Electronic Supplementary Material (ESI) for Environmental Science: Atmospheres.

Evaluating SOA formation from different sources of semi- and intermediate-volatility organic compounds from the Athabasca Oil Sands

Supplemental Information

SI Materials and Methods:

Measurements:

During the JOSM 2013 intensive field study, the National Research Council (NRC) Corvair 580 collected atmospheric and aircraft observations near the Oil Sands complexes in Northeastern Alberta. In this work, observations of VOCs, OA, BC, NO_x, and NO_y were used to constrain and evaluate the model, which have been described in previous work.[1–3] Briefly, VOCs were collected in stainless steel canisters during flight, and analyzed off-line in the National Air Pollution Survey (NAPS) laboratory by gas chromatography mass spectroscopy (GC-MS, Agilent, Santa Clara, CA, USA). OA, along with other non-refractory particulate matter, was measured on-line in a high-resolution time-of-flight aerosol mass spectrometer (HR-ToF-MS, Aerodyne Research Inc., Billerica, MA, USA). Refractory black carbon (BC in this work) was measured by a Single Particle Soot Photometer (SP2, Droplet Measurement Technologies, Boulder, CO, USA). NO, NO_2 and NO_y observations were collected using three separate instruments based on chemiluminescence with excess ozone (42iTL TRACE Level NO_x Analyzer, Thermo Fisher Scientific, Waltham, MA, USA) modified to measure with 1sec time resolution. The first instrument measured ambient NO directly; a second analyzer measured the sum of NO and large fraction of NO_2, selectively converted to NO using a photolytic converter (Air Quality Design Inc.). Ambient NO_2 was derived based on the readings of these two analysers and the photolytic converter efficiency determined through field calibrations (65% on average). A molybdenum converter heated at 325 °C was used to convert all oxidized nitrogen species (NO_y) to NO, followed by subsequent detection with a third chemiluminescencent analyser.

In 2018, the NRC Corvair 580 aircraft was jointly instrumented by ECCC and NRC with arrays of sensors for measurements of aircraft states (position, ground speed, roll, pitch etc), atmospheric states (temperature, pressure, winds etc) and atmospheric compositions (CO_2, NMOG, NO_x, NO_y, etc). The aircraft state parameters were measured using NovAtel GPS receivers and Inertial Measurement Units (IMU). The raw GPS data with added ground base station data were reprocessed to obtain sub-meter position accuracy. NO_x and NO_y were measured by the same methods as in 2013. A Picarro G2401-m Green House Gas (GHG) Cavity Ring-Down Spectrometer was used to measure carbon dioxide. Using the CRDS technique, these instruments measure concentrations of CO, CO_2, CH_4 and H_2O every two seconds. Calibrations were executed before, at points during, and after the study. In these calibrations, the Picarro instrument had their CO, CO_2 and CH_4 measurements compared to two gas cylinders containing a mixture of CO, CO_2 and CH_4 at two different concentrations (i.e. 389.680 ppm CO_2, 192.380 ppb CO, and 1.998 ppm CH_4 for the low concentration cylinder and 447.320 ppm CO_2, 486.910 ppb CO, and 2.979 ppm CH_4 for the high concentration cylinder). NMOG, measured as all organic carbon, excluding

methane and organic aerosols, was measured indirectly with an additional Piccaro G2401-m instrument, in a technique adapted from Veres et al. (2010).[4] In the NMOG Picarro, total NMOG, CH_4, and CO are converted to CO_2 over a platinum catalyst at 600 °C. The difference between CO_2 measurements with and without the platinum catalyst provides the concentration of NMOG + CH_4 + CO. NMOG concentrations are isolated by subtracting the measured CO and CH_4 from the Picarro without a platinum catalyst.

Ground-site measurements of BC, NO_x, NO_y, and OA were performed at an air measurement station adjacent to Fort Mackay (Ambient Measurement Station-13) during the 2013 intensive program, and was previously described in prior work.[5,6] Briefly, a soot-particle aerosol mass spectrometer (SP-AMS, Aerodyne Research Inc, Billerica, MA, USA) was used to measure non-refractory PM, including OA and BC. Positive matrix factorization (PMF) was utilized to identify the fraction of OA which is considered hydrocarbon-like (HOA). NO_y, NO_2, and NO were measured using the same principles (chemiluminescence with excess ozone and selective conversion to NO) as the 2013 and 2018 aircraft campaigns. Only two instruments (42iTL Thermo Fisher Scientific) were used and they were cycled between NO and NO_y, and NO and NO_2 respectively, which led to a time resolution of 1 min.

Photochemical Age. Photochemical age is a measure of the amount of photochemical oxidation experienced by a parcel of air.[7,8] It can be estimated from the ratio of the concentrations of two atmospheric species. The change in the ratio provides a quantitative metric for the integrated OH radical exposure, which can be expressed in units of time when assuming an average •OH concentration of 1.5 x 10^6 molec cm^{-3}.[9] This metric is then used when comparing the model to measurements of different species, including OA. In this work, photochemical age is determined using measurements of NO_x and NO_y at each transect (Equation S1):

$$Photochemical\ age\ (OH\ integrated\ days) = -1.61 * \log(\frac{NO_x}{NO_y}), \quad \text{(S1)}$$

NO_x is the sum of NO and NO_2, while NO_y is defined as all reactive oxidized nitrogen species, including NO_x and its reaction products. In the daytime, when the flights took place, the dominant loss process of NO_x (and production process of NO_Y) is assumed to be the reaction of NO_2 + OH → HNO_3. The NO_2 + OH reaction rate, and therefore photochemical age, are defined by that reaction.[10] NO_x and NO_y species were averaged within the plumes of each transect.

Gas Phase Oxidation:

The oxidation of precursor compounds is dynamically modeled using previously described rate constants and oxidized VOC yields (provided in Tables S2, S5, and S6). The P-VOC oxidation mechanisms explored in this work describe two methods for adjusting for biased-low laboratory-derived SOA yields. The first, baseline, method models the multi-generational aging of P-VOC oxidation products (VOC-sourced SVOCs, or V-SVOCs) as a source of increasing OA mass,[11] while the second parameterization adjusts for wall-losses in laboratory oxidation chambers to increase the rapid oxidation into lower volatility bins and increase OA mass.[12] The baseline oxidation of VOCs uses the Tsimipidi et al. framework (AGE), where VOCs are oxidized into four lumped volatility bins representing the products (c* = 1, 10, 100, and 1000 μg/m^3, where c* is the

effective saturation concentration).[11] The oxidized VOCs are further "aged" into a lower volatility bin (one bin at a time), at a rate of 1 x 10^{-11} cm^3 $molec^{-1}$ s^{-1} until reaching a minimum volatility of c* = 0.01. Only the gas-phase fraction of an organic species undergoes aging, while particle phase oxidation is not included in the model, since heterogeneous oxidation lifetimes are very long compared to the timescales explored in this study.[12–15] Alternatively, a more recent approach is to correct previously measured experimental chamber yields for the effect of vapor wall-losses in experimental chambers. This approach, developed in Ma et al. (WLC),[12] follows the oxidation of VOCs into four initial volatility bins as described for the AGE approach, but with a higher concentration of SVOCs in the lowest volatility bins. The WLC approach does not allow for multi-generational aging. The volatility bins and yields for the initial oxidation step for each precursor VOC are summarized in the supplementary materials (Tables S5 and S6).

In this work, the base oxidation parameterization treats P-IVOCs exclusively with multi-generational aging (SLOW), where the volatility bin of the emitted species is crucial to the oxidation process. This parameterization is an extension of the P-SVOC approach detailed below, and oxidation follows a "bin-hopping" approach where each oxidation reaction results in an order-of-magnitude decrease in volatility.[16] An alternative parameterization for SOA formation from P-IVOCs (I-SOA) is also used, in which the P-IVOCs are lumped into a single volatile species (FAST).[17] This lumped IVOC species is then oxidized to form less volatile products following a parameterization similar to that used for the VOCs in the AGE mechanism (see Figure 2B). This alternative parameterization for I-SOA formation, including yields, is based on the oxidation of diesel IVOC emissions as reported by Jathar et al.[18]

P-SVOCs (organic compounds with $10^{-2} < c^* < 10^3$, in µg m^{-3} at 298 K) in the model undergo •OH oxidation according to the aging framework reported initially by Robinson et al.[16] The parameterization for P-SVOC oxidation is abbreviated SLOW, and there is no alternative oxidation parameterization for P-SVOCs in this work. The multi-generational aging process is usually modelled as a decadal reduction of volatility with a rate that depends on the concentration of OH radicals and the aging rate-constant of the hydrocarbon substrate.[16,19] In many models, the multi-generational aging rate constant is estimated from the experimental rate constant of adding an alcohol moiety to a linear alkane via oxidation by an OH radical.[11,16,20] However, experimental data and structure-activity models demonstrate that the addition of oxygen species to an alkane can make the species more (or less) reactive.[15] In this work, we model each species' aging rate-constant as either a fixed value, or as a function of carbon and oxygen numbers and contrast the output with the fixed rate-constant commonly applied in VBS models. The base case SLOW oxidation rate for aging is 4 x 10^{-11} cm^3 $molec^{-1}$ s^{-1}, which is denoted as the FIXED scheme. The FIXED scheme uses the single oxidation rate constant for all species sourced from P-S/IVOCs. Following Donahue et al,[15] we can alternatively apply a vapour-phase rate constant that is calculated from the carbon and oxygen numbers of each model species:

$$k_{OH}^{vap} = 1.2\ x\ 10^{-12}(n_C + 9n_O - 10(O{:}C)^2)\ ; \qquad \text{(S1)}$$

Using a calculated rate-constant for P-S/IVOC oxidation is denoted as the VAR, or variable, scheme. However, the VAR scheme is only used for species which are also modeled with the SLOW parameterization of P-S/IVOC oxidation. Products of P-IVOC oxidation when the FAST

scheme is active, and all products of VOC oxidation do not use the VAR rates, even while the P-SVOC species in the same model run are modeled with VAR. As a result, all SVOC species sourced from VOCs use the standard oxidation rate-constant of 1 x 10^{-11} cm^3 $molec^{-1}$ s^{-1} while SVOC species sourced from P-IVOCs during the FAST parametrization use the rate constant of 4 x 10^{-11} cm^3 $molec^{-1}$ s^{-1}.

Further details describing the binned P-S/IVOC species and the corresponding ΔH_{vap} values, molecular weights, and carbon numbers are provided in Table S3.

Fragmentation. This work accounts for uncertainty in the processes which govern the loss of OA mass via fragmentation of SVOCs during oxidation and reduce the carbon number and size of SVOCs. This fragmentation occurs from carbon-carbon bond-breaking within the backbone of a hydrocarbon, which results in the formation of two (or more) hydrocarbons with fewer carbons than the parent molecule.[21] Often, though not always, this results in a net increase in volatility, since volatility is inversely proportional to carbon number. Fragmentation is more likely to occur in highly functionalized species, especially those with reactive carbonyl moieties.[22]

Fragmentation of semi-volatile species follows a modification of the 1-D model described by Koo et al.[23] For all VOC oxidation, VOC sourced SVOCs (V-SVOC) are assigned carbon numbers based on their volatility bin, as carbon number (n_c) tends to decrease with volatility due to fragmentation occurring with oxidation (n_c = 6.75, 7, 7.25, 7.5, 7.75, and 8 for c* = 0.01, 0.1, 1, 10, 100, and 1000). Multi-generation aging of V-SVOCs reduces the carbon number by a further 0.25 carbons per generation. For the SLOW parameters, P-SVOC sourced SVOCs (S-SVOCs) and P-IVOC sourced SVOCs (I-SVOCs) species initially have a carbon number that depends on the volatility bin (Table S3). S-SVOC species decrease their carbon number by 1.0 unit for each oxidation step. I-SVOC species decrease their carbon number by 0.5 units for each oxidation step. When P-IVOCs are oxidized via the FAST parameters, for the first-generation products the carbon number decreases with volatility (13.5, 14, 14.5, 15 for c* = 1, 10, 100, 1000), and subsequent multi-generation oxidation decreases carbon number by 0.5 for each oxidation step. The uncertainty of this fragmentation parameterization is evaluated in this work by activating the fragmentation effect or excluding it from the model. Additionally, a sensitivity study is run to determine the effect of increasing or decreasing the carbon loss during fragmentation (Figure 8B).

Fragmentation is accounted for during the conversion of molecular concentration to mass concentration. The effects of fragmentation counterbalance the increase in mass from the addition of oxygen due to oxidation by •OH during both the initial oxidation step and further multi-generational aging, following the work of Koo et al. in 3D models.[23]

Sample Calculation of SVOC physical characteristics:

Calculating the n_O, O:C, and MW of a C14 species with a saturation concentration of 10 µg m^{-3}, during Flight 19.

$$\log(c^*) = (25 - n_C) * 0.475 - n_O * 2.3 + \left(\frac{n_O * n_C}{n_O + n_C}\right) * 0.6\,; \qquad \text{(S2)}^{24}$$

$$0 = -2.3 * \boldsymbol{n_O^2} + (11.875 - 2.175 * n_C - \log(c^*)) * \boldsymbol{n_O} - (\log(c^*) * n_C + 11.875 * n_C - 0.475 * n_C^2)\ ; \quad \text{(S3)}$$

$$n_O = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\ ; \quad \text{(S4)}$$

$$a = -2.3\ ; b = 11.875\ - 2.175 * n_C - \log(c^*)$$

$$c = -\log(c^*) * n_C + 11.875 * n_C - 0.475 * n_C^2$$

$$n_O = \sout{-11.7}, 2.65 \quad \text{(S5)}$$

$$O{:}C = \frac{n_O}{n_C} = 0.17\ ; \quad \text{(S6)}$$

$$MW = n_C * 12.01 + n_O * 16.00 + n_H * 1.01\ ; \quad \text{(S7)}$$

$$MW = n_C * 12.01 + n_O * 16.00 + \left((1.74 - 0.4 * O{:}C) * n_C\right) * 1.01\ ; \quad \text{(S8)}$$

$$MW = 207.8 \quad \text{(S9)}$$

An example pathway of the successive aging of a C14.5 I-SVOC species with an initial saturation concentration of 100 µg m^{-3} into the saturation concentration bin of 0.01 µg m^{-3} is described in Table S7.

Table S1. Definitions for acronyms frequently used in this article.

1P	Parameterization for partitioning with a single aerosol phase
2P	Parameterization for partitioning with non-miscible polar and non-polar phases
20D	Parameterization for the volatility distribution of P-S/IVOCs for headspace emissions of Oil Sands earth at 20 degrees, according to Liggio et al (2016)
60D	Parameterization for the volatility distribution of P-S/IVOCs for headspace emissions of Oil Sands earth at 60 degrees, according to Liggio et al (2016)
ΔBC	Enhanced black carbon concentration over the background concentration of the flight
BEU	Parameterization for the volatility distribution of P-S/IVOCs for the emissions measured from the benchtop batch extraction unit described in Drollette et al (2020)
BIT	Parameterization for the volatility distribution of P-S/IVOCs for headspace emissions of lightly processed bitumen, according to Li et al (2019)
F_ON	Parameterization for fragmentation where carbon loss is described in Koo et al (2014)
F_OFF	Parameterization where fragmentation is not active
FIXED	Multigenerational aging rate constant for P-S/IVOCs is fixed at 4.0×10^{-11} cm^{-3} $molec^{-1}$ s^{-1}
FAST	Jathar et al, (2014) parameterization for secondary organic aerosol formation from P-IVOCs
WLC	Ma et al, (2017) parameterization for secondary organic aerosol formation from VOCs
ORE	Parameterization for the volatility distribution of P-S/IVOCs for headspace emissions of crushed raw ore, according to Li et al (2019)
P-S/IVOC	Primary semi-volatile and intermediate-volatility organic compounds
SLOW	Robinson et al, (2007) parameterization for secondary organic aerosol formation from P-S/IVOCs
S/I-SVOC	Secondary species formed from primary semi-volatile and intermediate volatility organic compounds
S/I-SOA	Secondary organic aerosols from P-S/IVOCs
AGE	Tsimpidi et al, (2010) parameterization for secondary organic aerosol formation from VOCs
V-SOA	Secondary organic aerosols sourced from VOCs
V-SVOC	Semi-volatile organic compounds sourced from VOCs
VAR	Multigenerational aging rate constant for P-S/IVOCs is variable depending on the species carbon number and oxygen number
VOC	Volatile organic compounds, with fewer than 10 carbons

Table S2. Anthropogenic VOC oxidation parameters used in the model. The multi-generational aging rate constant, following the initial oxidation, for VOC oxidation is 1.0 x 10^{-11} cm^{-3} $molec^{-1}$ s^{-1}. Oxidation products from aging have a volatility that is lower by one order of magnitude compared to the precursor species. All SOA from VOCs have ΔH_{vap} = 36 kJ mol^{-1}.[25]

Classification	Compound	k_{OH} (cm^3 $molec^{-1}$ s^{-1})	Molecular Weight (g mol^{-1})	ΔVOC/ΔBC μg m^{-3}/μg m^{-3} Flight 19	Flight 20
ALK4	Methylcyclopentane	5.68 x10^{-12}	84.16	3.237	6.146
	2,2,3-Trimethylbutane	3.81 x10^{-12}	100.21	0.844	0.146
ALK5	Cyclohexane	6.97 x10^{-12}	84.16	1.504	1.798
	Methylcyclohexane	9.64 x10^{-12}	98.19	3.511	4.376
	n-Heptane	6.76 x10^{-12}	100.21	7.141	17.519
	2-Methylhexane	6.89 x10^{-12}	100.21	1.652	4.348
	3-Methylhexane	7.17 x10^{-12}	100.21	2.837	7.192
	2,3-Dimethylpentane	7.15 x10^{-12}	100.21	1.356	2.125
	2,4-Dimethylpentane	4.77 x10^{-12}	100.21	0.430	3.763
	n-Octane	8.11 x10^{-12}	114.23	5.763	6.725
	3-Methylheptane	8.59 x10^{-12}	114.23	1.519	1.728
	2-Methylheptane	8.31 x10^{-12}	114.23	3.978	4.969
	2,2,4-Trimethylpentane	3.34 x10^{-12}	114.23	0.178	0.000
	2,3,4-Trimethylpentane	6.60 x10^{-12}	114.23	0.385	0.084
	n-Nonane	9.70 x10-^{12}	128.26	2.178	2.997
OLE1	Propene	2.63 x10^{-11}	42.08	0.622	0.404
	1-Butene	3.14 x10^{-11}	56.11	0.407	0.557
	1-Pentene	3.14 x10^{-11}	70.14	0.163	0.216
	2-Methyl-1-butene	6.10 x10^{-11}	70.14	0.081	0.111
	3-Methyl-1-butene	3.18 x10^{-11}	70.14	0.044	0.063
OLE2	1,3-Butadiene	6.66 x10^{-11}	54.09	0.030	0.014
	trans-2-Pentene	6.70 x10^{-11}	70.14	0.030	0.146
	cis-2-Pentene	6.50 x10^{-11}	70.14	0.022	0.070
	Styrene	5.80 x10^{-11}	104.15	0.022	0.014
ARO1	Toluene	5.63 x10^{-12}	92.14	4.074	5.010
	Ethylbenzene	7.00 x10^{-12}	106.17	1.133	1.171
	i-Propylbenzene	6.30 x10^{-12}	120.2	0.133	0.091
	n-Propylbenzene	5.80 x10^{-12}	120.2	0.244	0.216
	Benzene	1.22 x10^{-12}	78.11	1.748	0.780
ARO2	*o*-Ethyltoluene	1.19 x10^{-11}	120.2	0.274	0.376
	m/p-Ethyltoluene	1.52 x10^{-11}	120.2	0.563[a]	0.878[a]
	1,2,3-Trimethylbenzene	3.27 x10^{-11}	120.2	0.185	0.488
	1,2,4-Trimethylbenzene	3.25 x10^{-11}	120.2	0.326	0.725
	1,3,5-Trimethylbenzene	5.67 x10^{-11}	120.2	0.074	0.237
	m/p-Xylene	1.87 x10^{-11}	106.16	2.156[a]	3.638[a]
	o-Xylene	1.36 x10^{-11}	106.16	0.956	1.484
ISOP	Anthropogenic isoprene	1.00 x10^{-10}	68.12	5.896	4.697
IVOC	IVOCs	2.44 x10^{-11}	204.15	[b]	[b]

[a]Average of both emission ratios; [b]Applicable when IVOCs are treated as single VOC species, as discussed in Methods (FAST) section. The aging rate constant for IVOCs is 4.0 x 10^{-11} cm^{-3} $molec^{-1}$ s^{-1}, unless otherwise noted as discussed in Methods (VAR) section.

Table S3. Physical parameters of P-SVOCs for volatility bins of c* $\leq$ 1000, and P-IVOCs when using the SLOW parameterization. The aging rate constant for P/S-S/IVOC oxidation is 4.0 x 10^{-11} cm^{-3} $molec^{-1}$ s^{-1}, unless otherwise noted as discussed in Methods (VAR) section, and oxidation decreases volatility by one order of magnitude per generation.

Volatility Bin *c @ 298K ($\mu g\ m^{-3}$)**	**ΔH_{vap} ($kJ\ mol^{-1}$)**[16]	**Molecular Weight ($g\ mol^{-1}$)**[a]	**Average Carbon Number**	**Fraction (%) of total P-S/IVOC for each volatility distribution**			
				20D	**60D**	**ORE**	**BIT**
0.01	112	-	-	-	-	-	-
0.1	106	-	-	-	-	-	-
1	100	-	-	-	-	-	-
10	94	372	27.5	0.3	0.9	-	-
100	88	337	24	1.1	3.3	-	-
1000	82	295	21	2.0	9.1	-	-
10000	76	252	18	8.1	23.5	-	-
100000	70	210	15	48.0	40.0	3.88	6.33
1000000	64	177	12.5	38.2	21.3	18.1	18.1
10000000	58	147	10.5	1.2	0.3	78.0	75.6

[a]Molecular weight of species in the initial conditions; there is no molecular weight for species with no fraction of emitted P-S/IVOC.

Table S4. Biogenic VOC oxidation parameters used in the model. There is no multi-generational aging for biogenic species.

Classification	**Compound**	**k_{OH} ($cm^3\ molec^{-1}\ s^{-1}$)**	**Molecular Weight ($g\ mol^{-1}$)**	**Average [VOC] ($\mu g\ m^{-3}$)**[a]	
				Flight 19	**Flight 20**
ISOP	Biogenic isoprene	1.00 x10^{-10}	68.12	0.842	1.137
TERP	α-Pinene + β-pinene	9.82 x10^{-11}	136.24	1.024	1.025
SESQ[b]	Sesquiterpenes	9.82 x10^{-11}	204.15	0.127	0.081

[a]Ambient VOC concentrations for biogenic species discussed in Supplementary Information; [b]Calculated as a function of monoterpene concentration, as discussed in SI

Table S5. SOA yields from VOCs in High-NO_x conditions for AGE parameters.

Classification	Stoichiometric SOA yield, High-NO_x at 298 K (µg m^{-3})			
	1	**10**	**100**	**1000**
ALK4	0.000	0.038	0.000	0.000
ALK5	0.000	0.150	0.000	0.000
OLE1	0.001	0.005	0.038	0.150
OLE2	0.003	0.026	0.083	0.270
ARO1	0.003	0.165	0.300	0.435
ARO2	0.002	0.195	0.300	0.435
ISOP	0.001	0.023	0.015	0.000
TERP	0.012	0.122	0.201	0.500
SESQ	0.012	0.122	0.201	0.500
IVOC	0.044	0.071	0.410	0.300

Table S6. SOA yields from VOCs in High-NO_x conditions accounting for WLC parameters.

Classification	Stoichiometric SOA yield, High-NO_x at 298 K (µg m^{-3})			
	1	**10**	**100**	**1000**
ALK4	0.000	0.038	0.000	0.000
ALK5	0.157	0.000	0.000	0.000
OLE1	0.014	0.000	0.098	0.088
OLE2	0.052	0.000	0.183	0.157
ARO1	0.276	0.002	0.431	0.202
ARO2	0.310	0.000	0.420	0.209
ISOP	0.034	0.000	0.005	0.000
TERP	0.210	0.000	0.348	0.297
SESQ	0.210	0.000	0.348	0.297
IVOC	0.044	0.071	0.410	0.300

Table S7. Molecular properties described by the aging pathway of C14.5 alkanes as a sample of P-IVOC aging.

c* @ 300 K (µg m^{-3})	**n_C**	**n_O**	**O:C**	**MW (g mol^{-1})[a]**	**ΔH_{vap} (kJ mol^{-1})**[16]
100	14.50	1.69	0.12	203.0	88
10	14.25	2.30	0.16	209.7	94
1	14.00	2.90	0.21	216.2	100
0.1	13.75	3.48	0.25	222.5	106
0.01	13.50	4.05	0.30	228.7	112

[a]Assuming H:C = 1.74 – 0.4*O:C.

Table S8. Performance of all model iterations by root mean square error (μg m^{-3}). The best value for each volatility distribution is underlined. The best value for each combination of model parameterizations is bolded.

P-VOC Oxidation	P-IVOC Oxidation	Fragment-ation	# of phases	Aging Rate	Volatility Distribution			
					20D	60D	ORE	BIT
AGE	FAST	F_OFF	1P	FIXED	8.10	3.15	**3.07**	3.56
				VAR	7.87	3.35	**3.07**	3.56
			2P	FIXED	7.84	3.20	**3.13**	3.62
				VAR	7.68	3.46	**3.13**	3.62
		F_ON	1P	FIXED	6.86	**2.97**	3.07	3.56
				VAR	6.70	3.22	**3.07**	3.56
			2P	FIXED	6.65	**3.08**	3.13	3.62
				VAR	6.57	3.32	**3.13**	3.62
	SLOW	F_OFF	1P	FIXED	6.46	**1.70**	4.40	4.69
				VAR	10.54	**2.69**	3.16	3.84
			2P	FIXED	4.97	**1.86**	4.55	4.80
				VAR	10.57	**3.05**	3.23	3.89
		F_ON	1P	FIXED	4.84	**1.88**	4.75	4.94
				VAR	8.27	**2.23**	3.78	4.31
			2P	FIXED	2.94	**2.38**	4.79	4.96
				VAR	7.63	**2.56**	3.72	4.22
WLC	FAST	F_OFF	1P	FIXED	9.76	3.94	3.48	**3.27**
				VAR	9.51	3.97	3.48	**3.27**
			2P	FIXED	9.49	3.89	3.49	**3.30**
				VAR	9.31	4.00	3.49	**3.30**
		F_ON	1P	FIXED	8.58	3.29	3.48	**3.27**
				VAR	8.41	3.38	3.48	**3.27**
			2P	FIXED	8.36	**3.30**	3.49	**3.30**
				VAR	8.26	3.44	3.49	**3.30**
	SLOW	F_OFF	1P	FIXED	7.90	**3.03**	3.11	3.43
				VAR	12.09	4.05	**2.01**	2.64
			2P	FIXED	6.48	**2.83**	3.27	3.54
				VAR	12.14	4.30	**2.19**	2.75
		F_ON	1P	FIXED	6.04	**1.97**	3.49	3.70
				VAR	9.79	2.70	**2.49**	3.05
			2P	FIXED	4.07	**2.08**	3.53	3.72
				VAR	9.13	**2.97**	2.49	3.00
Average					7.93	**3.04**	3.36	3.64

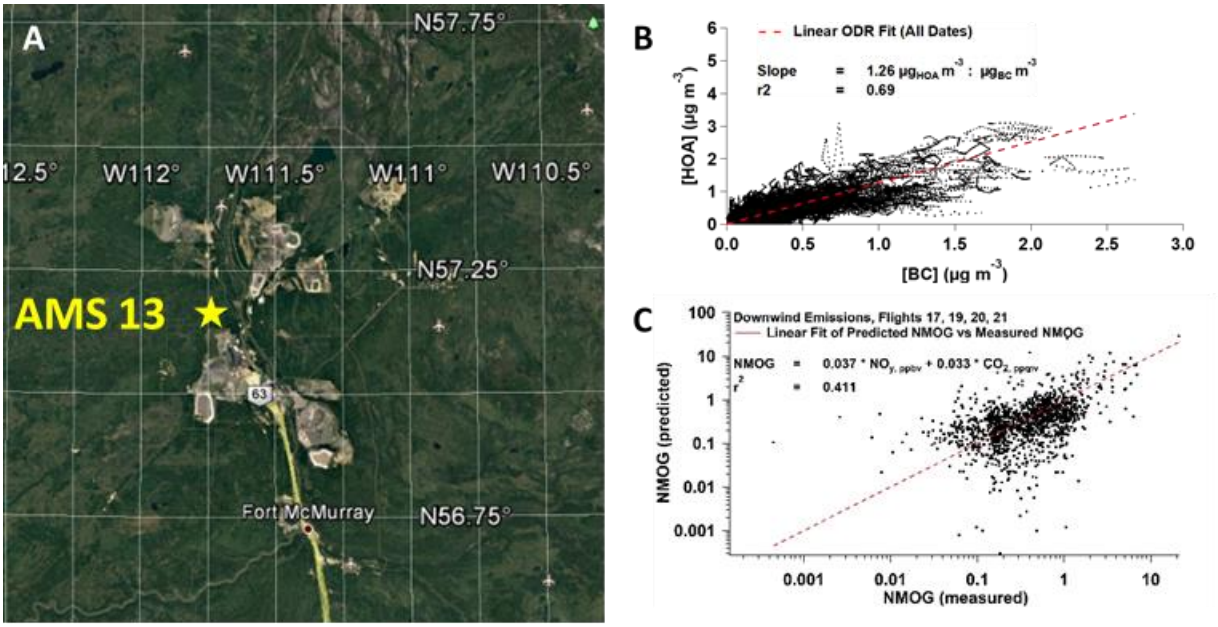

Figure S1: (A) The coordinates of the Oil Sands region and the ground station AMS 13, which is used to calculate the emission ratio of POA to ΔBC. The yellow star represents the location of AMS 13 in Fort Mackay, AB. **(B)** shows the correlation between HOA and BC concentrations at AMS 13, where the slope represents the emission ratio of POA/ΔBC. **(C)** shows the correlation between measured NMOG and NMOG predicted by a two-parameter linear fit between NMOG and NO_y + CO_2. NMOG, NO_y, and CO_2 mixing ratios are measured downwind of the Oil Sands facilities in 2018 during Flights 17, 19, 20, and 21.

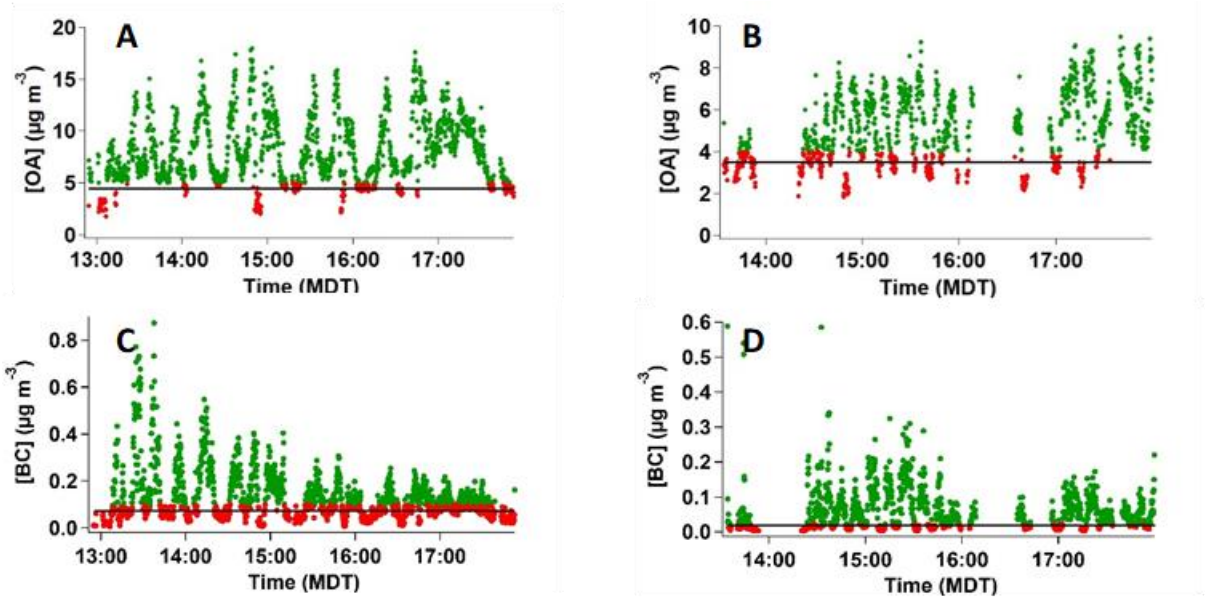

Figure S2. Time-series of OA for Flight 19 **(A)** and Flight 20 **(B)**. Background concentrations of OA are highlighted in red for each time series with the solid black line denoting the average concentration of OA for the background of each flight (Flight 19 = 4.5 µg m^{-3}; Flight 20 = 3.5 µg m^{-3}). Time series of BC for Flight 19 (**C**) and Flight 20 (**D**). Background concentrations of BC are highlighted in red for each time series with the solid black line denoting the average background concentration of each flight (Flight 19 = 0.077 µg m^{-3}; Flight 20 = 0.011 µg m^{-3})

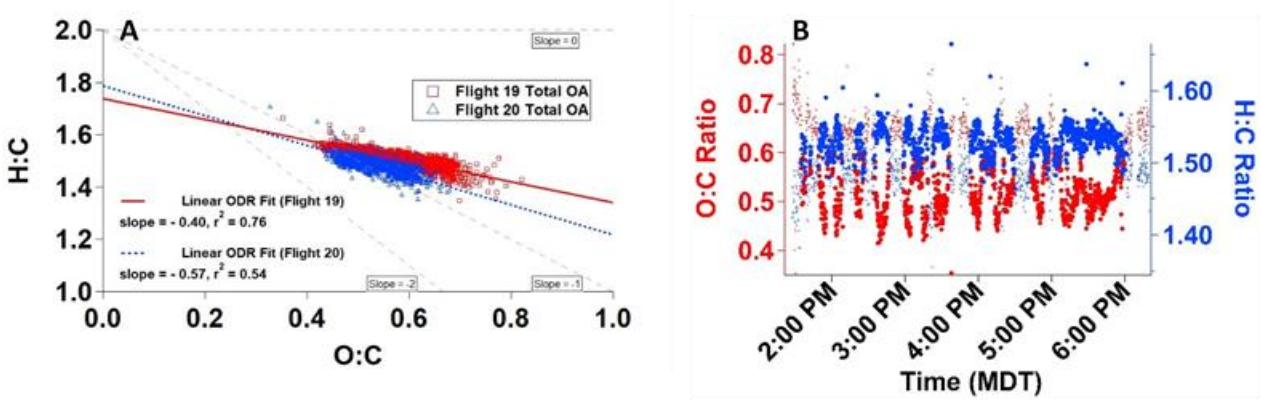

Figure S3. (A) Van Krevelen diagrams for Flights 19 and 20 are shown in red squares and blue triangles, respectively. Linear regression analyses correspond to all OA data including in- and out-of-plume data. Reference slopes, indicated by the dashed lines, describe the theoretical oxidation pathway of a methylene group being oxidized into a carbonyl or alcohol group (slopes = -2 or 0, respectively). **(B)** Time-series of O:C and H:C ratios within and outside of the emissions plume for Flight 19, for OA. Datapoints within the Oil Sands plume are denoted by filled markers, while background datapoints are denoted by smaller, hollow markers.

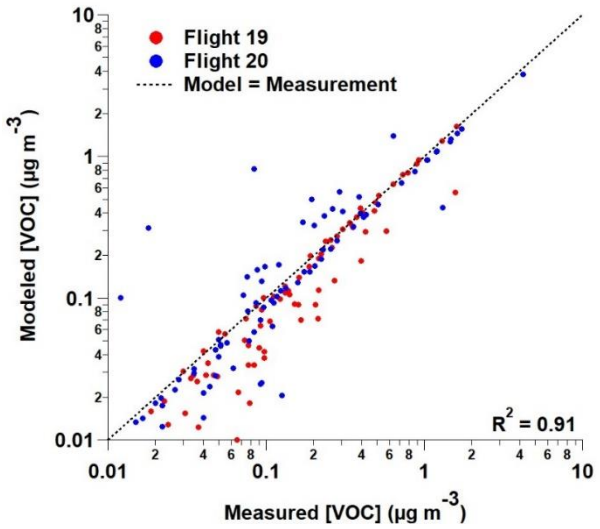

Figure S4. Model-measurement comparison of concentrations of VOCs between box model simulation and VOC cannisters acquired from Flights 19 (blue) and 20 (red). Each flight contains measurements from three transects and are compared to the model simulations at the same photochemical age. In Flight 19, VOC measurements are available from transects 2, 3, and 4. In Flight 20, VOC measurements are available from transects 1, 2, and 3.

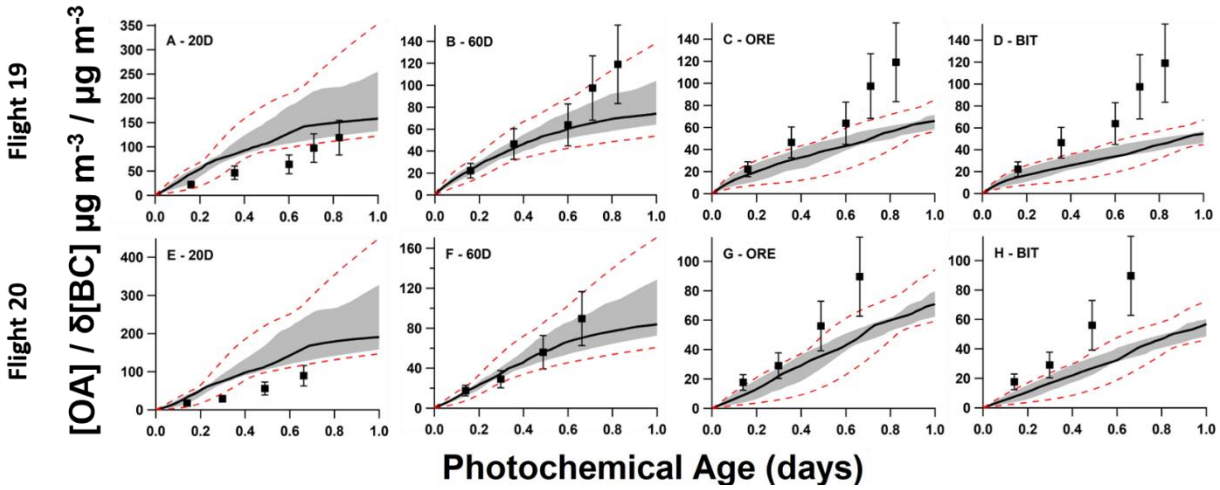

Figure S5: Ensemble model–measurement comparison of anthropogenic organic aerosol, normalized to black carbon, vs photochemical age, comparing the four evaluated volatility distributions of P-S/IVOCs. For model results, the median (50th percentile) line is depicted as a black line, surrounded by the 25th to 75th percentile range, shaded in grey. The 95th and 5th percentile results are depicted by the dashed lines. The top row represents ensemble results for Flight 19 in the 2013 JOSM campaign, while the bottom row represents results for Flight 20. All measurements, described with black squares are $\pm$ 30%, representing the measurement uncertainty. **A)** and **B)** use the volatility distributions previously described in Liggio et al. (2016) for the bitumen deposits heated to 20 and 60 ^{o}C, respectively. **C)** and **D)** use the volatility distributions of emissions from raw ore and bitumen, respectively, as described in Li et al. 2019. **E)** through **H)** represent the same ensemble runs as **A)** through **D)** for Flight 20.

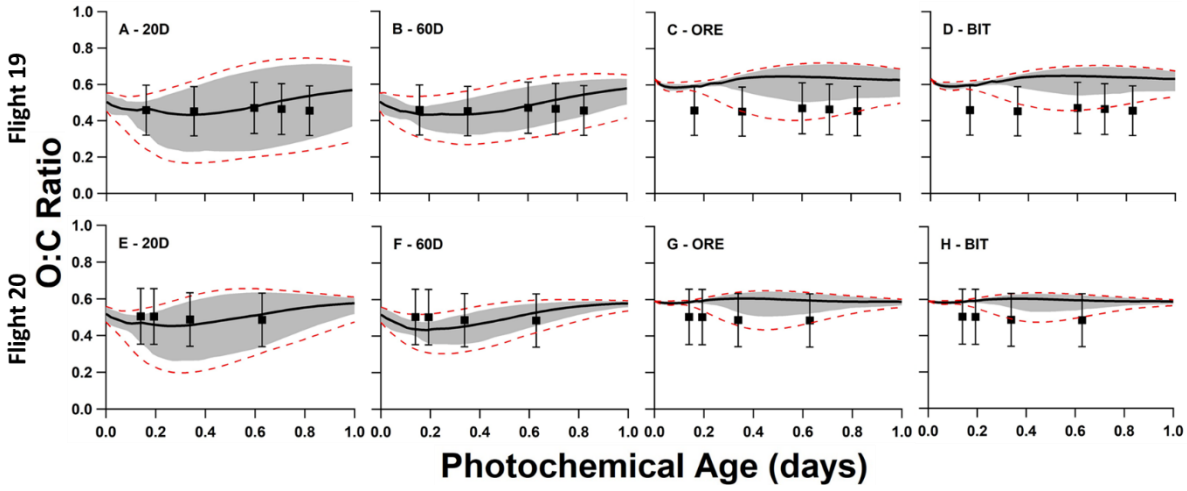

Figure S6. Ensemble model-measurement comparison of O:C ratio vs photochemical age, comparing the five evaluated volatility distributions of P-S/IVOCs. The median (50th percentile) line is depicted as a black line, surrounded by the 25th to 75th percentile range, shaded in grey. The 95th and 5th percentile results are described in the dashed lines. The top row represents ensemble results for Flight 19 in the 2013 JOSM campaign, while the bottom row represents results for Flight 20. The O:C is modeled using the total number of oxygen and carbon atoms for each component. All measurements, described with black squares are $\pm$ 30%, representing the measurement uncertainty. **A)** and **B)** use the volatility distributions previously described in Liggio et al, 2016, with the 20 and 60 ^{o}C distributions, respectively. **C)** and **D)** use the volatility distributions of emissions from raw ore and bitumen, respectively, as described in Li et al, 2019. **E)** through **H)** represent the same ensemble runs as **A)** through **D)**, respectively, for Flight 20.

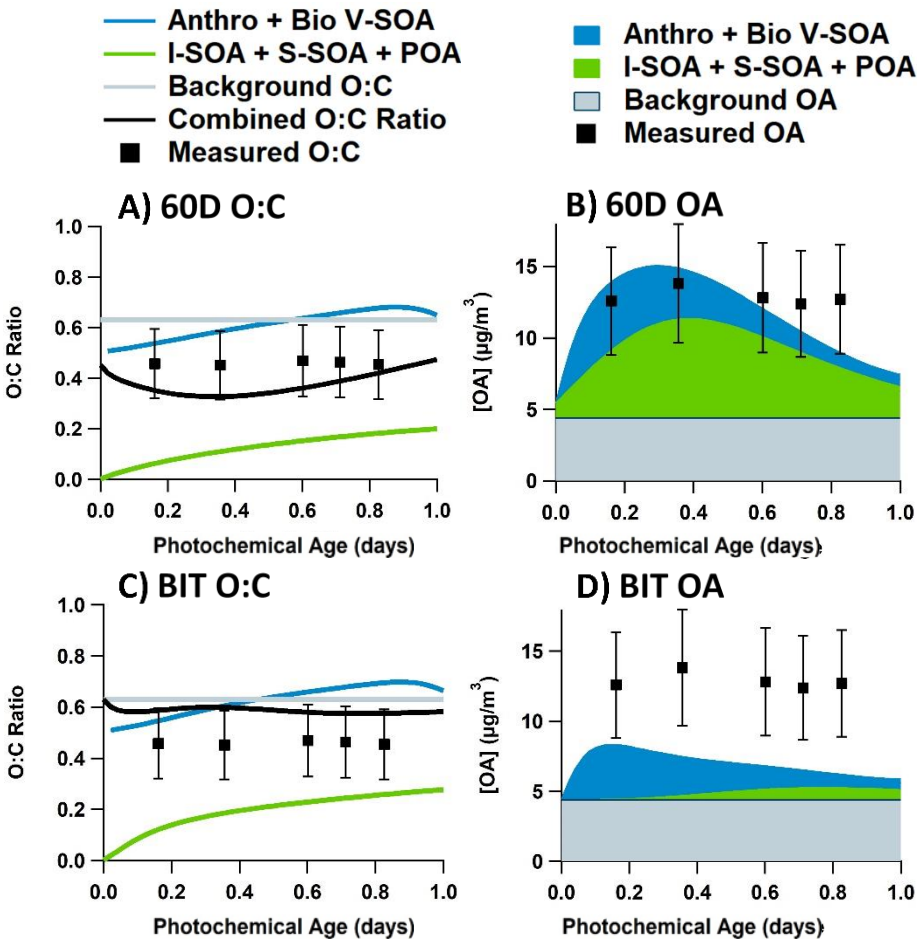

Figure S7. Evaluation of O:C ratios and OA formation and evolution for two volatility distributions used in the box model. Each figure uses the base case study (AGE + SLOW + 1P + F_ON + FIXED parameterizations). Model-measurement comparisons are shown for the **60D (A, B)** and **BIT (C, D)** volatility distributions during Flight 19. Each measurement data point (black square) represents an aircraft transect perpendicular to the plume direction. In all runs**,** S-SOA represents SOA sourced from P-SVOCs, I-SOA represents SOA sourced from P-IVOCs, V-SOA represents SOA from traditional VOCs from anthropogenic and biogenic sources. In **A)** and **C)** O:C is modeled using the number of oxygen and carbon atoms for each component, where the combined O:C ratio (solid black line) is calculated from the sum of all oxygen and carbon atoms of all components. Background OA represents the total OA in measurements from outside the plumes with a stable value of 4.5 µg m^{-3}. All measurements have a ± 30% measurement uncertainty.

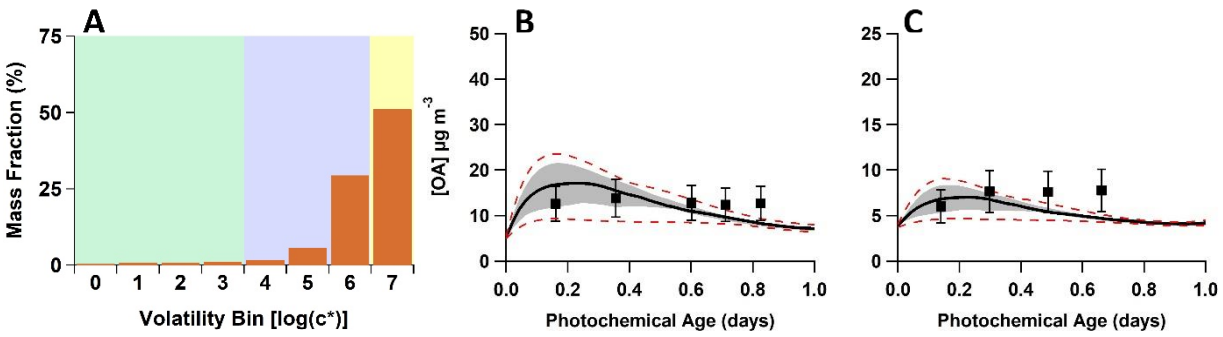

Figure S8. **(A)** P-S/IVOC distribution of BEU volatility distributions. P-S/IVOCs are binned into several saturation concentrations, c* at 298K, to account for the average volatility of the complex mixtures. **(B, C)** Ensemble model–measurement comparison of organic aerosol formation and evolution vs photochemical age for the BEU volatility distribution. For model results, the median (50th percentile) line is depicted as a black line, surrounded by the 25th to 75th percentile range, shaded in grey. The 95th and 5th percentile results are described in the dashed lines. **(B)** represents Flight 19, while **(C)** represents Flight 20. All measurements, described with black squares are $\pm$ 30%, representing the measurement uncertainty.

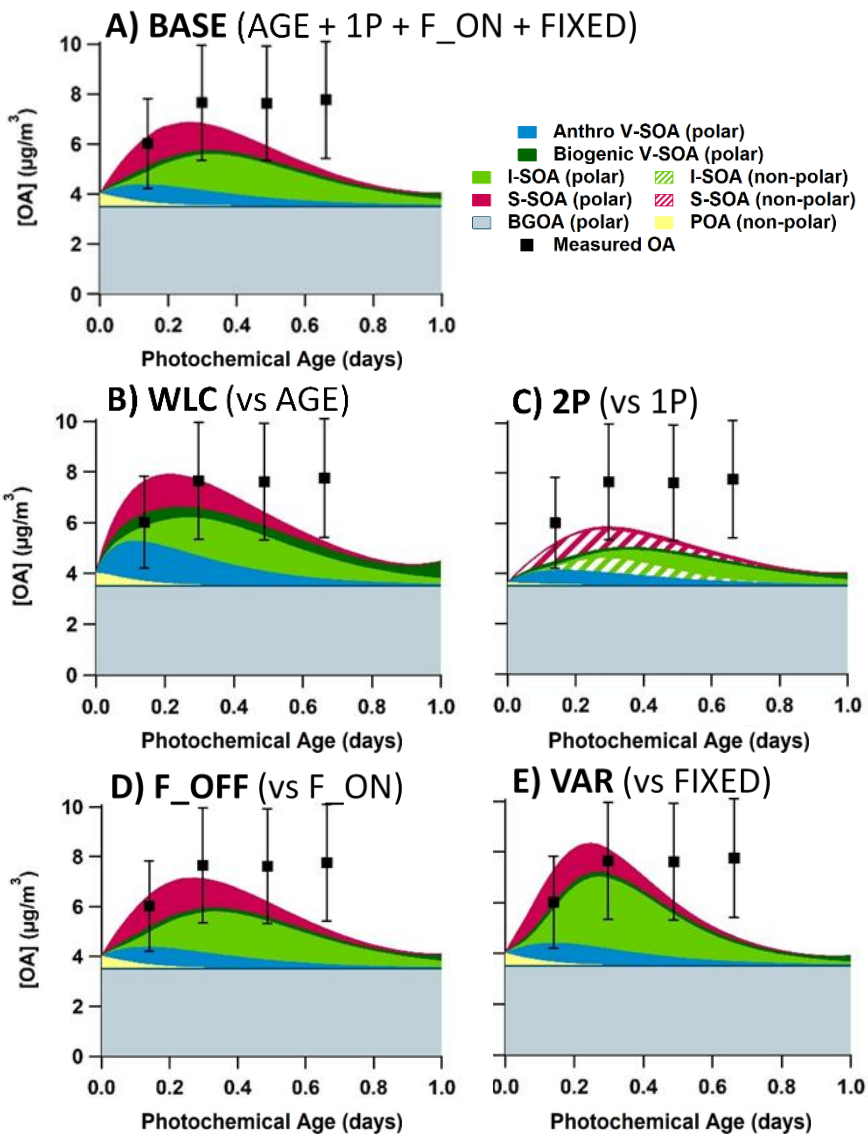

Figure S9: OA model-measurement comparisons of ensemble box model parameterizations compared to the base case **(A)**. All case studies in this figure evaluate Flight 20, with the 60D volatility distribution and the SLOW parameterization for P-IVOC oxidation. The base case uses the **AGE + 1P + F_ON + FIXED** parameterizations. In panels **(B)** through **(E)**, a single parameterization is altered, compared to the baseline. The P-VOC oxidation mechanism is changed with the **WLC** parameterization in **(B)**. Phase separation is evaluated with the **2P** parameterization in **(C)**. Fragmentation is evaluated with the **F_OFF** parameterization in **(D)**. S/IVOC oxidation rates are changed with the **VAR** parameterization in **(E)**. Each measurement data point represents an aircraft transect perpendicular to the plume direction. In all runs, S-SOA represents SOA sourced from P-SVOCs, I-SOA represents SOA sourced from P-IVOCs, V-SOA represents SOA from traditional VOCs from anthropogenic and biogenic sources. Background OA represents the total OA in measurements from outside the plumes with a stable value of 3.5 µg m^{-3} for Flight 20. All measurements have a $\pm$ 30% measurement uncertainty.

References:

1 S.-M. Li, A. Leithead, S. G. Moussa, J. Liggio, M. D. Moran, D. Wang, K. Hayden, A. Darlington, M. Gordon, R. Staebler, P. A. Makar, C. A. Stroud, R. McLaren, P. S. K. Liu, J. O. Brien, R. L. Mittermeier, J. Zhang, G. Marson, S. G. Cober, M. Wolde and J. J. B. Wentzell, Differences between measured and reported volatile organic compound emissions from oil sands facilities, *Proc. Natl. Acad. Sci.*, 2017, **114**, E3756–E3765.

2 J. Liggio, S.-M. Li, K. Hayden, Y. M. Taha, C. Stroud, A. Darlington, B. D. Drollette, M. Gordon, P. Lee, P. Liu, A. Leithead, S. G. Moussa, D. Wang, J. O'Brien, R. L. Mittermeier, J. R. Brook, G. Lu, R. M. Staebler, Y. Han, T. W. Tokarek, H. D. Osthoff, P. A. Makar, J. Zhang, D. L. Plata and D. R. Gentner, Oil sands operations as a large source of secondary organic aerosols, *Nature*, 2016, **534**, 91–94.

3 J. Zhang, M. D. Moran, Q. Zheng, P. A. Makar, P. Baratzadeh, G. Marson, P. Liu and S. M. Li, Emissions preparation and analysis for multiscale air quality modeling over the Athabasca Oil Sands Region of Alberta, Canada, *Atmos. Chem. Phys.*, 2018, **18**, 10459–10481.

4 P. Veres, J. B. Gilman, J. M. Roberts, W. C. Kuster, C. Warneke, I. R. Burling and J. De Gouw, Development and validation of a portable gas phase standard generation and calibration system for volatile organic compounds, *Atmos. Meas. Tech.*, 2010, **3**, 683–691.

5 A. K. Y. Lee, M. G. Adam, J. Liggio, S. M. Li, K. Li, M. D. Willis, J. P. D. Abbatt, T. W. Tokarek, C. A. Odame-Ankrah, H. D. Osthoff, K. Strawbridge and J. R. Brook, A large contribution of anthropogenic organo-nitrates to secondary organic aerosol in the Alberta oil sands, *Atmos. Chem. Phys.*, 2019, **19**, 12209–12219.

6 T. W. Tokarek, C. A. Odame-Ankrah, J. A. Huo, R. McLaren, A. K. Y. Y. Lee, M. G. Adam, M. D. Willis, J. P. D. D. Abbatt, C. Mihele, A. Darlington, R. L. Mittermeier, K. Strawbridge, K. L. Hayden, J. S. Olfert, E. G. Schnitzler, D. K. Brownsey, F. V. Assad, G. R. Wentworth, A. G. Tevlin, D. E. J. J. Worthy, S.-M. M. Li, J. Liggio, J. R. Brook and H. D. Osthoff, Principal component analysis of summertime ground site measurements in the Athabasca oil sands with a focus on analytically unresolved intermediate-volatility organic compounds, *Atmos. Chem. Phys.*, 2018, **18**, 17819–17841.

7 D. D. Parrish, A. Stohl, C. Forster, E. L. Atlas, D. R. Blake, P. D. Goldan, W. C. Kuster and J. A. de Gouw, Effects of mixing on evolution of hydrocarbon ratios in the troposphere, *J. Geophys. Res. Atmos.*, 2007, **112**, 1–17.

8 L. I. Kleinman, S. R. Springston, P. H. Daum, Y.-N. Lee, L. J. Nunnermacker, G. I. Senum, J. Wang, J. Weinstein-Lloyd, M. L. Alexander, J. Hubbe, J. Ortega, M. R. Canagaratna and J. Jayne, The time evolution of aerosol composition over the Mexico City plateau, *Atmos. Chem. Phys.*, 2008, **8**, 1559–1575.

9 P. F. Decarlo, I. M. Ulbrich, J. Crounse, B. De Foy, E. J. Dunlea, A. C. Aiken, D. Knapp, A. J. Weinheimer, T. Campos, P. O. Wennberg and J. L. Jimenez, Investigation of the sources and processing of organic aerosol over the Central Mexican Plateau from aircraft measurements during MILAGRO, *Atmos. Chem. Phys.*, 2010, **10**, 5257–5280.

10 J. C. Schroder, P. Campuzano-Jost, D. A. Day, V. Shah, K. Larson, J. M. Sommers, A. P. Sullivan, T. Campos, J. M. Reeves, A. Hills, R. S. Hornbrook, N. J. Blake, E. Scheuer, H. Guo, D. L. Fibiger, E. E. McDuffie, P. L. Hayes, R. J. Weber, J. E. Dibb, E. C. Apel, L. Jaeglé, S. S. Brown, J. A. Thornton and J. L. Jimenez, Sources and Secondary Production of Organic Aerosols in the Northeastern United States during WINTER, *J. Geophys. Res. Atmos.*, 2018, **123**, 7771–

7796.

11 A. P. Tsimpidi, V. A. Karydis, M. Zavala, W. Lei, L. Molina, I. M. Ulbrich, J. L. Jimenez and S. N. Pandis, Evaluation of the volatility basis-set approach for the simulation of organic aerosol formation in the Mexico City metropolitan area, *Atmos. Chem. Phys.*, 2010, **10**, 525–546.

12 P. K. Ma, Y. Zhao, A. L. Robinson, D. R. Worton, A. H. Goldstein, A. M. Ortega, J. L. Jimenez, P. Zotter, A. S. H. Prévôt, S. Szidat and P. L. Hayes, Evaluating the impact of new observational constraints on P-S/IVOC emissions, multi-generation oxidation, and chamber wall losses on SOA modeling for Los Angeles, CA, *Atmos. Chem. Phys.*, 2017, **17**, 9237–9259.

13 J. L. Jimenez, M. R. Canagaratna, N. M. Donahue, A. S. H. Prevot, Q. Zhang, J. H. Kroll, P. F. DeCarlo, J. D. Allan, H. Coe, N. L. Ng, A. C. Aiken, K. S. Docherty, I. M. Ulbrich, A. P. Grieshop, A. L. Robinson, J. Duplissy, J. D. Smith, K. R. Wilson, V. A. Lanz, C. Hueglin, Y. L. Sun, J. Tian, A. Laaksonen, T. Raatikainen, J. Rautiainen, P. Vaattovaara, M. Ehn, M. Kulmala, J. M. Tomlinson, D. R. Collins, M. J. Cubison, J. Dunlea, J. A. Huffman, T. B. Onasch, M. R. Alfarra, P. I. Williams, K. Bower, Y. Kondo, J. Schneider, F. Drewnick, S. Borrmann, S. Weimer, K. Demerjian, D. Salcedo, L. Cottrell, R. Griffin, A. Takami, T. Miyoshi, S. Hatakeyama, A. Shimono, J. Y. Sun, Y. M. Zhang, K. Dzepina, J. R. Kimmel, D. Sueper, J. T. Jayne, S. C. Herndon, A. M. Trimborn, L. R. Williams, E. C. Wood, A. M. Middlebrook, C. E. Kolb, U. Baltensperger and D. R. Worsnop, Evolution of Organic Aerosols in the Atmosphere, *Science (80-.).*, 2009, **326**, 1525–1529.

14 M. Shiraiwa, L. D. Yee, K. A. Schilling, C. L. Loza, J. S. Craven, A. Zuend, P. J. Ziemann and J. H. Seinfeld, Size distribution dynamics reveal particle-phase chemistry in organic aerosol formation., *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 11746–50.

15 N. M. Donahue, W. Chuang, S. A. Epstein, J. H. Kroll, D. R. Worsnop, A. L. Robinson, P. J. Adams and S. N. Pandis, Why do organic aerosols exist? Understanding aerosol lifetimes using the two-dimensional volatility basis set, *Environ. Chem.*, 2013, **10**, 151–157.

16 A. L. Robinson, N. M. Donahue, M. K. Shrivastava, E. a Weitkamp, A. M. Sage, A. P. Grieshop, T. E. Lane, J. R. Pierce and S. N. Pandis, Rethinking Organic Aerosols: Semivolatile Emissions and Photochemical Aging, *Science (80-.).*, 2007, **315**, 1259–1262.

17 S. H. Jathar, M. A. Miracolo, A. A. Presto, N. M. Donahue, P. J. Adams and A. L. Robinson, Modeling the formation and properties of traditional and non-traditional secondary organic aerosol: Problem formulation and application to aircraft exhaust, *Atmos. Chem. Phys.*, 2012, **12**, 9025–9040.

18 S. H. Jathar, T. D. Gordon, C. J. Hennigan, H. O. T. Pye, G. Pouliot, P. J. Adams, N. M. Donahue and A. L. Robinson, Unspeciated organic emissions from combustion sources and their influence on the secondary organic aerosol budget in the United States, *Proc. Natl. Acad. Sci.*, 2014, **111**, 10473–10478.

19 N. M. Donahue, A. L. Robinson, C. O. Stanier and S. N. Pandis, Coupled partitioning, dilution, and chemical aging of semivolatile organics, *Environ. Sci. Technol.*, 2006, **40**, 2635–2643.

20 R. Atkinson and J. Arey, Atmospheric Degradation of Volatile Organic Compounds Atmospheric Degradation of Volatile Organic Compounds, *Chem. Rev.*, 2003, **103**, 4605–4638.

21 J. H. Kroll and J. H. Seinfeld, Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere, *Atmos. Environ.*, 2008, **42**, 3593–3624.

22 H. J. Chacon-Madrid and N. M. Donahue, Fragmentation vs. functionalization: Chemical aging

and organic aerosol formation, *Atmos. Chem. Phys.*, 2011, **11**, 10553–10563.

23 B. Koo, E. Knipping and G. Yarwood, 1.5-Dimensional volatility basis set approach for modeling organic aerosol in CAMx and CMAQ, *Atmos. Environ.*, 2014, **95**, 158–164.

24 N. M. Donahue, S. A. Epstein, S. N. Pandis and A. L. Robinson, A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics, *Atmos. Chem. Phys.*, 2011, **11**, 3303–3318.

25 R. Volkamer, J. L. Jimenez, F. San Martini, K. Dzepina, Q. Zhang, D. Salcedo, L. T. Molina, D. R. Worsnop and M. J. Molina, Secondary organic aerosol formation from anthropogenic air pollution: Rapid and higher than expected, *Geophys. Res. Lett.*, 2006, **33**, 7–10.