We thank both reviewers and the editor for their time and effort reviewing this manuscript. All reviewer comments are reproduced below in ***bold, italicized font***. Our responses are shown in regular font. Changes to the text are indicated as underlined text for insertions or are ~~crossed out~~ for deletions. Line numbers given below are for the revised version with all markups shown. We numbered the reviewer comments for easier cross-referencing.

In the text below, we have responded to each of the concerns carefully and have made changes to the manuscript, where we believe they are warranted.

Anonymous Referee #1:

I went through the authors' responses and was still not very convinced by their rebuttal or still confused by their explanation about the PCA analysis. I noticed that the authors provided more evidence to confirm the unsolved peaks were from bitumen. However, this still cannot well explain the good correlation of gas-phase IVOCs with particle-phase LO-OOA (0.72, factor 5, Table 5) and poor correlation with particle-phase HOA (0.25, factor 5, Table 5). Their argument of "We do not agree, as such correlations could arise from a fraction of IVOCs partitioning to the aerosol phase, perhaps assisted by oxidation" can't convince me because the oxidative processes of particle-phase IVOCs will most likely destroy the correlation between gas-phase IVOCs and particle-phase LO-OOA, no need to mention that not only IVOCs from bitumen but also IVOCs from other sources can partition to the aerosol phase and be oxidized, which could further ruin the correlation between IVOCs vapor and LO-OOA. In contrast, with regard to the poor correlation between IVOCs vapor and HOA, the authors provided confused response "there are many heavy haulers emitting HOA driving all over the ground covered in bitumen that is dug up and carrying fresh bitumen. It would hence be surprising to not see a correlation between bitumen vapors and HOA". In fact, the correlation between bitumen vapors and HOA is poor (r=0.25). Should we be surprised about it or not?

This actually brought in another issue which I mentioned earlier i.e. collocation of factors. Because most pollutants used by the authors are gas-phase pollutants including IVOCs while HOA and LO-OOA are actually particle-phase pollutants. There is argument in the community that it may not be suitable to run PCA using parameters with different physical and chemical characteristics because of factor collocation problems. Hence, the confused PCA results at least for Factor 5 in Table 5 may be due to the collocation of more than one source. If this is true, the PCA results could mislead the readers. How will the PCA results be if the HOA, LO-OOA and/or PM10-1 are removed from the PCA running?

The reviewer is seeking clarification regarding (1) our interpretation of the relationship of IVOCs with particle-phase LO-OOA and HOA, and (2) the validity of using gas- and particle phase variables in the PCA.

Regarding question (1): The reviewer wonders about the relationship between the IVOCs observed and two of the factors observed by the aerosol mass spectrometer, the "hydrocarbon-like organic aerosol (HOA) ... included as a surrogate for fossil fuel combustion by vehicles (Jimenez et al., 2009) ... [and the] LO-OOA factor .... [which] appears to form rapidly after emission of precursors (Lee et al., 2018)." (lines 205-208 of the main manuscript).

The PCA shows that component 5 is strongly associated with IVOCs (r=0.74) and LO-OOA (r=0.72) but poorly with HOA (r=0.25). The HOA is associated mainly with component 2, a combustion source (likely diesel trucks).

A receptor-based analysis can only reveal such associations, but the physical interpretation(s) of these associations will always be less certain and, therefore, are presented in the "discussion" and not the "results" section. We interpret HOA in the conventional manner as a surrogate for fossil fuel combustion (Jimenez et al., 2009). The fact that the association of HOA with component 5 is poor suggests that IVOCs associated with this component are not dominated by diesel emissions and is consistent with bitumen vapors. We suggested (in our rebuttal letter) that it would have been reasonable to expect a correlation between bitumen vapors and HOA given that the trucks move bitumen on the mine sites; this turned out to be not (not strongly, anyhow) supported by the PCA.

We stated that "The correlation of LO-OOA with two of the three IVOC components in the main PCA and with PM_1 in the extended analysis is consistent with the high SOA formation potential of IVOCs", referring to the high SOA formation potential of IVOCs in general. Since the PCA does not give insight into the mechanism(s) of SOA formation or into the presence (or absence) of other organic compounds contributing to SOA formation, we have not commented on this issue in detail in this manuscript but note that the LO-OOA AMS factor will be examined in another manuscript (Lee et al., 2018), which was recently submitted to ACPD.

No changes were made to the manuscript in response to point (1), though we note that more detail is provided on the kinetics of IVOC oxidation in response to question (3) of reviewer #2 below.

Regarding point (2): The reviewer's concern is about using a combination of gas- and particle-phase variables in the analysis.

There have been many publications that have presented analyses with similar mixtures of variables (e.g., (Thurston and Spengler, 1985; Li et al., 1994; Statheropoulos et al., 1998)), so there is certainly precedence. Having said this, we thank the reviewer for the suggestion to perform a sensitivity run and have done so. The results have been added to the Supplemental information section as Table S-10 (reproduced below) and are presented as a 9-component solution, since the dust component associated with PM_{10-1} cannot be generated in the PCA when its main variable is removed.

The pattern in Table S-10 resembles that in Table 5 of the main manuscript, in that the same nine components emerged in both solutions with similar magnitude r values for each of the variables, including the IVOC signature. The only difference is that components 2 and 3 as well as 5 and 6 have traded places (i.e., the relative magnitudes of their eigenvalues, which were of similar magnitude in Table 5, have switched), which is inconsequential.

Table S-10. The pattern without aerosol variables after Varimax rotation with 9 components.

	1	2	3	4	5	6	7	8	9	Commu-nalities
Anthropogenic VOCs										
o-xylene	**0.88**	0.02	0.04	0.09	0.12	0.10	-0.03	0.17	**0.32**	0.94
1,2,3 - TMB	**0.94**	0.07	0.12	0.04	0.11	0.03	-0.02	0.18	-0.02	0.95
1,2,4 - TMB	**0.94**	0.01	0.11	0.10	0.08	0.08	-0.02	0.19	0.13	0.98
decane	**0.93**	-0.01	0.20	0.16	0.01	0.18	0.04	0.04	0.06	0.97
undecane	**0.89**	-0.07	0.26	0.25	-0.03	0.12	0.06	-0.04	-0.03	0.94
Biogenic VOCs										
α-pinene	-0.03	**0.97**	-0.08	-0.12	0.06	0.01	-0.08	0.02	0.00	0.98
ß-pinene	-0.02	**0.97**	-0.08	-0.12	0.05	0.00	-0.08	0.00	0.01	0.98
limonene	0.08	**0.92**	-0.04	-0.08	0.27	0.10	-0.11	0.03	-0.06	0.95
Combustion tracers										
NO_y	0.30	-0.25	**0.81**	0.23	-0.03	0.24	0.09	-0.06	0.03	0.92
rBC	**0.34**	0.04	**0.78**	0.08	0.12	**0.37**	0.11	0.13	-0.04	0.92
CO	**0.42**	0.04	0.16	0.03	0.10	0.05	0.05	**0.87**	-0.01	0.98
CO_2	0.10	**0.46**	0.06	-0.10	**0.84**	-0.02	-0.13	0.06	-0.05	0.96
Aerosol species										
pPAH	0.07	-0.07	**0.94**	-0.13	0.08	-0.06	0.11	0.12	0.03	0.95
Sulfur										
TS	0.26	-0.15	0.03	**0.94**	-0.05	0.03	-0.02	0.01	0.14	1.00
SO_2	0.12	-0.14	0.02	**0.98**	-0.05	-0.03	-0.03	0.02	-0.04	0.99
TRS	**0.59**	-0.07	0.04	0.14	-0.01	0.19	0.02	-0.02	**0.75**	0.97
Other										
IVOCs	**0.35**	0.13	**0.32**	-0.03	0.00	**0.84**	-0.03	0.05	0.15	0.98
NH_3	0.01	-0.23	0.21	-0.05	-0.10	-0.01	**0.94**	0.05	0.01	1.00
CH_4	**0.61**	0.09	**0.36**	-0.05	**0.59**	0.08	0.00	0.18	0.16	0.92
Eigenvalues	5.54	3.16	2.60	2.08	1.20	1.03	0.97	0.94	0.77	
% var.	29.15	16.63	13.68	10.96	6.33	5.40	5.11	4.96	4.03	
% Cum. var.	29.15	45.79	59.46	70.43	76.76	82.16	87.26	92.23	96.25	

This result is consistent with our assumption that the analytically unresolved IVOCs observed in this study were mainly of primary origin (assumed on the basis of the close proximity to sources and a bias of the measurement towards non-oxygenated hydrocarbons) and were likely due to bitumen vapors (on the basis of the similar response of the lab head space analysis to what was observed in ambient air (e.g., Figure 2 and Figure S-1) and that no other source that produces such a response is on hand).

We recognize that it is possible that there were minor contributions to the observed IVOCs by secondary processes. Other sources of IVOCs that contribute to SOA are, for example, diesel emissions (Zhao et al.,

2015). The AB oil sands region is somewhat unusual in this regard, since primary emissions of IVOCs are an unusually large contributor to SOA formation (Liggio et al., 2016; Li et al., 2017). Making a distinction between primary and secondary IVOCs would have required more advanced instrumentation, not on hand during this study. However, the consistency of the results of Table S-10 and Table 5 suggests that the inclusion of HOA, LO-OOA, and PM_{10-1} did not skew the PCA for the data in this paper, which should put the reviewer's concern to rest.

We have modified the paragraph starting on line 214 of the main manuscript as follows:

"To assess which components ~~have the greatest~~ impact ~~on~~ secondary product formation, a second PCA was performed which included variables mainly formed through atmospheric chemical processes and whose concentrations more strongly depend on air mass chemical age than those variables selected initially. In this PCA, odd oxygen ($O_x = O_3 + NO_2$), submicron aerosol $SO_4^{2-}{}_{(p)}$, $NO_3^-{}_{(p)}$, $NH_4^+{}_{(p)}$, a second, more-oxidized OOA factor (MO-OOA), and PM_1 volume were included, increasing the total number of variables to 28 (Table 4). Furthermore, since oxidation of IVOCs leads to formation of SOA (Robinson et al., 2007; Lee et al., 2018), and the photochemical conversion of IVOC to SOA may adversely affect the PCA, a PCA without secondary and aerosol variables is presented in the S.I. (Table S-10)."

We added the following to the supplementary information section (on line 281):

"**PCA without aerosol variables**

A sensitivity test was conducted by which all aerosol species were removed as variables. The results of this sensitivity test are shown in Table S-10 and are presented as a 9-component solution, since the dust component associated with PM_{10-1} (component 7 in Table 5) cannot be generated when its main variable is removed.

The pattern in Table S-10 resembles that in Table 5 of the main manuscript, in that the same nine components emerged in both solutions with similar magnitude r values for each of the variables, including the IVOC signature. The only difference is that components 2 and 3 as well as 5 and 6 have traded places (i.e., the relative magnitudes of their eigenvalues, which were similar in Table 5, have switched), which is inconsequential. Furthermore, the correlation coefficients in Table S-10 are of similar magnitude (i.e., within ± 0.1) as those in Table 5, which suggests that IVOC to SOA conversion does not adversely affect the PCA, likely because of the proximity of the receptor site to sources."

As such, a PCA by its very nature yields patterns, i.e., recurring structures, which are fusions of many single sources. Examples of factors/components comprised of numerous and somewhat dissimilar (point) sources are common in the literature; recurring examples include "biogenic" factors (originating from many different trees and tree species), "vehicle emission" or "traffic" factors (from a fleet of motor vehicles exhibiting a range of individual emission compositions), or "industrial sources" (from a range of activities) – see, for example (Bruno et al., 2001; Guo et al., 2004; Guo et al., 2007; Chavent et al., 2009; Lan et al., 2014; Cesari et al., 2016; Thurston and Spengler, 1985; Samara et al., 1994; Buhr et al., 1995). Emissions from these source types include a combination of trace gases (NO_x, SO_2, CO, CO_2, certain VOCs etc.) and particles (e.g., soot from a diesel vehicle, or OA from trees).

Anonymous Referee #2
Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

1. GC-ITMS: what type of column was used for measuring the IVOCs?

We added this information to the S.I. on line 76: "Operation, calibration and performance of this instrument have been described elsewhere (Tokarek et al., 2017; Liggio et al., 2016). Briefly, the GC was operated with 30 m (length) × 0.25 mm (inner diameter) × 0.25 µm (film thickness) DB-5MS analytical column with helium carrier gas. The GC-ITMS sampled from ...."

Would the IVOCs be limited to hydrocarbon species or would oxygenated species be detected too?

This is discussed on line 167 of the main manuscript:

"The use of a chromatographic column in this work biases the IVOC signal towards hydrocarbon-IVOCs, since oxygenated compounds (i.e., alcohols and acids) will not elute from the analytical column. Furthermore, the recovery of VOCs from the pre-concentration unit, while reproducible and likely complete for n-alkanes which bracket the bulk of IVOC emitted and whose calibration curves were linear, is not known for late-eluting compounds, but is assumed to be sufficiently reproducible to yield a semi-quantitative signal. "

No changes were made in response to this comment.

Depending on that answer, are the IVOC measurements discussed here primary and/or secondary species?

We believe that the IVOC observed the GC-ITMS are mainly primary in origin, as stated on line 160 of the main manuscript:

"In the interpretation of the integrated IVOC signal, it is assumed that it is of primary origin, i.e., emitted directly from point sources in the vicinity of the measurement site."

Please see also our response to reviewer #1, where we corroborate the notion that IVOCs are mainly primary.

No changes were made in response to this comment.

2. Line 160-167: This discussion would likely benefit from broadly citing and describing the IVOC work that has been done with fossil fuel sources other than aircraft. For example, Gentner et al. (PNAS, 2012) and Zhao et al. (ES&T, 2014, 2015, 2016) that look at IVOC emissions from evaporated fuel and mobile sources.

We thank the reviewer for alerting us to these papers. We agree that diesel engine IVOC emissions are more relevant than those from aircraft and have modified the text (on line 165) as follows:

"The IVOCs observed in this work likely encompass a portion of the total that is emitted. For example, IVOCs generated by combustion processes, such as ~~aircraft~~ diesel engine exhaust, are comprised of aliphatic alkanes, including cyclic and branched alkanes, and aromatics (Gentner et al., 2012; Zhao et al., 2015) ~~and oxygenated compounds (Cross et al., 2013)~~."

3. The manuscript at one point (line 173-175) says that species are conserved between emission and measurement but at other points allude to the role of secondary formation and aging (e.g., PCA with secondary pollutants). A short paragraph that discusses the transport times between emission and measurement (based on distances in Table 1 and average wind speeds), average OH concentrations, and kOH of IVOCs would be useful to understand the potential influence of photochemical aging on the measurements.

An excellent suggestion.

We have added the following text starting on line 116:

"A potentially important consideration is the photochemical aging of emissions between the points of emission and observation. During daytime, the average surface wind speed was 7.5 km hr^{-1} (2.1 m s^{-1}). The average transit times were 0.5 hr to the edge of the closest mining operation, 1.6 hr to the 12.2 km distant Mildred Lake Plant site, and 3.2 hr to the Muskeg River Mine site located 23.7 km upwind."

We do not have direct measurements for hydroxyl radical concentrations during this campaign; Liggio et al. (2016) reported an estimated mid-day [OH] of 7×10^6 molecules cm^{-3}.

To corroborate the above OH concentration and transport times, we calculated photochemical age using the method outlined by Borbon et al. (2013), but substituting n-decane for benzene since the latter was not quantified.

The following text was added the supplemental on line 516:

"**Estimate of photochemical age**

Photochemical age was calculated using the method outlined by Borbon et al. (2013), but substituting n-decane for benzene since the latter was not quantified. The photochemical age of an air mass, Δt was calculated from the observed concentrations of 124-trimethylbenzene (124TMB) and n-decane using:

$$\Delta t = \frac{1}{[OH]\times(k_{124TMB}-k_{decane})} \times \left[\ln\left(\frac{[124TMB]}{[decane]}\right)_{t=0} - \ln\left(\frac{[124TMB]}{[decane]}\right)\right] \quad \text{(S-1)}$$

where k_{124TMB} = 3.25×10^{-11} cm^3 molecule^{-1} s^{-1} and k_{decane} = 1.10×10^{-11} cm^3 molecule^{-1} s^{-1} are rate coefficients for reaction of OH with 124-TMB and n-decane (at 298 K), respectively, whose values were taken from Seinfeld and Pandis (2006). The ratio of [124TMB] to [decane] at the point of emission (time t = 0) was estimated from a plot of [124TMB] to [n-decane] (Figure S-13, left-hand side) and a straight-line fit to the nocturnal data (assumed to be unaffected by oxidation and shown in blue color). The slope of this line was 1.15±0.07 (r^2 = 0.84). Daytime data (color-coded by solar zenith angle, SZA) exhibit lower ratios of [124TMB]/[decane] as a result of the faster oxidation of 124TMB by OH.

Shown in Figure S-13 on the right-hand side is a plot of the photochemical age, calculated using equation (S-1) and an assumed [OH] of 7×10^6 molecules cm^{-3} taken from Liggio et al. (2016), as a function of SZA (filtered for peak OH of 11:00 and 16:00 local time). The error bars indicate ages calculated using emission factors of 1.08 and 1.22, respectively. The average (±1 standard deviation) photochemical age is 1.0±0.4 hr. This photochemical age applies mainly to component 1; we assume that the photochemical ages of sources associated with other components were similar.

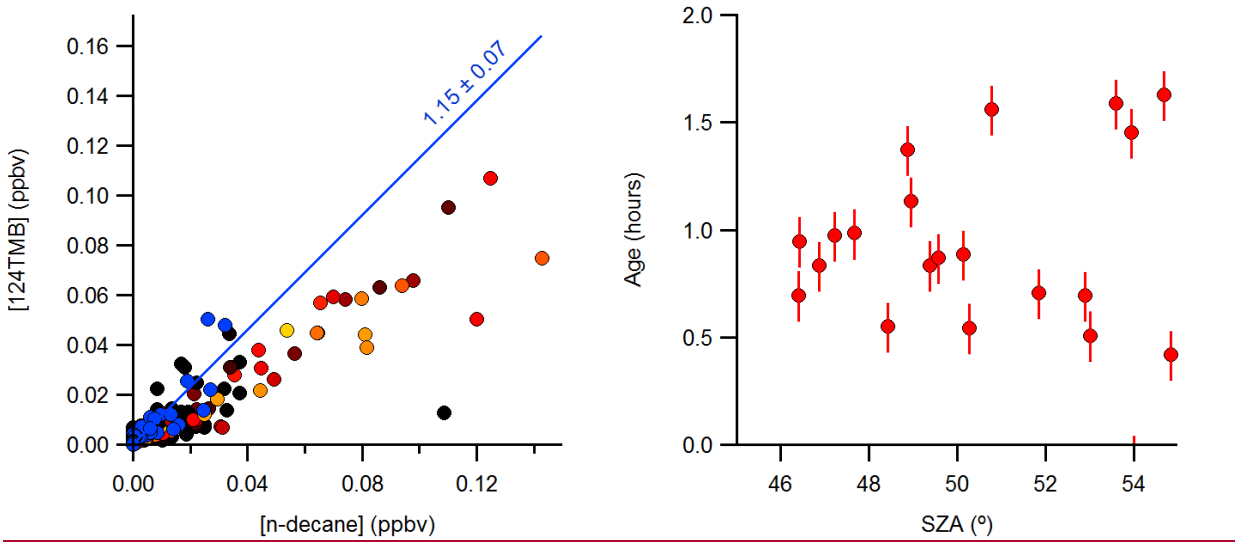


Figure S-13. (A) Plot of 124TMB mixing ratios against mixing ratios of n-decane, color-coded by solar zenith angle. The blue data points were collected at night. **(B)** Photochemical age calculating using equation S-1 plotted as a function of solar zenith angle.

In their analysis of IVOC photochemical aging, Zhao et al. (2014) estimated an average k_{OH} for diesel-exhaust IVOCs of 1.8×10^{-11} cm^3 molecule^{-1} s^{-1} (though their estimated rate coefficients varied and increased slightly with volatility bin between about 1 and 3 ×10^{-11} cm^3 molecule^{-1} s^{-1}). From this, we calculate a pseudo first-order lifetime of 130 min (2.17 hr) with respect to IVOC oxidation by OH during daytime. Using a photochemical age of 1.0±0.4 hr, we calculate that between 25% and 50% of the emitted IVOC is (potentially) oxidized during daytime. Photochemical aging will affect data collected during the daytime hours (from ~11 am to ~4 pm) or ~25% of the data (56 out of 218 data points) used in the PCA and likely resulted in partial conversion of IVOCs to SOA."

As stated in the main manuscript (line 168), oxygenated compounds (i.e., IVOC oxidation products) are unlikely to elute from the analytical column. Hence, the net effect is that the observed IVOC abundance

is attenuated during daytime, which is not expected to substantially impact the PCA as some, if not most, of the starting material will remain. In contrast, IVOC emissions at night or during the morning and late afternoon would be by-and-large unaffected by OH oxidation.

How and on what time scale this photo-oxidation chemistry will translate into SOA formation is a complex issue (details of which are beyond the scope of this manuscript), suffice to say that Figure 1 of Liggio et al. (2016) shows (some) SOA production on a 1 hr time scale, and that some of the SOA generated from IVOCs is included in the LO-OOA observed by AMS (Lee et al., 2018).

We have added the following paragraph to the main manuscript on line 632:

"Component 5 correlates strongly with LO-OOA (r = 0.72), which is likely generated in part by photochemical aging of IVOCs. A back-of-the-envelope calculation using a k_{OH} of 1.8×10^{-11} cm^3 $molecule^{-1}$ s^{-1} based on that used for diesel exhaust IVOCs (Zhao et al., 2014) and an estimated mid-day OH concentration of 7×10^6 molecules cm^{-3} (Liggio et al., 2016) gives a first-order lifetime of 130 min with respect to IVOC oxidation by OH during daytime. The photochemical age, estimated using relative concentrations of 124-TMB and n-decane and the method described by Borbon et al. (2013), during daytime was 1.0±0.4 hr; assuming similar photochemical ages, we estimate that between 25% and 50% of the emitted IVOC is (potentially) oxidized during daytime (see S. I.). This oxidation will contribute SOA growth (Kroll et al., 2011). Hence, we expect some formation and growth of organic aerosol associated with component 5."

We added the following paragraph to line 759 of the main manuscript:

"The relatively short distance to sources and young photochemical age suggests that IVOCs would experience a relatively small number of oxidation steps. Consistent with this interpretation, a correlation with the more-oxidized MO-OOA is not observed in component 5 (r = 0.10; Table 7). However, component 6, which is (poorly) anticorrelated with IVOCs (r = -0.23), is strongly correlated with MO-OOA (r = 0.92), consistent with the notion that this component is more photochemically processed and that IVOCs contribute to this SOA AMS factor."

4. Line 223: As the authors are aware, there are nuanced approaches to dealing with values below the detection limit based on the distribution of the data. See for example, https://www.tandfonline.com/doi/pdf/10.1080/08940630.1989.10466534

We thank the reviewer for bringing this reference (Cohen and Ryan, 1989) to our attention. Cohen and Ryan (1989) show numerical simulations that suggest that the log-fill-in method *may* provide more accurate results in some cases (depending on the geometric standard deviation and level of truncation of a variable) and suggest that a more tailored approach may benefit some analyses.

However, it is common practice in the literature (e.g., (Harrison et al., 1996; Polissar et al., 1998; Guo et al., 2004; Zhao et al., 2004; Mintz and McWhinney, 2008)) to use the "half of the LOD method" to replace values below an instrument's LOD with half of its value, which is what we followed.

We have expanded the references cited on lines 231 but have not made changes in direct response to this comment.

"When concentrations were below their respective limit of detection (LOD; values are given in Table 3), half the reported LOD was used to minimize bias (Harrison et al., 1996; Polissar et al., 1998; Zhao et al., 2004; Guo et al., 2004)."

5. Section 2.3.2: Does log-transforming the data instead of normalizing help with the PCA?

The reviewer asks an interesting and insightful question: whether a logarithmic transformation of the input data (in addition to, or as an alternative to, standardization to zero mean and unit variance to alleviate the effects of different measurement units) would improve the PCA.

We have chosen to follow common data pre-treatment practice in the field, which is to standardize the input to zero mean and unit variance (e.g., (Thurston and Spengler, 1985; Li et al., 1994; Statheropoulos et al., 1998; Bruno et al., 2001; Guo et al., 2004; Chan and Mozurkewich, 2007; Mintz and McWhinney, 2008; Jolliffe and Cadima, 2016)).

A logarithmic transformation is performed to reduce the influence of extreme values (or outliers) and may give more weight to lower concentration data points (Baxter, 1995) or when its based on a sound reason, such as taking a logarithm of an equilibrium constant, for which there are reasons nested in thermodynamics (Malinowski, 2002). The term "outlier" implies to us that a measurement value that was somehow made in error, which is not the case for this data set. Further, removing or dampening large concentration values is, in our opinion, not a good approach to pollution research.

No changes were made to the manuscript in response to the reviewer's question.

6. Line 285-286: The source signatures for a single source could also change with environmental conditions and hence result in varying association of the IVOCs with VOCs.

We thank the reviewer for this suggestion and have added the following text on line 291 of the manuscript:

"The IVOC magnitude also varied greatly and often increased and decreased in tandem with the other VOCs (e.g., on Aug 24, 16:30 UTC) but also increased independently from the other VOC abundances (e.g., on Aug 30, 01:20 UTC, and on the night of Aug 22). This behaviour suggests the presence of multiple sources with distinct signatures that are being sampled to a varying extent at different times or, perhaps, a single source whose emission profile varies. This, coupled with the intermittency of the highly elevated signals, presents an analysis problem frequently encountered in environmental analysis that is usually investigated through a factor or principal component analysis (Thurston et al., 2011; Guo et al., 2004)."

7. Line 328: Stating the coefficient of variation (CoV) would help compare the variability between pollutants better since the CoV is a normalized metric.

We modified the sentence in question as follows:

"Mixing ratios of SO_2 exhibited the most variability of all pollutants, as judged from the relative standard deviation of each of the measurements (Table 3)."

We made an analogous change to a sentence on line 360 of the main manuscript:

"Ammonia was not as variable as some of the other pollutants (e.g., the anthropogenic VOCs, sulfur species) as judged from its relative standard deviation (Table 3), which suggests a geographically more disperse source or sources similar to CO or CH_4, which have a "background"."

In Tables 3 and 4, we removed the "standard deviation" column and added a "relative standard deviation" column.

8. Line 578-582: Wouldn't photochemical oxidation make the IVOCs less prone to aerosol formation since they would be more likely to fragment?

The sentence in question reads: "whereas the IVOCs .... may have been ... oxidized to a greater extent and hence more prone to rapid aerosol formation."

Oxidation of organics is generally viewed as a source of atmospheric organic aerosol through secondary organic formation. This occurs because of increased functionality on the carbon skeleton, which reduces volatility (Kroll et al., 2011). The reviewer is correct that oxidation of hydrocarbons can also lead to their fragmentation. As pointed out by Kroll et al. (2011), small (four carbons or fewer) organic species are unlikely to contribute to aerosol formation, even though they might still form organic aerosol through oligomerization reactions. Oxidation reactions ("aging") of atmospheric organic aerosol are ultimately (after a few generations of oxidation) dominated by fragmentation reactions which then act as organic aerosol sinks, because oxidized organics may fragment and volatilize upon further oxidation (Kroll et al., 2011).

The IVOCs observed in this work are in the volatility range of C_{11} – C_{17} hydrocarbons; even if these were to fragment during the initial oxidation step, at least one of the fragments would contain more than four carbons (and likely also oxygen) and hence can contribute to aerosol formation.

No changes were made in response to the reviewer's question.

Also, would the GC-ITMS be able to detect them if the IVOCs were oxidized by the microbes? It's possible I have not understood the reasoning for differences in the IVOCs in components 1, 2, and 5 correctly. If that is the case, please consider rephrasing the last sections of that paragraph.

We do not believe that the GC-ITMS can differentiate between IVOCs oxidized photochemically or biochemically, though both microbial and photochemical oxidation are plausible. We have rephrased the paragraph as follows:

"One reason for the difference could be that the bitumen that is transported by the mining fleet is relatively freshly exposed, whereas the IVOCs released ~~by bitumen in~~from tailings ponds ~~has been processed by microbes and that released by~~ or from mine faces (component 5) may have been ~~photochemically~~ oxidized to a greater extent and hence more prone to rapid aerosol formation."

Literature cited above

Baxter, M. J.: Standardization and transfomration in principal component analysis, with applications to archaeometry, Appl. Stat.-J. R. Stat. Soc., 44, 513-527, 10.2307/2986142, 1995.

Borbon, A., Gilman, J. B., Kuster, W. C., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros, V., Lopez, M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D. D., and de Gouw, J. A.: Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude megacities: Observations versus emission inventories in Los Angeles and Paris, J. Geophys. Res.-Atmos., 118, 2041-2057, doi:10.1002/jgrd.50059, 2013.

Bruno, P., Caselli, M., de Gennaro, G., and Traini, A.: Source apportionment of gaseous atmospheric pollutants by means of an absolute principal component scores (APCS) receptor model, Fresenius J Anal Chem, 371, 1119-1123, 10.1007/s002160101084, 2001.

Buhr, M., Parrish, D., Elliot, J., Holloway, J., Carpenter, J., Goldan, P., Kuster, W., Trainer, M., Montzka, S., McKeen, S., and Fehsenfeld, F.: Evaluation of ozone precursor source types using principal component analysis of ambient air measurements in rural Alabama, J. Geophys. Res.-Atmos., 100, 22853-22860, doi:10.1029/95JD01837, 1995.

Cesari, D., Amato, F., Pandolfi, M., Alastuey, A., Querol, X., and Contini, D.: An inter-comparison of PM10 source apportionment using PCA and PMF receptor models in three European sites, Environm. Sci. Poll. Res., 23, 15133-15148, 10.1007/s11356-016-6599-z, 2016.

Chan, T. W., and Mozurkewich, M.: Application of absolute principal component analysis to size distribution data: identification of particle origins, Atmos. Chem. Phys., 7, 887-897, 10.5194/acp-7-887-2007, 2007.

Chavent, M., Guégan, H., Kuentz, V., Patouille, B., and Saracco, J.: PCA- and PMF-based methodology for air pollution sources identification and apportionment, Environmetrics, 20, 928-942, doi:10.1002/env.963, 2009.

Cohen, M. A., and Ryan, P. B.: Observations Less than the Analytical Limit of Detection: A New Approach, J. Air Waste Manag. Assoc., 39, 328-329, 10.1080/08940630.1989.10466534, 1989.

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.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions, Proceedings of the National Academy of Sciences, 109, 18318-18323, 10.1073/pnas.1212272109, 2012.

Guo, H., Wang, T., and Louie, P. K. K.: Source apportionment of ambient non-methane hydrocarbons in Hong Kong: Application of a principal component analysis/absolute principal component scores (PCA/APCS) receptor model, Environ. Pollut., 129, 489-498, 10.1016/j.envpol.2003.11.006, 2004.

Guo, H., So, K. L., Simpson, I. J., Barletta, B., Meinardi, S., and Blake, D. R.: C1–C8 volatile organic compounds in the atmosphere of Hong Kong: Overview of atmospheric processing and source apportionment, Atmos. Environm., 41, 1456-1472, 10.1016/j.atmosenv.2006.10.011, 2007.

Harrison, R. M., Smith, D. J. T., and Luhana, L.: Source Apportionment of Atmospheric Polycyclic Aromatic Hydrocarbons Collected from an Urban Location in Birmingham, U.K, Environm. Sci. Technol., 30, 825-832, 10.1021/es950252d, 1996.

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.: Evolution of Organic Aerosols in the Atmosphere, Science, 326, 1525-1529, 10.1126/science.1180353, 2009.

Jolliffe, I. T., and Cadima, J.: Principal component analysis: a review and recent developments, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374, 10.1098/rsta.2015.0202, 2016.

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.: Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol, Nature Chemistry, 3, 133, 10.1038/nchem.948, 2011.

Lan, C.-H., Huang, Y.-L., Ho, S.-H., and Peng, C.-Y.: Volatile organic compound identification and characterization by PCA and mapping at a high-technology science park, Environ. Pollut., 193, 156-164, 10.1016/j.envpol.2014.06.014, 2014.

Lee, A. K. Y., Adam, M. G., Liggio, J., Li, S.-M., Li, K., Willis, M. D., Abbatt, J. P. D., Tokarek, T. W., Odame-Ankrah, C. A., Huo, J. A., Osthoff, H. D., Strawbridge, K. B., and Brook, J. R.: A large contribution of anthropogenic organo-nitrates to secondary organic aerosol in Alberta oil sands, Atmos. Chem. Phys. Discuss., article number acp-2018-1177 2018.

Li, S.-M., Anlauf, K. G., Wiebe, H. A., and Bottenheim, J. W.: Estimating primary and secondary production of HCHO in eastern North America based on gas phase measurements and principal component analysis, Geophys. Res. Lett., 21, 669-672, doi:10.1029/94GL00643, 1994.

Li, S.-M., Leithead, A., Moussa, S. G., Liggio, J., Moran, M. D., Wang, D., Hayden, K., Darlington, A., Gordon, M., Staebler, R., Makar, P. A., Stroud, C. A., McLaren, R., Liu, P. S. K., O'Brien, J., Mittermeier, R. L., Zhang, J., Marson, G., Cober, S. G., Wolde, M., and Wentzell, J. J. B.: Differences between measured and reported volatile organic compound emissions from oil sands facilities in Alberta, Canada, Proceedings of the National Academy of Sciences, 114, E3756-E3765, 10.1073/pnas.1617862114, 2017.

Liggio, J., Li, S.-M., Hayden, K., Taha, Y. M., Stroud, C., Darlington, A., Drollette, B. D., Gordon, M., Lee, P., Liu, P., Leithead, A., Moussa, S. G., Wang, D., O'Brien, J., Mittermeier, R. L., Brook, J., Lu, G., Staebler, R., Han, Y., Tokarek, T. W., Osthoff, H. D., Makar, P. A., Zhang, J., Plata, D., and Gentner, D. R.: Oil Sands Operations as a Large Source of Secondary Organic Aerosols, Nature, 534, 91-94, 10.1038/nature17646, 2016.

Malinowski, E. R.: Factor analysis in chemistry, 3rd ed., Wiley, New York, 414 pp., 2002.

Mintz, R., and McWhinney, R. D.: Characterization of volatile organic compound emission sources in Fort Saskatchewan, Alberta using principal component analysis, J. Atmos. Chem., 60, 83-101, 10.1007/s10874-008-9110-5, 2008.

Polissar, A. V., Hopke, P. K., Paatero, P., Malm, W. C., and Sisler, J. F.: Atmospheric aerosol over Alaska: 2. Elemental composition and sources, J. Geophys. Res.-Atmos., 103, 19045-19057, doi:10.1029/98JD01212, 1998.

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.: Rethinking organic aerosols: Semivolatile emissions and photochemical aging, Science, 315, 1259-1262, 2007.

Samara, C., Kouimtzis, T., and Katsoulos, G. A.: Characterization of airborne particulate matter in Thessaloniki, Greece, Toxicological & Environmental Chemistry, 41, 221-232, 10.1080/02772249409357977, 1994.

Seinfeld, J. H., and Pandis, S. N.: Atmospheric chemistry and physics: from air pollution to climate change, 2nd ed., Wiley, Hoboken, N.J., 2006.

Statheropoulos, M., Vassiliadis, N., and Pappa, A.: Principal component and canonical correlation analysis for examining air pollution and meteorological data, Atmos. Environm., 32, 1087-1095, 10.1016/S1352-2310(97)00377-4, 1998.

Thurston, G. D., and Spengler, J. D.: A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston, Atmos.Environm. (1967), 19, 9-25, 10.1016/0004-6981(85)90132-5, 1985.

Thurston, G. D., Ito, K., and Lall, R.: A source apportionment of U.S. fine particulate matter air pollution, Atmos. Environ., 45, 3924-3936, 10.1016/j.atmosenv.2011.04.070, 2011.

Tokarek, T. W., Brownsey, D. K., Jordan, N., Garner, N. M., Ye, C. Z., Assad, F. V., Peace, A., Schiller, C. L., Mason, R. H., Vingarzan, R., and Osthoff, H. D.: Biogenic Emissions and Nocturnal Ozone Depletion Events at the Amphitrite Point Observatory on Vancouver Island, Atmosphere-Ocean, 1-12, 10.1080/07055900.2017.1306687, 2017.

Zhao, W., Hopke, P. K., and Karl, T.: Source Identification of Volatile Organic Compounds in Houston, Texas, Environm. Sci. Technol., 38, 1338-1347, 10.1021/es034999c, 2004.

Zhao, Y. L., Hennigan, C. J., May, A. A., Tkacik, D. S., de Gouw, J. A., Gilman, J. B., Kuster, W. C., Borbon, A., and Robinson, A. L.: Intermediate-Volatility Organic Compounds: A Large Source of Secondary Organic Aerosol, Environm. Sci. Technol., 48, 13743-13750, 10.1021/es5035188, 2014.

Zhao, Y. L., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A., and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol Production, Environm. Sci. Technol., 49, 11516-11526, 10.1021/acs.est.3b02841, 2015.

Principal component analysis of summertime ground site measurements in the Athabasca oil sands with a focus on analytically unresolved intermediate volatility organic compounds

Travis W. Tokarek[1], Charles A. Odame-Ankrah[1], Jennifer A. Huo[1], Robert McLaren[2], Alex K. Y. Lee[3, 4], Max G. Adam[4], Megan D. Willis[5], Jonathan P. D. Abbatt[5], Cristian Mihele[6], Andrea Darlington[6], Richard L. Mittermeier[6], Kevin Strawbridge[6], Katherine L. Hayden[6], Jason S. Olfert[7], Elijah. G. Schnitzler[8], Duncan K. Brownsey[1], Faisal V. Assad[1], Gregory R. Wentworth[5, a], Alex G. Tevlin[5], Douglas E. J. Worthy[6], Shao-Meng Li[6], John Liggio[6], Jeffrey R. Brook[6], and Hans D. Osthoff[1*]

[1] Department of Chemistry, University of Calgary, Calgary, Alberta, T2N 1N4, Canada

[2] Centre for Atmospheric Chemistry, York University, Toronto, Ontario, M3J 1P3, Canada

[3] Department of Civil and Environmental Engineering, National University of Singapore, Singapore 117576, Singapore

[4] NUS Environmental Research Institute, National University of Singapore, Singapore

[5] Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada

[6] Air Quality Research Division, Environment and Climate Change Canada, Toronto, Ontario, M3H 5T4, Canada

[7] Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, T6G 1H9, Canada

[8] Department of Chemistry, University of Alberta, Edmonton, Alberta, T6G 2G2, Canada

[a] Now at: Environmental Monitoring and Science Division, Alberta Environment and Parks, Edmonton, Alberta, T5J 5C6, Canada

* Corresponding author

for Atmos. Chem. Phys.

Abstract

In this paper, measurements of air pollutants made at a ground site near Fort McKay in the Athabasca oil sands region as part of a multi-platform campaign in the summer of 2013 are presented. The observations included measurements of selected volatile organic compounds (VOCs) by a gas chromatograph – ion trap mass spectrometer (GC-ITMS). This instrument observed a large, analytically unresolved hydrocarbon peak (with retention index between 1100 and 1700) associated with intermediate volatility organic compounds (IVOCs). However, the activities or processes that contribute to the release of these IVOCs in the oil sands region remain unclear.

Principal component analysis (PCA) with Varimax rotation was applied to elucidate major source types impacting the sampling site in the summer of 2013. The analysis included 28 variables, including concentrations of total odd nitrogen (NO_y), carbon dioxide (CO_2), methane (CH_4), ammonia (NH_3), carbon monoxide (CO), sulfur dioxide (SO_2), total reduced sulfur compounds (TRS), speciated monoterpenes (including α- and β-pinene and limonene), particle volume calculated from measured size distributions of particles less than 10 μm and 1 μm in diameter (PM_{10-1} and PM_1), particle-surface bound polycyclic aromatic hydrocarbons (pPAH), and aerosol mass spectrometer composition measurements, including refractory black carbon (rBC) and organic aerosol components. The PCA was complemented by bivariate polar plots showing the joint wind speed and direction dependence of air pollutant concentrations to illustrate the spatial distribution of sources in the area. Using the 95% cumulative percentage of variance criterion, ten components were identified and categorized by source type. These included emissions by wet tailings ponds, vegetation, open pit mining operations, upgrader facilities, and surface dust. Three components correlated with IVOCs, with the largest associated with surface mining and is likely caused by the unearthing and processing of raw bitumen.

1. Introduction

The Athabasca oil sands region of Northern Alberta, Canada, has seen extraordinary expansion of its oil sands production and processing facilities (CAPP, 2016) and associated emissions of air pollutants over the last several decades (Englander et al., 2013; Bari and Kindzierski, 2015). Air emissions from these facilities have been impacting surrounding communities, including the city of Fort McMurray and the community of Fort McKay (WBEA, 2013). To assess the impact of these emissions on human health, visibility, climate, and the ecosystems downwind, it is critical to obtain an understanding of the source types from all activities associated with oil sands operations (ECCC, 2016).

Prior to 2013, there had been only a single industry-independent study of trace gas emissions from the Athabasca oil sands mining operations (Simpson et al., 2010; Howell et al., 2014). The data showed elevated concentrations in n-alkanes (30% of the total quantified hydrocarbon emissions), cycloalkanes (49%), and aromatics (15%) in plumes from an oil sands surface mining facility intercepted from a single aircraft flight. These compounds are associated with oil and gas developments including mining, upgrading, and transportation of bitumen (Siddique et al., 2006). Specifically, these activities involve the use of naphtha, a complex mixture of aliphatic and aromatic hydrocarbons in the range of C_3 to C_{14} containing n-alkanes (e.g., n-heptane, n-octane, and n-nonane) and benzene, toluene, ethylbenzene, and xylenes (BTEX).

In August 2013, a comprehensive air quality study as a part of the Joint Oil Sands Monitoring (JOSM) plan (JOSM, 2012), referred to here as the 2013 JOSM intensive study was conducted. This study was performed in northern Alberta at two ground sites in and near Fort McKay in close proximity (as close as 3.5 km) to oil sands mining operations and from a National Research Council of Canada (NRC) Convair 580 research aircraft to characterize oil sands emissions and their downwind physical and chemical transformations (Gordon et al., 2015; Liggio et al., 2016; Li et al., 2017).

One ground site, located at the Wood Buffalo Environmental Association (WBEA) air monitoring station (AMS) 13 (Fig. 1), was equipped with a comprehensive set of instrumentation to measure concentrations of a wide range of trace gases and aerosols (Table 1), yielding a unique and new data set, parts of which are presented in this paper for the first time. As part of this effort, a gas chromatograph equipped with an ion trap mass spectrometer (GC-ITMS) was deployed at AMS 13. When air masses passing over regions with industrial activities were observed (as judged from a combination of local wind direction and tracer measurements), the total ion chromatogram showed an analytically unresolved hydrocarbon signal associated with intermediate volatile organic compounds (IVOCs) with saturation concentration (C^*) in the range 10^5 µg m^{-3}< C^* < 10^7 µg m^{-3} (Liggio et al., 2016).

Emission estimates for analytically unresolved hydrocarbons range from 5×10^6 kg $year^{-1}$ to 14×10^6 kg $year^{-1}$ for the two facilities that reported such emissions (Li et al., 2017). Using aircraft measurements during the 2013 study, Liggio et al. (2016) showed that IVOCs contributed to the majority of the observed secondary organic aerosol (SOA) mass production in a similar fashion as anthropogenic VOCs contributed to SOA production during the Deepwater Horizon oil spill (de Gouw et al., 2011) and rivaling the magnitude of SOA formation observed downwind of megacities (Liggio et al., 2016), though ultimately it has remained unclear which activities are associated with IVOC emissions.

In this paper, concurrent measurements of air pollutants at the AMS 13 ground site during the 2013 JOSM intensive study are presented. The analytically unresolved hydrocarbon signal was integrated and is presented as a time series and used as an input variable in a principal component analysis (PCA) to elucidate the origin of IVOCs in the Athabasca oil sands by association. The analysis presented here is a receptor analysis focusing on the normalized variability of pollutants impacting the AMS 13 ground site and hence does not constitute a comprehensive emission profile analysis of the oil sands facilities as a whole, for which aircraft-based measurements and/or direct plume or stack measurements are more suitable. PCA was chosen over the more popular positive matrix factorization (PMF) method (Paatero

and Tapper, 1994) because it yields a unique solution and is particularly suited as an exploratory tool for identification of components without *a priori* constraints (Jolliffe and Cadima, 2016). The PCA was complemented by bivariate polar plots (Carslaw and Ropkins, 2012; Carslaw and Beevers, 2013) to show the spatial distribution of sources in the region as a function of locally measured wind direction and speed. A second PCA was performed to investigate which components correlate with (and generate) secondary pollutants, i.e., pollutants that are formed by atmospheric processes. Potential sources and processes contributing to each of the components identified by PCA are discussed.

2. Experimental

2.1 Measurement location

Measurements of air pollutants were made at AMS 13 routine air monitoring station (Fig. 1), which is operated by WBEA. The site is located at 111.6423° W longitude and 57.1492° N latitude about 3 km from the southern edge of the community of Fort McKay, 300 m west from a public road, and 1 km west of the Athabasca river. The immediate vicinity of the site consisted of mixed-leaf boreal forest with a variety of tree species, including poplar, aspen, pine and spruce trees (Smreciu et al., 2013). The site was accessible via a gravel road; traffic on this road was restricted during the study period (August - September, 2013).

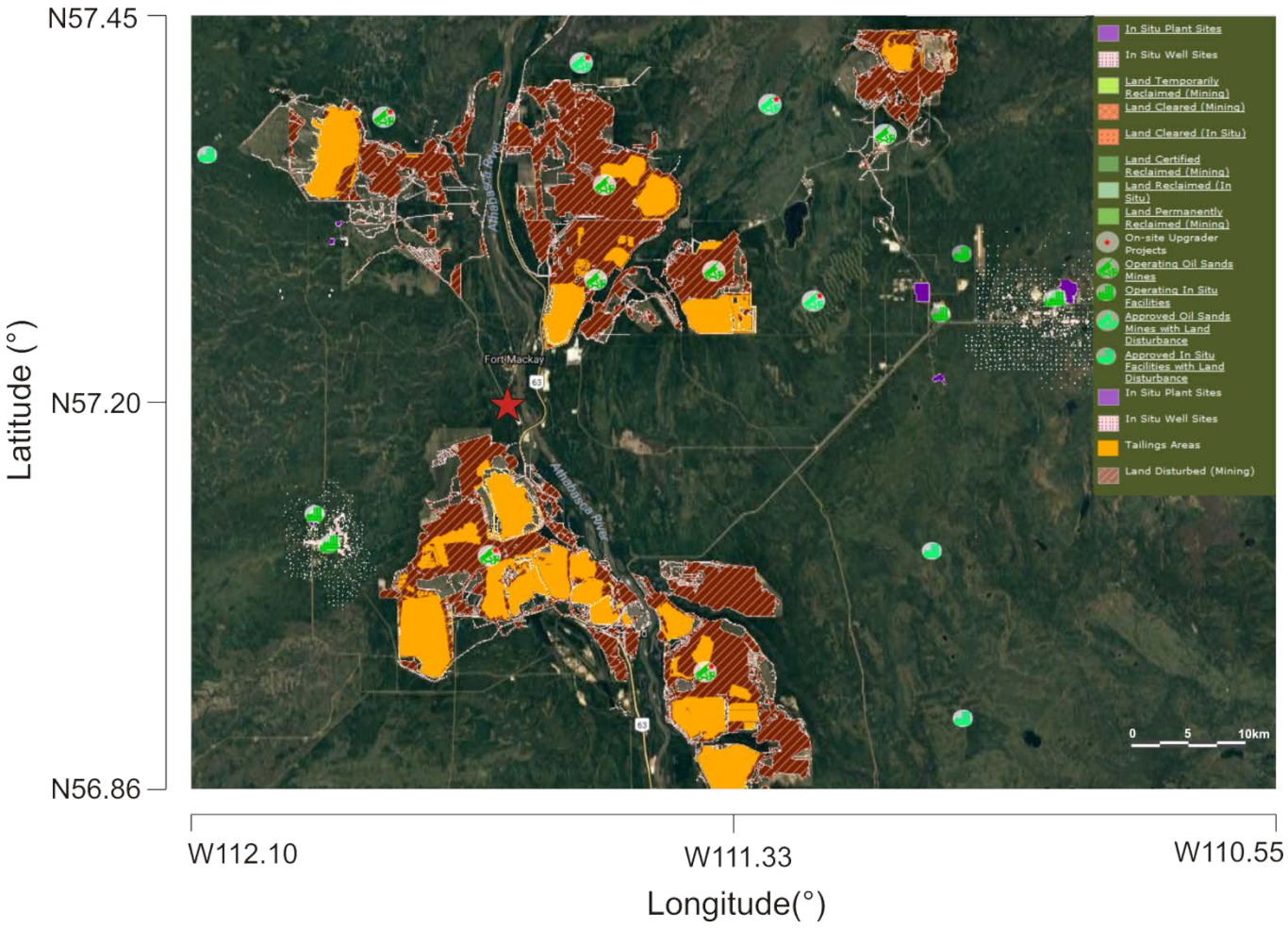
The site is impacted by emissions from nearby oil sands facilities (Table 1 and Fig. 1), including a large surface mining site operated by Syncrude Canada whose northeastern corner is located 3.5 km to the south of AMS 13 (and which is adjacent to the 5 km long Syncrude – Mildred Lake (SML) tailings pond) and from a large upgrader stack facility operated by Suncor Energy Inc. located to the Southeast. There are additional oil sands facilities operated (during the study period) by Canadian Natural Resources Limited, Imperial Oil, and Shell Canada to the North and Northeast. A potentially important

consideration is the photochemical aging of emissions between the points of emission and observation. During daytime, the average surface wind speed was 7.5 km hr^{-1} (2.1 m s^{-1}). The average transit times were 0.5 hr to the edge of the closest mining operation, 1.6 hr to the 12.2 km distant Mildred Lake Plant site, and 3.2 hr to the Muskeg River Mine site located 23.7 km upwind.

Table 1. Oil sands facilities located within 30 km of AMS 13. Distances were estimated using coordinates provided in the National Pollutant Release Inventory (NPRI, 2013) and do not account for the size of each facility whose boundaries may be considerably closer to (or further away from) AMS 13. PACPRM = Petroleum and coal products refining and manufacturing; OGPS = Oil and gas pipelines and storage.

Company	Name	Type	Direction	Distance (km)
Syncrude Canada Ltd.	Mildred Lake Plant Site	PACPRM	S	12.2
Athabasca Minerals Inc.	Susan Lake Gravel Pit	Mining and Quarrying	N	15.5
Syncrude Canada Ltd.	Aurora North Mine Site	PACPRM	NE	18.7
Suncor Energy	Suncor Energy Inc. Oil Sands	PACPRM	SE	19.4
Enbridge Pipelines Inc.	MacKay River Terminal	OGPS	WSW	19.7
Suncor Energy	MacKay River, In-Situ, Oil Sands Plant	PACPRM	WSW	19.9
Enbridge Pipelines Inc.	Athabasca Terminal	OGPS	SE	21.2
Williams Energy	Fort McMurray Hydrocarbon Liquids Extraction Facility	Conventional oil and gas extraction	SE	21.6
Canadian Natural Resources Limited	Horizon Oil Sands Processing Plant and Mine	PACPRM	NNW	21.8
Shell Canada Energy	Muskeg River Mine and Jackpine Mine	PACPRM	NNE	23.7

Figure 1. Map of oil sands facilities showing locations of surface mines and tailings ponds, downloaded from the Oil Sands Information Portal (Alberta, 2017). The red star indicates the location of AMS 13.



2.2 Instrumentation

A large number of instruments was deployed for this study; a partial list whose data were utilized in this manuscript is given in Table 2. Detailed descriptions of these instruments and operational aspects such as calibrations are given in the S.I. Sample observations of analytically unresolved hydrocarbons by GC-ITMS and how these data were used in the analysis are described in section 2.2.1 below.

135 **Table 2**. Instruments used to measure ambient gas-phase and aerosol species during the 2013 JOSM

136 intensive study at AMS 13.

Instrument and Model	Species measured	Time resolution	Reference
Picarro CRDS G2401	CO, CO_2, CH_4	1 min	(Chen et al., 2013; Nara et al., 2012)
Thermo Scientific, Model 42i	NO_y	10 s	(Tokarek et al., 2014; Odame-Ankrah, 2015)
Blue diode cavity ring-down spectroscopy	NO_2	1 s	(Paul and Osthoff, 2010; Odame-Ankrah, 2015)
Thermo Scientific Model 49i	O_3	10 s	(Tokarek et al., 2014; Odame-Ankrah, 2015)
Griffin/FLIR, model 450 GC-ITMS	VOCs	1 hr	(Tokarek et al., 2017b; Liggio et al., 2016)
Thermo Scientific CON101	TS	1 min	n/a
Thermo Scientific 43iTLE	SO_2	1 min	n/a
AIM-IC	$NH_{3(g)}$, $NH_4^+{}_{(p)}$	1 hr	(Markovic et al., 2012)
Aerodyne SP-AMS	rBC, $NH_4^+{}_{(p)}$, $SO_4^{2-}{}_{(p)}$, $NO_3^-{}_{(p)}$, $Cl^-{}_{(p)}$, organics	1-5 min (variable)	(Onasch et al., 2012)
TSI APS 3321	PM_{10-1} size distribution	5-6 min (variable)	(Peters and Leith, 2003)
TSI SMPS (3081 DMA, 3776 CPC)	PM_1 size distribution	6 min	(Wang and Flagan, 1990)
EcoChem Analytics PAS 2000CE	pPAH	1 min	(Wilson et al., 1994; Burtscher et al., 1982)

137

2.2.1 Analytically unresolved hydrocarbon signature

As previously reported (Liggio et al., 2016), the total ion chromatogram of the GC-ITMS occasionally showed elevated and analytically unresolved hydrocarbons in the volatility range of C_{11} – C_{17} with saturation vapor concentration (C^*) from 10^5 µg m^{-3}< C^* < 10^7 µg m^{-3}. An example is shown in Fig. 2.

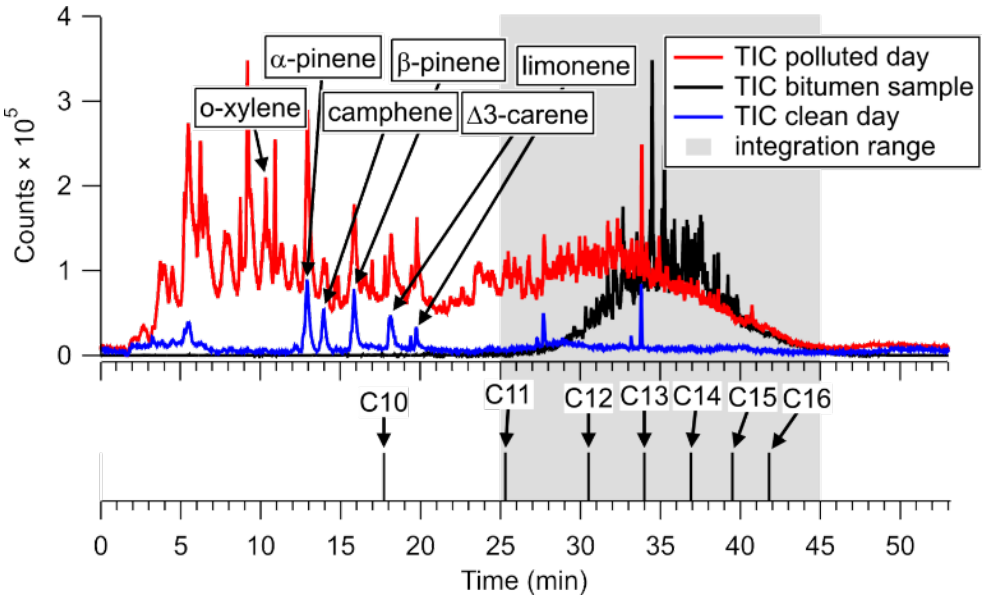


Figure 2. (**Top**) Total ion chromatograms of air samples collected on August 27, 2013 from 18:04 to 18:14 UTC (red) and on August 28, 2013 from 13:43 to 13:53 UTC (blue). The TIC of a head space sample of ground-up bitumen collected post-campaign is superimposed (black). The gray area indicates the range over which IVOC signal was integrated. (**Bottom**) Retention times of n-alkanes, determined after the measurement intensive by sampling a VOC mixture containing a C_{10} – C_{16} n-alkane ladder.

An offline analysis of the headspace above ground-up bitumen gave a similarly unresolved hydrocarbon signal (Fig. 2, black trace). In this particular case, the ambient air chromatogram also shows enhancements of lower molecular weight hydrocarbons (possibly from naphtha) that were not observed in the bitumen sample**.** The observed unresolved hydrocarbon feature is qualitatively similar to the "large chromatographic hump of unresolved complex mixtures" reported by Yang et al. (2011) during

their analysis of bitumen extracts.

The major ions contributing to the unresolved signals in Figure 2 are associated with alkanes (i.e., *m/z* 55, 57, 67, 69, etc. – see Fig. S-1). In contrast, counts at masses associated with aromatics (i.e., *m/z* 115, $C_9H_7^+$, and *m/z* 91, $C_7H_7^+$) as reported by Cross et al. (2013) were negligible in both the bitumen head space and polluted day samples. The ~~strong~~ resemblance of the unresolved hydrocarbon feature in ambient air with the bitumen head space sample both in terms of volatility (i.e., elution time) and electron impact mass fragmentation is consistent with bitumen as the source of IVOCs at this site.

In the interpretation of the integrated IVOC signal, it is assumed that it is of primary origin, i.e., emitted directly from point sources in the vicinity of the measurement site. For the PCA, the unresolved signal was integrated from a retention time of 25 min to 45 min (gray area in Fig. 2) in all ambient air chromatograms.

The IVOCs observed in this work likely encompass a portion of the total that is emitted. For example, IVOCs generated by combustion processes, such as ~~aircraft~~ diesel engine exhaust, are comprised of aliphatic alkanes, including cyclic and branched alkanes, and~~,~~ aromatics (Gentner et al., 2012; Zhao et al., 2015)~~and oxygenated compounds (Cross et al., 2013)~~. The use of a chromatographic column in this work biases the IVOC signal towards hydrocarbon-IVOCs, since oxygenated compounds (i.e., alcohols and acids) will not elute from the analytical column. Furthermore, the recovery of VOCs from the pre-concentration unit, while reproducible and likely complete for n-alkanes which bracket the bulk of IVOC emitted and whose calibration curves were linear, is not known for late-eluting compounds, but is assumed to be sufficiently reproducible to yield a semi-quantitative signal.

2.3 Principal Component Analysis

The PCA was carried out using the "Statistical Analysis System" (SAS™) Studio 3.4 software (SAS, 2015) using a method similar to that described by Thurston et al. (2011; 1985). The source-related components and their associated profiles are derived from the correlation matrix of the input trace constituents. This approach assumes that the total concentration of each "observable" (i.e., input variable) is made up of the sum of contributions from each of a smaller number of pollution sources and that variables are conserved between the points of emission and observation.

2.3.1 Selection of variables

22 variables whose ambient concentrations are dominated by primary emissions or which are formed very shortly after emission (such as the less oxidized oxygenated organic aerosol (LO-OOA) factor observed by the SP-AMS, see below) were included in the PCA (Table 3). These variables included CO_2, CH_4, NO_y, CO, and SO_2, which are known to be emitted in the oil sands region from stacks, the mine fleet and faces, tailings ponds, and by fugitive emissions (Percy, 2013). The median NO_x (= NO + NO_2) to NO_y ratio was 0.85, consistent with the close proximity of the measurement site to emission sources and limited chemical processing. Because NO_x constituted a large fraction of NO_y, its temporal variation was captured by the latter, and it was not included as a separate variable in the PCA.

For this work, mixing ratios of all non-methane hydrocarbons (NMHCs) that were quantified (i.e., o-xylene, the n-alkanes decane and undecane, the aromatics 1, 2, 3- and 1, 2, 4-trimethylbenzene (TMB), as well as limonene and α- and β-pinene) were included as variables. In addition, the aforementioned unresolved signal associated with IVOCs was included as a variable by integrating total GC-ITMS ion counts (*m/z* 50 –425) over a retention time range of 25-45 min (retention index range of 1100 to 1700).

Gas-phase ammonia was included as a variable because elevated reduced nitrogen concentrations have been observed in the region and were linked to the use of ammonia on an industrial scale, for example

as a floating agent and for hydrotreating (Bytnerowicz et al., 2010). Total sulfur and total reduced sulfur were added as tracers of upgrader stack SO_2 emissions and of "odours", believed to be emitted from oil sands tailings ponds which continue to be of concern in surrounding communities (Small et al., 2015; Percy, 2013; Holowenko et al., 2000).

Refractory black carbon was added as a variable since it is present in diesel truck exhaust and in biomass burning plumes and, hence, a combustion tracer (Wang et al., 2016; Briggs and Long). pPAHs were included because of their association with facility stack emissions and combustion particles in the area (Allen, 2008; Grimmer et al., 1987). Hydrocarbon-like organic aerosol (HOA) was included as a surrogate for fossil fuel combustion by vehicles (Jimenez et al., 2009). The LO-OOA factor was included as it ~~is unique to the Alberta oil sands and~~ appears to form rapidly after emission of precursors (Lee et al., 2018). Supermicron aerosol volume (PM_{10-1}, i.e., the volume of particles between PM_{10} and PM_1) was also included as a tracer of coarse particles from primary sources, which are expected to be dominated by dust emissions.

Table 3. Variables observed at the AMS 13 ground site during the 2013 JOSM campaign used for PCA.

Variable	Unit	Median[a]	Average[a,b]	~~Standard deviation~~R	LOD[~~ed~~]	Min.[a]	Max.[a]	Fraction <LOD
Anthropogenic VOCs								
o-xylene	pptv[~~fe~~]	5	30	~~69~~2.3	1	< LOD	635	10%
1,2,3 - TMB	pptv	1.7	4.3	~~7.9~~1.8	0.2	< LOD	67	27%
1,2,4 - TMB	pptv	2.1	7.7	1.9~~4.7~~	0.2	< LOD	107	8%
decane	pptv	0.5	8.5	2.1~~18.2~~	0.1	< LOD	125	44%
undecane	pptv	0.4	3.0	2.1~~6.3~~	0.1	< LOD	37	39%
Biogenic VOCs								
α-pinene	pptv	477	542	0.74~~401~~	1	19	1916	0%
ß-pinene	pptv	390	467	~~334~~0.72	1	18	1594	0%
limonene	pptv	150	179	0.88~~158~~	2	< LOD	711	1%
Combustion tracers								
NO_y	ppbv	1.79	4.00	1.4~~5.44~~	0.01	0.13	41.6	0%
rBC	µg m^{-3}	0.13	0.20	0.50~~10~~	0.02	< LOD	0.90	40%
CO	ppbv	117.6	120.0	0.15~~18.2~~	5.7[~~hg~~]	90.9	241.2	0%
CO_2	ppmv	420.2	433.2	~~39.5~~0.091	0.4[~~hg~~]	386.0	577.7	0%
Aerosol species								
pPAH	ng m^{-3}	1	2	~~2~~1	1[~~e~~]	< LOD	14	39%
PM_{10-1}	µm^3 cm^{-3}	11.2	14.4	~~12.9~~0.90	0.003	1.0	79.5	0%
HOA	µg m^{-3}	0.31	0.43	0.81~~35~~	N/A[~~fg~~]	0.04	2.32	N/A
LO-OOA	µg m^{-3}	1.19	2.00	1.1~~2.26~~	N/A[~~g~~]	0.11	15.6	N/A
Sulfur species								
Total sulfur (TS)	ppbv	0.22	1.41	~~4.27~~3.0	0.13	< LOD	33.3	35%
SO_2	ppbv	< LOD	1.0	4.0	0.2	< LOD	33.5	81%
Total reduced sulfur (TRS)	ppbv	0.26	0.38	~~1.05~~2.8	0.2	< LOD	14.8	81%
Other								
IVOCs	Counts × min	1.8×10^7	3.4×10^7	~~4.2×10^7~~1.2	N/A[~~g~~]	1.4×10^6	2.5×10^8	N/A
CH_4	ppbv	1999.2	2065.5	~~169.6~~0.082	1.8[~~hg~~]	1880	2959	0%
NH_3	µg m^{-3}	0.79	1.10	~~1.03~~0.94	0.05	0.06	5.75	39%

[a] Values were determined only from data points included in the PCA, not from the entire campaign.

[b] Average and relative standard deviation were calculated before zeros were replaced with 0.5×LOD.

[c] ~~Estimated.~~RSD = relative standard deviation

[~~de~~] LOD = limit of detection~~.~~

[~~ef~~] ppt = parts-per-trillion by volume (10^{-12})

[~~fg~~] N/A = data not available

[~~gh~~] calculated using 3 × standard deviation at ambient background levels

To assess which components ~~have the greatest~~ impact ~~on~~ secondary product formation, a second PCA was performed which included variables mainly formed through atmospheric chemical processes and whose concentrations more strongly depend on air mass chemical age than those variables selected initially. In this PCA, odd oxygen (O_x = O_3 + NO_2), submicron aerosol $SO_4^{2-}{}_{(p)}$, $NO_3^-{}_{(p)}$, $NH_4^+{}_{(p)}$, a second, more-oxidized OOA factor (MO-OOA), and PM_1 volume were included, increasing the total number of variables to 28 (Table 4). Furthermore, since oxidation of IVOCs leads to formation of SOA (Robinson et al., 2007; Lee et al., 2018), and the photochemical conversion of IVOC to SOA may adversely affect the PCA, a PCA without secondary and aerosol variables is presented in the S.I. (Table S-10).

Table 4. Variables added in the second PCA. Particle-phase concentrations, i.e., $SO_4^{2-}{}_{(p)}$, $NO_3^-{}_{(p)}$, $NH_4^+{}_{(p)}$ and MO-OOA were made by aerosol mass spectrometry and account for PM_1 only.

Variable	Unit	Median	Average	~~Standard deviatio~~RS	LOD	Min.	Max.
O_x	ppbv	7.35	11.1	~~10.6~~0.95	1	<LOD	41.1
$SO_4^{2-}{}_{(p)}$	μg m^{-3}	0.3	0.8	1.4~~1~~	0.1	<LOD	6.6
$NO_3^-{}_{(p)}$	μg m^{-3}	0.08	0.13	~~0.13~~1.0	0.01	0.01	0.72
$NH_4^+{}_{(p)}$	μg m^{-3}	0.13	0.28	~~0.37~~1.3	0.05	<LOD	2.21
MO-OOA	μg m^{-3}	1.65	1.83	~~0.960~~0.52	N/A	1.41×10^{-6}	4.65
PM_1 volume	μm^3 cm^{-3}	2.48	3.77	~~3.72~~0.99	N/A	0.35	20.9

2.3.2 Treatment of input data

Data used in the PCA were averaged to match the time resolution of the GC-ITMS VOC and IVOC measurements, i.e. over 10-minute-long periods (spaced ~ 1 hr apart) set by the start and stop times of the GC-ITMS pre-concentration period. When concentrations were below their respective limit of detection (LOD; values are given in Table 3), half the reported LOD was used to minimize bias (Harrison et al., 1996; Polissar et al., 1998; Zhao et al., 2004; Guo et al., 2004)~~(Harrison et al., 1996; Buhamra et al., 1998)~~. Prior to PCA, input variables were standardized to eliminate unit differences by subtracting the mean concentration $\overline{C_i}$ of pollutant *i* from the concentration of sample *k* ($C_{i,k}$) and dividing by the standard deviation (s_i) of all samples included in the PCA.

$$Z_{i,k} = \frac{C_{i,k} - \overline{C_i}}{s_i} \quad (1)$$

Here, $Z_{i,k}$ is the standardized pollutant concentration. In total, 218 data points from all identified species over the period of the campaign were used for the main PCA.

2.3.3 PCA solutions

In this work, the Varimax method (Kaiser, 1958) was used to rotate the loading matrix. This method is an orthogonal rotation (i.e., components are not expected to correlate) which minimizes the impact of high loadings, making the results easier to interpret (Kaiser, 1958). Several criteria (Table S-10) were considered for component selection: the latent root criterion, i.e., on the basis that rotated eigenvalues must be greater than unity, the (cumulative) percentage of variance criterion, where the extracted components accounts for >95% of the variance, and the Scree test (Fig. S-2) (Thurston and Spengler, 1985; Guo et al., 2004; Hair et al., 1998; Cattell, 1966). For the optimal solution presented in the main

manuscript, the 95% variance criterion was chosen, providing a 10-component solution for the PCA with only primary variables and an 11-component solution for the PCA with both primary and secondary variables. Components 1 through 4 were consistent regardless of the number of components retained. Solutions with fewer and more components are presented in the supplemental material section.

Time series of each of the components were calculated by multiplying the original standardized matrix by the rotated loading matrix and were used to generate bivariate polar plots (section 2.4).

2.4 Bivariate polar plots

The PCA was complemented by bivariate polar plots showing the wind speed and direction dependence of air pollutant concentrations. The use of these representations implies a linear relationship between local wind conditions and air mass origin, which may not be always the case (for example, during or after stagnation periods). In addition, local topography, such as the Athabasca river valley, complicates regional air flow patterns and limit the interpretability of polar plots in general and in particular to the E of AMS 13, where the river valley is located. The plots were generated with the Openair software package (Carslaw and Ropkins, 2012; Carslaw and Beevers, 2013) using the R programming language and the open-source software "RStudio: Integrated development environment for R" (RStudio Boston, 2017). The default setting (100) was used as the smoothing function.

3. Results

3.1. Overview of the data set

Time series of the 22 pollution tracers chosen for PCA -are presented in Fig. 3, grouped approximately by source type. Statistics of the data (i.e., median, average, maxima, minima, etc.) are summarized in Table 3.

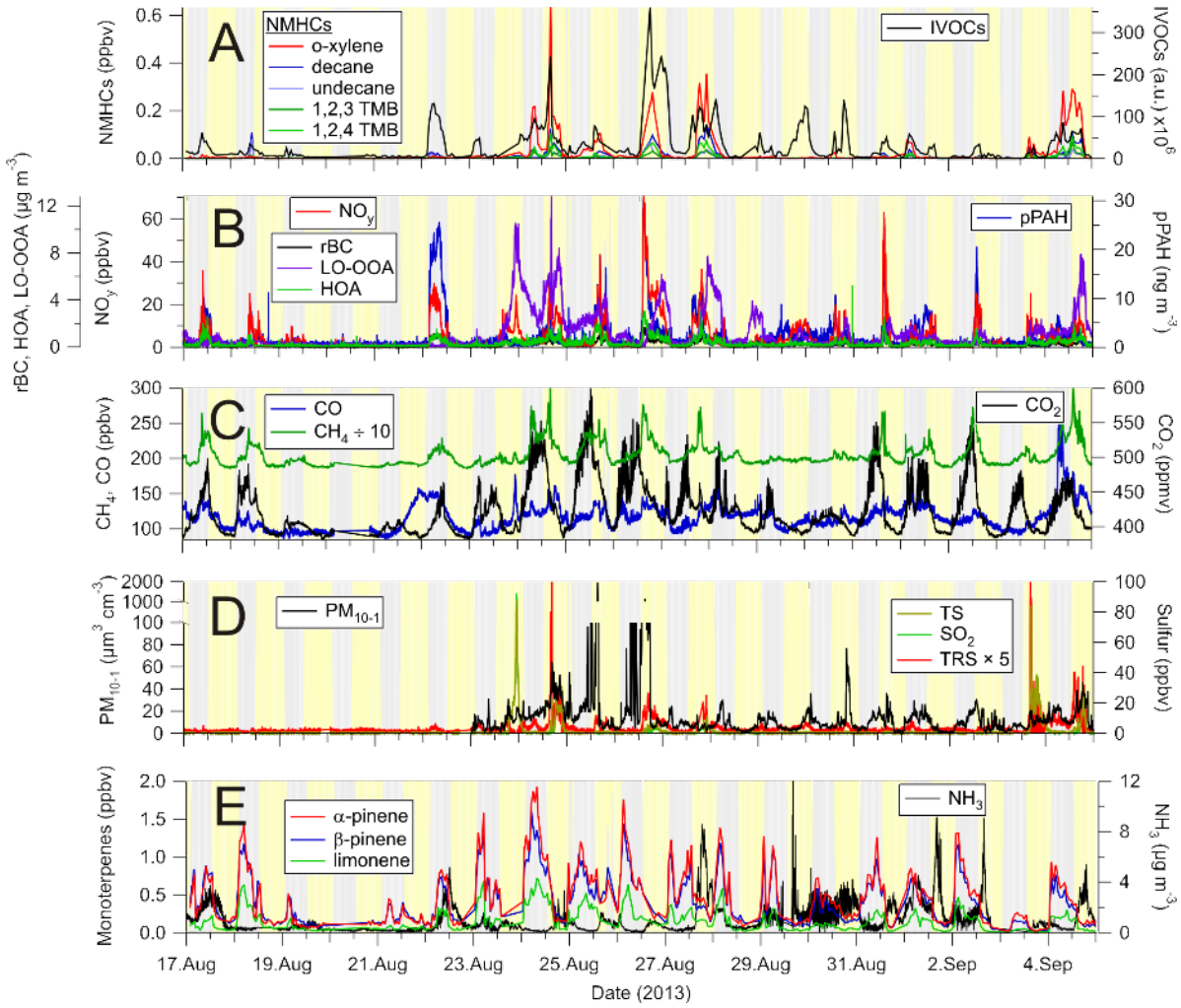


Figure 3. Time series of selected pollution tracers observed at the AMS 13 ground site in the Athabasca oil sands during the 2013 JOSM measurement intensive. The gray and yellow backgrounds represent night and day, respectively. (**A**) Selected non-methane hydrocarbons (NMHCs) and IVOCs. (**B**) Combustion product tracers: refractory black carbon (rBC), total odd nitrogen (NO_y) and particle surface bound polycyclic aromatic hydrocarbons (pPAH), and organic aerosol components: hydrocarbon-like organic aerosol (HOA) and less oxidized oxygenated organic aerosol (LO-OOA). (**C**) Methane (CH_4), carbon dioxide (CO_2) and monoxide (CO). (**D**) Total sulfur (TS), sulfur dioxide (SO_2), and total reduced sulfur (TRS) and PM_{10} particle volume. (**E**) Biogenic VOCs (α-pinene, ß-pinene and limonene) and ammonia (NH_3).

Time series of VOCs of primarily anthropogenic origin (i.e., o-xylene, 1, 2, 3- and 1, 2, 4-TMB, etc.) as well as the IVOC signature are shown in Fig. 3A. The abundances of these species, as well as the other compounds, were highly variable and varied as a function of time of day (i.e., boundary layer mixing height) and air mass origin, with higher VOC concentrations generally observed during daytime. The VOC concentrations varied between nearly pristine, remote conditions, with concentrations below detectable limits, to mixing ratios of aromatic species exceeding 100 pptv. The concentration range of o-xylene is within the extremes reported by WBEA in their 2013 annual report (WBEA, 2013), exemplifying that the data set is representative of typical pollutant levels in this region.

While there is some obvious covariance between variables (i.e., when the mixing ratios of one particular VOC increases, so do others), the ratios of hydrocarbons varied considerably. For example, on August 18, 10:50 UTC, the n-decane to o-xylene ratio was ~22:1, whereas on August 24, 07:40 UTC it was ~1:5.7. The IVOC magnitude also varied greatly and often increased and decreased in tandem with the other VOCs (e.g., on Aug 24, 16:30 UTC) but also increased independently from the other VOC abundances (e.g., on Aug 30, 01:20 UTC, and on the night of Aug 22). This behaviour suggests the presence of multiple sources with distinct signatures that are being sampled to a varying extent at different times or, perhaps, a single source whose emission profile varies. This, coupled with the intermittency of the highly elevated signals, presents an analysis problem frequently encountered in environmental analysis that is usually investigated through a factor or principal component analysis (Thurston et al., 2011; Guo et al., 2004).

Presented in Fig. 3B are the time series of NO_y, rBC and pPAH abundances, all of which are combustion byproducts. For example, rBC is emitted from combustion of fossil fuels, biofuels, open biomass burning, and burning of urban waste (Bond et al., 2004). Similar to the VOCs, the abundances of these species varied greatly, from very low, continental background levels (i.e., <100 pptv of NO_y, < LOD for rBC and pPAHs) to polluted concentrations (i.e., > 60 ppbv of NO_y, > 1 µg m^{-3} rBC, > 10 ng m^{-3} pPAHs)

characteristic of polluted urban and industrial areas. When high concentrations of NO_y were observed, its main component was NO_x (data not shown), which is a combustion byproduct usually associated with automobile exhaust. In the Alberta oil sands, emissions from off-road mining trucks as well as the upgrading processes are the main contributors to the NO_y burden (Percy, 2013; Watson et al., 2013).

Shown in Fig. 3C are the mixing ratios of the greenhouse gases CH_4 and CO_2 along with CO. Abundances of CO_2 were clearly attenuated by photosynthesis and respiration of the vegetation near the measurement site, as judged from the strong diurnal cycle in its concentration (not shown). Maxima typically occurred shortly after sunrise, coincident with the expected break-up of the nocturnal boundary layer. In addition to biogenic emissions from vegetation and soil, CO_2 originates from a variety of point and mobile sources in this region, including off-road mining trucks (Watson et al., 2013) and the extraction, upgrading, and refining of bitumen and on-road vehicle sources in the area (Nimana et al., 2015a, b). Concentrations of CO_2 spiked whenever these emissions were transported to the measurement site.

Concentrations of CH_4 also exhibit a diurnal cycle, with higher concentrations generally observed at night and peaking in the early morning hours. While CH_4 and CO_2 mixing ratios frequently correlated in plumes, their ratios were variable overall, suggesting they often originated from distinct sources. Potential methane point sources in the region include microbial production in tailings ponds (Siddique et al., 2012) and fugitive emissions associated with the mining and processing of bitumen (Johnson et al., 2016). Indeed, a recent analysis shows tailings ponds and open pit mining sources to be the largest sources of CH_4 in the region (Baray et al., 2018).

Similar to the anthropogenic VOCs, the abundances of CH_4 and CO_2 were highly variable and ranged from minima of 1.88 and 384 ppmv to maxima of 2.96 and 578 ppmv, corresponding to maximum enhancements of 1.63 and 1.47 relative to tropospheric global monthly means of 1.806±0.001 and

394.3±0.1 ppmv for July, 2013 (Dlugokencky, 2017b, a), respectively.

Mixing ratios of CO also varied with time but generally were not elevated greatly (median 118 ppbv) above background levels (minimum 91 ppbv), except for occasional spikes in concentration (Fig. 3C). Carbon monoxide is a tracer of biomass burning and fossil fuel combustion, in particular in automobiles with poorly performing or absent catalytic converters, but is also a byproduct of the oxidation of VOCs, in particular of methane and isoprene which are oxidized over a wide area upwind of AMS 13 (Miller et al., 2008).

Time series of sulfur species and PM_{10-1} volume are shown in Fig. 3D. The TS and SO_2 data are dominated by intermittent plumes containing SO_2 mixing ratios exceeding 5 ppbv. The highest mixing ratio observed was 92.5 ppbv (in between the preconcentration periods of the GC-ITMS). Mixing ratios of SO_2 exhibited the most variability of all pollutants, as judged from the relative standard deviation of each of the measurements (Table 3). TRS levels were generally small (< 1 ppbv) and variable, except for plumes; TRS abundances in plumes, however, are more uncertain since they were calculated by subtraction of two large numbers. When TS and SO_2 abundances were low (< 1 ppbv), TRS abundances were variable and occasionally exhibited spikes that did not show any obvious correlation with other variables, suggesting the presence of one or more distinct TRS sources. PM_{10} volume concentrations varied a lot as well and, just like TRS, did not show an obvious correlation with other variables. Fugitive dust emissions likely contributed to much of the PM_{10} volume in the Athabasca oil sands region (Wang et al., 2015).

Time series of monoterpene mixing ratios are shown in Fig. 3E. α-Pinene was generally the most abundant monoterpene, followed by β-pinene. Their ratio, averaged over the entire campaign was 1:0.85, though occasionally the α- to β-pinene ratio was below 1:2 (e.g., on Aug 28, 14:50 UTC and Sept 5, 12:40 UTC). Terpene mixing ratios were generally higher at night than during the day, with maxima of 1.9 and 1.6 ppbv, respectively, a diurnal pattern consistent with what has been observed at other forest

locations (Fuentes et al., 1996). Monoterpenes are emitted by plants via both photosynthetic and non-photosynthetic pathways (Fares et al., 2013; Guenther et al., 2012); at night, their emissions accumulate in a shallow nocturnal boundary layer, whereas during daytime, they are entrained aloft (above the canopy) and oxidized by the hydroxyl radical (OH) and O_3, which are more abundant during the day than at night (Fuentes et al., 1996). α- and β-pinene mixing ratios were lowest mid-day (median values at noon of 140 and 133 pptv, respectively). The largest daytime concentrations were observed on Aug 25, a cloudy day (as judged from spectral radiometer measurements of the NO_2 photolysis frequency): on this particular day, mixing ratios at noon were 687 and 850 pptv, respectively.

Also shown in Fig. 3E is the time series of ammonia. These data were dominated by spikes which were observed sporadically and did not correlate with other variables, suggesting the presence of nearby ammonia point sources. Ammonia was not as variable as some of the other pollutants (e.g., the anthropogenic VOCs, sulfur species) as judged from its relative standard deviation (Table 3), which suggests a geographically more disperse source or sources similar to CO or CH_4, which have a "background". This is consistent with a recent study by Whaley et al. (2018) that estimated over half (~57%) of the near-surface NH_3 during the study period originated from NH_3 bi-directional exchange (i.e. re-emission of NH_3 from plants and soils), with the remainder being from a mix of anthropogenic sources (~20%) and forest fires (~23%).

3.2. Principal component analysis

3.2.1. PCA with primary variables

The loadings of the optimum solution are presented in Table 5. The 10-component solution accounts for a cumulative variance of 95.5%. The communalities for the analysis, i.e., the fraction of total pollutant observations accounted for by the PCA are all greater than 85%, with the lowest communality obtained

for the IVOCs (0.86).

In the following, an overview of the observed components is presented. Associations with r>0.7, r>0.3, and r>0.2 are referred to as "strong", "weak", and "poor", respectively. Hypothesized identifications are given in section 4 and are summarized in Table 6 and Fig. 4.

The component accounting for most of the variance of the data, component 1, is strongly associated with the anthropogenic VOCs (r > 0.87), weakly associated with CH_4 (r = 0.59), TRS (r = 0.59), HOA (r = 0.40), LO-OOA (r = 0.45), CO (r = 0.41), and the IVOCs (r = 0.31), and poorly associated with NO_y (r = 0.27) and rBC (r = 0.30). Component 2 is strongly associated with the combustion tracers NO_y (r = 0.82), rBC (r = 0.77), HOA (r = 0.74), and pPAH (r = 0.94), weakly associated with CH_4 (r = 0.39) and IVOCs (r = 0.39), and poorly associated with ammonia (r = 0.20), and undecane and decane (r = 0.27 and 0.22, respectively). Component 3 is strongly associated (r > 0.9) with the biogenic VOCs and weakly associated with CO_2 (r = 0.48) and shows poor negative correlations with NO_y (r = -0.26) and ammonia (r = -0.24). Component 4 is strongly associated with SO_2 and TS (r = 0.97 and 0.93, respectively) and poorly with NO_y (r = 0.21) and LO-OOA (r = 0.28).

Components 1 through 4 emerged regardless of the number of components used to represent the data, whereas the structure of components 5 through 10 only fully emerged in the 10-component solution (see S.I.). Hence, components 6 through 10 are somewhat tentative as many (i.e., 7 – 9) are single variable components and have eigenvalues close to or below unity, i.e., account for less variance than any single variable. As a result, the interpretations of these components are subject to more uncertainty and are more speculative but are presented in the S.I. for the sake of completeness and transparency. For the purpose of this manuscript, this is inconsequential as components 6 – 10 are not associated with IVOCs.

Table 5. Loadings for the 10-factor, optimal solution (primary variables only). Coefficients with Pearson correlation coefficients r>0.3 are shown in bold font.

	1	2	3	4	5	6	7	8	9	10	Commu-nalities
Anthropogenic VOCs											
o-xylene	**0.88**	0.08	0.02	0.10	0.14	0.13	0.07	-0.04	0.16	**0.32**	0.95
1,2,3 - TMB	**0.93**	0.16	0.07	0.05	0.05	0.11	0.04	-0.02	0.18	-0.01	0.95
1,2,4 - TMB	**0.94**	0.14	0.01	0.10	0.11	0.08	0.07	-0.03	0.18	0.13	0.98
decane	**0.92**	0.22	-0.02	0.15	0.23	0.01	0.05	0.04	0.04	0.03	0.97
undecane	**0.87**	0.27	-0.08	0.23	0.20	-0.06	0.12	0.07	-0.04	-0.10	0.96
Biogenic VOCs											
α-pinene	-0.03	-0.08	**0.98**	-0.11	0.02	0.04	0.01	-0.08	0.02	0.01	0.98
ß-pinene	-0.02	-0.08	**0.98**	-0.12	0.02	0.03	0.02	-0.07	0.00	0.01	0.98
limonene	0.07	-0.03	**0.92**	-0.08	0.12	0.24	0.05	-0.11	0.03	-0.05	0.95
Combustion tracers											
NO_y	0.27	**0.82**	-0.26	0.21	0.22	-0.04	0.02	0.10	-0.08	0.01	0.92
rBC	**0.30**	**0.77**	0.03	0.05	**0.44**	0.10	0.09	0.13	0.12	-0.10	0.94
CO	**0.41**	0.18	0.04	0.02	0.09	0.09	0.08	0.06	**0.87**	-0.01	0.99
CO_2	0.09	0.08	**0.48**	-0.12	-0.03	**0.77**	0.25	-0.14	0.05	-0.08	0.95
Aerosol species											
pPAH	0.06	**0.94**	-0.07	-0.13	-0.11	0.07	0.01	0.13	0.10	0.04	0.95
PM_{10-1}	0.18	0.14	0.08	0.09	0.11	0.17	**0.93**	-0.03	0.07	0.08	0.98
HOA	**0.40**	**0.74**	0.02	0.12	0.25	0.15	0.23	-0.06	0.16	0.09	0.90
LO-OOA	**0.45**	0.11	0.12	0.28	**0.72**	0.05	0.25	0.00	0.10	0.04	0.91
Sulfur											
TS	0.25	0.04	-0.16	**0.93**	0.08	-0.05	0.07	-0.02	0.01	0.12	1.00
SO_2	0.12	0.03	-0.15	**0.97**	0.02	-0.04	0.03	-0.03	0.01	-0.05	0.99
TRS	**0.59**	0.04	-0.08	0.11	0.26	-0.04	0.16	0.04	-0.04	**0.71**	0.96
Other											
IVOCs	**0.31**	**0.39**	0.12	-0.08	**0.74**	-0.02	-0.02	-0.06	0.02	0.20	0.86
NH_3	0.01	0.20	-0.24	-0.05	-0.02	-0.08	-0.03	**0.94**	0.04	0.02	0.99
CH_4	**0.59**	**0.39**	0.10	-0.05	0.12	**0.59**	0.11	0.00	0.17	0.14	0.93
Eigenvalues	5.72	3.32	3.23	2.16	1.64	1.13	1.13	0.99	0.96	0.74	
% of variance	25.99	15.08	14.69	9.80	7.46	5.14	5.13	4.51	4.36	3.35	
Cumulative variance	25.99	41.07	55.76	65.56	73.02	78.16	83.30	87.81	92.17	95.52	

399 **Table 6**. Hypothesized identifications of principal components.

Component	Key observations	Possible source(s)	Relevant references
1	Enhancements of aromatics, n-alkanes, TRS, NO_y, rBC, HOA, LO-OOA, CO and CH_4	Wet tailings ponds and associated facilities	(Simpson et al., 2010; Small et al., 2015; Percy, 2013; Holowenko et al., 2000; Howell et al., 2014)
2	Enhancements of NO_y, rBC, pPAH and HOA due to engine exhaust	Mine fleet and operations	(Wang et al., 2016; Grimmer et al., 1987; Allen, 2008; Briggs and Long, 2016)
3	Enhancements of monoterpenes and CO_2, poor anticorrelation with NO_y and absence of anthropogenic VOCs	Biogenic emission and respiration	(Guenther et al., 2012; Helmig et al., 1999)
4	Enhancements of SO_2 and TS, poor correlation with NO_y and LO-OOA	Upgrader facilities	(Simpson et al., 2010; Kindzierski and Ranganathan, 2006)
5	Enhancements of IVOCs, rBC, LO-OOA, NO_y, and TRS	Surface exposed bitumen and hot-water based bitumen extraction	this work
6	Enhancements of CO_2 and CH_4, absence of combustion tracers	Mine face and soil	(Johnson et al., 2016; Rooney et al., 2012)
7	Enhancement of PM_{10-1}	Wind-blown dust	(Wang et al., 2015)
8	Enhancement of ammonia	Fugitive emissions from storage tanks and natural soil/plant emissions	(Bytnerowicz et al., 2010; Whaley et al., 2018)
9	Enhancement of CO	Incomplete hydrocarbon oxidation	(Marey et al., 2015)
10	Enhancements of TRS and o-xylene, poor association with CH_4	Composite tailings	(Small et al., 2015; Warren et al., 2016)

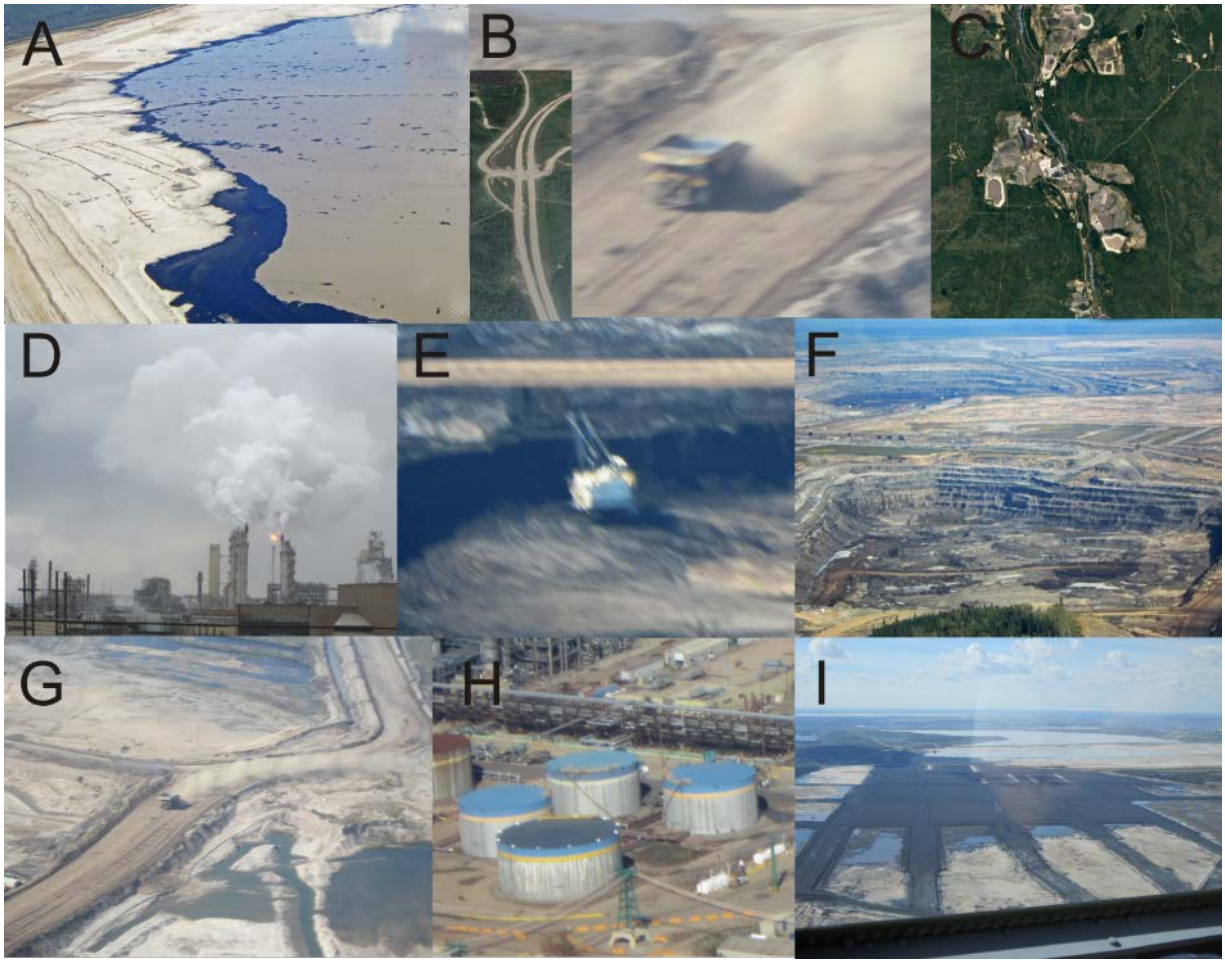


Figure 4. Images of likely sources associated with each of the principal components. From top left to bottom: **(A)** Wet tailings ponds (component 1). **(B)** Mine truck fleet and highway traffic emissions (component 2). **(C)** Biogenic emissions from vegetation (component 3). **(D)** Upgrader facilities (component 4). **(E)** Exposed bitumen on mined surfaces (component 5). **(F)** Fugitive greenhouse gas emissions from mine faces (component 6). **(G)** Wind-blown dust from exposed sand (component 7). **(H)** Fugitive emissions of ammonia from storage tanks (Component 8). **(I)** Composite (dry) tailings (component 10). No image is shown for production CO from oxidation of VOCs (component 9).

3.2.2. Extended PCA with added secondary variables

The loadings of the optimum solution that includes primary and secondary variables are shown in Table 7. In this 11-component solution, the 10 components originally identified were preserved, though their relative order was changed, with the upgrader component moving from the 4th to 2nd position. There was one new component (#6), which encompassed only secondary species, including MO-OOA (r = 0.92), O_x (r = 0.33), $NO_{3(p)}^-$ (r = 0.36), PM_1 (r = 0.31) and LO-OOA (r = 0.31).

$NH_{4(p)}^+$, $SO_{4(p)}^{2-}$, and $NO_{3(p)}^-$ are associated with the stack emissions component (#2, with r = 0.84, 0.84 and 0.44, respectively), which also weakly correlated with PM_1 (r = 0.44) and O_x (r = 0.36). The association of secondary variables with the primary components suggests rapid formation of these secondary products on a time scale that is similar to the transit time of the pollutants to the measurement site. PM_1 correlated strongly with the major IVOC component (component 5, r = 0.80), which also weakly associated with LO-OOA (r=0.66) and $NO_{3(p)}^-$ (r = 0.59), as well as $NH_{4(p)}^+$ and $SO_{4(p)}^{2-}$ (r = 0.32 and 0.33, respectively).

Table 7. Loadings for the 11-component solution with the inclusion of variables associated with secondary processes.

	1	2	3	4	5	6	7	8	9	10	11	Commu-nalities
Anthropogenic VOCs												
o-xylene	**0.89**	0.16	0.04	0.04	0.15	0.00	0.10	0.07	-0.04	0.17	0.24	0.94
1,2,3 - TMB	**0.91**	0.13	0.10	0.16	0.09	0.07	0.11	0.03	-0.03	0.16	-0.08	0.95
1,2,4 - TMB	**0.93**	0.19	0.02	0.13	0.13	0.05	0.06	0.07	-0.03	0.17	0.06	0.99
decane	**0.89**	0.25	0.00	0.22	0.26	0.05	-0.01	0.05	0.01	0.00	0.01	0.98
undecane	**0.81**	**0.35**	-0.08	0.27	0.21	0.15	-0.07	0.08	0.04	-0.12	-0.10	0.96
Biogenic VOCs												
α-pinene	0.00	-0.08	**0.98**	-0.07	0.05	0.03	0.01	0.01	-0.07	0.02	0.01	0.98
ß-pinene	0.01	-0.08	**0.98**	-0.08	0.05	0.05	0.01	0.03	-0.06	0.01	0.02	0.98
limonene	0.11	-0.02	**0.92**	-0.02	0.14	0.09	0.21	0.02	-0.10	0.02	-0.03	0.95
Combustion tracers												
NO_y	0.23	0.20	-0.27	**0.82**	0.21	-0.06	-0.07	0.03	0.10	-0.10	0.01	0.92
rBC	0.22	0.15	0.05	**0.80**	**0.43**	0.15	0.10	0.05	0.09	0.07	0.00	0.95
CO	**0.40**	0.09	0.08	0.20	0.09	0.22	0.08	0.06	0.03	**0.83**	-0.02	0.97
CO_2	0.12	-0.07	**0.50**	0.08	-0.03	0.09	**0.75**	0.28	-0.12	0.03	-0.08	0.95
Aerosol species												
pPAH	0.06	-0.10	-0.06	**0.93**	-0.07	-0.06	0.07	0.03	0.15	0.13	-0.05	0.94
PM_{10-1}	0.19	0.16	0.08	0.16	0.13	0.08	0.18	**0.91**	-0.03	0.05	0.07	0.99
PM_1	0.24	**0.44**	0.00	0.17	**0.70**	**0.31**	-0.06	0.11	-0.04	0.07	-0.14	0.90
$NH_4^+{}_{(p)}$	0.28	**0.84**	0.02	0.12	**0.32**	0.22	0.06	0.07	-0.04	0.14	-0.04	0.97
$SO_4^{2-}{}_{(p)}$	0.29	**0.84**	0.03	0.12	**0.33**	0.19	0.06	0.06	-0.05	0.12	-0.05	0.97
$NO_3^-{}_{(p)}$	**0.30**	**0.44**	0.09	0.23	**0.59**	**0.36**	0.08	0.15	-0.13	0.02	0.24	0.92
HOA	**0.37**	0.18	0.02	**0.77**	0.25	0.10	0.10	0.18	-0.08	0.13	0.14	0.93
LO-OOA	**0.37**	**0.40**	0.12	0.16	**0.66**	**0.31**	0.03	0.12	-0.06	0.00	0.27	0.97
MO-OOA	0.10	0.15	0.09	0.00	0.10	**0.92**	0.05	0.07	0.10	0.16	-0.03	0.95
Sulfur												
TS	0.27	**0.90**	-0.20	0.03	0.04	-0.04	-0.09	0.07	0.00	-0.04	0.18	0.98
SO_2	0.09	**0.96**	-0.19	0.02	-0.03	-0.01	-0.08	0.03	-0.02	-0.03	0.00	0.98
TRS	**0.65**	0.14	-0.10	0.05	0.23	-0.08	-0.07	0.17	0.06	-0.04	**0.63**	0.95
Other												
IVOCs	**0.34**	-0.01	0.12	**0.33**	**0.80**	-0.23	-0.02	0.02	0.02	0.06	0.06	0.94
NH_3	-0.03	-0.08	-0.22	0.21	-0.04	0.09	-0.07	-0.03	**0.93**	0.02	0.02	0.99
O_x	0.07	**0.36**	**-0.62**	0.01	0.27	**0.33**	**-0.41**	-0.07	-0.03	-0.14	0.12	0.91
CH_4	**0.60**	0.00	0.14	**0.42**	0.10	0.08	**0.57**	0.08	-0.04	0.13	0.16	0.94
Eigenvalues	5.85	4.30	3.71	3.51	2.78	1.58	1.24	1.09	1.01	0.94	0.75	
% of variance	20.90	15.34	13.25	12.52	9.92	5.65	4.43	3.88	3.59	3.37	2.66	
Cumulative variance	20.90	36.24	49.49	62.02	71.94	77.59	82.03	85.90	89.50	92.87	95.53	

3.3 Bivariate polar plots

Bivariate polar plots were generated for all components and their dominant, associated variables and are shown in the supplemental material section (Figs. S2-S11). Winds were predominantly from the SW but were also observed often from the S and N. Fig. 5A shows the plot for IVOCs. The highest concentrations were observed when the local wind direction was from the NE, where several facilities including the Aurora North, Musket River and Jackpine mines and large swaths of disturbed and cleared land are located in close proximity to each other (Table 1 and Fig. 1). The second highest IVOC signal intensity was observed when local wind direction was from the SSE.

The bivariate polar plots of the 3 components associated with IVOCs are shown in Fig. 5B-D. These components are associated with winds from the NE, E, SE and S at low to moderate speeds (1-3 m s^{-1}). Component 5 (Fig. 5B) was the most strongly correlated with IVOCs and shows the most spatial overlap with the distribution of the IVOC source; however, the intensities differ owing to the association of component 5 with other variables such rBC and LO-OOA.

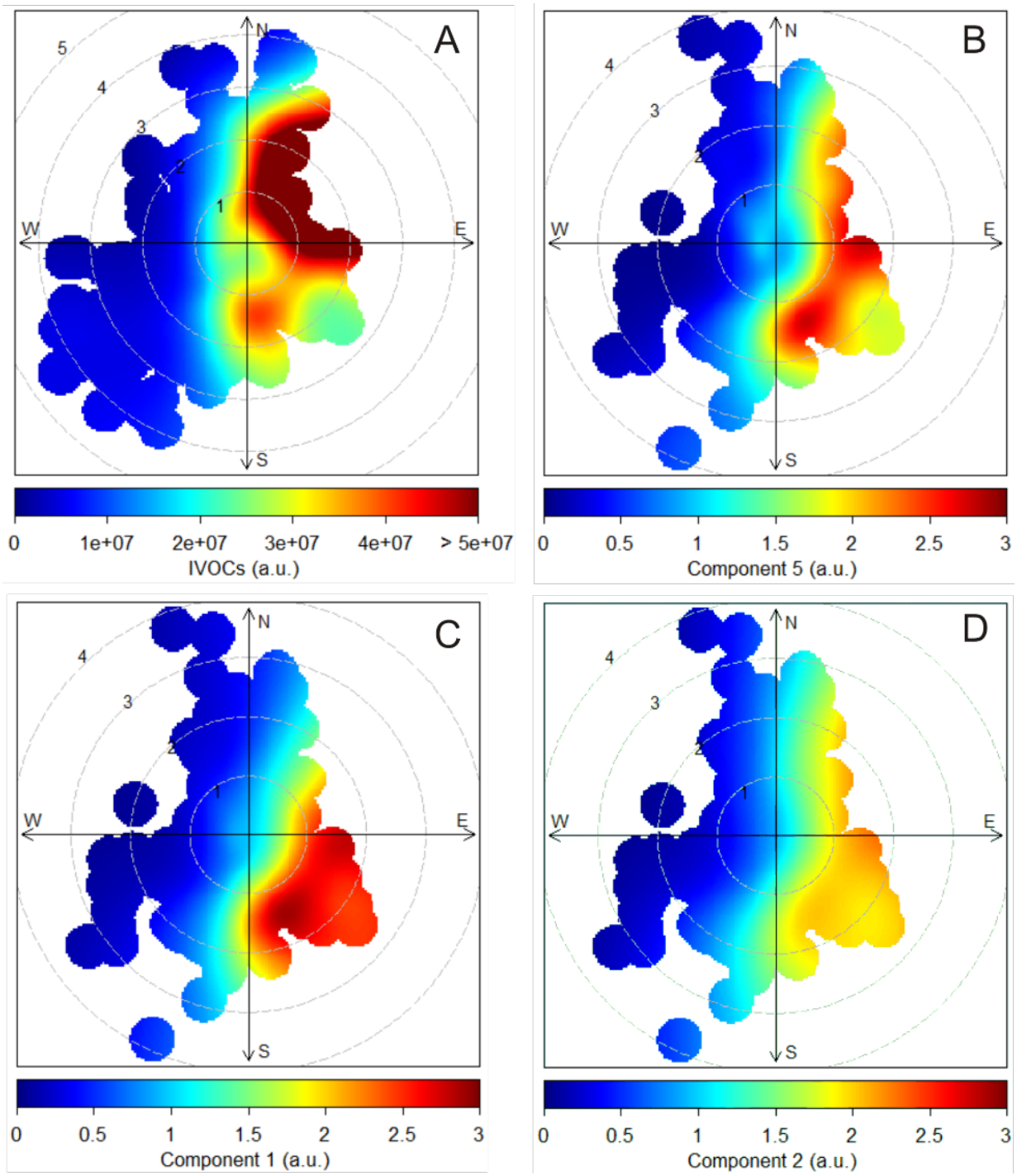


Figure 5. Bivariate polar plots related to IVOCs: **(A)** IVOCs from the complete data set. **(B)** Component 5 extracted from the main PCA (Table 5). **(C)** Component 1 extracted from the main PCA. **(D)** Component 2 extracted from the main PCA. Wind direction is binned into 10° intervals and wind direction into 30° intervals. The polar axis indicates wind speed (m s^{-1}). a.u. = arbitrary units.

4. Discussion

This work has added to the relatively few data sets of pollutants in the Athabasca oil sands region, one of the largest emitters of airborne pollutants in Canada (NPRI, 2013), that are available in the open literature. Earlier source apportionment studies in the region investigated ground level O_3 and $PM_{2.5}$ (Cho et al., 2012), examined VOCs (Bari and Kindzierski, 2018; Bari et al., 2016) and $PM_{2.5}$ (Bari and Kindzierski, 2017; Landis et al., 2017) impacting the nearby communities of Fort McKay and Fort McMurray, or investigated pollutants such as PAHs as they affect sediments (Jautzy et al., 2013) or lichens (Landis et al., 2012). The measurement suite in this work encompassed a larger variety of collocated analytical instruments closer to oil sands mining operations than these earlier studies and included a first, direct observation of airborne IVOCs, that is unique to this area and we have not observed elsewhere where we have made GC-ITMS measurements, i.e., in Calgary and on Vancouver Island (Tokarek et al., 2017a).

The main objective of this work was to elucidate the origin of the IVOC signature observed at the AMS 13 ground site downwind from the AB oil sands mining operations (Fig. 2) through a PCA. The optimum solution identified 10 components, of which three were associated with the IVOC signature: 1, 2, and 5 (Table 5). Tentative assignments of these components to source types in the oil sands are given in Table 6 and are discussed below.

Emission inventories show that the facilities that process the mined bitumen are by far the largest anthropogenic point sources in the oil sands region (NPRI, 2013), consistent with recent aircraft measurements (Baray et al., 2018; Howell et al., 2014; Li et al., 2017; Simpson et al., 2010) which have shown substantial emissions of NO_y, SO_2, CO, VOCs, CO_2, and CH_4, from these facilities and associated mining activities. No single component correlates with all of these variables, suggesting that the PCA is able to distinguish between source types within the facilities such as tailings ponds (component 1), stack

emissions (component 4), and mining (component 2).

Close-up overflights (Howell et al., 2014; Li et al., 2017; Baray et al., 2018) were able to spatially resolve various oil sands facility emission sources (i.e., tailings ponds from upgraders, fluid coking reactors, hydrocrackers and –treaters); the PCA presented in this manuscript is not expected to do this in all cases because some emissions would have frequently merged into a single plume by the time of observation at AMS 13; unless their emissions vary considerably in time, these sources could be interpreted as originating from a single source in the PCA.

The discussion below focuses on components that are associated with IVOCs (section 4.1), followed by those that are not (section 4.2). The PCAthat included 6 secondary products is discussed in section 4.3. Components which are not associated with IVOCs and have only tentatively been identified (i.e., components 6 – 10) are discussed in the S.I.

4.1 Components associated with IVOCs

4.1.1. Component 1: Tailings ponds (wet tailings)

Component 1 is strongly associated with anthropogenic VOCs ($r > 0.87$) and weakly with TRS ($r = 0.59$), and CH_4 ($r = 0.59$). These pollutants originate from tailings ponds (Small et al., 2015), though it is unclear from this analysis how large a source tailings ponds are compared to fugitive emissions of these pollutants from the nearby processing (e.g., bitumen separation and mining) facilities.

Tailings ponds cover large areas of land and are used to slowly (on a time scale of years to decades) separate solid components, or tailings, from water used in bitumen extraction. Residual bitumen often floats to the top of the settling basins. Most tailings ponds are "wet" (as they contain residual naphtha that is used as a diluent during the transfer of tailings to the ponds) and emit VOCs, CH_4, and CO_2 (Small et al., 2015). The presence of o-xylene, TMB and the n-alkanes in component 1 is consistent with the

fugitive release of VOCs from residual naphtha, which contains these compounds (Siddique et al., 2008; Siddique et al., 2011; Small et al., 2015). Furthermore, the observation of TRS and CH_4 from this source is consistent with the presence of anaerobic sulfur reducing bacteria and methanogens within the ponds, which degrade not only the residual bitumen (Holowenko et al., 2000; Percy, 2013; Quagraine et al., 2005) but also the various components of naphtha (Shahimin and Siddique, 2017; Small et al., 2015). Overall, tailings ponds emissions explain much of the TRS and CH_4 concentration variability in this data set (Table 5) and in a recent aircraft study (Baray et al., 2018).

While component 1 correlates with CH_4 ($r = 0.59$), it does not correlate with CO_2 ($r = 0.09$). Emissions of CH_4 from tailings ponds due to methanogenic bacterial activity are well-documented (Small et al., 2015; Yeh et al., 2010) and hence the correlation with CH_4 is not unexpected. On the other hand, the lack of correlation with CO_2 seems inconsistent with emission inventories that generally present tailings ponds as large CO_2 sources (Small et al., 2015). One plausible explanation is that tailings ponds are a relatively small CO_2 source overall in the region and that other, larger CO_2 sources and sinks (such as photosynthesis and respiration by the vegetation surrounding the site) dominate the variance impacting the PCA results. It may also indicate that, at least on aggregate and for the particular ponds detected in this work, the emissions are in a regime where the release of CH_4 dominates over CO_2, i.e., the ponds have, perhaps, become more anoxic than believed to be the case in previous studies and hence emit more CH_4 (Holowenko et al., 2000). For example, Small et al. (2015) showed that older tailings ponds (those without the addition of fresh froth or thickening treatments) tended to emit more CH_4, while newer ponds are associated with higher VOC emissions. It is likely that component 1 is dominated by the nearest pond (the Mildred Lake settling basin, 6 – 11 km SSE of AMS 13) and other tailings in the SE where the majority of air samples originated from. The Mildred Lake settling basin is one of the oldest in the region and is still actively being used; the correlation with CH_4 and VOC emissions is hence expected.

Component 1 is also associated with NO_y, rBC, CO, and HOA, though these correlations are relatively

513 modest (r = 0.27, 0.30, 0.41, and 0.40, respectively). These species typically originate from combustion
514 sources, such as generators, motor vehicles, including diesel powered engines powering generators or
515 pumps; it is not obvious if and to what extent these are operated on or near tailings ponds, though.
516 Satellite observations have shown elevated concentrations of NO_2 above on-site upgrader facilities,
517 likely a result of emissions from extraction and transport sources (McLinden et al., 2012). In addition,
518 one of the major highways of the region is located adjacent to the Mildred Lake settling basin and other
519 major ponds in the region; highway traffic emissions (of CO, NO_y, rBC, and HOA) may hence also be
520 partially included in component 1.

521 The bivariate polar plot shows that component 1 was observed when local wind speeds were from the
522 SE and E of the measurement site (Fig. 5C), which is consistent with the notion that the Mildred Lake
523 settling basin and emissions along Highway 63 and, potentially, more distant facilities are sources
524 contributing to this component.

525 Component 1 is associated with the IVOC signature, though to a lesser degree than components 2 and 5.
526 The association of the IVOC signal with component 1 is slightly poorer (r = 0.31) than the association
527 with component 2 (r = 0.39), but significantly poorer than component 5 (r = 0.74). One possible
528 explanation for the association of IVOCs with tailings ponds vapor is the presence of bitumen in the
529 ponds that was not separated from the sand during the separation stage (Holowenko et al., 2000). This
530 semi-processed bitumen would be expected to emit the same IVOC vapors to those that were observed
531 in the lab (Fig. 2). Tailings ponds contain anywhere from 0.5% - 5% residual bitumen by weight
532 (Chalaturnyk et al., 2002; Holowenko et al., 2000; Penner and Foght, 2010). As illustrated in Fig. 4A,
533 some of this material floats on the ponds' surfaces, where IVOCs can partition to the air. Emission of
534 IVOCs from bitumen floating on tailings ponds would be a function of many variables (e.g., diluent
535 composition, extraction methodology, settling rate, temperature, etc.) and is thus not expected to be as
536 persistent as CH_4 partitioning from the ponds to the above air or from exposed bitumen on the mine

surface, leading to a lower overall correlation.

Component 1 is also weakly associated with the less oxidized oxygenated organic aerosol factor, LO-OOA (r = 0.45). Liggio et al. (2016) found that the observed secondary organic aerosol is dominated by an OOA factor whose mass spectrum was similar to those of aerosols formed from oxidized bitumen vapours. The organic aerosol budget in this study was also dominated by an OOA factor, the LO-OOA (Lee et al., 2018). The association of LO-OOA with component 1 is thus consistent with its association with IVOCs.

4.1.2. Component 2: Mine fleet and vehicle emissions

Component 2 strongly correlates with NO_y (r = 0.82), rBC (r = 0.77), pPAH (r = 0.94), and HOA (r = 0.74), which suggests a combustion source such as diesel engines. In the AB oil sands, there is a sizeable off-road mining truck fleet consisting of heavy aggregate haulers. In addition, there are diesel engine sources associated with generators, pumps and land moving equipment, i.e., graders, dozers, hydraulic excavators, and electric rope shovels (Watson et al., 2013; Wang et al., 2016). Most of these non-road applications have been exempt from highway fuel taxes, on-road fuel formulation requirements and after-engine exhaust treatment (Watson et al., 2013). Emissions from the hauler fleet and the stationary sources would fit the profile of component 2. Other diesel engines operated in the region include a commuter bus fleet, pickup and delivery trucks, tractor-trailers, and privately owned diesel powered automobiles used to commute from the work sites to the major residential areas around Fort McMurray, whose emissions are likely captured by component 2 as well, though the magnitude of these relative to the mining truck fleet is not known. Consistent with component 2 being associated with an anthropogenic source is its poor correlation with undecane (r =0.27), likely arising from fugitive fuel emissions.

The bivariate polar plot (Fig. 5D) for component 2 and NO_y in particular (Fig. S-4A) match the location of

Highway 63 which crosses the river to the SE of AMS 13 and bends to the E and is indicative of a line source. At the same time, some of the largest mining operations in the region, the Susan Lake Gravel Pit, Aurora North, Muskeg river, and Millennium mines are located to the NE and SE of AMS 13 as well. NO_y, rBC, and HOA (Fig. S-4A, B and D) all appear to have dominating point sources to the S and E when wind speeds are 1-2 m s^{-1}. These directions are the same as the Fort McKay industrial park to the E and the Syncrude Mildred Lake facility parking lot to the S which would have a higher concentration of vehicles emitting these pollutants in a smaller area, whose emissions would be in addition to those from industrial activities.

Component 2 is associated with the IVOCs signature and CH_4 (both r = 0.39). The mining activities bring bitumen to the surface; similar to what we had observed in lab experiments (Fig. 2, black trace), the surface exposure of bitumen during mining and on-site processing is expected to be associated with fugitive emissions of CH_4 (Johnson et al., 2016) and IVOCs.

Fine-fraction particle-surface bound PAHs (pPAH) are associated strongly with component 2, but no other components. Measurements of individual PAHs in snow and moss downwind from the oil sands facilities have identified multiple sources of PAHs in the Athabasca oil sands, which include wind-blown petroleum coke dust (also referred to as petcoke for short), a carbonaceous residual product from the upgrading of crude petroleum that is stockpiled on mine sites, and emissions from fine tailings, oil sands ore, and naturally exposed bitumen (Zhang et al., 2016; Jautzy et al., 2015; Parajulee and Wania, 2014). Given this diversity of known sources, the associations of PAHs with only a single component is surprising, though indicates that emissions from the mining fleet (which would include diesel and, perhaps, wind-blown emissions from petcoke that is being transported) gave rise to most of the variability in surface-bound PAH concentrations in this data set. The petcoke emissions identified in the studies mentioned above are likely mainly associated with larger, supermicron sized particles, whose PAH content would not be detected by the pPAH measurement in this data set.

Component 2 is not associated with LO-OOA (r = 0.11), even though IVOCs are associated with this component. This feature may indicate that the IVOCs emitted in component 2 are qualitatively different from those emitted by components 1 and 5, in that they are less likely to yield organic aerosol on the time scale of transport from emission to observation. One reason for the difference could be that the bitumen that is transported by the mining fleet is relatively freshly exposed, whereas the IVOCs released from~~by bitumen in~~ tailings ponds ~~has been processed by microbes and that released by~~ or from mine faces (component 5) may have been ~~photochemically~~ oxidized to a greater extent and hence more prone to rapid aerosol formation.

There is no association of component 2 with CO_2 (r = -0.08). This is somewhat unexpected as the trucks are expected to release CO_2 (Wang et al., 2016) but could be due to significantly larger CO_2 sources in the area dominating the observed CO_2 variability at AMS 13 (e.g., components 3 and 6). Furthermore, one would expect an association of non-road mining truck emissions with aromatics and alkanes. Component 2 exhibited only poor correlations with decane (r = 0.22) and undecane (r = 0.27) and no correlation with o-xylene (r = 0.08), suggesting that other components (i.e., component 1) explained most of the variability of their concentrations at this site.

4.1.3. Component 5: Surface-exposed bitumen and hot-water bitumen extraction

Component 5 correlates more strongly with the IVOCs (r = 0.74) than with any other component and correlates strongly with LO-OOA (r = 0.72), weakly with rBC (r = 0.44), and poorly with HOA (r = 0.25), NO_y (r = 0.22), decane (r = 0.23), undecane (r = 0.20), and TRS (r = 0.26). We interpret this profile as emissions from surface-exposed bitumen which outgases IVOCs.

One possibility is that these emissions occur on mine faces, where previously unexposed bitumen is brought to the surface as a result of mining. Only a relatively small portion of the mine faces is actively

607 mined; those parts give rise to rBC and NO_y emissions from combustion engines in heavy haulers or
608 generators powering equipment. The poor association of component 5 with TRS could be due to sulfur
609 reducing bacteria found on the surface of bitumen. However, most of the variability of TRS at AMS 13 is
610 attributed to composite or “dry” tailings ponds given their more conducive environment to microbial
611 activity.

612 Component 5 does not correlate with CO_2 ($r = -0.03$) or with CH_4 ($r = 0.12$), which is somewhat at odds
613 with the notion of mine faces as the main source of IVOCs. The mine faces give rise to substantial
614 fugitive emissions of CO_2 and CH_4 (Johnson et al., 2016) – these emissions are likely captured by
615 component 6 in this analysis (see S.I.). It is unclear to what extent these greenhouse gases are released
616 relatively quickly from "hot spots" (i.e., from a small number of locations) through surface cracks and
617 fissures or by slow release from new material that is exposed and then releases greenhouse gases
618 during material handling, transport and processing (Johnson et al., 2016). IVOCs from surface-exposed
619 bitumen are likely released by the latter mechanism and are temperature-dependent. If the mine faces
620 are indeed the main IVOC source, the analysis results presented here suggest that the IVOCs emissions
621 from surface-exposed bitumen on mine faces are decoupled from CH_4 emissions in time and appear as a
622 distinct component and hence corroborate the "hot spots" or fast release hypothesis, though clearly,
623 more work is needed to characterize greenhouse gas emissions from oil sands mine faces.

624 The association of IVOCs with component 5 may also be a result of fugitive emissions during the hot
625 water-based extraction of bitumen sand slurries during the separation phase of bitumen treatment.
626 Generally, bitumen is extracted in a weak alkaline environment by aeration of the solution to optimize
627 the separation of sand and bitumen (Masliyah et al., 2004). Unrecovered bitumen and naphtha then end
628 up in tailings. The recovered bitumen and naphtha are moved to upgrader facilities where they undergo
629 further treatment (such as coking or hydrotreatment). The magnitude of fugitive emissions during these
630 downstream extraction processes could be large, considering the bitumen is heated and actively

aerated. Future work should investigate IVOC fluxes near extraction plants and on mine faces.

Component 5 correlates strongly with LO-OOA (r = 0.72), which is likely generated in part by photochemical aging of IVOCs. A back-of-the-envelope calculation using a k_{OH} of 1.8×10^{-11} cm^3 $molecule^{-1}$ s^{-1} based on that used diesel exhaust IVOCs (Zhao et al., 2014) and an estimated mid-day OH concentration of 7×10^6 molecules cm^{-3} (Liggio et al., 2016) gives a first-order lifetime of 130 min with respect to IVOC oxidation by OH during daytime. The photochemical age, estimated using relative concentrations of 124-TMB and n-decane and the method described by Borbon et al. (2013), during daytime was 1.0±0.4 hr; assuming similar photochemical ages, we estimate that between 25% and 50% of the emitted IVOC is (potentially) oxidized during daytime (see S. I.).. This oxidation will contribute SOA growth (Kroll et al., 2011). Hence, we expect some formation and growth of organic aerosol associated with component 5.

Finally, it is conceivable that a "natural" background of IVOCs exists in the region (since bitumen can be found at or near the surface in many parts of the region); such a natural background would also be included in component 5. However, this "natural" bitumen would have been exposed at the surface for geological time scales and, unlike unexposed, buried bitumen, likely would have lost most of its volatile content over that period. Furthermore, the mine faces occupy large swaths of land in the region (as evident from satellite imagery). Thus, the IVOCs emissions are more likely due to anthropogenic activity than due to a natural phenomenon.

4.2. Components not associated with IVOCs

4.2.1. Component 3: Biogenic emissions and respiration

Component 3 is strongly correlated with the monoterpenes α-pinene (r = 0.98), ß-pinene (r = 0.98) and

limonene (r = 0.92) and is hence identified as a biogenic emissions source. This component is also weakly associated with CO_2 (r = 0.48).

At AMS 13, CO_2 and the monoterpenes exhibit a very similar diurnal cycle: they are present in higher concentrations during the night than during the day (Fig. 3) due to a decrease in the boundary layer height (BLH) at night coupled with plant respiration of CO_2 and non-photochemical emission of monoterpenes (Fares et al., 2013; Guenther et al., 2012). During the day, mixing ratios of CO_2 are lower due to plant uptake and photosynthesis, and mixing ratios of terpenes are lower due to higher mixing heights and vertical entrainment and due to oxidation by O_3 and OH (Fuentes et al., 1996). Hence, the PCA gives a *positive* correlation of monoterpenes with CO_2 even though the physical processes, photosynthesis and respiration, work in opposite direction.

The bivariate polar plots (Fig. S-5A-C) show that the monoterpenes and CO_2 were observed in highest concentrations when the wind speeds were low (< 1 m s^{-1}), consistent with formation of a stable nocturnal boundary layer.

To corroborate this interpretation, the PCA was repeated with BLH estimated by a light detection and ranging (LIDAR) instrument (Strawbridge et al., in prep.) added as a variable (Table S-9 in the S.I.). Since BLH is not "emitted" by any source, it appears as a single variable component (r = 0.90). The only other component that BLH (anti)correlates with is the biogenic component 3 (r = -0.35).

The dominant monoterpene species observed was α-pinene, followed by ß-pinene and limonene, though occasionally there was twice as much ß-pinene than α-pinene in the sampled air. Some variability of this ratio is expected since emission factors vary considerably between tree species (Geron et al., 2000) which are not homogeneously distributed throughout the region (e.g., Fig. S1 of Rooney et al. (2012)).

Simpson et al. (2010) observed enhancements of α-pinene and, to a greater extent, β-pinene over the

oil sands (up to 217 pptv and 610 pptv) compared to background levels of 20±7 and 84±24 pptv, respectively, during mid-day overflights (which occurred between 11:00 and 13:00 local time). Similar enhancements were also reported by Li et al. (2017) who observed emissions of biogenic hydrocarbons in the four facilities sampled, three of which showed a higher β- than α-pinene concentration. The PCA (Table 5) showed no significant correlation of α- and β-pinene with any of the anthropogenic components, which implies that the biogenic source strength is simply too large for any anthropogenic emissions of terpenes to be picked up in the analysis, especially considering that terpenes are relatively short-lived.

The biogenic source shows poor anticorrelations with NO_y (r = -0.26) and NH_3 (r = -0.24). Many NO_y species (i.e., NO_2, HONO, peroxycarboxylic nitric anhydrides or PAN, and HNO_3) deposit to the forest canopy (Hsu et al., 2016; Min et al., 2014; Fenn et al., 2015); at night, when mixing heights are lower, their concentrations are expected to decrease faster than during the day and are thus out of phase with the CO_2 and terpene concentrations. The poor anticorrelation with NH_3 likely arises because the NH_3 emissions from plants are mainly stomatal and scale with temperature and are hence larger during the day than at night, anticorrelated with the terpene source (Whaley et al., 2018).

4.2.2 Component 4: Upgrader emissions

Component 4 is strongly correlated with SO_2 (r = 0.97) and total sulfur (r = 0.93). By far the largest source of SO_2 in the region are upgrader facilities, which emit as much as 6×10^7 kg annually according to emission inventories (ECCC, 2013). Significant SO_2 emissions from upgrader facilities have recently been confirmed by aircraft studies (Simpson et al., 2010; Howell et al., 2014; Liggio et al., 2016). Component 4 is also poorly correlated with NO_y (r = 0.21) but not with rBC (r = 0.05), consistent with a non-sooty (i.e., lean) combustion source such as upgrader stacks. Strong enhancements in SO_2 were only observed intermittently as "spikes", which is expected when sampling emissions from relatively few and discrete

699 point sources.

700 Component 4 isnot associated with CO_2 (r = -0.12), even though inventories indicate that the upgrading
701 facilities are the largest CO_2 source in the region (Furimsky, 2003; Englander et al., 2013; Yeh et al.,
702 2010). In this data set, the lack of correlation of component 4 with CO_2 (and to some extent with PM_{10-1}
703 as well) likely arises mainly from a sampling bias as stack emissions were only observed during daytime,
704 likely due to diurnal variability of the atmospheric boundary layer structure as explained below.

705 Most of the variability in CO_2 concentration at AMS 13 is due to surface-based sources that originate
706 from large areas, especially biogenic processes (photosynthesis during the day and respiration at night,
707 component 3) and anthropogenic surface sources such as those captured by component 6 (section
708 4.2.3). Other anthropogenic pollutants, such as SO_2, NO_y, and CH_4, are not subject to large biogenically
709 driven processes and are less affected than CO_2.

710 In contrast to surface sources, emissions from the > 100 m tall stacks are comparatively undersampled
711 and observed mainly during daytime, when vertical mixing brings elevated plumes to the surface, yet
712 CO_2 concentrations are generally much lower than during the night due to uptake by vegetation. At
713 night, pollutants emitted from stacks are injected above the likely very shallow nocturnal surface layer
714 and were hence not observed at the surface. Vertical profile measurements of SO_2 stack plumes by a
715 Pandora spectral sun photometer at Fort McKay during daytime have shown considerable vertical
716 gradients and only occasional transport of SO_2 all the way to the surface (Fioletov et al., 2016).

717 The association of component 4 with CO_2 is negative because the stack emission source is observed only
718 during the day when the large biogenic sink dominates and effectively masks the relatively small
719 increase due to anthropogenic CO_2. In contrast, background concentrations of SO_2 are comparatively
720 low, and the increase in SO_2 concentrations is readily picked up the PCA.

721 It would be interesting to conduct a future study in winter when biogenic activities decrease; a

wintertime PCA of surface measurements might be able to associate CO_2 enhancements with upgraders, though boundary layer mixing heights would decrease as well, which would make a PCA using surface data even more challenging.

Component 4 does not correlate with PM_{10-1} volume ($r = 0.09$). It is clear that the emitted SO_2 will contribute to secondary aerosol formation downwind, such that a correlation of stack emissions with PM_{10-1} volume might be expected. However, these secondary contributions will likely mostly be in the submicron aerosol fraction, which adds relatively little to PM_{10-1} volume. Further, PM_{10-1} volume is dominated by coarse particles from other primary sources, mostly wind-blown emission of sand from the mine surfaces, roadways and, perhaps, bioaerosol (component 7, see S.I.). These effects make PM_{10-1} volume from stacks appear comparatively small, such that the variability of the larger, surface-based sources likely masks the contribution of stacks emissions to PM_{10-1} variability.

The bivariate polar plot of component 4 (Fig. S-6D) shows that the largest magnitudes were observed when local winds were from the SE. The corresponding plot of SO_2 (Fig. S-6A) reveals two more distinct sources: a larger one from the E and a smaller one from the SSE. However, only two facilities (Sunrise and Firebag) are located to the E at relatively large distances of 37 km and 47 km respectively. The largest known upgraders and SO_2 sources in the area (i.e., upgraders located at the Mildred Lake and Suncor base plants) are located to the S and SE of AMS 13. Considering that the stack emissions are only observed intermittently, we speculate that there exists a mesoscale transport pattern in the Athabasca river valley which channel emissions, such that the local wind direction and speed may be misleading as to the true location of these sources. For more extensive data sets, such phenomena may very well average out but perhaps did not in this case.

4.3. Extended PCA with added secondary variables

744 The extended analysis (Table 7) qualitatively preserves the structure (with the exception of an added
745 "Aged" component, # 6) of the original 10-component solution but allows an assessment of which
746 components most result in formation of secondary products such as SOA, which has implications for
747 health (Bernstein, 2004) and climate (Charlson et al., 1992). Secondary products vary considerably as a
748 function of air mass chemical age (which depends, amongst other components, on time of day and
749 synoptic conditions, including wind speed) and are hence expected to add considerable noise and
750 scatter to the results leading to lower correlations. On the other hand, the distance between the
751 measurement site and sources is fixed, such that this variability should average out over time. This
752 indeed appears to have happened in this data set in spite of the relatively low sample size.

753 The analysis indicates that the component with the strongest IVOC source (Component 5) also has the
754 highest association with PM_1 (r = 0.70; Table 7). Aircraft measurements combined with a modelling
755 study have required a group of IVOC hydrocarbons to explain the significant SOA formation and growth
756 downwind of the oil sands region (Liggio et al., 2016). The association of IVOCs with PM_1 volume is
757 consistent with the hypothesis that oxidation of IVOCs observed at AMS 13 leads to SOA generation and
758 appears to have a significant impact on the variation in PM_1 mass.

759 The relatively short distance to sources and young photochemical age suggests that IVOCs would
760 experience a relatively small number of oxidation steps. Consistent with this interpretation, a correlation
761 with the more-oxidized MO-OOA is not observed in component 5 (r = 0.10; Table 7). However,
762 component 6, which is (poorly) anticorrelated with IVOCs (r = -0.23), is strongly correlated with MO-OOA
763 (r = 0.92), consistent with the notion that this component is more photochemically processed and that
764 IVOCs contribute to this SOA AMS factor.

765 The second component influencing PM_1 is that from stack emissions (Component 4 in the primary PCA;
766 Component 2 in the secondary PCA) (Tables 5 and 7). It is well established that the oxidation of SO_2 to

767 sulfate will lead to formation of fine particulate matter. This apparently occurs, at least partially, on the
768 time scale between the point of emission and the AMS 13 site (assuming a wind speed of 3 m/s and a
769 distance of 11 km, the transit time is 1 hour), though some fraction of SO_4^{2-}(p) is likely directly emitted.

770

771 **5. Summary and conclusions**

772 A PCA was applied to continuous measurements of 22 primary pollutant tracers at the AMS 13 ground
773 site in the Athabasca oil sands during the 2013 JOSM intensive study to elucidate the origins of airborne
774 analytically unresolved hydrocarbons that were observed by GC-ITMS. The analysis identified 10
775 components. Three components correlated with the IVOC signature and were tentatively assigned to
776 mine faces and, potentially, hot-water bitumen extraction facilities, the mine hauler fleet, and wet
777 tailings ponds emissions. All three are anthropogenic activities that involve the handling of raw bitumen,
778 i.e., the unearthing, mining and transport of crude bitumen, and the disposal of processed material that
779 contains residual bitumen in wet tailings ponds. The PCA results are consistent with our previous
780 interpretation that the unresolved hydrocarbons originate from bitumen based on the similarity of the
781 chromatograms with those obtained in a head space vapor analysis of ground-up bitumen in the
782 laboratory.

783 Liggio et al. (2016) showed that these hydrocarbons constitute a group of IVOCs in the saturation vapor
784 concentration (C^*) range 10^5 μg m^{-3}< C^* < 10^7 μg m^{-3} that contribute significantly to secondary organic
785 aerosol formation and growth downwind of the oil sands facilities. The correlation of LO-OOA with two
786 of the three IVOC components in the main PCA and with PM_1 in the extended analysis is consistent with
787 the high SOA formation potential of IVOCs and suggests that further differentiation may be needed and
788 stresses the need for IVOCs to be routinely monitored. In particular, direct measurements of emissions
789 throughout the processing of raw bitumen are needed to pinpoint source contributions more accurately

and aid in the development of potential mitigation strategies.

The PCA in this study suffered from several limitations. For instance, PCA does not provide insight into emission factors of individual facilities, though it does capture what conditions change ambient concentrations the most. Further, the receptor nature of PCA did not always discern between large source areas that may have many individual point sources coming together at the point of observation. For example, component 1 contains an obvious tailings pond signature because of its high correlation with anthropogenic VOCs, methane and TRS, but also includes several combustion sources, making interpretation of this IVOC source location more challenging. A longer continuous data set with a greater number of variables would have perhaps been able to resolve these different sources, including the various tailings ponds, of which there are 19 in the region, all with slightly different emission profiles (Small et al., 2015) .

Another limitation is the bias of this (and most) ground site data set towards surface-based emissions and the undersampling of stack emissions. Facility stacks were only observed in the daytime because at night the mixing height is so low that the stacks are emitting directly into the residual layer. These emissions could be quantified using aircraft based platforms (Howell et al., 2014; Li et al., 2017; Baray et al., 2018). The PCA struggled most with the allocation of greenhouse gases. Mixing ratios of CO_2, in particular, were difficult to reconcile in this analysis due to a high background and large attenuation by biogenic activity and boundary layer meteorology. Forests greatly affected CO_2 levels in the region because it is taken up during the day when plants are photosynthetically active and emitted at night when plants undergo cellular respiration. This CO_2 source and sink appears to dominate the PCA, effectively masking relatively small emissions from tailings ponds, facilities, and tail pipes in particular from the mine hauling fleet.

Finally, there is a need for improved monitoring methods for IVOCs. For instance, future studies should

focus on characterizing the VOCs in the above mentioned volatility range using a greater mass and time resolution instrument, such as a time-of-flight mass spectrometer (TOF-MS) or higher resolution separation methods (e.g., multi-dimensional gas chromatography), and also include measurement of speciated aerosol organic composition by, for example, thermal desorption aerosol GC (TAG) analysis (Williams et al., 2006). Future studies should also investigate how IVOC volatility distributions vary with source type and chemical age.

Acknowledgments

Funding for this study was provided by Environment and Climate Change Canada and the Canada-Alberta Oil Sands Monitoring program. The GC-ITMS used in this work was purchased using funds provided by the Canada Foundation for Innovation and matching funds by the Alberta government. TWT, JAH, DKB, FVA and GRW acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Training Experience Program (CREATE) program Integrating Atmospheric Chemistry and Physics from Earth to Space (IACPES).

826 **6. References**

827 Alberta, G. o.: Government of Alberta, oil sands information portal: http://osip.alberta.ca, access: 23-
828 FEB-2017, 2017.
829 Allen, E. W.: Process water treatment in Canada's oil sands industry: II. A review of emerging
830 technologies, J. Environ. Eng. Sci., 7, 499-524, 10.1139/s08-020, 2008.
831 Baray, S., Darlington, A., Gordon, M., Hayden, K. L., Leithead, A., Li, S. M., Liu, P. S. K., Mittermeier, R. L.,
832 Moussa, S. G., O'Brien, J., Staebler, R., Wolde, M., Worthy, D., and McLaren, R.: Quantification of
833 methane sources in the Athabasca Oil Sands Region of Alberta by aircraft mass balance, Atmos.
834 Chem. Phys., 18, 7361-7378, 10.5194/acp-18-7361-2018, 2018.
835 Bari, M., and Kindzierski, W. B.: Fifteen-year trends in criteria air pollutants in oil sands communities of
836 Alberta, Canada, Environ. Int., 74, 200-208, 10.1016/j.envint.2014.10.009, 2015.
837 Bari, M. A., Kindzierski, W. B., and Spink, D.: Twelve-year trends in ambient concentrations of volatile
838 organic compounds in a community of the Alberta Oil Sands Region, Canada, Environ. Int., 91, 40-50,
839 10.1016/j.envint.2016.02.015, 2016.
840 Bari, M. A., and Kindzierski, W. B.: Ambient fine particulate matter (PM2.5) in Canadian oil sands
841 communities: Levels, sources and potential human health risk, Sci. Tot. Environm., 595, 828-838,
842 10.1016/j.scitotenv.2017.04.023, 2017.
843 Bari, M. A., and Kindzierski, W. B.: Ambient volatile organic compounds (VOCs) in communities of the
844 Athabasca oil sands region: Sources and screening health risk assessment, Environ. Pollut., 235, 602-
845 614, 10.1016/j.envpol.2017.12.065, 2018.
846 Bernstein, D. M.: Increased mortality in COPD among construction workers exposed to inorganic dust,
847 Eur. Resp. J., 24, 512-512, 10.1183/09031936.04.00044504, 2004.

848 Bond, T. C., Streets, D. G., Yarber, K. F., Nelson, S. M., Woo, J. H., and Klimont, Z.: A technology-based
849 global inventory of black and organic carbon emissions from combustion, J. Geophys. Res., 109,
850 D14203, 10.1029/2003JD003697, 2004.
851 Briggs, N. L., and Long, C. M.: Critical review of black carbon and elemental carbon source
852 apportionment in Europe and the United States, Atmos. Environ., 144, 409-427,
853 10.1016/j.atmosenv.2016.09.002, 2016.
854 Burtscher, H., Scherrer, L., Siegmann, H. C., Schmidtott, A., and Federer, B.: Probing aerosols by
855 photoelectric charging, J. Appl. Phys., 53, 3787-3791, 10.1063/1.331120, 1982.
856 Bytnerowicz, A., Fraczek, W., Schilling, S., and Alexander, D.: Spatial and temporal distribution of
857 ambient nitric acid and ammonia in the Athabasca Oil Sands Region, Alberta, J. Limnol., 69, 11-21,
858 10.3274/jl10-69-s1-03, 2010.
859 CAPP, C. A. f. P. P.: The facts on Canada's oil sands: http://www.capp.ca/publications-and-
860 statistics/publications/296225, access: April 20, 2017, 2016.
861 Carslaw, D. C., and Ropkins, K.: openair - An R package for air quality data analysis, Environ. Modell.
862 Softw., 27-28, 52-61, 10.1016/j.envsoft.2011.09.008, 2012.
863 Carslaw, D. C., and Beevers, S. D.: Characterising and understanding emission sources using bivariate
864 polar plots and k-means clustering, Environ. Modell. Softw., 40, 325-329,
865 10.1016/j.envsoft.2012.09.005, 2013.
866 Cattell, R. B.: The Scree Test For The Number Of Factors, Multivariate Behavioral Research, 1, 245-276,
867 10.1207/s15327906mbr0102_10, 1966.
868 Chalaturnyk, R. J., Don Scott, J., and Ozum, B.: Management of oil sands tailings, Petroleum Science and
869 Technology, 20, 1025-1046, 10.1081/lft-120003695, 2002.

870 Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakley, J. A., Hansen, J. E., and Hofmann, D. J.:
871 Climate forcing by anthropogenic aerosols, Science, 255, 423-430, 10.1126/science.255.5043.423
872 1992.
873 Chen, H., Karion, A., Rella, C. W., Winderlich, J., Gerbig, C., Filges, A., Newberger, T., Sweeney, C., and
874 Tans, P. P.: Accurate measurements of carbon monoxide in humid air using the cavity ring-down
875 spectroscopy (CRDS) technique, Atmos. Meas. Tech., 6, 1031-1040, 10.5194/amt-6-1031-2013, 2013.
876 Cho, S., Morris, R., McEachern, P., Shah, T., Johnson, J., and Nopmongcol, U.: Emission sources
877 sensitivity study for ground-level ozone and PM2.5 due to oil sands development using air quality
878 modelling system: Part II – Source apportionment modelling, Atmos. Environm., 55, 542-556,
879 10.1016/j.atmosenv.2012.02.025, 2012.
880 Cross, E. S., Hunter, J. F., Carrasquillo, A. J., Franklin, J. P., Herndon, S. C., Jayne, J. T., Worsnop, D. R.,
881 Miake-Lye, R. C., and Kroll, J. H.: Online measurements of the emissions of intermediate-volatility and
882 semi-volatile organic compounds from aircraft, Atmos. Chem. Phys., 13, 7845-7858, 10.5194/acp-13-
883 7845-2013, 2013.
884 de Gouw, J. A., Middlebrook, A. M., Warneke, C., Ahmadov, R., Atlas, E. L., Bahreini, R., Blake, D. R.,
885 Brock, C. A., Brioude, J., Fahey, D. W., Fehsenfeld, F. C., Holloway, J. S., Le Henaff, M., Lueb, R. A.,
886 McKeen, S. A., Meagher, J. F., Murphy, D. M., Paris, C., Parrish, D. D., Perring, A. E., Pollack, I. B.,
887 Ravishankara, A. R., Robinson, A. L., Ryerson, T. B., Schwarz, J. P., Spackman, J. R., Srinivasan, A., and
888 Watts, L. A.: Organic Aerosol Formation Downwind from the Deepwater Horizon Oil Spill, Science,
889 331, 1295-1299, 10.1126/science.1200320, 2011.
890 Dlugokencky, E.: Trends in atmospheric methane: www.esrl.noaa.gov/gmd/ccgg/trends_ch4/, access:
891 April 11, 2017, 2017a.
892 Dlugokencky, E.: Trends in atmospheric carbon dioxide: www.esrl.noaa.gov/gmd/ccgg/trends/, access:
893 April 11, 2017, 2017b.

ECCC: National pollutant release inventory (NPRI): http://open.canada.ca/data/en/dataset/e40099ae-b116-4c48-9475-f3806fe5a6a6, access: October 5, 2016, 2013.

ECCC: Joint oil sands monitoring program emissions inventory compilation report, Environment and Climate Change Canada, Downsview, 2016.

Englander, J. G., Bharadwaj, S., and Brandt, A. R.: Historical trends in greenhouse gas emissions of the Alberta oil sands (1970-2010), Environm. Res. Lett., 8, 044036, 10.1088/1748-9326/8/4/044036, 2013.

Fares, S., Schnitzhofer, R., Jiang, X., Guenther, A., Hansel, A., and Loreto, F.: Observations of Diurnal to Weekly Variations of Monoterpene-Dominated Fluxes of Volatile Organic Compounds from Mediterranean Forests: Implications for Regional Modeling, Environm. Sci. Technol., 47, 11073-11082, 10.1021/es4022156, 2013.

Fenn, M. E., Bytnerowicz, A., Schilling, S. L., and Ross, C. S.: Atmospheric deposition of nitrogen, sulfur and base cations in jack pine stands in the Athabasca Oil Sands Region, Alberta, Canada, Environ. Pollut., 196, 497-510, 10.1016/j.envpol.2014.08.023, 2015.

Fioletov, V. E., McLinden, C. A., Cede, A., Davies, J., Mihele, C., Netcheva, S., Li, S. M., and O'Brien, J.: Sulfur dioxide (SO_2) vertical column density measurements by Pandora spectrometer over the Canadian oil sands, Atmospheric Measurement Techniques, 9, 2961-2976, 10.5194/amt-9-2961-2016, 2016.

Fuentes, J. D., Wang, D., Neumann, H. H., Gillespie, T. J., DenHartog, G., and Dann, T. F.: Ambient biogenic hydrocarbons and isoprene emissions from a mixed deciduous forest, J. Atmos. Chem., 25, 67-95, 10.1007/BF00053286, 1996.

Furimsky, E.: Emissions of carbon dioxide from tar sands plants in Canada, Energy Fuels, 17, 1541-1548, 10.1021/ef0301102, 2003.

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.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions, Proceedings of the National Academy of Sciences, 109, 18318-18323, 10.1073/pnas.1212272109, 2012.

Geron, C., Rasmussen, R., Arnts, R. R., and Guenther, A.: A review and synthesis of monoterpene speciation from forests in the United States, Atmos. Environm., 34, 1761-1781, 10.1016/S1352-2310(99)00364-7, 2000.

Gordon, M., Li, S. M., Staebler, R., Darlington, A., Hayden, K., O'Brien, J., and Wolde, M.: Determining air pollutant emission rates based on mass balance using airborne measurement data over the Alberta oil sands operations, Atmos. Meas. Tech., 8, 3745-3765, 10.5194/amt-8-3745-2015, 2015.

Grimmer, G., Brune, H., Deutschwenzel, R., Dettbarn, G., Jacob, J., Naujack, K. W., Mohr, U., and Ernst, H.: Contribution of polycyclic aromatic-hydrocarbons and nitro-derivatives to the carcinogenic impact of diesel-engine exhaust condensate evaluated by implantation into the lungs of rats, Cancer Lett., 37, 173-180, 10.1016/0304-3835(87)90160-1, 1987.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geosci. Model Dev., 5, 1471-1492, 10.5194/gmd-5-1471-2012, 2012.

Guo, H., Wang, T., and Louie, P. K. K.: Source apportionment of ambient non-methane hydrocarbons in Hong Kong: Application of a principal component analysis/absolute principal component scores (PCA/APCS) receptor model, Environ. Pollut., 129, 489-498, 10.1016/j.envpol.2003.11.006, 2004.

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C.: Multivariate data analysis, in, 7th edition ed., Prentice-Hall, Upper Saddle River, NJ, pp. 108 -110, 1998.

941 Harrison, R. M., Smith, D. J. T., and Luhana, L.: Source Apportionment of Atmospheric Polycyclic
942 Aromatic Hydrocarbons Collected from an Urban Location in Birmingham, U.K, Environm. Sci.
943 Technol., 30, 825-832, 10.1021/es950252d, 1996.
944 Helmig, D., Klinger, L. F., Guenther, A., Vierling, L., Geron, C., and Zimmerman, P.: Biogenic volatile
945 organic compound emissions (BVOCs) I. Identifications from three continental sites in the U.S,
946 Chemosphere, 38, 2163-2187, 10.1016/S0045-6535(98)00425-1, 1999.
947 Holowenko, F. M., MacKinnon, M. D., and Fedorak, P. M.: Methanogens and sulfate-reducing bacteria in
948 oil sands fine tailings waste, Canadian Journal of Microbiology, 46, 927-937, 10.1139/cjm-46-10-927,
949 2000.
950 Howell, S. G., Clarke, A. D., Freitag, S., McNaughton, C. S., Kapustin, V., Brekovskikh, V., Jimenez, J. L.,
951 and Cubison, M. J.: An airborne assessment of atmospheric particulate emissions from the processing
952 of Athabasca oil sands, Atmos. Chem. Phys., 14, 5073-5087, 10.5194/acp-14-5073-2014, 2014.
953 Hsu, Y. M., Bytnerowicz, A., Fenn, M. E., and Percy, K. E.: Atmospheric dry deposition of sulfur and
954 nitrogen in the Athabasca Oil Sands Region, Alberta, Canada, Sci. Tot. Environm., 568, 285-295,
955 10.1016/j.scitotenv.2016.05.205, 2016.
956 Jautzy, J., Ahad, J. M. E., Gobeil, C., and Savard, M. M.: Century-Long Source Apportionment of PAHs in
957 Athabasca Oil Sands Region Lakes Using Diagnostic Ratios and Compound-Specific Carbon Isotope
958 Signatures, Environm. Sci. Technol., 47, 6155-6163, 10.1021/es400642e, 2013.
959 Jautzy, J. J., Ahad, J. M. E., Gobeil, C., Smirnoff, A., Barst, B. D., and Savard, M. M.: Isotopic Evidence for
960 Oil Sands Petroleum Coke in the Peace-Athabasca Delta, Environm. Sci. Technol., 49, 12062-12070,
961 10.1021/acs.est.5b03232, 2015.
962 Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F.,
963 Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A.
964 L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A.,

965 Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R.,
966 Cubison, M. J., E., Dunlea, J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K.,
967 Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell,
968 L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina,
969 K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C.,
970 Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of Organic Aerosols
971 in the Atmosphere, Science, 326, 1525-1529, 10.1126/science.1180353, 2009.
972 Johnson, M. R., Crosland, B. M., McEwen, J. D., Hager, D. B., Armitage, J. R., Karimi-Golpayegani, M., and
973 Picard, D. J.: Estimating fugitive methane emissions from oil sands mining using extractive core
974 samples, Atmos. Environm., 144, 111-123, 10.1016/j.atmosenv.2016.08.073, 2016.
975 Jolliffe, I. T., and Cadima, J.: Principal component analysis: a review and recent developments,
976 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,
977 374, 10.1098/rsta.2015.0202, 2016.
978 Kaiser, H. F.: The varimax criterion for analytic rotation in factor-analysis, Psychometrika, 23, 187-200,
979 10.1007/bf02289233, 1958.
980 Kindzierski, W. B., and Ranganathan, H. K. S.: Indoor and outdoor SO2 in a community near oil sand
981 extraction and production facilities in northern Alberta, J. Environ. Eng. Sci., 5, S121-S129,
982 10.1139/s06-022, 2006.
983 Landis, M. S., Pancras, J. P., Graney, J. R., Stevens, R. K., Percy, K. E., and Krupa, S.: Chapter 18 - Receptor
984 Modeling of Epiphytic Lichens to Elucidate the Sources and Spatial Distribution of Inorganic Air
985 Pollution in the Athabasca Oil Sands Region, in: Developments in Environmental Science, edited by:
986 Percy, K. E., Elsevier, 427-467, 2012.
987 Landis, M. S., Pancras, J. P., Graney, J. R., White, E. M., Edgerton, E. S., Legge, A., and Percy, K. E.: Source
988 apportionment of ambient fine and coarse particulate matter at the Fort McKay community site, in

989 the Athabasca Oil Sands Region, Alberta, Canada, Sci. Tot. Environm., 584, 105-117,
990 10.1016/j.scitotenv.2017.01.110, 2017.
991 Lee, A. K. Y., Adam, M. G., Liggio, J., Li, S.-M., Li, K., Willis, M. D., Abbatt, J. P. D., Tokarek, T. W., Odame-
992 Ankrah, C. A., Huo, J. A., Osthoff, H. D., Strawbridge, K. B., and Brook, J. R.: A large contribution of
993 anthropogenic organo-nitrates to secondary organic aerosol in Alberta oil sands, Atmos. Chem. Phys.
994 Discuss., 2018.
995 Li, S.-M., Leithead, A., Moussa, S. G., Liggio, J., Moran, M. D., Wang, D., Hayden, K., Darlington, A.,
996 Gordon, M., Staebler, R., Makar, P. A., Stroud, C. A., McLaren, R., Liu, P. S. K., Oâ€™Brien, J.,
997 Mittermeier, R. L., Zhang, J., Marson, G., Cober, S. G., Wolde, M., and Wentzell, J. J. B.: Differences
998 between measured and reported volatile organic compound emissions from oil sands facilities in
999 Alberta, Canada, Proceedings of the National Academy of Sciences, 114, E3756-E3765,
1000 10.1073/pnas.1617862114, 2017.
1001 Liggio, J., Li, S.-M., Hayden, K., Taha, Y. M., Stroud, C., Darlington, A., Drollette, B. D., Gordon, M., Lee, P.,
1002 Liu, P., Leithead, A., Moussa, S. G., Wang, D., O'Brien, J., Mittermeier, R. L., Brook, J., Lu, G., Staebler,
1003 R., Han, Y., Tokarek, T. W., Osthoff, H. D., Makar, P. A., Zhang, J., Plata, D., and Gentner, D. R.: Oil
1004 Sands Operations as a Large Source of Secondary Organic Aerosols, Nature, 534, 91-94,
1005 10.1038/nature17646, 2016.
1006 Marey, H. S., Hashisho, Z., Fu, L., and Gille, J.: Spatial and temporal variation in CO over Alberta using
1007 measurements from satellites, aircraft, and ground stations, Atmos. Chem. Phys., 15, 3893-3908,
1008 10.5194/acp-15-3893-2015, 2015.
1009 Markovic, M. Z., VandenBoer, T. C., and Murphy, J. G.: Characterization and optimization of an online
1010 system for the simultaneous measurement of atmospheric water-soluble constituents in the gas and
1011 particle phases, J. Environ. Monit., 14, 1872-1884, 10.1039/C2EM00004K, 2012.

1012 Masliyah, J., Zhou, Z. J., Xu, Z. H., Czarnecki, J., and Hamza, H.: Understanding water-based bitumen
1013 extraction from athabasca oil sands, Can. J. Chem. Eng., 82, 628-654, 10.1002/cjce.5450820403,
1014 2004.
1015 McLinden, C. A., Fioletov, V., Boersma, K. F., Krotkov, N., Sioris, C. E., Veefkind, J. P., and Yang, K.: Air
1016 quality over the Canadian oil sands: A first assessment using satellite observations, Geophys. Res.
1017 Lett., 39, 8, 10.1029/2011gl050273, 2012.
1018 Miller, S. M., Matross, D. M., Andrews, A. E., Millet, D. B., Longo, M., Gottlieb, E. W., Hirsch, A. I., Gerbig,
1019 C., Lin, J. C., Daube, B. C., Hudman, R. C., Dias, P. L. S., Chow, V. Y., and Wofsy, S. C.: Sources of carbon
1020 monoxide and formaldehyde in North America determined from high-resolution atmospheric data,
1021 Atmos. Chem. Phys., 8, 7673-7696, 10.5194/acp-8-7673-2008, 2008.
1022 Min, K. E., Pusede, S. E., Browne, E. C., LaFranchi, B. W., and Cohen, R. C.: Eddy covariance fluxes and
1023 vertical concentration gradient measurements of NO and NO2 over a ponderosa pine ecosystem:
1024 observational evidence for within-canopy chemical removal of NOx, Atmos. Chem. Phys., 14, 5495-
1025 5512, 10.5194/acp-14-5495-2014, 2014.
1026 Nara, H., Tanimoto, H., Tohjima, Y., Mukai, H., Nojiri, Y., Katsumata, K., and Rella, C. W.: Effect of air
1027 composition (N_2, O_2, Ar, and H_2O) on CO_2 and CH_4 measurement by wavelength-scanned cavity ring-
1028 down spectroscopy: calibration and measurement strategy, Atmos. Meas. Tech., 5, 2689-2701,
1029 10.5194/amt-5-2689-2012, 2012.
1030 Nimana, B., Canter, C., and Kumar, A.: Energy consumption and greenhouse gas emissions in the
1031 recovery and extraction of crude bitumen from Canada's oil sands, Appl. Energy, 143, 189-199,
1032 10.1016/j.apenergy.2015.01.024, 2015a.
1033 Nimana, B., Canter, C., and Kumar, A.: Energy consumption and greenhouse gas emissions in upgrading
1034 and refining of Canada's oil sands products, Energy, 83, 65-79, 10.1016/j.energy.2015.01.085, 2015b.

NPRI: Detailed facility information: http://www.ec.gc.ca/inrp-npri/donnees-data/index.cfm?do=facility_information&lang=En&opt_npri_id=0000002274&opt_report_year=2013, access: April 13, 2017, 2013.

Odame-Ankrah, C. A.: Improved detection instrument for nitrogen oxide species, Ph.D., Chemistry, University of Calgary, http://hdl.handle.net/11023/2006, 10.5072/PRISM/26475, Calgary, 2015.

Onasch, T. B., Trimborn, A., Fortner, E. C., Jayne, J. T., Kok, G. L., Williams, L. R., Davidovits, P., and Worsnop, D. R.: Soot Particle Aerosol Mass Spectrometer: Development, Validation, and Initial Application, Aerosol Sci. Technol., 46, 804-817, 10.1080/02786826.2012.663948, 2012.

Paatero, P., and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111-126, doi:10.1002/env.3170050203, 1994.

Parajulee, A., and Wania, F.: Evaluating officially reported polycyclic aromatic hydrocarbon emissions in the Athabasca oil sands region with a multimedia fate model, Proceedings of the National Academy of Sciences, 111, 3344-3349, 10.1073/pnas.1319780111, 2014.

Paul, D., and Osthoff, H. D.: Absolute Measurements of Total Peroxy Nitrate Mixing Ratios by Thermal Dissociation Blue Diode Laser Cavity Ring-Down Spectroscopy, Anal. Chem., 82, 6695-6703, 10.1021/ac101441z, 2010.

Penner, T. J., and Foght, J. M.: Mature fine tailings from oil sands processing harbour diverse methanogenic communities, Canadian Journal of Microbiology, 56, 459-470, 10.1139/w10-029, 2010.

Percy, K. E.: Ambient Air Quality and Linkage to Ecosystems in the Athabasca Oil Sands, Alberta, Geosci. Can., 40, 182-201, 10.12789/geocanj.2013.40.014, 2013.

Peters, T. M., and Leith, D.: Concentration measurement and counting efficiency of the aerodynamic particle sizer 3321, J. Aerosol Sci., 34, 627-634, 10.1016/s0021-8502(03)00030-2, 2003.

1058 Polissar, A. V., Hopke, P. K., Paatero, P., Malm, W. C., and Sisler, J. F.: Atmospheric aerosol over Alaska:
1059 2. Elemental composition and sources, J. Geophys. Res.-Atmos., 103, 19045-19057,
1060 doi:10.1029/98JD01212, 1998.
1061 Quagraine, E. K., Headley, J. V., and Peterson, H. G.: Is biodegradation of bitumen a source of recalcitrant
1062 naphthenic acid mixtures in oil sands tailing pond waters?, J. Environ. Sci. Health Part A-Toxic/Hazard.
1063 Subst. Environ. Eng., 40, 671-684, 10.1081/ese-200046637, 2005.
1064 Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop, A. P., Lane,
1065 T. E., Pierce, J. R., and Pandis, S. N.: Rethinking organic aerosols: Semivolatile emissions and
1066 photochemical aging, Science, 315, 1259-1262, 2007.
1067 Rooney, R. C., Bayley, S. E., and Schindler, D. W.: Oil sands mining and reclamation cause massive loss of
1068 peatland and stored carbon, Proc. Natl. Acad. Sci. U.S.A., 109, 4933-4937, 10.1073/pnas.1117693108,
1069 2012.
1070 RStudio Boston, M.: Integrated development environment for R, 2017.
1071 Shahimin, M. F. M., and Siddique, T.: Sequential biodegradation of complex naphtha hydrocarbons
1072 under methanogenic conditions in two different oil sands tailings, Environ. Pollut., 221, 398-406,
1073 10.1016/j.envpol.2016.12.002, 2017.
1074 Siddique, T., Fedorak, P. M., and Foght, J. M.: Biodegradation of short-chain n-alkanes in oil sands
1075 tailings under methanogenic conditions, Environ. Sci. Technol., 40, 5459-5464, 10.1021/es060993m,
1076 2006.
1077 Siddique, T., Gupta, R., Fedorak, P. M., MacKinnon, M. D., and Foght, J. M.: A first approximation kinetic
1078 model to predict methane generation from an oil sands tailings settling basin, Chemosphere, 72,
1079 1573-1580, 10.1016/j.chemosphere.2008.04.036, 2008.

1080 Siddique, T., Penner, T., Semple, K., and Foght, J. M.: Anaerobic Biodegradation of Longer-Chain n-
1081 Alkanes Coupled to Methane Production in Oil Sands Tailings, Environm. Sci. Technol., 45, 5892-5899,
1082 10.1021/es200649t, 2011.
1083 Siddique, T., Penner, T., Klassen, J., Nesbo, C., and Foght, J. M.: Microbial Communities Involved in
1084 Methane Production from Hydrocarbons in Oil Sands Tailings, Environ. Sci. Technol., 46, 9802-9810,
1085 10.1021/e53022024, 2012.
1086 Simpson, I. J., Blake, N. J., Barletta, B., Diskin, G. S., Fuelberg, H. E., Gorham, K., Huey, L. G., Meinardi, S.,
1087 Rowland, F. S., Vay, S. A., Weinheimer, A. J., Yang, M., and Blake, D. R.: Characterization of trace
1088 gases measured over Alberta oil sands mining operations: 76 speciated C2–C10 volatile organic
1089 compounds (VOCs), CO_2, CH_4, CO, NO, NO_2, NO_y, O_3 and SO_2, Atmos. Chem. Phys., 10, 11931-11954,
1090 10.5194/acp-10-11931-2010, 2010.
1091 Small, C. C., Cho, S., Hashisho, Z., and Ulrich, A. C.: Emissions from oil sands tailings ponds: Review of
1092 tailings pond parameters and emission estimates, Journal of Petroleum Science and Engineering, 127,
1093 490-501, 10.1016/j.petrol.2014.11.020, 2015.
1094 Thurston, G. D., and Spengler, J. D.: A quantitative assessment of source contributions to inhalable
1095 particulate matter pollution in metropolitan Boston, Atmos. Environ., 19, 9-25, 10.1016/0004-
1096 6981(85)90132-5, 1985.
1097 Thurston, G. D., Ito, K., and Lall, R.: A source apportionment of U.S. fine particulate matter air pollution,
1098 Atmos. Environ., 45, 3924-3936, 10.1016/j.atmosenv.2011.04.070, 2011.
1099 Tokarek, T. W., Huo, J. A., Odame-Ankrah, C. A., Hammoud, D., Taha, Y. M., and Osthoff, H. D.: A gas
1100 chromatograph for quantification of peroxycarboxylic nitric anhydrides calibrated by thermal
1101 dissociation cavity ring-down spectroscopy, Atmos. Meas. Tech., 7, 3263-3283, 10.5194/amt-7-3263-
1102 2014, 2014.

1103 Tokarek, T. W., Brownsey, D. K., Jordan, N., Garner, N. M., Ye, C. Z., Assad, F. V., Peace, A., Schiller, C. L.,
1104 Mason, R. H., Vingarzan, R., and Osthoff, H. D.: Biogenic emissions and nocturnal ozone depletion
1105 events at the Amphitrite Point Observatory on Vancouver Island, Atmosphere-Ocean, 55, 121-132,
1106 10.1080/07055900.2017.1306687, 2017a.
1107 Tokarek, T. W., Brownsey, D. K., Jordan, N., Garner, N. M., Ye, C. Z., Assad, F. V., Peace, A., Schiller, C. L.,
1108 Mason, R. H., Vingarzan, R., and Osthoff, H. D.: Biogenic Emissions and Nocturnal Ozone Depletion
1109 Events at the Amphitrite Point Observatory on Vancouver Island, Atmosphere-Ocean, 1-12,
1110 10.1080/07055900.2017.1306687, 2017b.
1111 Wang, S. C., and Flagan, R. C.: Scanning electrical mobility spectrometer, Aerosol Sci. Technol., 13, 230-
1112 240, 10.1080/02786829008959441, 1990.
1113 Wang, X. L., Chow, J. C., Kohl, S. D., Percy, K. E., Legge, A. H., and Watson, J. G.: Characterization of
1114 PM2.5 and PM10 fugitive dust source profiles in the Athabasca Oil Sands Region, J. Air Waste Manag.
1115 Assoc., 65, 1421-1433, 10.1080/10962247.2015.1100693, 2015.
1116 Wang, X. L., Chow, J. C., Kohl, S. D., Percy, K. E., Legge, A. H., and Watson, J. G.: Real-world emission
1117 factors for Caterpillar 797B heavy haulers during mining operations, Particuology, 28, 22-30,
1118 10.1016/j.partic.2015.07.001, 2016.
1119 Warren, L. A., Kendra, K. E., Brady, A. L., and Slater, G. F.: Sulfur Biogeochemistry of an Oil Sands
1120 Composite Tailings Deposit, Front. Microbiol., 6, 14, 10.3389/fmicb.2015.01533, 2016.
1121 Watson, J., Chow, J., Wang, X., Zielinska, B., Kohl, S., and Gronstal, S.: Characterization of real-world
1122 emissions from nonroad mining trucks in the Athabasca Oil Sands Region during September, 2009,
1123 2013.
1124 WBEA: WBEA annual report 2013, Wood Buffalo Environmental Association, 2013.
1125 Whaley, C. H., Makar, P. A., Shephard, M. W., Zhang, L., Zhang, J., Zheng, Q., Akingunola, A., Wentworth,
1126 G. R., Murphy, J. G., Kharol, S. K., and Cady-Pereira, K. E.: Contributions of natural and anthropogenic

sources to ambient ammonia in the Athabasca Oil Sands and north-western Canada, Atmos. Chem. Phys., 18, 2011-2034, 10.5194/acp-18-2011-2018, 2018.

Williams, B. J., Goldstein, A. H., Kreisberg, N. M., and Hering, S. V.: An in-situ instrument for speciated organic composition of atmospheric aerosols: Thermal Desorption Aerosol GC/MS-FID (TAG), Aerosol Sci. Technol., 40, 627-638, 10.1080/02786820600754631, 2006.

Wilson, N. K., Barbour, R. K., Chuang, J. C., and Mukund, R.: Evaluation of a real-time monitor for fine particle-bound PAH in air, Polycycl. Aromat. Compd., 5, 167-174, 10.1080/10406639408015168, 1994.

Yang, C., Wang, Z., Yang, Z., Hollebone, B., Brown, C. E., Landriault, M., and Fieldhouse, B.: Chemical Fingerprints of Alberta Oil Sands and Related Petroleum Products, Environmental Forensics, 12, 173-188, 10.1080/15275922.2011.574312, 2011.

Yeh, S., Jordaan, S. M., Brandt, A. R., Turetsky, M. R., Spatari, S., and Keith, D. W.: Land Use Greenhouse Gas Emissions from Conventional Oil Production and Oil Sands, Environm. Sci. Technol., 44, 8766-8772, 10.1021/es1013278, 2010.

Zhang, Y., Wang, Y., Chen, G., Smeltzer, C., Crawford, J., Olson, J., Szykman, J., Weinheimer, A. J., Knapp, D. J., Montzka, D. D., Wisthaler, A., Mikoviny, T., Fried, A., and Diskin, G.: Large vertical gradient of reactive nitrogen oxides in the boundary layer: Modeling analysis of DISCOVER-AQ 2011 observations, J. Geophys. Res.-Atmos., 121, 1922-1934, 10.1002/2015jd024203, 2016.

Zhao, W., Hopke, P. K., and Karl, T.: Source Identification of Volatile Organic Compounds in Houston, Texas, Environm. Sci. Technol., 38, 1338-1347, 10.1021/es034999c, 2004.

Zhao, Y. L., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A., and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol Production, Environm. Sci. Technol., 49, 11516-11526, 10.1021/acs.est.3b02841, 2015.

1151

1152

Supplementary information for

Principal component analysis of summertime ground site measurements in the Athabasca oil sands with a focus on analytically unresolved intermediate volatility organic compounds

Travis W. Tokarek[1], Charles A. Odame-Ankrah[1], Jennifer A. Huo[1], Robert McLaren[2], Alex K. Y. Lee[3, 4], Max G. Adam[4], Megan D. Willis[5], Jonathan P. D. Abbatt[5], Cristian Mihele[6], Andrea Darlington[6], Richard L. Mittermeier[6], Kevin Strawbridge[6], Katherine L. Hayden[6], Jason S. Olfert[7], Elijah. G. Schnitzler[8], Duncan K. Brownsey[1], Faisal V. Assad[1], Gregory R. Wentworth[5, a], Alex G. Tevlin[5], Douglas E. J. Worthy[6], Shao-Meng Li[6], John Liggio[6], Jeffrey R. Brook[6], and Hans D. Osthoff[1*]

[1] Department of Chemistry, University of Calgary, Calgary, Alberta, T2N 1N4, Canada

[2] Centre for Atmospheric Chemistry, York University, Toronto, Ontario M3J 1P3, Canada

[3] Department of Civil and Environmental Engineering, National University of Singapore, Singapore 117576, Singapore.

[4] NUS Environmental Research Institute, National University of Singapore, Singapore

[5] Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada

[6] Air Quality Research Division, Environment and Climate Change Canada, Toronto, Ontario, M3H 5T4, Canada

[7] Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, T6G 1H9, Canada

[8] Department of Chemistry, University of Alberta, Edmonton, Alberta, T6G 2G2, Canada

[a] Now at: Environmental Monitoring and Science Division, Alberta Environment and Parks, Edmonton, Alberta, T5J 5C6, Canada

* Corresponding author

Table of contents

Descriptions of instrumentation used

A Griffin 450 gas chromatograph equipped with a cylindrical ion trap mass spectrometer and electron impact ionization (GC-ITMS) was used to quantify selected VOCs including o-xylene, decane, undecane, 1,2,3- and 1,2,4-trimethylbenzene (TMB), and several monoterpenes (i.e., α-pinene, ß-pinene and limonene). The GC-ITMS primary responsibility was the quantification of monoterpenes. The remaining VOCs quantified were chosen because (a) they sufficiently resolved on the analytical column, and (b) response factors could be determined, either because the compounds of interest were part of the VOC standard used in the field (such as the aromatics o-xylene, 1,2,3- and 1,2,4-TMB, see below) or relative response factors were determined post-campaign. Operation, calibration and performance of this instrument have been described elsewhere (Tokarek et al., 2017; Liggio et al., 2016). Briefly, the GC was operated with 30 m (length) × 0.25 mm (inner diameter) × 0.25 µm (film thickness) DB-5MS analytical column with helium carrier gas. The GC-ITMS sampled from a 3.6 m long stainless-steel inlet with an o.d. of 0.635 cm from a height of 5 m above ground. A 1 m long section of the inlet was heated to 110 °C and optimized to remove interference due to O_3 while avoiding decomposition of alkenes (Tokarek et al., 2017). The GC oven was programmed as follows: hold at 40 °C for 3.00 min, heat at 1.5 °C min^{-1} to 70° C (reached at 23.00 min), heat at 5° C min^{-1} to 200 °C (reached at 49.00 min) and hold for 4 min (total 53.00 min). This was followed by a 5 min recovery time to allow the oven and pre-concentration trap to cool back to 40 °C. The ion trap mass spectrometer was set to an *m/z* range of 50-425. After data reduction, the GC-ITMS generated 10-minute average concentrations of each VOC quantified every hour.

During the campaign, the GC-ITMS was calibrated in the field using an IONICON VOC standard (Table S-1) containing (in addition to VOCs that the GC-ITMS did not detect) α-pinene and o-xylene at mixing ratios of ~ 1 ppmv and an uncertainty of 5% and 6%, respectively. A commercial calibrator assembly (IONICON, GCU Standard) was used to deliver diluted calibration mixtures. The instrument responses to

the VOC standards were linear ($R^2 > 0.99$). The GC-ITMS was calibrated for other VOCs offline relative to α-pinene. In the field, there was no noticeable carry-over (i.e., memory effects) of IVOCs, which was occasionally evaluated by flooding the inlet with purified, VOC-free air.

Matrices of ions plotted against retention times for the total ion chromatograms (shown in Figure 2 in the main manuscript) are shown in Fig. S-1. In both cases, the greatest intensity is with masses are associated with alkanes (i.e., *m/z* 55, 57, 67, 69, etc.).

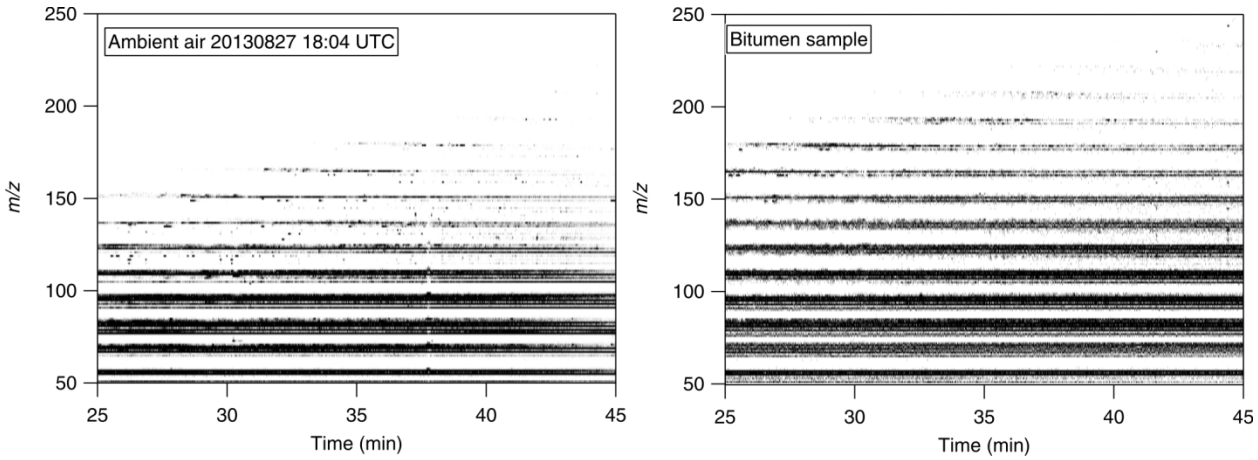


Figure S-1. Scatter of ions as a function of retention time for the total ion chromatograms shown in Figure 2 of the main manuscript. Darker pixels represent a higher intensity than lighter pixels.

Mixing ratios of carbon monoxide (CO), carbon dioxide (CO_2) and methane (CH_4) in ambient air were quantified using a commercial cavity ring-down spectrometer (Picarro G2401) (Nara et al., 2012; Chen et al., 2013). Ambient air was sampled from a height of 10 m through 0.635 cm outer diameter (o.d.) perfluoroalkoxyalkane (PFA) Teflon™ tubing and a 47 mm diameter, 1 µm pore filter at a flow rate of ~0.5 L min^{-1}. A scrubber ($MgClO_4$) was installed at the base of the sample line to remove water from the air. Operating procedures developed for Canada's greenhouse gas network of monitors across all stations in Canada by the Climate Division of ECCC were followed (ECCC, 2013a).The cavity ring-down spectrometer was calibrated every few days with calibrated standard gas mixtures (Scott-Marrin); a

target background mixture (CO_2 at a mixing ratio of 379.5 parts-per-million by volume (ppmv), CH_4 at 1.976 ppmv and CO at 198.4 parts-per-billion by volume (ppbv)) and a working mixture (CO_2 = 452.15 ppmv, CH_4 = 2.988 ppmv, CO = 494.5 ppbv). The estimated precision of 1 min data was $\pm$0.12 ppmv, $\pm$0.6 ppbv, and $\pm$1.89 ppbv for CO_2, CH_4 and CO respectively, while the estimated accuracy was < 1 ppmv, < 3 ppbv, and < 4 ppbv, respectively.

Mixing ratios of total odd nitrogen ($NO_y \equiv NO + NO_2 + \Sigma PAN + \Sigma AN + HNO_3 + HONO + 2N_2O_5 + ClNO_2 + ...$) were measured by a chemiluminescence analyzer equipped with a heated Molybdenum converter (Thermo 42i) as described elsewhere (Tokarek et al., 2014; Odame-Ankrah, 2015).

The total sulfur (TS) measurements were conducted using a thermal oxidizer (Thermo Scientific Model CON101) to convert TS to SO_2 and detected using a pulsed-fluorescence analyzer (Thermo Scientific, Model 43iTLE). SO_2 was measured directly with a second analyzer (Thermo Scientific, Model 43iTLE). Total reduced sulfur (TRS) mixing ratios were calculated by subtracting mixing ratios of SO_2 from TS.

Concentrations of particle-surface bound polycyclic aromatic hydrocarbons (pPAH) were measured using a photoelectric aerosol sensor (EcoChem Analytics, Model PAS 2000CE) (Wilson et al., 1994; Burtscher et al., 1982).

Two soot-particle aerosol mass spectrometers (SP-AMS, Aerodyne Research, Inc.) (Onasch et al., 2012) measured non-refractory PM_1 components. Both SP-AMS were high resolution time-of-flight aerosol mass spectrometers (HR-ToF-AMS) fitted with a diode pumped Nd:YAG 1064 nm laser vaporizer; one SP-AMS had its oven removed to measure black carbon containing particles only using the laser. Direct calibrations of rBC using mono-disperse "Regal Black" (Cabot Corp. R400) particles were carried out three times during the 2013 JOSM intensive study. Positive Matrix Factorization (PMF) was performed to identify the potential sources of organic aerosol as described in the companion study (Adam et al., in prep). Factors associated with primary aerosol, i.e., hydrocarbon-like organic aerosol (HOA), a less

oxidized oxygenated organic aerosol factor (LO-OOA) and measured refractory black carbon (rBC) were added as variables for PCA analysis. Mass spectra associated with LO-OOA exhibited H/C, O/C and N/C ratios of ~1.62, ~0.36, and ~0.004, respectively; while the O/C and N/C ratios are similar to HOA, the H/C ratio of LO-OOA more resembles the more oxidized OOA factor (MO-OOA) (Adam et al., in prep.).

Particle volumes were calculated (assuming spherical particle shapes) from sub- and super-micron size distributions acquired using a scanning mobility particle sizer (SMPS, TSI with a differential mobility analyzer model 3081 and condensation particle counter model 3776; PM_1) and a 0.071 cm impactor over the size range of 13.6 nm to 736.5 nm and an Aerodynamic Particle Sizer (APS, TSI 3321; PM_{10-1}) over the size range 1.04 µm to 10.4 µm, respectively. Both instruments were operated at ambient relative humidity. The SMPS sampled through conductive silicon tubing to minimize wall losses due to wall charges. The APS was operated from a container located on top of the trailer and sampled from a 1.6 m tall, ½ o.d. aluminum tube whose tip was bent into a U-shape.

An ambient ion monitor – ion chromatograph (AIM-IC) (Markovic et al., 2012) was used to measure hourly averaged gas-phase NH_3 and $PM_{2.5}$ particle-phase (i.e., of particles < 2.5 µm diameter) NH_4^+ concentrations. High time-resolution particle-phase NH_4^+ measurements made by the SP-AMS were scaled by interpolated phase ratios observed by AIM-IC to calculate gas-phase NH_3 concentrations at high time resolution. This approach assumes the same phase ratios for $PM_{2.5}$ as for PM_1.

Determination of optimum PCA solution

The full component pattern (before component removal, with rotation, i.e., showing 22 components for 22 variables) obtained for this data set is shown in Table S-2. A common challenge in PCA is the determination of the maximum number of components to retain in the analysis. Several criteria are used for this purpose: the latent root criterion, where only components with eigenvalues greater than 1 are considered significant, the 5% variance criterion, where the last component selected accounts for only a small portion (<5%) of the variance, the 95% cumulative percentage of variance criterion, where the extracted components account for at least 95% of the total variance, and the Scree test. In the latter, the eigenvalues are plotted against the number of components in the order of extraction (Fig. S-2); to avoid including too many components with unique variance, the number of acceptable components is located at the point where this plot becomes horizontal. The latent root criterion is most commonly used, but tends to extract too few components when the number of variables is < 20 (Hair et al., 1998). The Scree test, on the other hand, often requires "some art in administering it" (Cattell, 1966), i.e., is subjective, though generally results in the inclusion of two or three more components than the latent root criterion (Hair et al., 1998).

The maximum component number for each criterion are summarized in Table S-~~9~~11. The Scree test plot (Fig. S-1) shows two plateaus where the slope becomes approximately horizontal: The first is located at N = 5 and the second at N = 12. The latent root criterion and the <5% variance method suggests a 7-component solution, whereas the >95% percentage of variance criterion suggests using a 10-component solution. Hair et al. (1998) recommend to examine component solutions with differing numbers of components to evaluate which best represents the structure of the variables. In the following, solutions are presented in ascending order of extracted components.

5-component solution

As a first attempt at interpretation of the PCA, the first cut-off of the Scree test criterion was chosen (N = 5 variables). The results (after Varimax rotation) are presented in Table S-3.

The 5-component solution accounts for a cumulative variance of 81.0 % after rotation. Communalities for the analysis, i.e., the fraction of total pollutant observations accounted for by the PCA (Otto, 2007), are greater than 70% for 18 variables. The lowest communalities were obtained for gas-phase ammonia (0.40), CO (0.48) and PM_{10-1} (0.51). TRS and the IVOCs were also relatively poorly represented (0.63 and 0.73, respectively). All eigenvalues are greater than 1.

The component accounting for most of the variance of the data, component 1, is strongly associated with all of the anthropogenic VOCs (with correlations of $r > 0.8$) and TRS ($r = 0.76$), weakly associated with CH_4 ($r = 0.62$), HOA ($r = 0.44$), LO-OOA ($r = 0.59$), IVOCs ($r = 0.47$), and CO ($r = 0.53$), and poorly associated with NO_y and TS ($r = 0.25$ and $r = 0.28$, respectively). Component 1 is consistent with tailings ponds emissions with potentially small contributions from nearby facilities (interpreted from weak and poor correlations with rBC ($r = 0.33$) and NO_y ($r = 0.25$)), which would otherwise remain unexplained.

Component 2 is strongly associated with the combustion tracers NO_y ($r = 0.83$), rBC ($r = 0.89$) and pPAH ($r = 0.83$) and weakly associated with IVOCs ($r = 0.61$), gas-phase ammonia ($r = 0.34$), undecane ($r = 0.31$), and CH_4 (0.38), but poorly and not with CO or CO_2 ($r = 0.19$ and 0.06, respectively); this component is identified as mine fleet emissions. Component 3 is strongly associated ($r > 0.9$) with the biogenic VOCs and weakly ($r = 0.55$) associated with CO_2 and is identified as a biogenic component.

Component 4 is strongly associated with SO_2 and TS ($r = 0.93$ and 0.91, respectively) and is consistent with emissions from upgrader facilities. These four components persisted, with little variation, in all solutions with a greater number of selected components (see below).

Component 5 is strongly associated with CO_2 ($r = 0.71$), and weakly associated with PM_{10-1} ($r = 0.57$), CH_4 ($r = 0.53$) and CO ($r = 0.40$). We are not aware of a source type that would fit this profile, i.e., combine

198 this particular set of pollutants without also being associated with NO_y (r = 0.02). This suggests that this
199 component is an artifact arising from an insufficient number of components used in the analysis and
200 motivates the inclusion of more components.
201
202 **6-component solution**
203 A 6-component solution is shown in Table S-4. Satisfying the percentage of variance criterion of the last
204 component accounting for less than 5% of the variance (4.6% in this case, Table S-2) was selected.
205 This solution accounts for a total variance of 85.23%. The first four components are essentially
206 unchanged from the 5-component solution (with the exception of LO-OOA in component 2 becoming
207 more poorly correlated (r = 0.22)). Component 5 is strongly associated with IVOCs (r = 0.70) and weakly
208 associated with LO-OOA (r = 0.60), and TRS (r = 0.56). Component 6 is strongly associated with PM_{10-1} (r
209 = 0.81) and weakly associated with CO_2 (r = 0.62), CH_4 (r = 0.41), HOA (r = 0.30) and NH_3 (r = 0.36) and,
210 unlike the 5-component solution, not associated with CO.
211
212 **7-component solution**
213 Next, the latent root criterion gives a 7-component solution. The PCA results (after Varimax rotation) are
214 presented in Table S-5. The seven components account for a cumulative variance of 88.7% after
215 rotation. Communalities for the analysis are all greater than 60%, with the lowest communality obtained
216 for CO (0.61). All eigenvalues are greater than 1.
217 Components 1 through 4 have the same associations with similar r values as those in the 5-component
218 analysis, with the only significant exception a poorer association (r = 0.20) of component 2 with gas-
219 phase ammonia.
220 The identifications of components 5 through 7 of the 7-component solution are murky at best.
221 Component 5 is weakly associated with TRS (r = 0.56) and IVOCs (r = 0.66). Component 6 is strongly

associated with PM_{10-1} volume (r = 0.89), and weakly with CO_2 (r = 0.54), and CH_4 (r = 0.36) and appears to be combination of a dust component with a source of greenhouse gases, whereas component 7 is strongly associated with gas-phase ammonia (r = 0.82) and poorly associated with CO (r = 0.29). Both appear to be amalgamations of distinct sources and suggest that too few components were selected. Hair et al. (1998) note that the latent root criterion has a tendency to extract a conservative number of components if the number of variables is < 20, close to the 22 variables in this analysis, consistent with what is observed here. Hence, the 7-component solution is sub-optimal.

8-component solution

An 8-component solution is presented in Table S-6. Not satisfying any criterion, it is included here for the sake of completeness. Owing to the inclusion of an additional component, the cumulative variance improved to 91.6%. The greatest improvement was seen for CO, gas-phase ammonia, as well as the IVOCs, whose communalities increased from 0.61, 0.91, and 0.80 (for the 7-component solution) to 0.96, 0.96 and 0.84, respectively.

The main effect of the inclusion of an additional component was the separation of component 7 into two distinct components: one of these was strongly associated with gas-phase ammonia (r = 0.92), and the other was strongly associated with CO (r = 0.85). A considerable fraction of the CO observed in the region is generated as a byproduct of the photochemical oxidation of hydrocarbons (Shephard et al., 2015); component 8 appears to capture this source, whereas component 1 captures the anthropogenic emissions. The area near the oil sands mining operations is enriched in ammonia, which originates from multiple sources: it is used as a floating agent to separate and recover bitumen from tar and is generated during bitumen upgrading (called hydrotreating) in which N is removed as NH_3 and can be present as a contaminant in tailing ponds. Other sources, such as agricultural activities, biological decay processes, and smoldering fires are relatively minor in the region (Bytnerowicz et al., 2010). The poor

246 association of component 2 with ammonia (r = 0.22) may capture the use of ammonia as a floating
247 agent, whereas component 8 embodies the remaining sources.
248 Component 5 is strongly associated with IVOCs (r = 0.71), and weakly associated with LO-OOA (r = 0.65)
249 and TRS (r = 0.40). It is unclear if these variables originate from the same source or are forced together
250 as a result of having chosen too few components. Considering that component 7 is split when an
251 additional component is used (see below), the latter is more likely. Component 6 remains strongly
252 associated with PM_{10-1} volume (r = 0.89), and weakly associated with CO_2 (r = 0.53), and CH_4 (r = 0.35)
253 and is difficult to interpret. Because of the unclear classification of components 5 through 8, the 8-
254 component solution is rejected.
255
256 **9-component solution**
257 A 9-component solution is presented in Table S-7. Components 1 through 8 describe sources that are
258 qualitatively similar to those provided by the 8-component solution. Component 9 is strongly associated
259 with TRS (r = 0.71) and poorly associated with o-xylene (r = 0.30); its profile is consistent with tailings
260 ponds emission, where the presence of naphtha as a diluent gives rise to BTEX emissions and bacteria
261 produce reduced sulfur compounds (Small et al., 2015; Warren et al., 2016). Component 6 is strongly
262 associated with PM_{10-1} (r = 0.89) and weakly associated with CO_2 (r = 0.54) and CH_4 (r = 0.41). We have
263 decided to reject this solution on the basis that < 95 % cumulative variance is observed.
264
265 **10-component solution**
266 Next, a 10-component solution with cumulative variance of 95.5%, satisfying the 95% criterion, was
267 considered. With this solution, all communalities are >0.85 (Table 3). Component 6 is strongly associated
268 with CO_2 (r = 0.77) and weakly associated with CH_4 (r = 0.59) but is not associated with other combustion
269 tracers and is identified as inactive open-pit mines (see main text). Component 7 is strongly correlated

with PM_{10-1} (r = 0.93) and is identified as wind-blown dust. Component 8 and 9 are strongly associated with a single variable each, gas-phase ammonia (r = 0.94) and CO (r = 0.87), respectively. Component 10 is strongly associated with TRS (r = 0.71) and weakly associated with o-xylene (r = 0.32). Overall, this component is most consistent with a tailings ponds source, where the presence of naphtha as diluent gives rise to BTEX emissions, and sulfur-reducing bacteria are at work (Small et al., 2015; Warren et al., 2016). Overall, the 10-component solution was judged to be optimal.

11-component solution

The 11-component analysis is presented in Table S-8. Component 10 is now strongly associated with LO-OOA (r = 0.72) and weakly with rBC (r = 0.34) and has a low eigenvalue of 0.87. This solution is therefore rejected as we believe it contains too many components.

PCA without aerosol variables

A sensitivity test was conducted by which all aerosol species were removed as variables. The results of this sensitivity test are shown in Table S-10 and are presented as a 9-component solution, since the dust component associated with PM_{10-1} (component 7 in Table 5) cannot be generated when its main variable is removed.

The pattern in Table S-10 resembles that in Table 5 of the main manuscript, in that the same nine components emerged in both solutions with similar magnitude r values for each of the variables, including the IVOC signature. The only difference is that components 2 and 3 as well as 5 and 6 have traded places (i.e., the relative magnitudes of their eigenvalues, which were similar in Table 5, have switched), which is inconsequential. Furthermore, the correlation coefficients in Table S-10 are of similar magnitude (i.e., within $\pm$0.1) as those in Table 5, which suggests that IVOC to SOA conversion does not adversely affect the PCA, likely because of the proximity of the receptor site to sources.

Discussion of low-eigenvalue components

Component 6: A non-combustion source of CO_2 and CH_4

Component 6 of the analysis has a strong association with the greenhouse gases CO_2 ($r = 0.77$) and a weak association with CH_4 ($r = 0.59$) but is not associated with tracers of combustion (i.e., NO_y, pPAH, rBC) or naphtha (i.e., anthropogenic VOCs).

A significant amount of carbon is stored in bitumen, which, on geological time scales, conduces formation of CO_2 and CH_4 (i.e., natural gas) reservoirs and pools. When bitumen is mined, substantial emissions of CO_2 and, in particular, of CH_4 occur (Johnson et al., 2016). It is unclear, though, to what extent these greenhouse gases are released from "hot spots" (i.e., from a small number of locations) through surface cracks and fissures in the mine faces, or from new material that is exposed and then releases greenhouse gases during material handling, transport and processing (Johnson et al., 2016). The PCA analysis presented here would be more consistent with the "hot spots" hypothesis since component 6 is not associated with NOy, PAHs, or CO, which are expected to be emitted by the Diesel machinery involved in surface mining (i.e., active disturbance of the bitumen).

Another potential source contribution to component 6 is the degradation of peat and surface soil. Peatland soils, as they occur in the boreal forest surrounding the AMS 13 site, have long been recognized as important contributors to greenhouse gas fluxes and may also be contributing to component 6 (Miller et al., 2014; Gorham, 1991; Warner et al., 2017). The fixation and/or release of CO_2 as well as consumption and/or production of CH_4 through root, anaerobic and aerobic microbial respiration are dependent on soil conditions such as water table position, temperature, soil pH, and plant community composition (Yavitt et al., 2005; Oertel et al., 2016; Whalen, 2005). Emissions from peat and surface soil that was stripped as part of surface mining is expected to release between

317 1.1×10^{10} and 4.7×10^{10} kg stored carbon (Rooney et al., 2012), though it is unclear on what time scale this
318 release will occur. Some of this historical peat material is used for land reclamation. However, a
319 preliminary assessment of greenhouse gas fluxes from such a site gave no indication of significant
320 emissions, at least in the short term (Nwaishi et al., 2016). The bivariate polar plot shows that
321 component 6 is associated with no particular wind direction but with relatively low wind speeds (<
322 1.5 m/s; Figure S-7C), consistent with a dispersed surface source. Further, when variables associated
323 with secondary processes were added to the analysis (Table 7), component 6 anticorrelates with O_x
324 ($r = -0.41$). Dry deposition is a significant O_3 and NO_2, and therefore O_x, loss process (Wesely and Hicks,
325 2000; Zhang et al., 2002).

326 Overall, we have too little information to constrain soil fluxes for this data set. Considering the large CH_4
327 and CO_2 concentrations observed in this study, it is more likely that anthropogenic sources dominate
328 over natural soil emissions (Thompson et al., 2017). Future field campaigns at AMS 13 would benefit
329 from N_2O measurements to constrain contributions of natural sources to greenhouse gas
330 concentrations, such as those produced by microbes in water-logged soil.

331

332 **Component 7: Wind-blown dust**

333 Component 7 is correlated with PM_{10-1} ($r = 0.93$) and poorly with CO_2 ($r = 0.25$), HOA ($r = 0.23$), and LO-
334 OOA ($r = 0.25$). In the Athabasca oil sands region, surface mining has created large portions of land
335 whose surface is void of vegetation and is covered by sand and soil particles, which are readily
336 suspended by wind and vehicle traffic. Other mining activities add to the PM_{10-1} emissions, including
337 combustion processes, tailings sands, and mine haul roads, though the contributions of each of these to
338 the overall PM_{10-1} burden is uncertain (Wang et al., 2015). Recently, Phillips-Smith et al. investigated
339 metal species found in $PM_{2.5}$ aerosol at AMS 13 and found haul road dust and soil from mine faces to be

important sources of $PM_{2.5}$ (Phillips-Smith et al., 2017) and, likely, PM_{10-1} as well. The very poor associations of this component with CO_2 and CH_4 and lack of association with NO_y ($r = 0.02$) suggest contributions of open mine face soil in addition to dust suspended by vehicles travelling on unpaved roads.

The size range captured by PM_{10-1} may also include bioaerosol, including bacteria, fungal spores and plant pollen, which constitute the "natural" background aerosol over vegetated continental regions, typically contributing a few μg m^{-3} of aerosol mass (Huffman et al., 2010). Considering the large PM_{10-1} volumes observed in this work (Table 3), the contribution of bioaerosol is likely minor.

Component 8: Ammonia

Component 8 is a single variable component strongly associated with NH_3 ($r = 0.94$) but with no other variables.

Bytnerowicz et al. (2010) reported larger concentrations of NH_3 in the oil sands region than the provincial average. More recently, Shephard et al. (2015) reported enhancements of NH_3 in the general area as judged from satellite observations. Both studies hence suggest the existence of anthropogenic sources, though Shephard et al. (2015) speculated that biomass burning can contribute to the ammonia burden in the region. A recent modelling study by Whaley et al. (2018) estimated that around half of near-surface NH_3 during the study was likely from bi-directional exchange (i.e., re-emission from soil and plants).

In the oil sands, NH_3 is used as a floating agent for the separation and recovery of bitumen from tar, during bitumen upgrading in a process called "hydrotreating", and in tailing ponds, which, on occasion, have been contaminated with NH_3 to such a degree that they outgas it (Bytnerowicz et al., 2010).

Ammonia is also used for flue gas de-sulfurization by Syncrude; emission inventories (NPRI, 2013; ECCC, 2013b) suggest their fugitive emissions are the largest anthropogenic source in the region, though it is not clear if all sources are accurately inventoried.

The lack of association of ammonia with other variables in this component and the bivariate polar plots (Figure S-9) are consistent with an NH_3-specific source profile, such as fugitive emissions from one or more point sources that emit independently from other activities (i.e., ammonia storage tanks) and natural emissions from soil and trees (Whaley et al., 2018).

Component 9: Incomplete hydrocarbon oxidation

Component 9 is another single variable component and strongly correlates with CO (r = 0.87).The conventional interpretation of CO is as a byproduct of incomplete VOC oxidation, as it is found in fossil fuel combustion exhaust or in biomass burning plumes. Component 9, however, is not associated with NO_y (r = -0.08) or CO_2 (r = 0.05), which rules out this conventional interpretation.

Recently, Marey et al. (2015) examined the spatial distribution of CO in Northern Alberta using a combination of satellite and ground station data and found that most CO is derived from anthropogenic sources, biomass burning and the photochemical oxidation of methane and other VOCs. During the 2013 JOSM study, there was no obvious (i.e., tracer) evidence for fire emissions impacting the measurements at AMS 13 (Phillips-Smith et al., 2017), though an impact from distant sources (such as fires located 1,000s of km upwind in British Columbia or Washington State) cannot be entirely ruled out. We therefore interpret component 9 as a VOC oxidation product component.

Component 10: Dry tailings

Component 10 is strongly associated with TRS (r = 0.71) and weakly with o-xylene (r = 0.32) and poorly with IVOCs (r= 0.20). This component is qualitatively similar to component 1, in that the presence of o-xylene suggests emission of naphtha, and the presence of TRS suggests anaerobic sulfur reducing bacteria and methanogens as they occur in tailings ponds (Holowenko et al., 2000; Percy, 2013; Quagraine et al., 2005). However, the absence of correlations with NO_y, rBC, and CO suggests that this source is not in spatial proximity with a continuously operating combustion source. The much poorer correlations of o-xylene, CH_4, and IVOCs than for component 1 suggests that this component is much more "aged", i.e., emits less naphtha and bitumen.

As part of the reclamation process, tailings ponds in the Alberta oil sands region are converted into "composite tailings", which consist of a consolidated alkaline, saline mixture of processed sand, residual bitumen, clay fines, and gypsum ($CaSO_4$). This mixture settles and releases water, forming shallow pools of surface water (Figure 4J). Due to intensive microbial activity, composite tailings deposits are strong sources of H_2S and, likely, other reduced sulfur species (Warren et al., 2016; Bradford et al., 2017). Composite tailings are a source consistent with the emission profile of component 10. The association with TRS is explained by its production from biological activity and the presence of IVOCs by outgassing from the residual bitumen. Syncrude (the company operating closest to AMS 13) has been undertaking a pilot scale wetland reclamation project in the Athabasca Oil Sands Region to allow the development of a fen wetland above composite tailings (Bradford et al., 2017). Component 10 is hence interpreted as a dry tailings pond component, though the confidence in this interpretation is somewhat marginal as judged, for example, from the low eigenvalue of 0.74.

Bivariate polar plots

Bivariate polar plots map a surface using wind direction and wind speed and then model pollutant concentrations. While PCA is good at showing the temporal distribution of sources, bivariate polar plots help to show the spatial distribution of sources.

Figure S-3 shows a sample of variables associated with component 1. This component appears to dominate when winds are from the SSE and E and of moderate wind speeds (2-3 m/s).

Figure S-4 shows a sample of dominant variables associated with component 2. This component appears to dominate when winds are from the E at low wind speeds (1-2 m/s). The map appears to track the location of the Athabasca river and highway 63, corroborating that this source is from vehicular emissions.

Figure S-5 shows a sample of dominant variables associated with component 3. This component appears to dominate when winds are stagnant and local. This is unsurprising because biogenic emissions are expected to be emitted in great concentrations locally since our site is surrounded on all sides by forest.

Figure S-6 shows a sample of dominant variables associated with component 4 (or with component 2 in the secondary processes PCA). This component appears to dominate when winds are moderate (2-3 m/s) and from the SE and E.

Figure S-7 shows a sample of dominant variables associated with component 5. This component appears to dominate when winds are from the E at moderate wind speeds (2-3 m/s).

Figure S-8 shows a sample of dominant variables associated with component 6. This component appears to dominate when winds are stagnant and local. This suggests that this source is biogenic and may be due to emissions from trees.

Figure S-9 shows a sample of dominant variables associated with component 7. This component appears to dominate when winds are from the SE and E at moderate wind speeds (1-3 m/s).

Figure S-10 shows a sample of dominant variables associated with component 8. This component does

not appear to have a specific direction associated with it and is observed in all directions. This component is observed when winds are at moderate to high speeds (2-4 m/s).

Figure S-11 shows a sample of dominant variables associated with component 9. This component is observed when winds are from the S, SE, and E. This component is observed when winds are at low to moderate speeds (1-3 m/s).

Figure S-12 shows a sample of dominant variables associated with component 10. This component is observed when winds are from the SSE. This component is observed when winds are around 1.5 m/s. This source is very likely a point source due to its consistency with wind direction and speed.

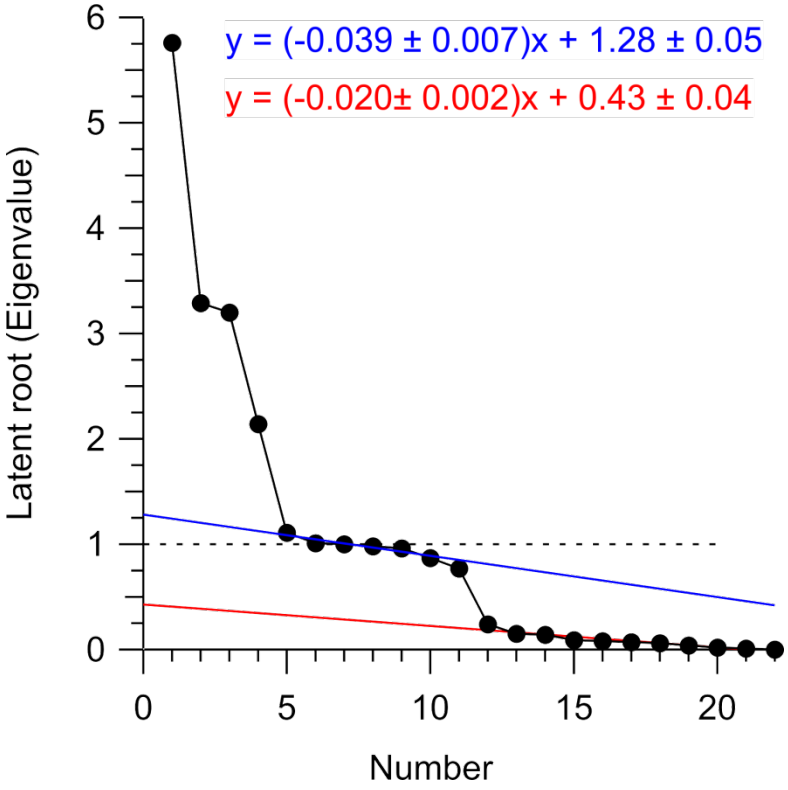


Figure S-2. Scree plot used to consider the number of components to retain. Dashed line represents the latent root criterions (eigenvalues > 1). Blue line represents the first instance of eigenvalues becoming

horizontal. Red line represents the second instance of eigenvalues becoming horizontal.

Table S-1. Ionimed Analytical GCU Standard.

Compound	Volume mixing ratio (ppmv)	Uncertainty (%)
Formaldehyde	1.01	±8
Methanol	1.01	±8
Acetonitrile	1.01	±6
Acetaldehyde	1.01	±5
Ethanol	1.01	±8
Acrolein	0.98	±5
Acetone	1.02	±5
Isoprene	0.99	±5
Crotonaldehyde	0.92	±6
2-Butanone	1.01	±5
Benzene	1.01	±5
Toluene	1.02	±5
o-xylene	1.03	±6
Chlorobenzene	1.02	±5
α-pinene	0.93	±5
1,2, Dichlorobenzene	1.03	±7
1,2,4-Trichlorobenzene	1.01	±9

444 **Table S-2.** The component pattern after Varimax rotation. Correlations greater than 0.30 or less than -0.30 are bolded.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Anthropogenic VOCs																						
o-xylene	**0.89**	0.07	0.03	0.10	0.09	0.06	-0.03	0.11	0.15	0.06	0.26	0.11	0.00	0.04	-0.08	0.24	0.00	-0.01	-0.02	0.00	-0.01	0.00
1,2,3 - TMB	**0.94**	0.15	0.07	0.06	0.05	0.08	-0.01	0.06	0.17	0.01	-0.01	0.02	-0.04	0.01	-0.08	-0.14	0.03	0.00	-0.09	0.00	-0.07	0.00
1,2,4 - TMB	**0.94**	0.13	0.01	0.11	0.08	0.05	-0.02	0.09	0.17	0.04	0.12	0.03	-0.01	0.02	-0.06	0.03	0.01	0.00	-0.04	0.00	0.09	0.00
decane	**0.91**	0.22	-0.02	0.15	0.05	0.00	0.04	0.15	0.05	0.15	0.07	-0.02	0.04	-0.01	0.07	-0.03	-0.03	0.02	0.16	0.00	-0.01	0.00
undecane	**0.85**	0.27	-0.08	0.23	0.08	-0.03	0.06	0.05	0.00	0.20	0.01	-0.10	0.09	0.00	0.26	-0.05	-0.02	0.02	0.02	0.00	0.00	0.00
Biogenic VOCs																						
α-pinene	-0.03	-0.08	**0.98**	-0.11	0.02	0.05	-0.08	0.02	0.01	0.00	-0.01	0.02	0.00	0.00	-0.01	0.01	-0.08	-0.01	0.00	0.09	0.00	0.00
ß-pinene	-0.02	-0.08	**0.97**	-0.12	0.01	0.05	-0.08	0.01	0.01	0.01	0.01	-0.02	0.00	-0.02	0.02	0.00	-0.10	0.01	0.01	-0.09	0.00	0.00
limonene	0.08	-0.02	**0.92**	-0.08	0.06	0.23	-0.11	0.08	0.02	0.07	-0.05	0.02	-0.03	0.03	-0.02	0.00	0.23	0.00	-0.01	0.00	0.00	0.00
Combustion tracers																						
NO_y	0.26	**0.80**	-0.25	0.21	0.03	-0.05	0.10	0.19	-0.04	0.07	0.04	0.01	**0.34**	0.00	0.04	0.00	-0.01	-0.01	0.01	0.00	0.00	0.00
rBC	**0.31**	**0.80**	0.03	0.05	0.08	0.07	0.11	0.24	0.12	**0.34**	-0.03	0.02	-0.02	-0.05	0.02	-0.01	0.00	0.22	0.01	0.00	0.00	0.00
CO	**0.41**	0.18	0.04	0.02	0.08	0.07	0.05	0.03	**0.88**	0.06	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CO_2	0.10	0.09	**0.46**	-0.12	0.23	**0.82**	-0.14	-0.03	0.07	0.00	-0.05	0.02	-0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Aerosol species																						
pPAH	0.07	**0.94**	-0.08	-0.11	0.02	0.06	0.14	0.00	0.09	-0.15	0.00	0.01	-0.14	-0.07	-0.02	0.00	0.00	-0.10	0.00	0.00	0.00	0.00
PM_{10-1}	0.18	0.13	0.07	0.10	**0.94**	0.16	-0.03	0.04	0.07	0.10	0.08	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HOA	**0.41**	**0.74**	0.02	0.11	0.21	0.11	-0.04	0.13	0.15	0.19	0.10	0.05	0.01	**0.35**	0.00	0.01	0.01	-0.01	0.00	0.00	0.00	0.00
LO-OOA	**0.45**	0.17	0.13	0.25	0.19	0.01	-0.04	0.28	0.10	0.73	0.16	0.02	0.01	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sulfur																						
TS	0.25	0.04	-0.16	**0.94**	0.07	-0.05	-0.02	0.02	0.01	0.09	0.14	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01
SO_2	0.11	0.02	-0.15	**0.98**	0.04	-0.04	-0.03	-0.02	0.01	0.05	-0.05	-0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01
TRS	**0.57**	0.05	-0.08	0.11	0.14	-0.05	0.03	0.16	-0.01	0.13	**0.77**	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Other																						
IVOCs	**0.34**	**0.34**	0.12	-0.03	0.05	-0.02	-0.02	**0.84**	0.04	0.18	0.13	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NH_3	0.01	0.19	-0.23	-0.04	-0.03	-0.09	**0.95**	-0.01	0.04	-0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CH_4	**0.60**	**0.39**	0.10	-0.05	0.14	0.44	0.00	0.06	0.16	0.08	0.09	**0.46**	0.01	0.03	-0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Eigenvalues	5.76	3.29	3.20	2.14	1.11	1.01	1.00	0.98	0.96	0.87	0.77	0.24	0.15	0.14	0.09	0.08	0.07	0.06	0.04	0.02	0.01	0.00
% of variance	26.17	14.97	14.55	9.75	5.06	4.60	4.53	4.46	4.34	3.94	3.51	1.09	0.69	0.63	0.42	0.37	0.33	0.28	0.17	0.08	0.06	0.00
% cum. var.	26.17	41.14	55.69	65.43	70.49	75.09	79.62	84.07	88.42	92.36	95.87	96.97	97.66	98.29	98.71	99.08	99.41	99.69	99.86	99.94	100.0	100

445

Table S-3. The pattern after Varimax rotation with 5 components selected.

	1	2	3	4	5	Communalities
Anthropogenic VOCs						
o-xylene	**0.94**	0.08	0.03	0.09	0.15	0.93
1,2,3 - TMB	**0.90**	0.14	0.04	0.01	0.23	0.89
1,2,4 - TMB	**0.95**	0.14	-0.01	0.09	0.18	0.97
decane	**0.91**	0.27	-0.03	0.16	0.05	0.93
undecane	**0.82**	**0.31**	-0.10	0.26	0.05	0.84
Biogenic VOCs						
α-pinene	-0.03	-0.05	**0.94**	-0.15	0.04	0.91
ß-pinene	-0.02	-0.06	**0.94**	-0.15	0.03	0.90
limonene	0.07	0.02	**0.94**	-0.10	0.18	0.93
Combustion tracers						
NO_y	0.25	**0.83**	-0.29	0.22	0.02	0.89
rBC	**0.33**	**0.89**	0.04	0.07	0.13	0.92
CO	**0.53**	0.19	-0.02	-0.08	**0.40**	0.48
CO_2	0.07	0.06	**0.55**	-0.13	**0.71**	0.83
Aerosol species						
pPAH	0.01	**0.83**	-0.20	-0.20	0.27	0.84
PM_{10-1}	0.21	0.19	0.15	0.29	**0.57**	0.51
HOA	**0.44**	**0.75**	0.03	0.15	**0.32**	0.88
LO-OOA	**0.59**	**0.37**	0.28	0.41	-0.06	0.74
Sulfur						
TS	0.28	0.04	-0.18	**0.91**	0.01	0.94
SO_2	0.10	0.01	-0.18	**0.93**	0.05	0.91
TRS	**0.76**	0.10	-0.04	0.17	-0.13	0.63
Other						
IVOCs	**0.47**	**0.61**	0.28	0.03	-0.25	0.73
NH_3	0.02	**0.34**	**-0.47**	-0.21	-0.14	0.40
CH_4	**0.62**	**0.38**	0.12	-0.09	**0.53**	0.84
Eigenvalues	6.42	3.79	3.56	2.34	1.70	
% of variance	29.20	17.25	16.20	10.63	7.74	
Cumulative variance	29.20	46.45	62.65	73.28	81.01	

448 **Table S-4.** The pattern after Varimax rotation with 6 components selected.

	1	2	3	4	5	6	Communalities
Anthropogenic VOCs							
o-xylene	**0.92**	0.04	0.01	0.08	0.24	0.13	0.93
1,2,3 - TMB	**0.94**	0.16	0.08	0.07	0.06	0.05	0.92
1,2,4 - TMB	**0.95**	0.13	0.00	0.11	0.16	0.08	0.97
decane	**0.88**	0.23	-0.02	0.18	0.28	0.00	0.94
undecane	**0.79**	0.28	-0.09	0.29	0.25	0.00	0.85
Biogenic VOCs							
α-pinene	-0.02	-0.09	**0.96**	-0.12	0.04	0.01	0.95
ß-pinene	-0.02	-0.10	**0.96**	-0.12	0.05	0.01	0.94
limonene	0.09	-0.01	**0.95**	-0.09	0.04	0.15	0.95
Combustion tracers							
NO_y	0.22	**0.81**	-0.27	0.23	0.25	0.00	0.89
rBC	**0.31**	**0.85**	0.07	0.08	0.28	0.10	0.92
CO	**0.64**	0.29	0.09	0.02	-0.23	0.09	0.56
CO_2	0.17	0.12	**0.57**	-0.17	-0.23	**0.62**	0.84
Aerosol species							
pPAH	0.09	**0.90**	-0.10	-0.14	-0.10	0.05	0.87
PM_{10-1}	0.19	0.14	0.03	0.12	0.17	**0.81**	0.76
HOA	**0.44**	**0.73**	0.04	0.13	0.22	**0.30**	0.88
LO-OOA	**0.46**	0.22	0.17	**0.32**	**0.60**	0.21	0.79
Sulfur							
TS	0.25	0.02	-0.18	**0.92**	0.12	0.06	0.97
SO_2	0.10	0.03	-0.15	**0.97**	-0.02	0.02	0.98
TRS	**0.62**	-0.04	-0.17	0.06	**0.56**	0.14	0.75
Other							
IVOCs	**0.31**	**0.43**	0.17	-0.06	**0.70**	0.02	0.80
NH_3	0.06	**0.41**	**-0.36**	-0.11	-0.11	**-0.36**	0.46
CH_4	**0.68**	**0.42**	0.15	-0.10	0.00	**0.41**	0.84
Eigenvalues	6.09	3.60	3.47	2.25	1.76	1.58	
% of variance	27.70	16.38	15.78	10.22	7.98	7.18	
Cumulative variance	27.70	44.07	59.85	70.08	78.06	85.23	

449

Table S-5. The pattern after Varimax rotation with 7 components selected.

	1	2	3	4	5	6	7	Commu-nalities
Anthropogenic VOCs								
o-xylene	**0.93**	0.07	0.01	0.08	0.21	0.12	-0.05	0.93
1,2,3 - TMB	**0.94**	0.18	0.08	0.06	0.03	0.03	-0.03	0.93
1,2,4 - TMB	**0.96**	0.15	0.00	0.11	0.14	0.07	-0.02	0.98
decane	**0.88**	0.26	-0.02	0.18	0.26	0.00	0.01	0.94
undecane	**0.79**	**0.30**	-0.09	0.28	0.23	0.00	0.03	0.85
Biogenic VOCs								
α-pinene	-0.03	-0.09	**0.96**	-0.11	0.04	0.00	-0.05	0.96
ß-pinene	-0.02	-0.10	**0.96**	-0.11	0.06	0.01	-0.05	0.95
limonene	0.09	0.01	**0.95**	-0.08	0.03	0.12	-0.13	0.95
Combustion tracers								
NO_y	0.22	**0.83**	-0.28	0.22	0.20	-0.02	0.06	0.91
rBC	**0.31**	**0.85**	0.07	0.08	0.25	0.11	0.13	0.92
CO	**0.62**	0.22	0.13	0.03	-0.21	0.19	0.29	0.61
CO_2	0.17	0.17	**0.56**	-0.18	-0.28	**0.54**	-0.27	0.84
Aerosol species								
pPAH	0.08	**0.89**	-0.10	-0.15	-0.14	0.03	0.15	0.88
PM_{10-1}	0.18	0.12	0.06	0.13	0.18	**0.89**	0.01	0.89
HOA	**0.44**	**0.76**	0.03	0.12	0.16	0.27	-0.03	0.89
LO-OOA	**0.46**	0.23	0.18	**0.33**	**0.59**	0.25	0.01	0.81
Sulfur								
TS	0.25	0.04	-0.18	**0.92**	0.12	0.06	-0.02	0.97
SO_2	0.10	0.04	-0.15	**0.97**	-0.02	0.02	-0.03	0.98
TRS	**0.62**	-0.03	-0.16	0.07	**0.56**	0.19	0.02	0.77
Other								
IVOCs	**0.32**	**0.48**	0.16	-0.07	**0.66**	0.00	-0.05	0.80
NH_3	0.01	0.20	-0.26	-0.06	0.00	-0.05	**0.89**	0.91
CH_4	**0.68**	**0.45**	0.15	-0.11	-0.05	**0.36**	-0.10	0.85
Eigenvalues	6.09	3.65	3.42	2.23	1.61	1.46	1.03	
% of variance	27.70	16.60	15.57	10.14	7.33	6.66	4.70	
Cumulative variance	27.70	44.30	59.86	70.01	77.34	84.00	88.70	

Table S-6. The pattern after Varimax rotation with 8 components selected.

	1	2	3	4	5	6	7	8	Communalities
Anthropogenic VOCs									
o-xylene	**0.93**	0.08	0.03	0.07	0.14	0.13	-0.03	0.11	0.93
1,2,3 - TMB	**0.91**	0.19	0.09	0.06	-0.01	0.02	-0.04	0.21	0.93
1,2,4 - TMB	**0.95**	0.16	0.01	0.11	0.09	0.07	-0.03	0.18	0.98
decane	**0.90**	0.26	-0.01	0.17	0.19	0.01	0.03	0.08	0.95
undecane	**0.82**	**0.30**	-0.07	0.28	0.15	0.01	0.06	0.03	0.87
Biogenic VOCs									
α-pinene	-0.04	-0.10	**0.97**	-0.11	0.06	0.00	-0.05	0.00	0.96
ß-pinene	-0.02	-0.10	**0.96**	-0.11	0.06	0.00	-0.04	-0.01	0.96
limonene	0.07	0.00	**0.95**	-0.08	0.06	0.11	-0.14	0.05	0.95
Combustion tracers									
NO_y	0.25	**0.83**	-0.26	0.22	0.19	-0.01	0.10	-0.08	0.92
rBC	0.28	**0.83**	0.06	0.07	**0.33**	0.10	0.09	0.17	0.93
CO	**0.42**	0.18	0.04	0.01	0.07	0.11	0.06	**0.85**	0.96
CO_2	0.13	0.19	**0.58**	-0.17	-0.28	**0.53**	-0.27	0.10	0.86
Aerosol species									
pPAH	0.06	**0.91**	-0.08	-0.14	-0.11	0.03	0.16	0.06	0.89
PM_{10-1}	0.18	0.12	0.07	0.12	0.16	**0.89**	0.01	0.06	0.89
HOA	**0.42**	**0.75**	0.03	0.12	0.21	0.26	-0.06	0.15	0.89
LO-OOA	**0.46**	0.19	0.15	**0.30**	**0.65**	0.24	-0.04	0.15	0.87
Sulfur									
TS	0.28	0.03	-0.18	**0.92**	0.09	0.07	-0.01	-0.02	0.97
SO_2	0.11	0.04	-0.15	**0.97**	-0.01	0.02	-0.04	0.03	0.98
TRS	**0.72**	-0.03	-0.13	0.06	**0.40**	0.23	0.11	-0.20	0.80
Other									
IVOCs	**0.35**	**0.43**	0.13	-0.09	**0.71**	0.00	-0.07	0.01	0.84
NH_3	0.01	0.22	-0.24	-0.05	-0.05	-0.04	**0.92**	0.05	0.96
CH_4	**0.65**	**0.47**	0.17	-0.11	-0.08	**0.35**	-0.10	0.16	0.86
Eigenvalues	5.99	3.58	3.40	2.20	1.52	1.43	1.03	1.00	
% of variance	27.23	16.28	15.44	10.01	6.90	6.52	4.70	4.53	
Cumulative variance	27.23	43.51	58.95	68.96	75.86	82.37	87.08	91.61	

Table S-7. The pattern after Varimax rotation with 9 components selected.

	1	2	3	4	5	6	7	8	9	Communalities
Anthropogenic VOCs										
o-xylene	**0.89**	0.09	0.03	0.09	0.11	0.11	-0.05	0.16	**0.30**	0.95
1,2,3 - TMB	**0.93**	0.16	0.08	0.05	0.05	0.05	-0.02	0.17	-0.02	0.94
1,2,4 - TMB	**0.94**	0.15	0.01	0.11	0.11	0.08	-0.02	0.18	0.12	0.98
decane	**0.91**	0.22	-0.03	0.16	0.25	0.04	0.05	0.03	0.03	0.97
undecane	**0.85**	0.25	-0.10	0.25	0.24	0.06	0.10	-0.05	-0.08	0.94
Biogenic VOCs										
α-pinene	-0.04	-0.09	**0.97**	-0.10	0.05	0.00	-0.06	0.02	0.02	0.97
ß-pinene	-0.03	-0.10	**0.97**	-0.11	0.05	0.00	-0.05	0.00	0.02	0.96
limonene	0.09	-0.01	**0.94**	-0.09	0.08	0.13	-0.13	0.03	-0.06	0.95
Combustion tracers										
NO_y	0.26	**0.82**	-0.26	0.22	0.22	0.00	0.11	-0.09	0.02	0.92
rBC	**0.31**	**0.79**	0.04	0.05	**0.41**	0.12	0.12	0.12	-0.10	0.94
CO	**0.42**	0.19	0.05	0.02	0.08	0.10	0.06	**0.87**	-0.02	0.98
CO_2	0.16	0.17	**0.56**	-0.18	-0.22	**0.56**	-0.25	0.06	-0.17	0.86
Aerosol species										
pPAH	0.06	**0.93**	-0.07	-0.12	-0.11	0.01	0.14	0.09	0.03	0.93
PM_{10-1}	0.16	0.11	0.06	0.12	0.17	**0.89**	0.02	0.06	0.10	0.89
HOA	**0.41**	**0.75**	0.03	0.13	0.23	0.25	-0.06	0.16	0.08	0.90
LO-OOA	**0.46**	0.14	0.13	0.28	0.70	0.26	-0.01	0.10	0.05	0.90
Sulfur										
TS	0.25	0.04	-0.17	**0.93**	0.08	0.06	-0.02	0.00	0.13	0.99
SO_2	0.11	0.03	-0.15	**0.97**	0.01	0.02	-0.03	0.01	-0.05	0.99
TRS	**0.59**	0.05	-0.09	0.11	0.24	0.14	0.04	-0.04	**0.71**	0.96
Other										
IVOCs	**0.32**	**0.41**	0.12	-0.09	**0.70**	-0.01	-0.08	0.02	0.20	0.84
NH_3	0.01	0.21	-0.24	-0.05	-0.04	-0.04	**0.93**	0.04	0.01	0.97
CH_4	**0.65**	**0.47**	0.16	-0.10	-0.06	**0.36**	-0.10	0.17	0.07	0.86
Eigenvalues	5.84	3.44	3.37	2.19	1.58	1.47	1.03	0.95	0.74	
% of variance	26.54	15.63	15.30	9.98	7.19	6.66	4.69	4.31	3.38	
Cumulative variance	26.54	42.17	57.47	67.44	74.63	81.29	85.98	90.29	93.67	

Table S-8. The factor pattern after Varimax rotation with 11 factors selected.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	Factor 11	Commu-nalities
Anthropogenic VOCs												
o-xylene	**0.88**	0.08	0.03	0.10	0.13	0.07	-0.04	0.17	0.11	0.04	**0.32**	0.95
1,2,3 - TMB	**0.94**	0.16	0.07	0.05	0.11	0.05	-0.01	0.18	0.06	-0.01	-0.02	0.96
1,2,4 - TMB	**0.94**	0.15	0.01	0.11	0.08	0.08	-0.02	0.18	0.09	0.03	0.13	0.99
decane	**0.92**	0.24	-0.02	0.15	0.00	0.05	0.04	0.04	0.13	0.16	0.05	0.97
undecane	**0.87**	0.29	-0.08	0.22	-0.06	0.09	0.05	-0.05	0.03	0.22	-0.05	0.96
Biogenic VOCs												
α-pinene	-0.03	-0.08	**0.98**	-0.11	0.04	0.01	-0.08	0.02	0.02	0.00	0.00	0.98
ß-pinene	-0.02	-0.08	**0.98**	-0.12	0.02	0.02	-0.07	0.00	0.00	0.02	0.01	0.98
limonene	0.08	-0.02	**0.93**	-0.08	0.24	0.05	-0.11	0.03	0.09	0.06	-0.05	0.95
Combustion tracers												
NO_y	0.27	**0.83**	-0.26	0.21	-0.04	0.03	0.10	-0.08	0.18	0.07	0.01	0.92
rBC	**0.30**	**0.81**	0.04	0.04	0.09	0.07	0.12	0.12	0.23	**0.34**	-0.05	0.95
CO	**0.41**	0.19	0.04	0.02	0.08	0.08	0.05	**0.87**	0.03	0.06	-0.01	0.99
CO_2	0.09	0.08	**0.48**	-0.12	**0.77**	0.25	-0.14	0.05	-0.04	-0.01	-0.09	0.95
Aerosol species												
pPAH	0.06	**0.93**	-0.07	-0.12	0.07	0.02	0.14	0.10	-0.02	-0.20	0.01	0.95
PM_{10-1}	0.18	0.14	0.08	0.10	0.17	**0.94**	-0.03	0.07	0.04	0.09	0.08	1.00
HOA	**0.40**	**0.77**	0.03	0.11	0.14	0.20	-0.07	0.16	0.09	0.19	0.13	0.92
LO-OOA	**0.45**	0.19	0.13	0.25	0.03	0.19	-0.04	0.11	0.27	**0.72**	0.16	0.98
Sulfur												
TS	0.26	0.05	-0.16	**0.93**	-0.05	0.07	-0.02	0.01	0.02	0.08	0.13	0.26
SO_2	0.12	0.03	-0.15	**0.98**	-0.04	0.04	-0.03	0.01	-0.02	0.05	-0.05	0.12
TRS	**0.58**	0.06	-0.08	0.10	-0.04	0.14	0.03	-0.03	0.16	0.13	**0.74**	0.58
Other												
IVOCs	**0.34**	**0.37**	0.13	-0.03	-0.01	0.05	-0.03	0.03	**0.82**	0.18	0.12	1.00
NH_3	0.01	0.20	-0.24	-0.04	-0.08	-0.03	**0.94**	0.04	-0.02	-0.01	0.01	1.00
CH_4	**0.59**	**0.40**	0.10	-0.06	**0.59**	0.10	0.00	0.17	0.05	0.07	0.15	0.93
Eigenvalues	5.75	3.43	3.24	2.13	1.12	1.10	0.99	0.96	0.93	0.87	0.79	
% var.	26.14	15.61	14.72	9.66	5.11	4.99	4.51	4.35	4.22	3.95	3.60	
% Cum. var.	26.14	41.74	56.46	66.12	71.23	76.22	80.73	85.09	89.31	93.26	96.86	

Table S-9. The pattern with mixing height included after Varimax rotation with 10 components.

	1	2	3	4	5	6	7	8	9	10	Commu-nalities
Anthropogenic VOCs											
o-xylene	**0.89**	0.04	0.03	0.10	0.29	0.10	-0.01	-0.06	0.17	0.16	0.95
1,2,3 - TMB	**0.94**	0.17	0.10	0.04	-0.04	0.01	-0.03	-0.04	0.17	0.06	0.95
1,2,4 - TMB	**0.94**	0.13	0.03	0.10	0.11	0.09	-0.02	-0.06	0.18	0.07	0.98
decane	**0.92**	0.25	0.03	0.15	0.11	0.11	0.01	0.04	0.03	-0.04	0.97
undecane	**0.87**	**0.31**	-0.05	0.23	-0.03	0.17	0.03	0.10	-0.05	-0.11	0.96
Biogenic VOCs											
α-pinene	-0.02	-0.08	**0.96**	-0.10	0.03	0.01	-0.05	-0.11	0.01	0.02	0.96
ß-pinene	-0.01	-0.08	**0.96**	-0.11	0.04	0.02	-0.05	-0.12	-0.01	0.02	0.96
limonene	0.11	0.02	**0.95**	-0.08	0.04	0.02	-0.11	-0.06	0.03	0.12	0.96
Combustion tracers											
NO_y	0.21	**0.86**	-0.25	0.21	0.11	0.06	0.10	0.01	-0.08	-0.02	0.92
rBC	0.29	**0.89**	0.12	0.02	0.10	0.19	0.03	0.09	0.10	-0.01	0.95
CO	**0.43**	0.20	0.04	0.01	0.00	0.08	0.02	0.05	**0.86**	0.07	0.98
CO_2	0.15	0.17	**0.56**	-0.13	-0.12	0.13	-0.14	-0.12	0.08	**0.68**	0.91
Aerosol species											
pPAH	0.01	**0.86**	-0.09	-0.13	-0.08	-0.03	0.23	-0.18	0.12	0.17	0.90
PM_{10-1}	**0.31**	0.22	0.05	0.15	0.12	**0.88**	0.04	-0.01	0.08	0.10	0.97
HOA	**0.45**	**0.79**	0.02	0.14	0.16	0.15	-0.02	-0.02	0.16	0.10	0.93
LO-OOA	**0.52**	**0.30**	0.21	0.26	**0.37**	**0.36**	-0.17	0.27	0.04	-0.18	0.88
Sulfur											
TS	0.27	0.06	-0.16	**0.93**	0.10	0.10	-0.03	0.04	-0.01	-0.02	1.00
SO_2	0.11	0.05	-0.14	**0.97**	-0.06	0.06	-0.05	0.05	0.01	-0.05	0.99
TRS	**0.64**	0.01	-0.12	0.11	**0.63**	0.20	0.09	-0.04	-0.09	0.12	0.91
Other											
IVOCs	0.28	**0.50**	0.22	-0.07	**0.66**	0.08	-0.13	0.04	0.06	-0.18	0.87
NH_3	0.00	0.22	-0.20	-0.07	-0.03	0.03	**0.92**	0.11	0.02	-0.06	0.96
CH_4	**0.64**	**0.43**	0.15	-0.07	0.09	0.10	0.00	-0.07	0.17	**0.50**	0.92
Mixing height	-0.04	-0.07	**-0.35**	0.07	0.01	0.00	0.12	**0.90**	0.04	-0.07	0.96
Eigenvalues	6.06	3.81	3.51	2.16	1.19	1.12	1.03	1.02	0.94	0.92	
% var.	26.35	16.57	15.27	9.39	5.19	4.86	4.48	4.42	4.08	3.98	
% Cum. var.	26.35	42.92	58.19	67.58	72.77	77.63	82.11	86.53	90.61	94.60	

Table S-10. The pattern without aerosol variables after Varimax rotation with 9 components.

	1	2	3	4	5	6	7	8	9	Commu-nalities
Anthropogenic VOCs										
o-xylene	**0.88**	0.02	0.04	0.09	0.12	0.10	-0.03	0.17	**0.32**	0.94
1,2,3 - TMB	**0.94**	0.07	0.12	0.04	0.11	0.03	-0.02	0.18	-0.02	0.95
1,2,4 - TMB	**0.94**	0.01	0.11	0.10	0.08	0.08	-0.02	0.19	0.13	0.98
decane	**0.93**	-0.01	0.20	0.16	0.01	0.18	0.04	0.04	0.06	0.97
undecane	**0.89**	-0.07	0.26	0.25	-0.03	0.12	0.06	-0.04	-0.03	0.94
Biogenic VOCs										
α-pinene	-0.03	**0.97**	-0.08	-0.12	0.06	0.01	-0.08	0.02	0.00	0.98
ß-pinene	-0.02	**0.97**	-0.08	-0.12	0.05	0.00	-0.08	0.00	0.01	0.98
limonene	0.08	**0.92**	-0.04	-0.08	0.27	0.10	-0.11	0.03	-0.06	0.95
Combustion tracers										
NO_y	0.30	-0.25	**0.81**	0.23	-0.03	0.24	0.09	-0.06	0.03	0.92
rBC	**0.34**	0.04	**0.78**	0.08	0.12	**0.37**	0.11	0.13	-0.04	0.92
CO	**0.42**	0.04	0.16	0.03	0.10	0.05	0.05	**0.87**	-0.01	0.98
CO_2	0.10	**0.46**	0.06	-0.10	**0.84**	-0.02	-0.13	0.06	-0.05	0.96
Aerosol species										
pPAH	0.07	-0.07	**0.94**	-0.13	0.08	-0.06	0.11	0.12	0.03	0.95
Sulfur										
TS	0.26	-0.15	0.03	**0.94**	-0.05	0.03	-0.02	0.01	0.14	1.00
SO_2	0.12	-0.14	0.02	**0.98**	-0.05	-0.03	-0.03	0.02	-0.04	0.99
TRS	**0.59**	-0.07	0.04	0.14	-0.01	0.19	0.02	-0.02	**0.75**	0.97
Other										
IVOCs	**0.35**	0.13	**0.32**	-0.03	0.00	**0.84**	-0.03	0.05	0.15	0.98
NH_3	0.01	-0.23	0.21	-0.05	-0.10	-0.01	**0.94**	0.05	0.01	1.00
CH_4	**0.61**	0.09	**0.36**	-0.05	**0.59**	0.08	0.00	0.18	0.16	0.92
Eigenvalues	5.54	3.16	2.60	2.08	1.20	1.03	0.97	0.94	0.77	
% var.	29.15	16.63	13.68	10.96	6.33	5.40	5.11	4.96	4.03	
% Cum. var.	29.15	45.79	59.46	70.43	76.76	82.16	87.26	92.23	96.25	

462 **Table S-~~10~~11.** Criteria for number of components extracted by PCA.

Criterion	Number of components extracted	% Variance explained (after rotation)
Scree test 1	5	81.0%
< 5% variance	6	85.2%
Latent root	7	88.7%
≥ 95% cumulative variance	10	95.5%
Scree test 2	11	96.9%

463

464

465 **Table S-12~~1~~.** Association of IVOCs with relevant components.

# of components in solution	Oil sands surface mining facilities (Component 1)	Mine fleet and operations (Component 2)	Mine face (Component 5)
5	0.47	0.61	n/a
6	0.31	0.43	n/a
7	0.32	0.48	0.66
8	0.35	0.43	0.71
9	0.32	0.41	0.70
10	0.31	0.39	0.74
11	0.34	0.37	n/a

466

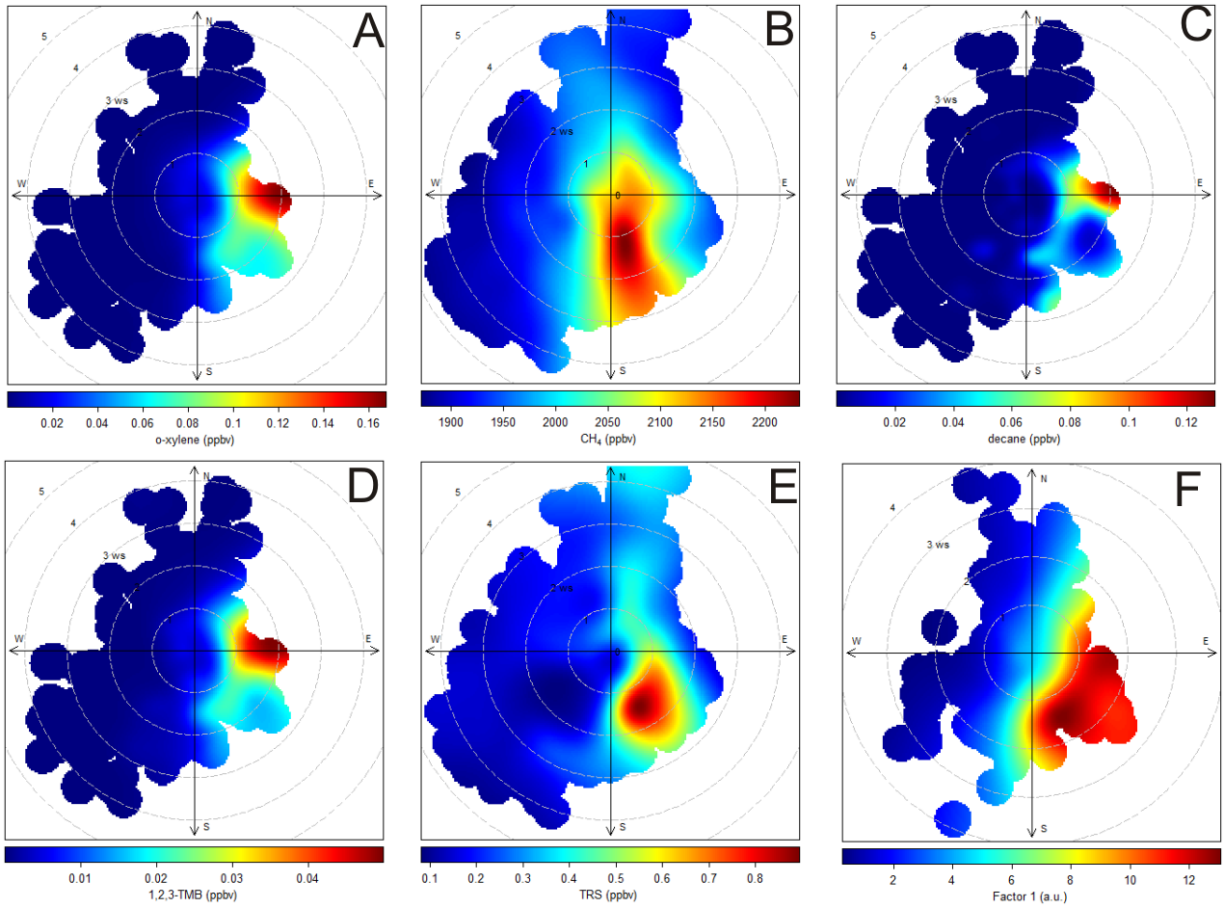


Figure S-3. Bivariate polar plots associated with component 1 for the optimum primary pollutant solution (Table 5.). **(A)** o-xylene, **(B)** CH_4, **(C)** decane, **(D)** 1, 2, 3-TMB, **(E)** TRS, **(F)** and component 1.

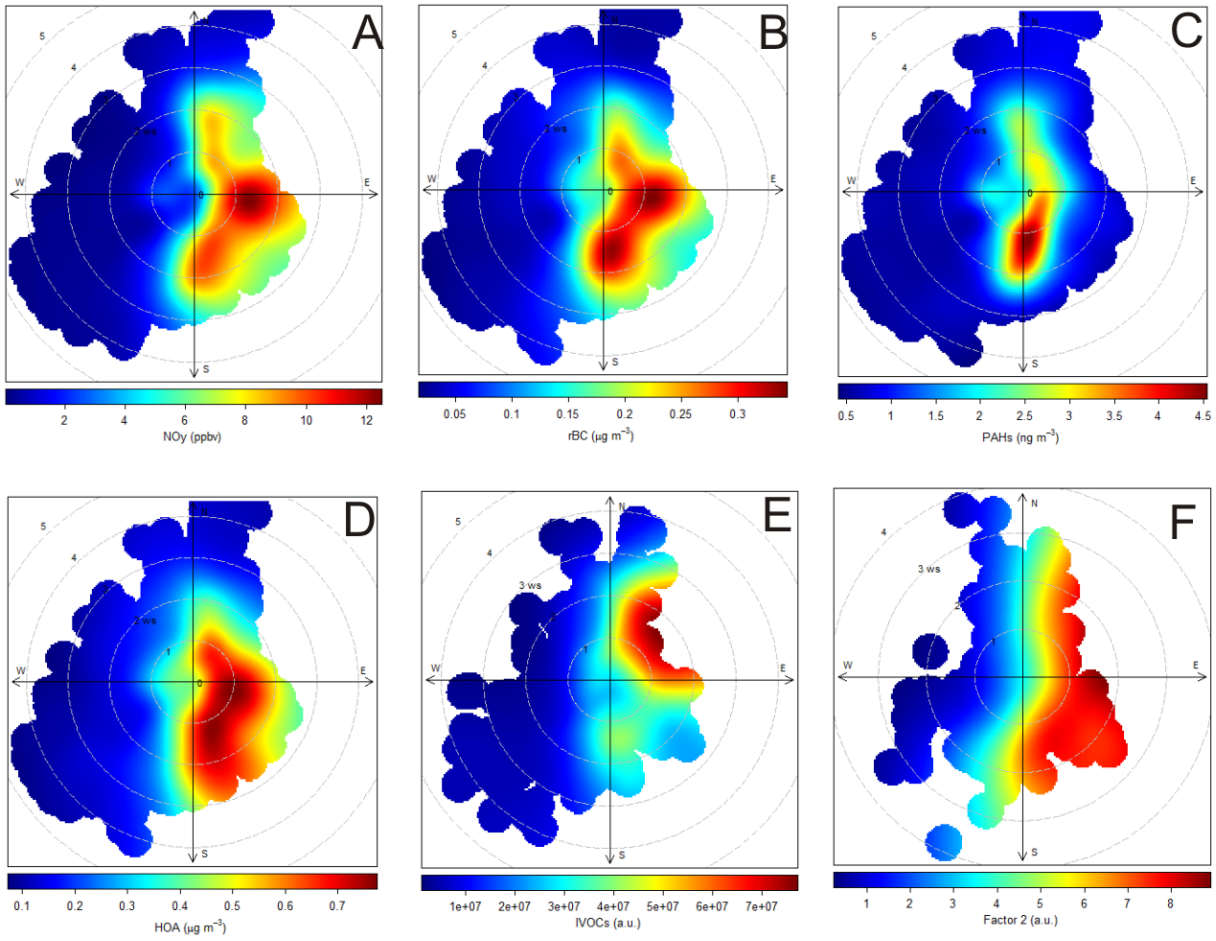


Figure S-4. Bivariate polar plots associated with component 2 for the optimum primary pollutant solution (Table 5.). **(A)** NO_y, **(B)** rBC, **(C)** PAHs, **(D)** HOA, **(E)** IVOCs, **(F)** and component 2.

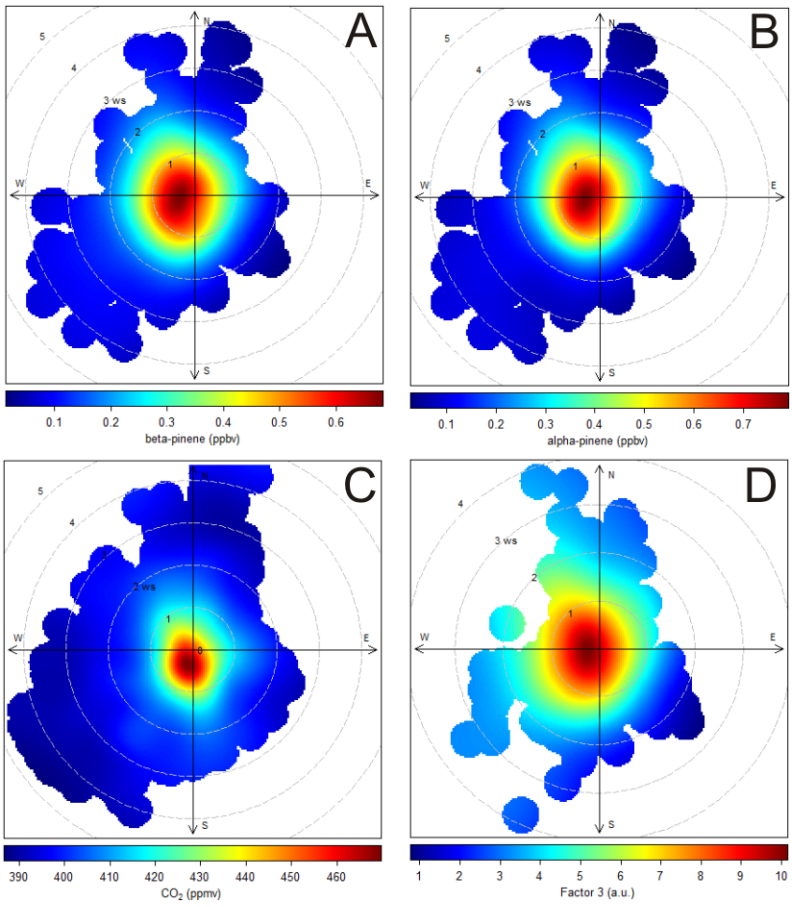


Figure S-5. Bivariate polar plots associated with component 3 for the optimum primary pollutant solution (Table 5.). **(A)** ß-pinene, **(B)** α-pinene, **(C)** CO_2, **(D)** and component 3.

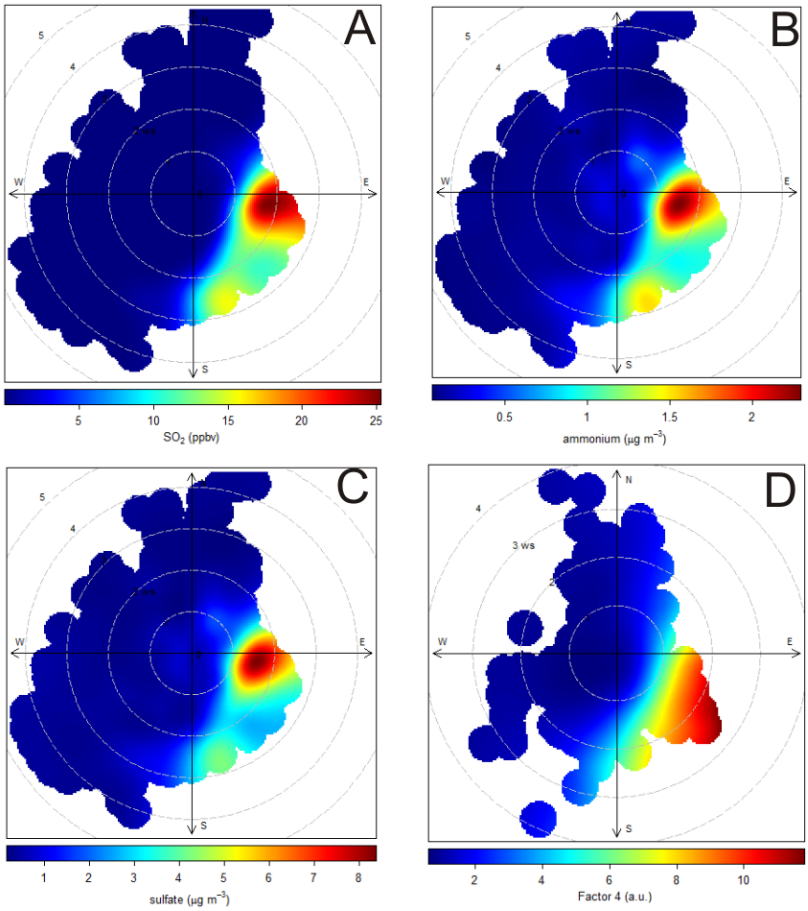


Figure S-6. Bivariate polar plots associated with component 4 for the optimum secondary pollutant solution (Table 7). **(A)** SO_2, **(B)** $NH_4^+{}_{(p)}$, **(C)** $SO_4^{2-}{}_{(p)}$, **(D)** and component 4.

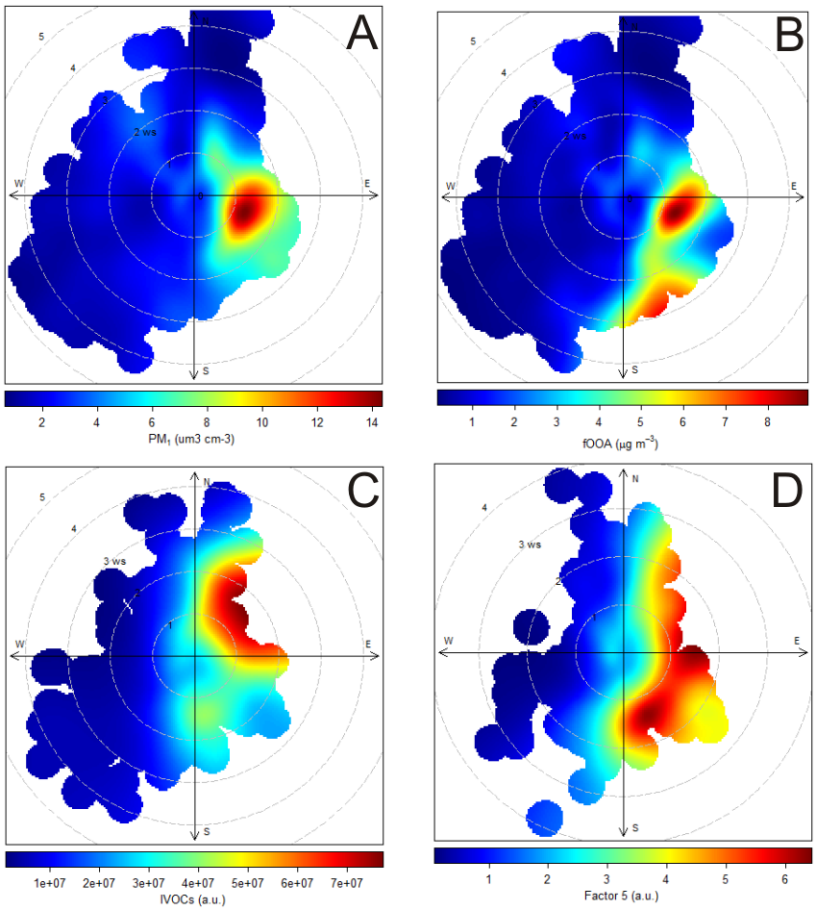


Figure S-7. Bivariate polar plots associated with component 5 for the optimum secondary pollutant solution (Table 7). **(A)** PM_1 (11-component solution), **(B)** LO-OOA, **(C)** IVOCs, and **(D)** component 5.

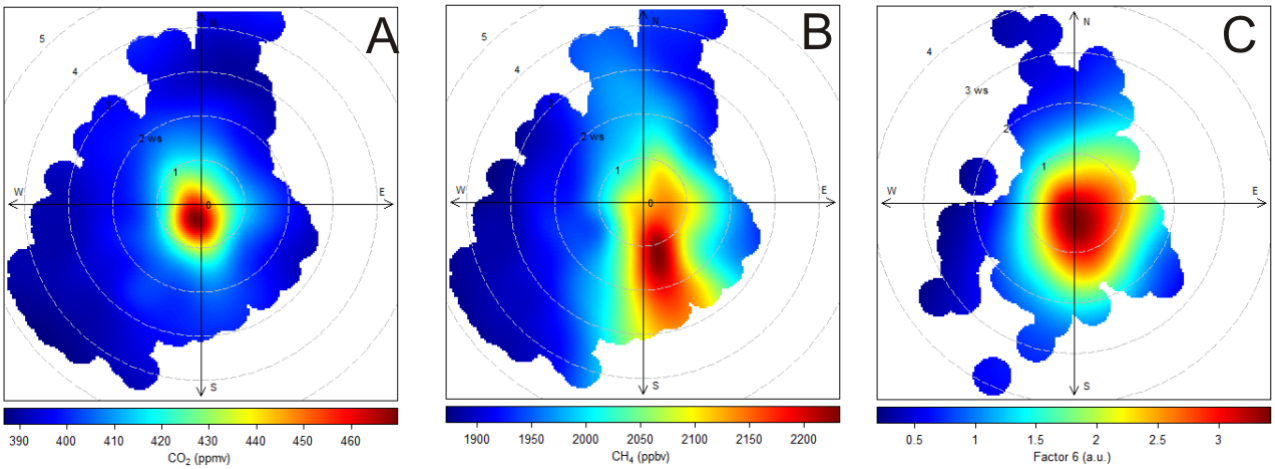


Figure S-8. Bivariate polar plots associated with component 6 for the optimum primary pollutant solution (Table 5). **(A)** CO_2, **(B)** CH_4, and **(C)** component 6.

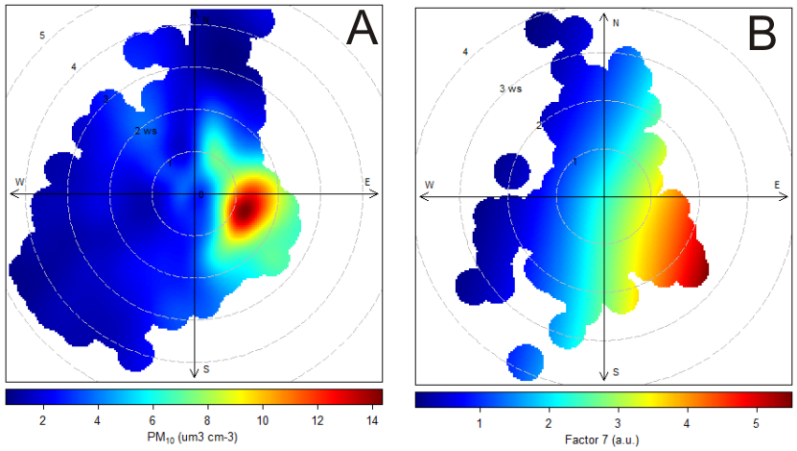


Figure S-9. Bivariate polar plots associated with component 7 for the optimum primary pollutant solution (Table 5). **(A)** PM_{10-1}, **(B)** and component 7.

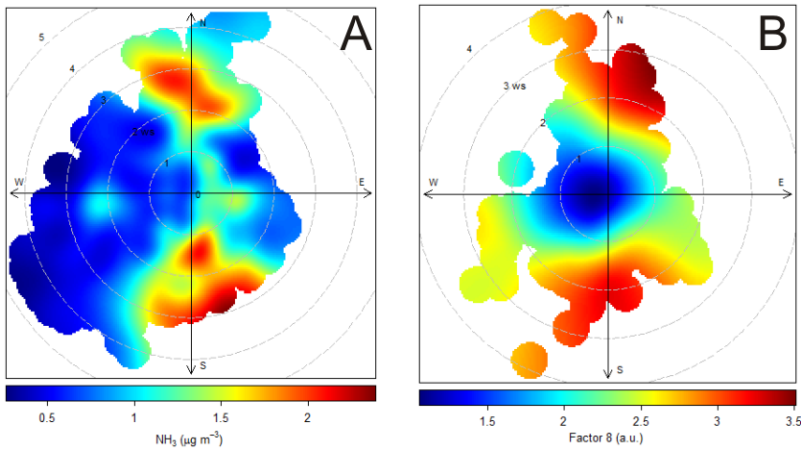


Figure S-10. Bivariate polar plots associated with component 8 for the optimum primary pollutant solution (Table 5). **(A)** NH_3, **(B)** and component 8.

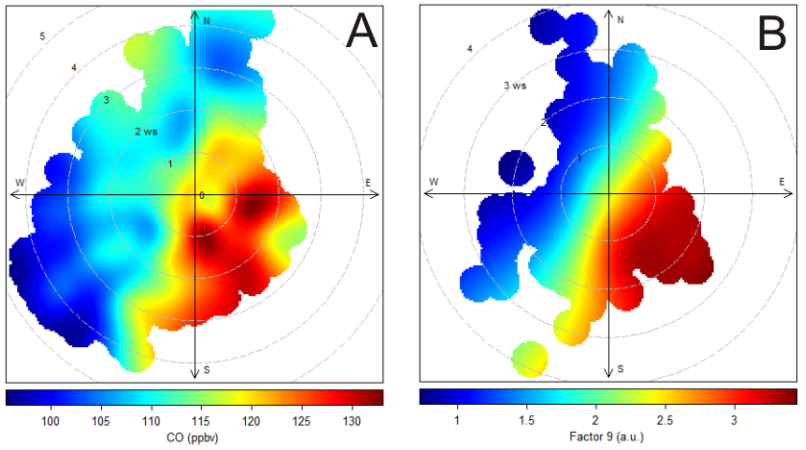


Figure S-11. Bivariate polar plots associated with component 9 for the optimum primary pollutant solution (Table 5). **(A)** CO, and **(B)** component 9.

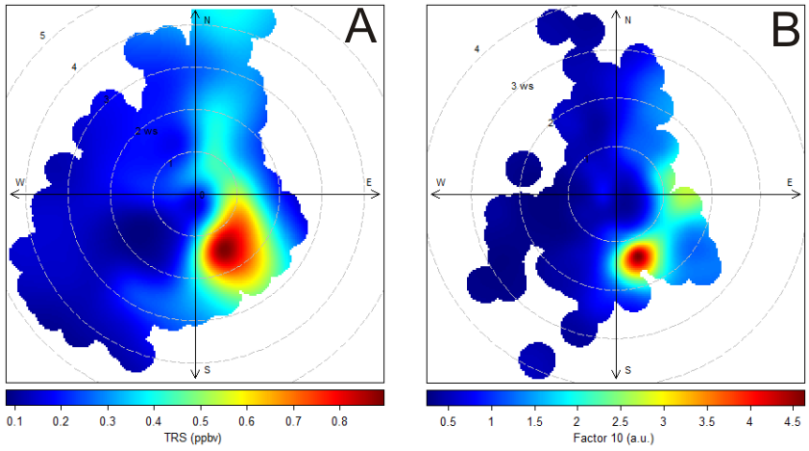


Figure S-12. Bivariate polar plots associated with component 10 for the optimum primary pollutant solution (Table 5). **(A)** TRS, **(B)** and component 10.

Estimate of photochemical age

Photochemical age was calculated using the method outlined by Borbon et al. (2013), but substituting n-decane for benzene since the latter was not quantified. The photochemical age of an air mass, Δt was calculated from the observed concentrations of 124-trimethylbenzene (124TMB) and n-decane using:

$$\Delta t = \frac{1}{[OH]\times(k_{124TMB}-k_{decane})} \times \left[\ln\left(\frac{[124TMB]}{[decane]}\right)_{t=0} - \ln\left(\frac{[124TMB]}{[decane]}\right)\right] \quad \text{(S-1)}$$

where k_{124TMB} = 3.25×10^{-11} cm^3 molecule^{-1} s^{-1} and k_{decane} = 1.10×10^{-11} cm^3 molecule^{-1} s^{-1} are rate coefficients for reaction of OH with 124-TMB and n-decane (at 298 K), respectively, whose values were taken from Seinfeld and Pandis (2006). The ratio of [124TMB] to [decane] at the point of emission (time t = 0) was estimated from a plot of [124TMB] to [n-decane] (Figure S-13, left-hand side) and a straight-line fit to the nocturnal data (assumed to be unaffected by oxidation and shown in blue color). The slope of this line was 1.15±0.07 (r^2 = 0.84). Daytime data (color-coded by solar zenith angle, SZA) exhibit lower ratios of [124TMB]/[decane] as a result of the faster oxidation of 124TMB by OH.

Shown in Figure S-13 on the right-hand side is a plot of the photochemical age, calculated using equation (S-1) and an assumed [OH] of 7×10^6 molecules cm^{-3} taken from Liggio et al. (2016), as a function of SZA (filtered for peak OH of 11:00 and 16:00 local time). The error bars indicate ages calculated using emission factors of 1.08 and 1.22, respectively. The average (±1 standard deviation) photochemical age is 1.0±0.4 hr. This photochemical age applies mainly to component 1; we assume that the photochemical ages of sources associated with other components were similar.

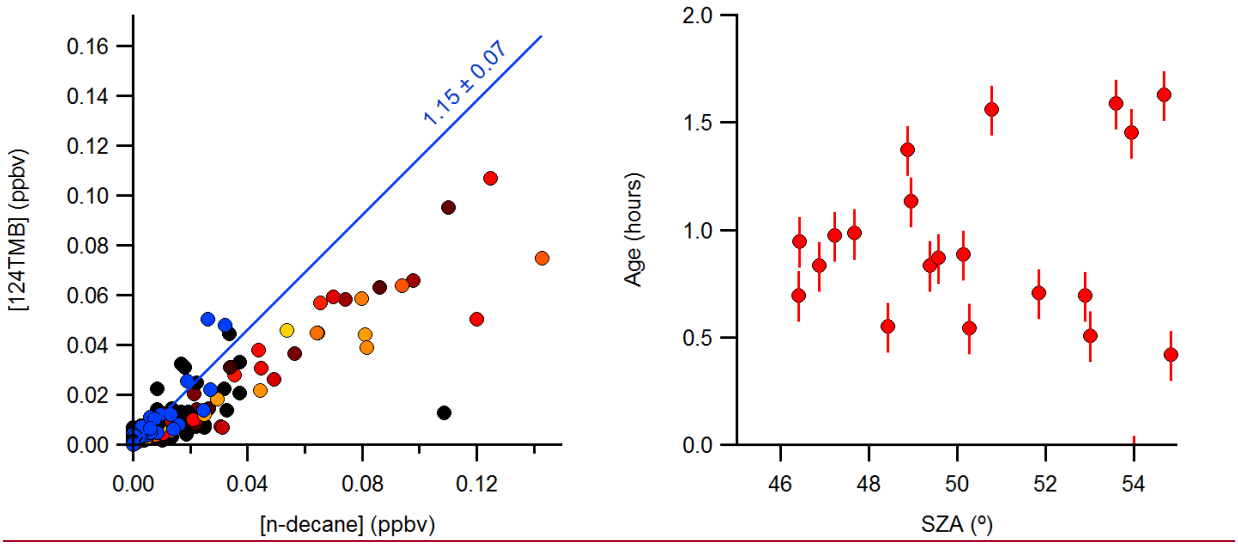


Figure S-13. (A) Plot of 124TMB mixing ratios against mixing ratios of n-decane, color-coded by solar zenith angle. The blue data points were collected at night. **(B)** Photochemical age calculating using equation S-1 plotted as a function of solar zenith angle.

In their analysis of IVOC photochemical aging, Zhao et al. (2014) estimated an average k_{OH} for diesel-exhaust IVOCs of 1.8×10^{-11} cm^3 molecule^{-1} s^{-1} (though their estimated rate coefficients varied and increased slightly with volatility bin between about 1 and 3 $\times10^{-11}$ cm^3 molecule^{-1} s^{-1}). From this, we calculate a pseudo first-order lifetime of 130 min (2.17 hr) with respect to IVOC oxidation by OH during daytime. Using a photochemical age of 1.0±0.4 hr, we calculate that between 25% and 50% of the emitted IVOC is (potentially) oxidized during daytime. Photochemical aging will affect data collected during the daytime hours (from ~11 am to ~4 pm) or ~25% of the data (56 out of 218 data points) used in the PCA and likely resulted in partial conversion of IVOCs to SOA.

References

Borbon, A., Gilman, J. B., Kuster, W. C., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros, V., Lopez, M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D. D., and de Gouw, J. A.: Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude megacities: Observations versus emission inventories in Los Angeles and Paris, J. Geophys. Res.-Atmos., 118, 2041-2057, doi:10.1002/jgrd.50059, 2013.

Bradford, L. M., Ziolkowski, L. A., Goad, C., Warren, L. A., and Slater, G. F.: Elucidating carbon sources driving microbial metabolism during oil sands reclamation, Journal of Environmental Management, 188, 246-254, 10.1016/j.jenvman.2016.11.029, 2017.

Burtscher, H., Scherrer, L., Siegmann, H. C., Schmidtott, A., and Federer, B.: Probing aerosols by photoelectric charging, J. Appl. Phys., 53, 3787-3791, 10.1063/1.331120, 1982.

Bytnerowicz, A., Fraczek, W., Schilling, S., and Alexander, D.: Spatial and temporal distribution of ambient nitric acid and ammonia in the Athabasca Oil Sands Region, Alberta, J. Limnol., 69, 11-21, 10.3274/jl10-69-s1-03, 2010.

Cattell, R. B.: The Scree Test For The Number Of Factors, Multivariate Behavioral Research, 1, 245-276, 10.1207/s15327906mbr0102_10, 1966.

Chen, H., Karion, A., Rella, C. W., Winderlich, J., Gerbig, C., Filges, A., Newberger, T., Sweeney, C., and Tans, P. P.: Accurate measurements of carbon monoxide in humid air using the cavity ring-down spectroscopy (CRDS) technique, Atmos. Meas. Tech., 6, 1031-1040, 10.5194/amt-6-1031-2013, 2013.

ECCC: Measurement instrumentation: carbon dioxide: https://www.ec.gc.ca/mges-ghgm/default.asp?lang=En&n=7903528C-1, access: April 25, 2017, 2013a.

ECCC: National pollutant release inventory (NPRI): http://open.canada.ca/data/en/dataset/e40099ae-b116-4c48-9475-f3806fe5a6a6, access: October 5, 2016, 2013b.

574 Gorham, E.: Northern peatlands - role in the carbon-cycle and probable responses to climatic warming,
575 Ecol. Appl., 1, 182-195, 10.2307/1941811, 1991.
576 Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C.: Multivariate data analysis, in, 7th edition ed.,
577 Prentice-Hall, Upper Saddle River, NJ, pp. 108 -110, 1998.
578 Holowenko, F. M., MacKinnon, M. D., and Fedorak, P. M.: Methanogens and sulfate-reducing bacteria in
579 oil sands fine tailings waste, Canadian Journal of Microbiology, 46, 927-937, 10.1139/cjm-46-10-927,
580 2000.
581 Huffman, J. A., Treutlein, B., and Pöschl, U.: Fluorescent biological aerosol particle concentrations and
582 size distributions measured with an Ultraviolet Aerodynamic Particle Sizer (UV-APS) in Central
583 Europe, Atmos. Chem. Phys., 10, 3215-3233, 10.5194/acp-10-3215-2010, 2010.
584 Johnson, M. R., Crosland, B. M., McEwen, J. D., Hager, D. B., Armitage, J. R., Karimi-Golpayegani, M., and
585 Picard, D. J.: Estimating fugitive methane emissions from oil sands mining using extractive core
586 samples, Atmos. Environm., 144, 111-123, 10.1016/j.atmosenv.2016.08.073, 2016.
587 Liggio, J., Li, S.-M., Hayden, K., Taha, Y. M., Stroud, C., Darlington, A., Drollette, B. D., Gordon, M., Lee, P.,
588 Liu, P., Leithead, A., Moussa, S. G., Wang, D., O'Brien, J., Mittermeier, R. L., Brook, J., Lu, G., Staebler,
589 R., Han, Y., Tokarek, T. W., Osthoff, H. D., Makar, P. A., Zhang, J., Plata, D., and Gentner, D. R.: Oil
590 Sands Operations as a Large Source of Secondary Organic Aerosols, Nature, 534, 91-94,
591 10.1038/nature17646, 2016.
592 Marey, H. S., Hashisho, Z., Fu, L., and Gille, J.: Spatial and temporal variation in CO over Alberta using
593 measurements from satellites, aircraft, and ground stations, Atmos. Chem. Phys., 15, 3893-3908,
594 10.5194/acp-15-3893-2015, 2015.
595 Markovic, M. Z., VandenBoer, T. C., and Murphy, J. G.: Characterization and optimization of an online
596 system for the simultaneous measurement of atmospheric water-soluble constituents in the gas and
597 particle phases, J. Environ. Monit., 14, 1872-1884, 10.1039/C2EM00004K, 2012.

598 Miller, S. M., Worthy, D. E. J., Michalak, A. M., Wofsy, S. C., Kort, E. A., Havice, T. C., Andrews, A. E.,
599 Dlugokencky, E. J., Kaplan, J. O., Levi, P. J., Tian, H. Q., and Zhang, B. W.: Observational constraints on
600 the distribution, seasonality, and environmental predictors of North American boreal methane
601 emissions, Glob. Biogeochem. Cycle, 28, 146-160, 10.1002/2013gb004580, 2014.
602 Nara, H., Tanimoto, H., Tohjima, Y., Mukai, H., Nojiri, Y., Katsumata, K., and Rella, C. W.: Effect of air
603 composition (N_2, O_2, Ar, and H_2O) on CO_2 and CH_4 measurement by wavelength-scanned cavity ring-
604 down spectroscopy: calibration and measurement strategy, Atmos. Meas. Tech., 5, 2689-2701,
605 10.5194/amt-5-2689-2012, 2012.
606 NPRI: Detailed facility information: http://www.ec.gc.ca/inrp-npri/donnees-
607 data/index.cfm?do=facility_information&lang=En&opt_npri_id=0000002274&opt_report_year=2013
608 , access: April 13, 2017, 2013.
609 Nwaishi, F., Petrone, R. M., Macrae, M. L., Price, J. S., Strack, M., and Andersen, R.: Preliminary
610 assessment of greenhouse gas emissions from a constructed fen on post-mining landscape in the
611 Athabasca oil sands region, Alberta, Canada, Ecol. Eng., 95, 119-128, 10.1016/j.ecoleng.2016.06.061,
612 2016.
613 Odame-Ankrah, C. A.: Improved detection instrument for nitrogen oxide species, Ph.D., Chemistry,
614 University of Calgary, http://hdl.handle.net/11023/2006, 10.5072/PRISM/26475, Calgary, 2015.
615 Oertel, C., Matschullat, J., Zurba, K., Zimmermann, F., and Erasmi, S.: Greenhouse gas emissions from
616 soils A review, Chem Erde-Geochem., 76, 327-352, 10.1016/j.chemer.2016.04.002, 2016.
617 Onasch, T. B., Trimborn, A., Fortner, E. C., Jayne, J. T., Kok, G. L., Williams, L. R., Davidovits, P., and
618 Worsnop, D. R.: Soot Particle Aerosol Mass Spectrometer: Development, Validation, and Initial
619 Application, Aerosol Sci. Technol., 46, 804-817, 10.1080/02786826.2012.663948, 2012.
620 Percy, K. E.: Ambient Air Quality and Linkage to Ecosystems in the Athabasca Oil Sands, Alberta, Geosci.
621 Can., 40, 182-201, 10.12789/geocanj.2013.40.014, 2013.

622 Phillips-Smith, C., Jeong, C. H., Healy, R. M., Dabek-Zlotorzynska, E., Celo, V., Brook, J. R., and Evans, G.:
623 Sources of Particulate Matter in the Athabasca Oil Sands Region: Investigation through a Comparison
624 of Trace Element Measurement Methodologies, Atmos. Chem. Phys. Discuss., 2017, 1-34,
625 10.5194/acp-2016-966, 2017.
626 Quagraine, E. K., Headley, J. V., and Peterson, H. G.: Is biodegradation of bitumen a source of recalcitrant
627 naphthenic acid mixtures in oil sands tailing pond waters?, J. Environ. Sci. Health Part A-Toxic/Hazard.
628 Subst. Environ. Eng., 40, 671-684, 10.1081/ese-200046637, 2005.
629 Rooney, R. C., Bayley, S. E., and Schindler, D. W.: Oil sands mining and reclamation cause massive loss of
630 peatland and stored carbon, Proc. Natl. Acad. Sci. U.S.A., 109, 4933-4937, 10.1073/pnas.1117693108,
631 2012.
632 Seinfeld, J. H., and Pandis, S. N.: Atmospheric chemistry and physics: from air pollution to climate
633 change, 2nd ed., Wiley, Hoboken, N.J., 2006.
634 Shephard, M. W., McLinden, C. A., Cady-Pereira, K. E., Luo, M., Moussa, S. G., Leithead, A., Liggio, J.,
635 Staebler, R. M., Akingunola, A., Makar, P., Lehr, P., Zhang, J., Henze, D. K., Millet, D. B., Bash, J. O.,
636 Zhu, L., Wells, K. C., Capps, S. L., Chaliyakunnel, S., Gordon, M., Hayden, K., Brook, J. R., Wolde, M.,
637 and Li, S. M.: Tropospheric Emission Spectrometer (TES) satellite observations of ammonia,
638 methanol, formic acid, and carbon monoxide over the Canadian oil sands: validation and model
639 evaluation, Atmospheric Measurement Techniques, 8, 5189-5211, 10.5194/amt-8-5189-2015, 2015.
640 Small, C. C., Cho, S., Hashisho, Z., and Ulrich, A. C.: Emissions from oil sands tailings ponds: Review of
641 tailings pond parameters and emission estimates, Journal of Petroleum Science and Engineering, 127,
642 490-501, 10.1016/j.petrol.2014.11.020, 2015.
643 Thompson, R. L., Sasakawa, M., Machida, T., Aalto, T., Worthy, D., Lavric, J. V., Myhre, C. L., and Stohl, A.:
644 Methane fluxes in the high northern latitudes for 2005-2013 estimated using a Bayesian atmospheric
645 inversion, Atmos. Chem. Phys., 17, 3553-3572, 10.5194/acp-17-3553-2017, 2017.

646 Tokarek, T. W., Huo, J. A., Odame-Ankrah, C. A., Hammoud, D., Taha, Y. M., and Osthoff, H. D.: A gas
647 chromatograph for quantification of peroxycarboxylic nitric anhydrides calibrated by thermal
648 dissociation cavity ring-down spectroscopy, Atmos. Meas. Tech., 7, 3263-3283, 10.5194/amt-7-3263-
649 2014, 2014.
650 Tokarek, T. W., Brownsey, D. K., Jordan, N., Garner, N. M., Ye, C. Z., Assad, F. V., Peace, A., Schiller, C. L.,
651 Mason, R. H., Vingarzan, R., and Osthoff, H. D.: Biogenic Emissions and Nocturnal Ozone Depletion
652 Events at the Amphitrite Point Observatory on Vancouver Island, Atmosphere-Ocean, 1-12,
653 10.1080/07055900.2017.1306687, 2017.
654 Wang, X. L., Chow, J. C., Kohl, S. D., Percy, K. E., Legge, A. H., and Watson, J. G.: Characterization of
655 PM2.5 and PM10 fugitive dust source profiles in the Athabasca Oil Sands Region, J. Air Waste Manag.
656 Assoc., 65, 1421-1433, 10.1080/10962247.2015.1100693, 2015.
657 Warner, D. L., Villarreal, S., McWilliams, K., Inamdar, S., and Vargas, R.: Carbon Dioxide and Methane
658 Fluxes From Tree Stems, Coarse Woody Debris, and Soils in an Upland Temperate Forest, Ecosystems,
659 10.1007/s10021-016-0106-8, 2017.
660 Warren, L. A., Kendra, K. E., Brady, A. L., and Slater, G. F.: Sulfur Biogeochemistry of an Oil Sands
661 Composite Tailings Deposit, Front. Microbiol., 6, 14, 10.3389/fmicb.2015.01533, 2016.
662 Wesely, M. L., and Hicks, B. B.: A review of the current status of knowledge on dry deposition, Atmos.
663 Environm., 34, 2261-2282, 10.1016/S1352-2310(99)00467-7, 2000.
664 Whalen, S. C.: Biogeochemistry of methane exchange between natural wetlands and the atmosphere,
665 Environ. Eng. Sci., 22, 73-94, 10.1089/ees.2005.22.73, 2005.
666 Whaley, C. H., Makar, P. A., Shephard, M. W., Zhang, L., Zhang, J., Zheng, Q., Akingunola, A., Wentworth,
667 G. R., Murphy, J. G., Kharol, S. K., and Cady-Pereira, K. E.: Contributions of natural and anthropogenic
668 sources to ambient ammonia in the Athabasca Oil Sands and north-western Canada, Atmos. Chem.
669 Phys., 18, 2011-2034, 10.5194/acp-18-2011-2018, 2018.

Wilson, N. K., Barbour, R. K., Chuang, J. C., and Mukund, R.: Evaluation of a real-time monitor for fine particle-bound PAH in air, Polycycl. Aromat. Compd., 5, 167-174, 10.1080/10406639408015168, 1994.

Yavitt, J. B., Williams, C. J., and Wieder, R. K.: Soil chemistry versus environmental controls on production of CH_4 and CO_2 in northern peatlands, Eur. J. Soil Sci., 56, 169-178, 10.1111/j.1365-2389.2004.00657.x, 2005.

Zhang, L. M., Brook, J. R., and Vet, R.: On ozone dry deposition - with emphasis on non-stomatal uptake and wet canopies, Atmos. Environm., 36, 4787-4799, 10.1016/s1352-2310(02)00567-8, 2002.

Zhao, Y. L., Hennigan, C. J., May, A. A., Tkacik, D. S., de Gouw, J. A., Gilman, J. B., Kuster, W. C., Borbon, A., and Robinson, A. L.: Intermediate-Volatility Organic Compounds: A Large Source of Secondary Organic Aerosol, Environm. Sci. Technol., 48, 13743-13750, 10.1021/es5035188, 2014.