whole ASOM sample is dominated by CHO (1367 formulas, 43%) and CHOS (1038 formulas, 33%) compounds,
328 followed by CHON (419 formulas, 13%), CHONS (257 formulas, 8.2%) and CHOP(N,S) (71 formulas, 2.3%, Fig. 3b). The formulas
329 present in ASOM, but not in WSOM (i.e., WIOM_{acn}), shows a similar atomic distribution to the WSOM formulas (Table 1) with CHO
330 (428 formulas, 40%) and CHOS (387 formulas, 36%) formulas being most abundant, but there are more CHONS (136 formulas, 13%)
331 than CHON (101 formulas, 9.4%) formulas.

332 There are 1710 WIOM formulas found in either pyridine or acetonitrile (and not water), and 400 of those formulas are
333 common between the two organic solvents. The formulas common between WIOM_{pyr} and WIOM_{acn} consist mostly of CHOS (225
334 formulas, 56%) and CHO (95 formulas, 24%) compounds followed by CHONS (62 formulas, 16%) and a few CHON (16 formulas,
335 4.0%) and CHOP(N,S) (2 formulas, 0.5%) compounds. The 630 formulas unique to pyridine (i.e., not found in acetonitrile or water)
336 are dominated by sulfur-containing formulas with CHONS (249 formulas, 40%) and CHOS (217 formulas, 34%), and also contains
337 formulas with CHON (85 formulas, 14%), CHO (60 formulas, 10%), and CHOP(N,S) (19, 3.0%). The 680 formulas unique to
338 acetonitrile are dominated by formulas with only CHO (333 formulas, 49%) followed by CHOS (162 formulas, 24%), CHON (85
339 formulas, 13%), CHONS (74 formulas, 11%), and CHOP(N,S) (26 formulas, 3.8%).

340 The relative distributions of formulas for each compound class based on molecular structure (as determined by AI_{mod} values as
341 described above) is shown in Fig. 3c. WSOM and ASOM contain mostly olefinic compounds (60% in each case), and have
342 significant contributions from aliphatic compounds (36% and 37%, respectively). Contrarily, PSOM is dominated by aliphatic

343 formulas (58%) with a smaller amount of olefinic compounds (39%), and this predominance is consistent with the ^1H NMR results. In
344 all three extracts, the contributions from aromatic and condensed aromatic compounds are small ($\leq 3.0\%$ and $\leq 0.6\%$, respectively). The
345 relative amount of each molecular structure type based on aromaticity index for the WIOM_{pyr} and WIOM_{acn} are not significantly
346 different than whole PSOM and ASOM, respectively. Unless the WIOM contains a significant portion of non-ionizable (by ESI)
347 aromatic compounds, the WIOM analyzed in this study may not absorb light as efficiently as the WSOM. This is surprising because
348 we expect the insoluble OM (presumably fossil-derived, Wozniak et al., 2012b) to be aromatic in nature. However, all (13 of 13)
349 condensed aromatic structures and most (52 of 67) aromatic formulas identified in the PSOM are not found in WSOM, suggesting that
350 pyridine may be selective for certain aromatic and condensed aromatic compounds. We speculate that the aromatic character in these
351 samples is low due to a lack of a strong combustion source. Unfortunately we cannot verify true aromatic content using these methods
352 due to the signal from pyridine in the aromatic region of the ^1H NMR spectrum.

353 Each molecular formula type (e.g., CHO, etc.) was plotted on a van Krevelen diagram based on its presence in WSOM,
354 WIOM_{pyr} , or WIOM_{acn} (Fig. 4). Phosphorus-containing formulas were omitted from the figure due to their low number frequency and
355 relatively low spectral intensity. Molecular formulas that appear in WSOM and either PSOM or ASOM were removed from the
356 diagrams for PSOM and ASOM (Fig. 4e-l), so that only the formulas unique to each of the organic solvents (i.e., WIOM_{pyr} and
357 WIOM_{acn}) are shown. The lines on each diagram indicate the type of molecular structure (i.e., aliphatic, olefinic, aromatic, and
358 condensed aromatic) based on AI_{mod} values. Nearly all of the CHO formulas in the WSOM (Fig. 4a) have O/C ratios between 0.2 and
359 0.8 and H/C ratios between 0.95 and 2.00. Formulas within this region of the diagram are typical of many types of natural OM

360 samples, and have previously been attributed to lignin (e.g., Sleighter and Hatcher, 2007) and carboxylic-rich alicyclic molecules
361 (CRAM; e.g., Stubbins et al., 2010), or SOA (e.g., Wozniak et al., 2008). While lignin itself is not likely to be highly abundant in the
362 atmosphere, the compounds found in this region of the van Krevelen diagram could be molecularly similar to lignin (i.e., a branched
363 polymer-containing aromatic rings and various oxygenated functional groups) or derivatives of lignin. Several of the compounds
364 (~33%) within this region meet the operational definition of CRAM (Hertkorn et al., 2006), and could represent CRAM-like structures
365 (i.e., carboxylated and fused alicyclic rings). A previous study by Wozniak et al. (2008) identified formulas in this region as being
366 consistent with those produced through laboratory aging reactions (e.g., pinene ozonolysis), but FTICR-MS cannot provide direct
367 evidence that the compounds identified in this study are secondary in nature.

368 The CHON formulas in WSOM (Fig. 4b) are localized to O/C ratios between 0.2 and 0.8 and H/C ratios between 1.00 and
369 1.75. The formulas in this region above an H/C ratio of 1.50 have previously been attributed to peptides; however, 97% of these
370 formulas have only one nitrogen and cannot truly be peptides because multiple nitrogen atoms would be required. Additionally, all of
371 these CHON formulas have an O/N ratio of at least 3, which suggests that the functionality of the nitrogen may be that of a nitrate
372 group (NO₃). We also recognize the possibility of reduced nitrogen functional groups (e.g., amines and azo compounds) present in
373 molecules containing other ionizable functional groups (e.g., carboxylic acid), but we are unable to differentiate the two possibilities
374 with this method of analysis. WSOM formulas containing both sulfur and nitrogen (CHONS, Fig. 4c) are predominantly aliphatic with
375 relatively high O/C ratios (>0.5). Most (93%) of these formulas have sufficient oxygen atoms (≥ 7) to contain at least one nitrate and
376 one sulfate functional group, indicating that many of these formulas could be nitrooxyorganosulfates which have been identified in

377 previous ambient atmospheric samples (e.g., Surratt et al., 2007; LeClair et al., 2012). CHOS formulas are the second most dominant
378 formula type in the WSOM, and these formulas separate into two distinct regions of the van Krevelen diagram (Fig. 4d). One region
379 contains mostly aliphatic and some olefinic formulas with O/C ratios greater than 0.25 and H/C ratios greater than 1.3. All of the
380 CHOS formulas in this region have O/S of at least 4, suggesting that they are saturated organosulfates or organosulfates with a few
381 double bonds. The other distinct CHOS region contains olefinic and aromatic formulas with a low O/C (<0.4) and H/C ratios less than
382 1.4. Most (70%) of these formulas have O/S ratios of at least 4, indicating that they could be aromatic organosulfates or
383 organosulfates with multiple unsaturations. The CHOS formulas that do not have sufficient O atoms (O/S <4) to be organosulfates
384 must contain a reduced sulfur functional group (e.g., sulfonates and thiols). Organic compounds with reduced forms of sulfur,
385 specifically thia arenes, have been identified in anthropogenic aerosol emissions sources and are known to be toxic (Eastmond et al.,
386 1984; Rogge et al., 1993c). Sulfonates are ubiquitous in detergents and personal care products (Debelius et al., 2008; Lara-Martín et
387 al., 2006), and have been previously identified in atmospheric OM (Altieri et al., 2009). It is important to recognize that atomic ratios
388 do not confirm the presence of organosulfate or organonitrates; however, these structures have been confirmed in other studies of
389 atmospheric OM and it is reasonable to suggest their presence in these samples. A study by LeClair et al. (2012) of atmospheric
390 WSOM using FTICR-MS and collision induced dissociation provides direct evidence for neutral losses of HNO₃ from 63% of
391 detected CHON compounds and 33% of the detected CHONS compounds, and for neutral losses of SO₃ from 85% of detected CHOS
392 compounds and 42% of the detected CHONS compounds. Neutral losses of HNO₃ and SO₃ were interpreted in that study as indicative
393 of organonitrates and organosulfates, respectively. While differences in sample type and instrumentation limit making a direct

394 quantitative comparison here, their results provide good support for the presence of organonitrates, organosulfates and
395 nitrooxyorganosulfates in atmospheric WSOM including these samples. Reduced forms of nitrogen and sulfur are likely also present
396 as evidenced by the formulas with O/S less than 4 and the neutral loss evidence in the LeClair et al. (2012) study which indicates
397 reduced forms must make up a portion of the CHON and CHOS compounds.

398 As stated previously, PSOM is able to dissolve or detect 1030 compounds that WSOM does not (table 1). The characteristics
399 for all formulas in $WIOM_{pyr}$ indicate a lower average O/C (0.29) than WSOM (0.46), a higher average H/C (1.54) than WSOM (1.45),
400 and lower average AI_{mod} (0.11) than WSOM (0.15, table 1). The CHO formulas found in the $WIOM_{pyr}$ plot in a region of higher H/C
401 and lower O/C than those identified in the WSOM (Fig. 4e), implying that they are, on average, more aliphatic and less oxidized than
402 the CHO compounds found in WSOM. The CHON compounds in $WIOM_{pyr}$ show variable H/C and O/C ratios (Fig. 4f). Most of
403 these compounds are either aliphatic or olefinic with O/N ratios of at least 3, indicating they may be organonitrates which have been
404 identified previously in atmospheric WSOM (e.g., LeClair et al., 2012). The CHONS formulas make up a large portion of the
405 $WIOM_{pyr}$ formulas (30%) and they plot into two distinct regions on the van Krevelen diagram (Fig. 4g). The first region contains a
406 majority of the formulas, and they are largely aliphatic molecules with O/C ratios greater than 0.5 and H/C ratios greater than 1.25,
407 and the second region contains mostly aromatic formulas with O/C ratios less than 0.4 and H/C ratios less than 1.25. Many (70%) of
408 the CHONS formulas contain enough oxygen atoms to contain at least one sulfate and one nitrate functional group, and are potentially
409 nitrooxyorganosulfates. Most (81%) of the compounds in the high O/C and high H/C region contain sufficient oxygen, but only a few
410 (6%) of the formulas in the low O/C and low H/C region have sufficient oxygen to be nitrooxyorganosulfates. The CHONS formulas

411 that do not have sufficient oxygen to form nitrooxyorganosulfates must contain at least one sulfur or nitrogen atom present in a
412 reduced form. CHOS formulas make up the largest fraction (43%) of $WIOM_{pyr}$, and the bulk (96%) of those formulas are aliphatic or
413 olefinic. Like WSOM, the CHOS formulas in $WIOM_{pyr}$ plot in two distinct regions of the van Krevelen diagram (Fig. 4h). The first
414 region contains mostly aliphatic formulas spanning the entire O/C range between 0.03 and 1.2 and have H/C ratios greater than 1.2.
415 The second region contains mostly olefinic and aromatic compounds with O/C ratios less than 0.5 and H/C ratios less than 1.5. The
416 average O/C ratio for CHOS compounds in $WIOM_{pyr}$ is nearly the same as that of the WSOM (0.49 ± 0.31 and 0.47 ± 0.23 ,
417 respectively), but the standard deviation for $WIOM_{pyr}$ is greater. This larger deviation indicates that the $WIOM_{pyr}$ CHOS compounds
418 are more diverse than those in WSOM. Most (80%) of the CHOS formulas in the $WIOM_{pyr}$ have O/S ratios of at least 4, suggesting
419 possible organosulfates. Formulas with O/S ratios less than 4 indicate the presence of a more reduced form of sulfur within the
420 structure. The average AI_{mod} of compounds with more reduced forms of sulfur is much greater than that of compounds with O/S
421 sufficient to be organosulfates (0.38 and 0.04, respectively). The major difference between $WIOM_{pyr}$ and WSOM is the increased
422 detection of aliphatic organosulfates and nitrooxyorganosulfates in the $WIOM_{pyr}$, suggesting that pyridine is a more suitable solvent
423 than water for detecting these compounds in aerosol OM using ESI-FTICR-MS.

424 In spite of having many general molecular formula similarities to WSOM, the characteristics for all formulas in $WIOM_{acn}$
425 (1080) indicate a lower average O/C (0.39) than WSOM (0.43), a higher average H/C (1.53) than WSOM (1.45), and an identical
426 average AI_{mod} (0.15). These characteristics suggest that the formulas unique to ASOM are less oxygenated and/or have longer carbon
427 chains. This is clearly shown by the many formulas that plot on the left side (low O/C) of the van Krevelen diagram for $WIOM_{acn}$

428 (Fig. 3i-1). The average number of carbon atoms per formula is slightly larger at 24 carbon atoms for WIOM_{acn} and 22 carbon atoms
429 for WSOM. Both WIOM_{acn} and WSOM are dominated by CHO formulas, but the CHO compounds in the WIOM_{acn} are localized to
430 the upper left corner (low O/C and high H/C) of the diagram (Fig. 4i). Essentially all (99%) of these formulas are aliphatic or olefinic.
431 There are a small number of CHON formulas (101) found in WIOM_{acn}, and most of these formulas have higher H/C (mean value of
432 1.66), and are split between a lower O/C (<0.6) and a high O/C (>0.8). Nearly all (97%) of the CHON formulas have an O/N ratio
433 greater than 3 indicating the possibility of these compounds to contain nitrate as part of their structure. The CHONS formulas are also
434 relatively scarce, and they plot in two separate areas on the van Krevelen diagram (Fig. 4k). Roughly half of the formulas plot below
435 O/C ratio of 0.4 and the other half above 0.4. All of the formulas with O/C > 0.4 have sufficient oxygen atoms (at least 7) to form at
436 least one nitrate and one sulfate group as is found in nitrooxyorganosulfates, but could still contain reduced N and S functional groups.
437 However, only 24% (17 of 70) of the CHONS formulas in the lower O/C region have enough oxygen atoms to form
438 nitrooxyorganosulfates, which indicates the presence of more reduced N- or S-containing functional groups. Like WSOM, CHOS
439 formulas are the second most abundant molecule type in the WIOM_{acn} comprising 387 of the 1080 formulas (36%). These formulas
440 are split into two regions of the diagram (Fig. 4l), where one region is dominated by lower O/C and H/C ratios and contains mostly
441 olefinic and aromatic compounds. The other region is predominantly aliphatic compounds covering the full range of O/C ratios. A
442 majority (76%) of all of the CHOS formulas have sufficient oxygen to form organosulfates (O/S \geq 4). All of the formulas in the higher
443 H/C and aliphatic region contain sufficient oxygen to form organosulfates, but more than half (55%) of the formulas in the low H/C
444 and low O/C region have O/S ratios less than 4. These formulas in the low O/C and low H/C region have multiple unsaturations and

445 have more reduced forms of sulfur in their structure. The major differences between $WIOM_{acn}$ and WSOM are in the high H/C and
446 low O/C region, suggesting that acetonitrile is a more suitable solvent than water for detecting less polar aerosol OM compounds (i.e.,
447 fewer oxygen and heteroatoms and/or larger carbon chains) by ESI-FTICR-MS.

448 The WIOM compounds have chemical characteristics distinct from those in WSOM. Both organic solvents extracted
449 compounds that were more aliphatic in nature than those found in the WSOM, as indicated by the lower O/C ratios and longer carbon
450 chain lengths (for ASOM) of the CHO formulas. Previous work shows that aliphatic components of aerosol WSOM are quantitatively
451 small, implying that they are, as expected, largely water-insoluble (Wozniak et al., 2012b). Radiocarbon signatures indicate that the
452 water-insoluble components are mostly fossil-derived (Wozniak et al., 2012b), but can also come from contemporary plant material
453 (Rogge et al., 1993b). The aliphatic and olefinic hydrocarbon material that is released during fossil fuel combustion can be
454 functionalized through various atmospheric oxidation reactions, thus, increasing its polarity and water-solubility.

455 Nearly 75% of the formulas unique to PSOM include sulfur as part of the molecular formula, indicating a potential selectivity
456 for organosulfates and nitrooxyorganosulfates. This selectivity may be due, in part, to the increased efficiency of aliphatic species
457 (over unsaturated species) to undergo reactions with SO_x , and that pyridine may actually be selective for the aliphatic portion rather
458 than the actual sulfate functional group. Sulfur species, especially SO_x , are well known to play important role in atmospheric aging
459 reactions. Organosulfates are formed in the atmosphere through the acid-catalyzed ring-opening reaction of epoxides in the presence
460 of acidic sulfate seed aerosols (Minerath and Elrod, 2009), and these organosulfates can undergo nighttime NO_3 radical oxidation and
461 photo-oxidation to form nitrooxyorganosulfates (Surratt et al., 2008). These compounds have been identified in ambient atmospheric

462 OM (e.g., Altieri et al., 2012; Mazzoleni et al., 2012; Mitra et al., 2013; Reemtsma et al., 2006; Schmitt-Kopplin et al., 2010; Wozniak
463 et al., 2008). Sulfate is emitted from numerous anthropogenic and biogenic sources and is ubiquitously in the atmosphere. The
464 aerosol OM at this sampling site was in proximity to a known SO_x emission source (coal-fired power plant), providing ample
465 opportunity for atmospheric aging reactions with sulfate to occur and form the observed organosulfate compounds. Organosulfates are
466 very polar in nature which may increase the ability of aerosol OM to act as cloud condensation nuclei and, therefore, have an indirect
467 radiative effect (Hallquist et al., 2009).

468

469 **4 Conclusions**

470 Pyridine and acetonitrile are suitable solvents for analyzing organic aerosols using ESI-FTICR-MS and identify a molecularly-
471 unique portion of aerosol OM when compared to the water-soluble fraction. While the water-soluble fraction is of paramount
472 importance due to the implications those compounds have on environmental processes, such as cloud condensation nuclei formation
473 and mobility in watersheds, analysis of only the WSOM omits a large portion of OM (as much as 90%). The PSOM and ASOM
474 fractions are expected to be more lipophilic and are likely to be more interactive with biological tissues in the environment.
475 Acetonitrile was able to analyze several unique water-insoluble molecular formulas; however, the elemental distributions and formula
476 types (e.g., aliphatic) of these compounds were very similar to those of WSOM. Pyridine elucidated a molecularly unique and,
477 therefore, complementary set of chemical formulas than those in either ASOM or WSOM. PSOM has a stronger preference for
478 extracting or analyzing aliphatic sulfur-containing formulas, which are important aerosol components due to their abundance and

479 hygroscopicity which allows them to act as cloud condensation nuclei and impact climate via the indirect effect. Because WIOM has
480 been found to contain the majority of fossil fuel derived OM, analysis of the WIOM, such as OM found in ASOM and PSOM,
481 provides clues to the molecular composition of the fossil material present in aerosols emitted from anthropogenic sources and
482 therefore, insights into its potential fates and impacts in the environment.

483

484 **Acknowledgments**

485 The work presented in this manuscript was funded by the Frank Batten endowment to P.G. Hatcher. A.S. Willoughby was funded by
486 the U.S. EPA STAR graduate fellowship program, Grand FP 91736001. The EPA has not officially endorsed this publication and the
487 views expressed herein may not reflect the views of the EPA. The authors acknowledge assistance from the ODU COSMIC facility
488 for FTICR-MS and NMR analyses. We thank Dr. Rachel Sleighter for useful discussions about the data.

489

490

491

492 **References**

493

494 Alfarra, M., Paulsen, D., Gysel, M., Garforth, A., Dommen, J., Prevot, A., Worsnop, D., Baltensperger, U., and Coe, H.: A mass
495 spectrometric study of secondary organic aerosols formed from the photooxidation of anthropogenic and biogenic precursors in a
496 reaction chamber, *Atmos. Chem. Phys.*, 6, 5279-5293, 2006.

497

498 Altieri, K., Hastings, M., Peters, A., and Sigman, D.: Molecular characterization of water soluble organic nitrogen in marine rainwater
499 by ultra-high resolution electrospray ionization mass spectrometry, *Atmos. Chem. Phys.*, 12, 3557-3571, 2012.

500

501 Altieri, K., Turpin, B., and Seitzinger, S.: Oligomers, organosulfates, and nitrooxy organosulfates in rainwater identified by ultra-high
502 resolution electrospray ionization FT-ICR mass spectrometry, *Atmos. Chem. Phys.*, 9, 2533-2542, 2009.

503

504 Andreae, M. and Crutzen, P.: Atmospheric aerosols: Biogeochemical sources and role in atmospheric chemistry, *Science*, 276, 1052-
505 1058, 1997.

506

507 Andreae, M. and Gelencsér, A.: Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols, *Atmos. Chem.*
508 *Phys.*, 6, 3131-3148, 2006.

509

510 Bateman, A., Nizkorodov, S., Laskin, J., and Laskin, A.: High-Resolution Electrospray Ionization Mass Spectrometry Analysis of
511 Water-Soluble Organic Aerosols Collected with a Particle into Liquid Sampler, *Anal. Chem.*, 82, 8010-8016, 2010.

512

513 Bateman, A. P., Walser, M. L., Desyaterik, Y., Laskin, J., Laskin, A., and Nizkorodov, S. A.: The effect of solvent on the analysis of
514 secondary organic aerosol using electrospray ionization mass spectrometry, *Environ. Sci. Technol.*, 42, 7341-7346, 2008.

515

516 Bond, T., Doherty, S., Fahey, D., Forster, P., Berntsen, T., DeAngelo, B., Flanner, M., Ghan, S., Kärcher, B., and Koch, D.: Bounding
517 the role of black carbon in the climate system: A scientific assessment, *J. Geophys. Res.-Atmos.*, 118, 1-173, 2013.

518

519 Booth, B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., and Bellouin, N.: Aerosols implicated as a prime driver of twentieth-
520 century North Atlantic climate variability, *Nature*, 484, 228-232, 2012.

521

522 Crutzen, P. J. and Andreae, M. O.: Biomass burning in the tropics: Impact on atmospheric chemistry and biogeochemical cycles,
523 *Science*, 250, 1669-1678, 1990.

524
525 Debelius, B., Forja, J., Del Valls, A., and Lubián, L.: Effect of linear alkylbenzene sulfonate (LAS) and atrazine on marine
526 microalgae, *Mar. Pollut. Bull.*, 57, 559-568, 2008.
527
528 Decesari, S., Mircea, M., Cavalli, F., Fuzzi, S., Moretti, F., Tagliavini, E., and Facchini, M.: Source attribution of water-soluble
529 organic aerosol by nuclear magnetic resonance spectroscopy, *Environ. Sci. Technol.*, 41, 2479-2484, 2007.
530
531 Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G.: A simple and efficient method for the solid-phase extraction of dissolved organic
532 matter (SPE-DOM) from seawater, *Limnol. Oceanogr-Meth.*, 6, 230-235, 2008.
533
534 Eastmond, D. A., Booth, G. M., and Lee, M. L.: Toxicity, accumulation, and elimination of polycyclic aromatic sulfur heterocycles in
535 *Daphnia magna*, *Arch. Environ. Con. Tox.*, 13, 105-111, 1984.
536
537 Hallquist, M., Wenger, J., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N., George, C., and
538 Goldstein, A.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, *Atmos. Chem. Phys.*,
539 9, 5155-5236, 2009.
540
541 Heaton, K., Sleighter, R., Hatcher, P., Hall IV, W., and Johnston, M.: Composition Domains in Monoterpene Secondary Organic
542 Aerosol, *Environ. Sci. Technol.*, 43, 6950-6955, 2009.
543
544 Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Kettrup, A., and Hedges, J. I.: Characterization
545 of a major refractory component of marine dissolved organic matter, *Geochim. Cosmochim. Ac.*, 70, 2990-3010, 2006.
546
547 Hockaday, W., Grannas, A., Kim, S., and Hatcher, P.: Direct molecular evidence for the degradation and mobility of black carbon in
548 soils from ultrahigh-resolution mass spectral analysis of dissolved organic matter from a fire-impacted forest soil, *Org. Geochem.*, 37,
549 501-510, 2006.
550
551 Jacobson, M.: Physically-based treatment of elemental carbon optics: Implications for global direct forcing of aerosols, *Geophys. Res.*
552 *Lett.*, 27, 217-220, 2000.
553
554 Jimenez, J., Canagaratna, M., Donahue, N., Prevot, A., Zhang, Q., Kroll, J., DeCarlo, P., Allan, J., Coe, H., and Ng, N.: Evolution of
555 organic aerosols in the atmosphere, *Science*, 326, 1525-1529, 2009.
556

557 Kamga, A. W., Behar, F., and Hatcher, P. G.: Quantitative Analysis of Long Chain Fatty Acids Present in a Type I Kerogen Using
558 Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Compared with BF₃/MeOH
559 Methylation/GC-FID, *J. Am. Soc. Mass Spectrom.*, 25, 880-890, 2014.
560

561 Kanakidou, M., Seinfeld, J., Pandis, S., Barnes, I., Dentener, F., Facchini, M., Van Dingenen, R., Ervens, B., Nenes, A., Nielsen, C.,
562 Swietlicki, E., Putaud, J., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G., Winterhalter, R., Myhre, C., Tsigaridis, K., Vignati, E.,
563 Stephanou, E., and Wilson, J.: Organic aerosol and global climate modelling: a review, *Atmos. Chem. Phys.*, 5, 1053-1123, 2005.
564

565 Kendrick, E.: A Mass Scale Based on CH₂= 14.0000 for High Resolution Mass Spectrometry of Organic Compounds, *Anal. Chem.*,
566 35, 2146-2154, 1963.
567

568 Kleefeld, S., Hoffer, A., Krivacsy, Z., and Jennings, S.: Importance of organic and black carbon in atmospheric aerosols at Mace
569 Head, on the West Coast of Ireland (53° 19'N, 9° 54'W), *Atmos. Environ.*, 36, 4479-4490, 2002.
570

571 Koch, B. and Dittmar, T.: From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter, *Rapid*
572 *Commun. Mass Sp.*, 20, 926-932, 2006.
573

574 Lara-Martín, P. A., Gómez-Parra, A., and González-Mazo, E.: Development of a method for the simultaneous analysis of anionic and
575 non-ionic surfactants and their carboxylated metabolites in environmental samples by mixed-mode liquid chromatography–mass
576 spectrometry, *J. Chromatogr. A*, 1137, 188-197, 2006.
577

578 Laskin, J., Laskin, A., Roach, P., Slysz, G., Anderson, G., Nizkorodov, S., Bones, D., and Nguyen, L.: High-Resolution Desorption
579 Electrospray Ionization Mass Spectrometry for Chemical Characterization of Organic Aerosols, *Anal. Chem.*, 82, 2048-2058, 2010.
580

581 LeClair, J. P., Collett, J. L., and Mazzoleni, L. R.: Fragmentation analysis of water-soluble atmospheric organic matter using
582 ultrahigh-resolution FT-ICR mass spectrometry, *Environ. Sci. Technol.*, 46, 4312-4322, 2012.
583

584 Lin, P., Rincon, A. G., Kalberer, M., and Yu, J. Z.: Elemental Composition of HULIS in the Pearl River Delta Region, China: Results
585 Inferred from Positive and Negative Electrospray High Resolution Mass Spectrometric Data, *Environ. Sci. Technol.*, 46, 7454-7462,
586 2012.
587

588 Mayol-Bracero, O., Guyon, P., Graham, B., Roberts, G., Andreae, M., Decesari, S., Facchini, M., Fuzzi, S., and Artaxo, P.: Water-
589 soluble organic compounds in biomass burning aerosols over Amazonia 2. Apportionment of the chemical composition and
590 importance of the polyacidic fraction, *J. Geophys. Res.*, 107, 8091, doi: 10.1029/2001JD000522, 2002.
591

592 Mazzoleni, L., Ehrmann, B., Shen, X., Marshall, A., and Collett Jr, J.: Water-soluble atmospheric organic matter in fog: exact masses
593 and chemical formula identification by ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry, *Environ.*
594 *Sci. Technol.*, 44, 3690-3697, 2010.
595

596 Mazzoleni, L. R., Saranjampour, P., Dalbec, M. M., Samburova, V., Hallar, A. G., Zielinska, B., Lowenthal, D. H., and Kohl, S.:
597 Identification of water-soluble organic carbon in non-urban aerosols using ultrahigh-resolution FT-ICR mass spectrometry: organic
598 anions, *Environ. Chem.*, 9, 285-297, 2012.
599

600 McKee, G. A. and Hatcher, P. G.: Alkyl amides in two organic-rich anoxic sediments: A possible new abiotic route for N
601 sequestration, *Geochim. Cosmochim. Ac.*, 74, 6436-6450, 2010.
602

603 Minerath, E. C. and Elrod, M. J.: Assessing the potential for diol and hydroxy sulfate ester formation from the reaction of epoxides in
604 tropospheric aerosols, *Environ. Sci. Technol.*, 43, 1386-1392, 2009.
605

606 Mitra, S., Wozniak, A. S., Miller, R., Hatcher, P. G., Buonassissi, C., and Brown, M.: Multiproxy probing of rainwater dissolved
607 organic matter (DOM) composition in coastal storms as a function of trajectory, *Mar. Chem.*, 154, 67-76, 2013.
608

609 Moretti, F., Tagliavini, E., Decesari, S., Facchini, M., Rinaldi, M., and Fuzzi, S.: NMR determination of total carbonyls and carboxyls:
610 a tool for tracing the evolution of atmospheric oxidized organic aerosols, *Environ. Sci. Technol.*, 42, 4844-4849, 2008.
611

612 Ramanathan, V., Crutzen, P., Kiehl, J., and Rosenfeld, D.: Aerosols, climate, and the hydrological cycle, *Science*, 294, 2119-2124,
613 2001.
614

615 Reemtsma, T., These, A., Venkatachari, P., Xia, X., Hopke, P., Springer, A., and Linscheids, M.: Identification of fulvic acids and
616 sulfated and nitrated analogues in atmospheric aerosol by electrospray ionization Fourier transform ion cyclotron resonance mass
617 spectrometry, *Anal. Chem.*, 78, 8299-8304, 2006.
618

619 Reinhardt, A., Emmenegger, C., Gerrits, B., Panse, C., Dommen, J., Baltensperger, U., Zenobi, R., and Kalberer, M.: Ultrahigh mass
620 resolution and accurate mass measurements as a tool to characterize oligomers in secondary organic aerosols, *Anal. Chem.*, 79, 4074-
621 4082, 2007.

622

623 Robertson, J. and O'Reilly, E.: Electronic and atomic structure of amorphous carbon, *Phys. Rev. B*, 35, 2946-2957, 1987.

624

625 Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R.: Sources of fine organic aerosol. 2. Noncatalyst
626 and catalyst-equipped automobiles and heavy-duty diesel trucks, *Environ. Sci. Technol.*, 27, 636-651, 1993a.

627

628 Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R.: Sources of fine organic aerosol. 4. Particulate
629 abrasion products from leaf surfaces of urban plants, *Environ. Sci. Technol.*, 27, 2700-2711, 1993b.

630

631 Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R.: Sources of fine organic aerosol. 5. Natural gas
632 home appliances, *Environ. Sci. Technol.*, 27, 2736-2744, 1993c.

633

634 Salmon, E., Behar, F., and Hatcher, P. G.: Molecular characterization of Type I kerogen from the Green River Formation using
635 advanced NMR techniques in combination with electrospray ionization/ultrahigh resolution mass spectrometry, *Org. Geochem.*, 42,
636 301-315, 2011.

637

638 Sanders, J. K. and Hunter, B. K.: *Modern NMR spectroscopy: a guide for chemists*, Oxford University Press Inc., New York, NY,
639 1993.

640

641 Schmitt-Kopplin, P., Gelencser, A., Dabek-Zlotorzynska, E., Kiss, G., Hertkorn, N., Harir, M., Hong, Y., and Gebefugi, I.: Analysis of
642 the Unresolved Organic Fraction in Atmospheric Aerosols with Ultrahigh-Resolution Mass Spectrometry and Nuclear Magnetic
643 Resonance Spectroscopy: Organosulfates As Photochemical Smog Constituents, *Anal. Chem.*, 82, 8017-8026, 2010.

644

645 Shakya, K. M., Place, P. F., Griffin, R. J., and Talbot, R. W.: Carbonaceous content and water-soluble organic functionality of
646 atmospheric aerosols at a semi-rural New England location, *J. Geophys. Res.-Atmos.*, 117, D03301, doi:10.1029/2011JD016113,
647 2012.

648

649 Sleighter, R. and Hatcher, P.: The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the
650 molecular characterization of natural organic matter, *J. Mass Spectrom.*, 42, 559-574, 2007.

651

652 Sleighter, R., McKee, G., Liu, Z., and Hatcher, P.: Naturally present fatty acids as internal calibrants for Fourier transform mass
653 spectra of dissolved organic matter, *Limnol. Oceanogr.-Meth.*, 6, 246-253, 2008.

654

655 Sleighter, R. L., Chen, H., Wozniak, A. S., Willoughby, A. S., Caricasole, P., and Hatcher, P. G.: Establishing a measure of
656 reproducibility of ultrahigh-resolution mass spectra for complex mixtures of natural organic matter, *Anal. Chem.*, 84, 9184-9191,
657 2012.

658

659 Stubbins, A., Raitso, D. E., Hinrichsen, H. H., Douglass, J. G., Rose, A. L., Hedley, J., Finlay, K., Dong, H. P., Archer, S. D., and
660 Blukacz, E.: Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as
661 revealed by ultrahigh precision mass spectrometry, *Limnol. Oceanogr.*, 55, 1467-1477, 2010.

662

663 Stubbins, A., Hood, E., Raymond, P. A., Aiken, G. R., Sleighter, R. L., Hernes, P. J., Butman, D., Hatcher, P. G., Striegl, R. G., and
664 Schuster, P.: Anthropogenic aerosols as a source of ancient dissolved organic matter in glaciers, *Nat. Geosci.*, 5, 198-201, 2012.

665

666 Sullivan, A., Weber, R., Clements, A., Turner, J., Bae, M., and Schauer, J.: A method for on-line measurement of water-soluble
667 organic carbon in ambient aerosol particles: Results from an urban site, *Geophys. Res. Lett.*, 31, L13105, doi:10.1029/2004GL019681,
668 2004.

669

670 Surratt, J. D., Gómez-González, Y., Chan, A. W. H., Vermeulen, R., Shahgholi, M., Kleindienst, T. E., Edney, E. O., Offenberg, J. H.,
671 Lewandowski, M., and Jaoui, M.: Organosulfate formation in biogenic secondary organic aerosol, *J. Phys. Chem. A*, 112, 8345-8378,
672 2008.

673

674 Surratt, J. D., Kroll, J. H., Kleindienst, T. E., Edney, E. O., Claeys, M., Sorooshian, A., Ng, N. L., Offenberg, J. H., Lewandowski, M.,
675 and Jaoui, M.: Evidence for organosulfates in secondary organic aerosol, *Environ. Sci. Technol.*, 41, 517-527, 2007.

676

677 Szidat, S., Jenk, T., Gaggeler, H., Synal, H., Fisseha, R., Baltensperger, U., Kalberer, K., Samburova, V., Wacker, L., and Saurer, M.:
678 Source apportionment of aerosols by ^{14}C measurements in different carbonaceous particle fractions, *Radiocarbon*, 46, 475-484, 2004.

679

680 Wozniak, A., Bauer, J., Sleighter, R., Dickhut, R., and Hatcher, P.: Technical Note: Molecular characterization of aerosol-derived
681 water soluble organic carbon using ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass
682 spectrometry, *Atmos. Chem. Phys.*, 8, 5099-5111, 2008.

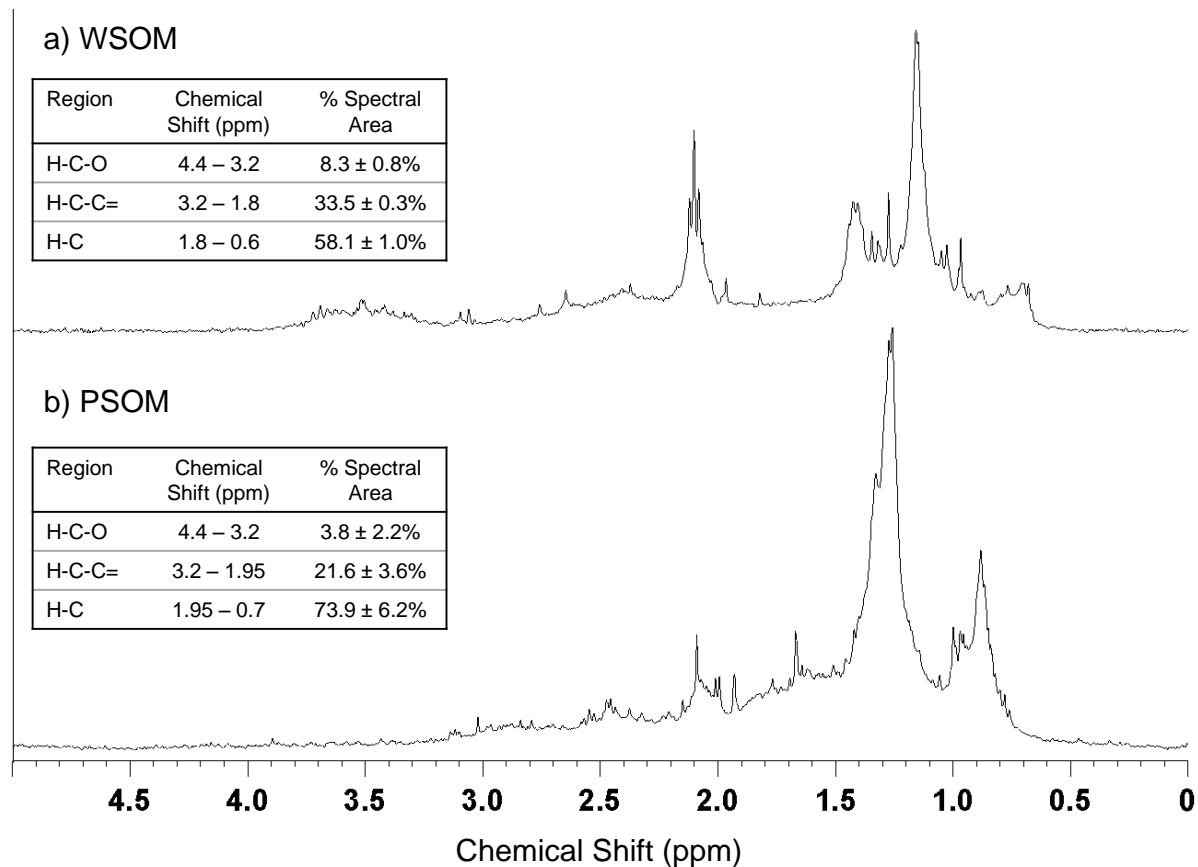
683

684 Wozniak, A. S., Bauer, J. E., and Dickhut, R. M.: Characteristics of water-soluble organic carbon associated with aerosol particles in
685 the eastern United States, *Atmos. Environ.*, 46, 181-188, 2012a.
686
687 Wozniak, A. S., Bauer, J. E., Dickhut, R. M., Xu, L., and McNichol, A. P.: Isotopic characterization of aerosol organic carbon
688 components over the eastern United States, *J. Geophys. Res.*, 117, D13303, doi:10.1029/2011JD017153, 2012b.
689
690 Wu, Z., Jernström, S., Hughey, C. A., Rodgers, R. P., and Marshall, A. G.: Resolution of 10 000 compositionally distinct components
691 in polar coal extracts by negative-ion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Energ.*
692 *Fuel*, 17, 946-953, 2003.
693
694 Zappoli, S., Andracchio, A., Fuzzi, S., Facchini, M., Gelencser, A., Kiss, G., Krivacsy, Z., Molnar, A., Meszaros, E., and Hansson, H.
695 C.: Inorganic, organic and macromolecular components of fine aerosol in different areas of Europe in relation to their water solubility,
696 *Atmos. Environ.*, 33, 2733-2743, 1999.
697

698 **Table 1.** Molecular formula distributions for each solvent extract based on elemental composition with number average
 699 characteristics. The values for PSOM and ASOM do not include formulas common with WSOM, thus represent characteristics for
 700 WIOM.
 701

Solvent Extract	Elemental Composition	# of Formulas	% of Formulas	Number Averages		
				O/C	H/C	AI _{mod}
WSOM	CHO	1563	46%	0.43 ± 0.15	1.43 ± 0.24	0.17 ± 0.14
	CHON	671	20%	0.47 ± 0.23	1.41 ± 0.20	0.16 ± 0.16
	CHONS	214	6.3%	0.71 ± 0.21	1.65 ± 0.20	0.02 ± 0.10
	CHOS	868	26%	0.47 ± 0.23	1.46 ± 0.35	0.15 ± 0.21
	CHOP(N,S)	80	2.4%	0.39 ± 0.09	1.54 ± 0.14	0.09 ± 0.10
	Total	3396	100%	0.46 ± 0.19	1.45 ± 0.27	0.15 ± 0.17
PSOM (WIOM)	CHO	155	15%	0.29 ± 0.20	1.59 ± 0.36	0.13 ± 0.17
	CHON	101	9.8%	0.54 ± 0.32	1.53 ± 0.42	0.19 ± 0.21
	CHONS	311	30%	0.64 ± 0.23	1.52 ± 0.28	0.07 ± 0.18
	CHOS	442	43%	0.49 ± 0.31	1.54 ± 0.38	0.10 ± 0.17
	CHOP(N,S)	21	2.0%	0.49 ± 0.18	1.43 ± 0.50	0.23 ± 0.29
	Total	1030	100%	0.51 ± 0.29	1.54 ± 0.35	0.11 ± 0.18
ASOM (WIOM)	CHO	428	40%	0.25 ± 0.16	1.67 ± 0.24	0.10 ± 0.12
	CHON	101	9.4%	0.48 ± 0.35	1.66 ± 0.42	0.12 ± 0.22
	CHONS	136	13%	0.45 ± 0.25	1.27 ± 0.29	0.25 ± 0.25
	CHOS	387	36%	0.49 ± 0.32	1.42 ± 0.36	0.16 ± 0.20
	CHOP(N,S)	28	2.6%	0.42 ± 0.20	1.49 ± 0.25	0.14 ± 0.19
	Total	1080	100%	0.39 ± 0.28	1.53 ± 0.35	0.15 ± 0.19

702

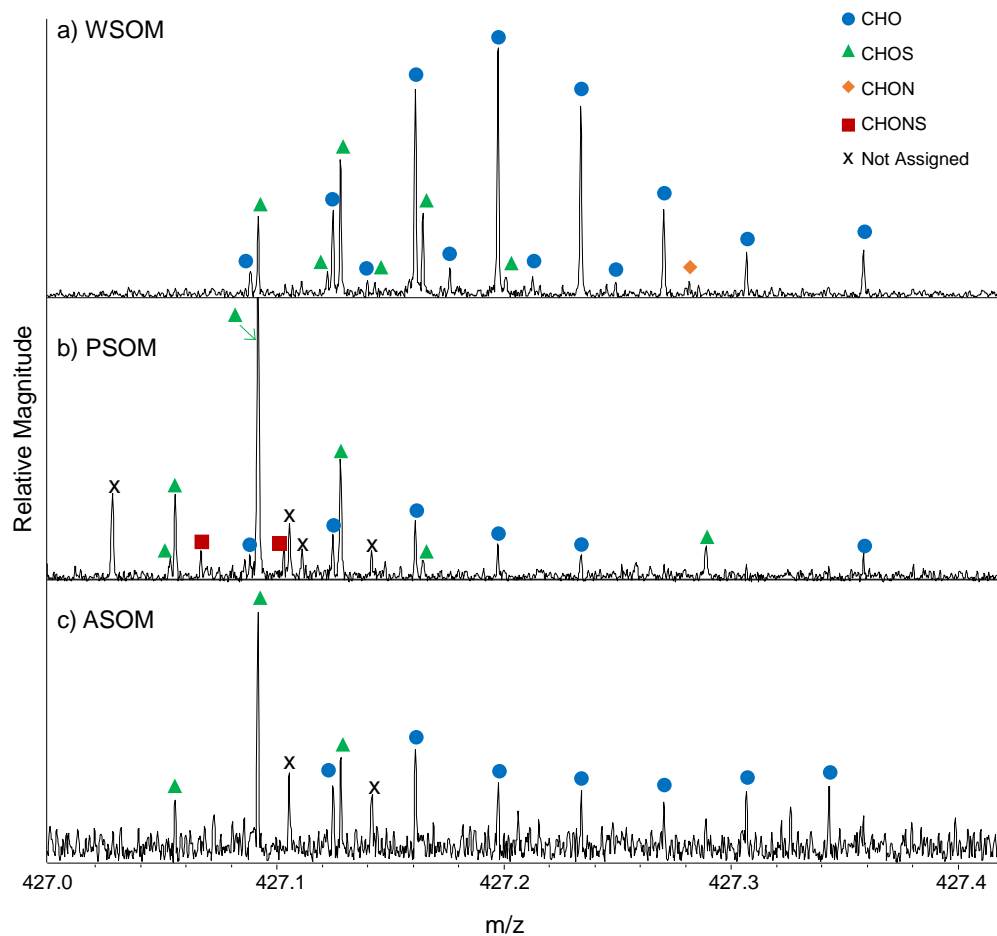


704

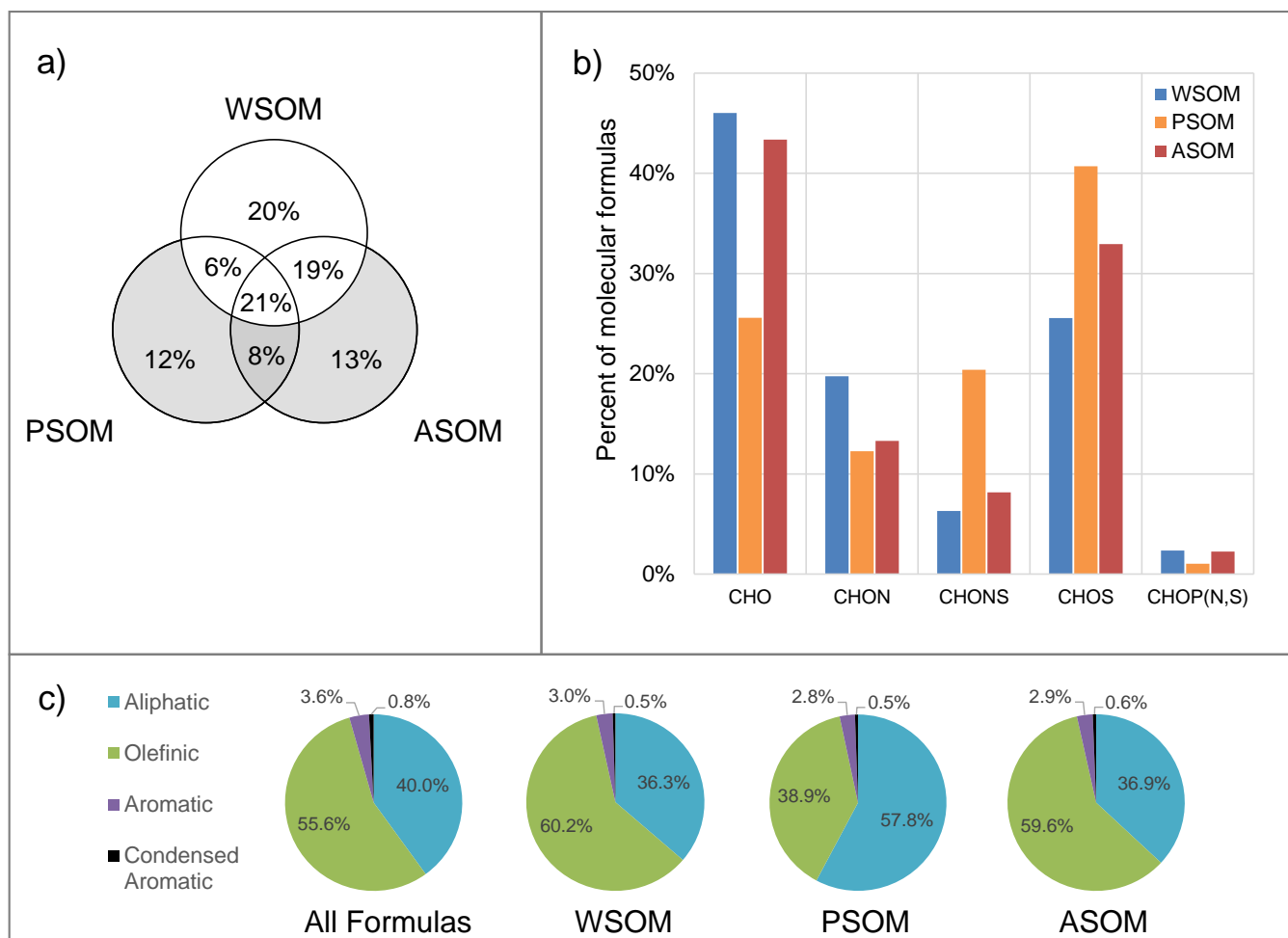
705 **Figure 1.** Expanded ^1H NMR spectra between 0-5 ppm for a) WSOM and b) PSOM for the aerosol particulate sample collected 25-26
 706 June 2013. The table insets give the chemical shifts and average relative intensities (normalized to total intensity between 0 and 4.4
 707 ppm) and standard deviations for the three aerosol samples in the major proton regions, including aliphatic (H-C), unsaturated (H-C-
 708 C=), and oxygenated aliphatics (H-C-O).

709

710

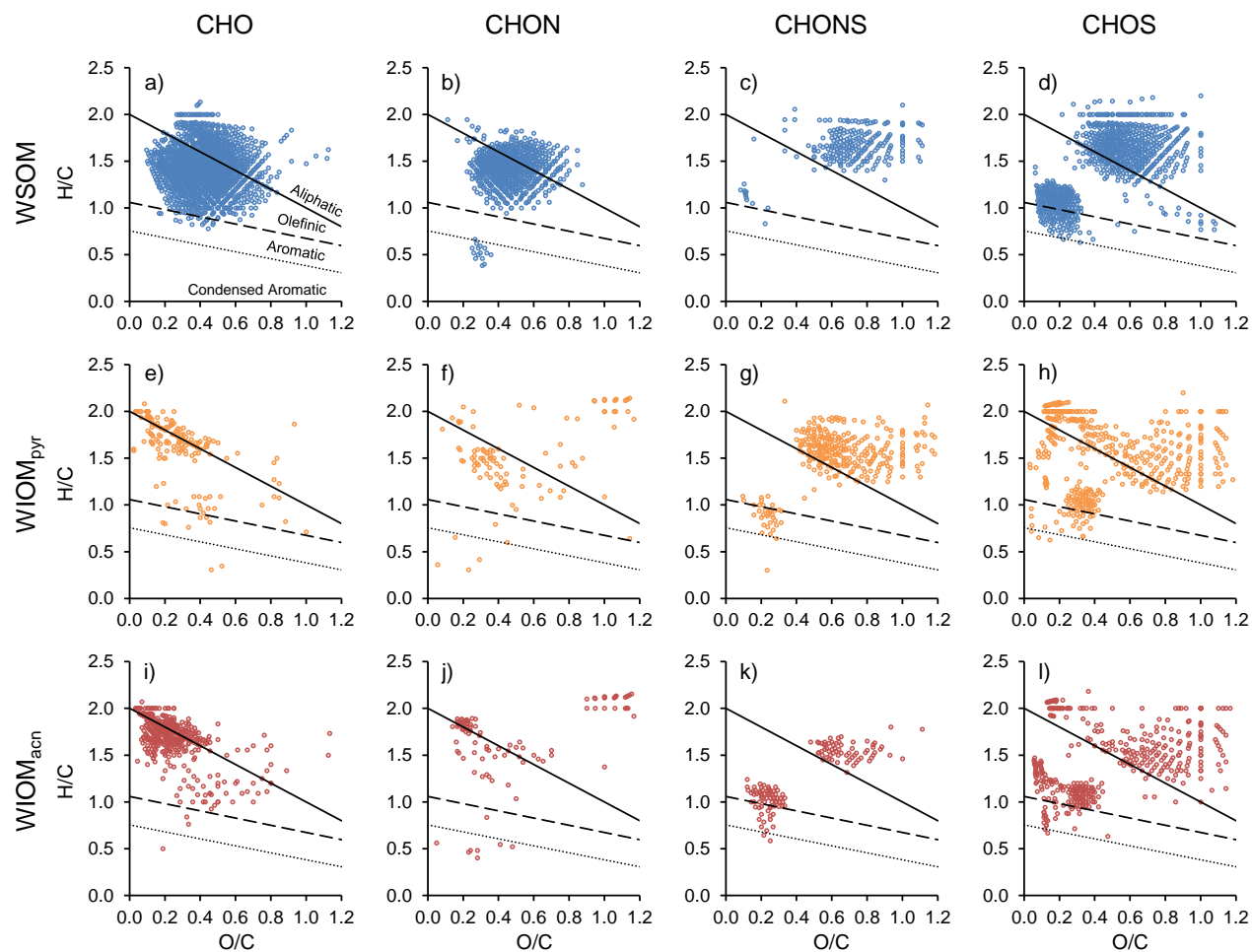


711
 712 **Figure 2.** FTICR mass spectra expanded for m/z at a nominal mass of 427 for a) WSOM, b) PSOM and c) ASOM. Peaks with $S/N \geq 3$
 713 have a colored shape above the peak to indicate the elemental makeup of the assigned molecular formula. Blue circles represent CHO
 714 formulas, green triangles represent CHOS formulas, orange diamonds represent CHON formulas, red squares represent CHONS
 715 formulas, and a black “x” denotes m/z that did not have a formula assigned under the chosen criteria.
 716
 717



718
719

720 **Figure 3.** The a) Venn diagram showing the relative distributions of all molecular formulas (5106 total) present in any of the three
721 solvents, and grey areas represent WIOM. Percentages in areas of overlap are percentages of molecular formulas that appear in both
722 (or all three) of those samples. Percentages in areas of no overlap are molecular formulas that are unique to that individual sample.
723 The b) histogram of the fractional contributions (%) of molecular formulas from various elemental combinations to the total for all
724 formulas (5106 total), WSOM formulas (3396 total), PSOM formulas (2397 total), or ASOM formulas (3152 total). The c) pie charts
725 showing the fractional contributions (%) of molecular structure classes as determined by AI_{mod} calculations.



726
727

728 **Figure 4.** van Krevelen diagrams for molecular formulas assigned to the a-d) WSOM, e-h) WIOM_{pyr}, and i-l) WIOM_{acn} extracts. Any
729 formula present in WSOM has been removed from the WIOM_{pyr} and WIOM_{acn} plots. Each diagram is plotted based on elemental
730 content of each molecular formula (CHO, CHON, CHONS, and CHOS). The labeled regions in a) WSOM CHO formulas correspond
731 to their aromaticity based on AI_{mod} and these regions are consistent for all of the diagrams. Formulas above the solid line are aliphatic,
732 just below the solid line are olefinic, below the dashed line are aromatic, and below the dotted line are condensed aromatic.