Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

Atmospheric Chemistry and Physics Discussions

Air Quality Predictions using Measurement-Derived Organic Gaseous and Particle Emissions in a Petrochemical-Dominated Region

Craig A Stroud[1], Paul A Makar[1], Junhua Zhang[1], Michael D. Moran[1], Ayodeji Akingunola[1], Shao-Meng Li[1], Amy Leithead[1], Katherine Hayden[1], and May Siu[2]

[1]Air Quality Research Division, Environment and Climate Change Canada, 4905 Dufferin Street, Toronto, Ontario, M3H 5T4, Canada
[2]Air Quality Research Division, Environment and Climate Change Canada, 335 River Road, Ottawa, Ontario, K1V 1C7, Canada

Corresponding author: Craig A. Stroud (craig.stroud@canada.ca)

Abstract

This study assesses the impact of revised volatile organic compound (VOC) and organic aerosol (OA) emissions estimates in the GEM-MACH (Global Environmental Multiscale–Modelling Air Quality and CHemistry) chemical transport model, driven with two different emissions input datasets, using observations from the 2013 Joint Oil Sands Monitoring (JOSM) intensive field study. The first emissions dataset (base-case run) makes use of regulatory-reported VOC and particulate matter emissions data for the large oil sands mining facilities in northeastern Alberta, Canada, while the second emissions dataset (sensitivity run) uses emissions estimates based on box-flight aircraft observations around specific facilities (Li *et al*., 2017, Zhang *et al*., 2017) and a mass-balance analysis (Gordon *et al*., 2015) to derive total facility emission rates. The preparation of model-ready emissions files for the base-case and sensitivity run is described in an accompanying paper by Zhang *et al*. (2017).

The large increases in VOC and OA emissions in the revised emissions data set for four large oil sands mining facilities were found to improve the modeled VOC and OA concentration maxima in plumes from these facilities, as shown with the 99th percentile statistic and illustrated by case studies. The results show that the VOC emission speciation profile from each oil sand facility is unique and different from standard petrochemical-refinery emission speciation profiles

used for other regions in North America. A feedback between larger long-chain alkane emissions and higher secondary organic aerosol (SOA) concentrations was found to be significant for some facilities and improved OA predictions for those plumes. The use of the revised emissions data resulted in a large improvement of the model OA bias; however, the decrease in OA correlation coefficient suggests the need for further improvements to model organic aerosol emissions and formation processes. Including intermediate volatile organic compound (IVOC) emissions as precursors to SOA and spatially allocating more PM_1 POA emissions (primary organic aerosol of 1.0 μm or less in diameter) to mine-face locations are both recommended to improve OA bias and correlation further. A systematic bias in the background OA was also predicted on most flights, likely due to under-predictions in biogenic SOA formation. Overall, the weight of evidence suggests that the new aircraft-observation-derived organic emissions help to constrain better the fugitive organic emissions, which are a challenge to estimate in the creation of bottom-up emission inventories. This work shows that the use of facility-specific emissions, based on direct observations, rather than generic emission factors and speciation profiles can result in improvements to model predictions of VOCs and OA. Emissions estimation techniques, such as those used to construct the inventories in our study, may therefore have beneficial impacts when applied to other regions with large sources of VOCs and OA**.**

1 Introduction

Chemical transport models (CTMs) are useful tools to support clean energy policy decisions because they can be used to assess the impact of past and future pollutant emission changes on air quality (e.g., Schultz *et al*., 2003; Kelly *et al*., 2012; Rouleau *et al*., 2013; Lelieveld *et al*., 2015). CTMs can also be run in forecast mode with their output being used to support air quality forecasts (Moran *et al*., 2010; Chai *et al*., 2013). CTMs require pollutant emission inputs,

typically at hourly intervals, at the model grid spatial resolution (Dickson and Oliver, 1991; Houyoux *et al*., 2003; Pouliot *et al*., 2012, 2015; Zhang *et al*., 2017). The pollutant emission input files are based on the processing of emission inventories compiled for all emission sectors, usually at some geopolitical spatial resolution (e.g., county, province/state, or country), and may thus require the application of spatial disaggregation factor fields to allocate emissions to the model grid. North American emission inventories are typically derived from bottom-up approaches, where representative pollutant emission factors (e.g., pollutant mass emission per volume of fuel burned) are multiplied by activity factors (e.g., volume of fuel burned per unit time). In developed countries, industrial facilities are usually required to report estimates of their pollutant emissions to national inventories such as the National Pollutant Release Inventory (NPRI) in Canada (https://www.canada.ca/en/environment-climate-change/services/national-pollutant-release-inventory.html) and the National Emissions Inventory (NEI) in the United States (https://www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei.html). Updates of these inventories occur under a regulatory framework on a regular basis. However, reporting requirements may be limited to aggregated mass emissions on an annual basis (e.g., a total bulk mass of VOC emitted rather than a detailed and observation-based emissions of individual speciated VOCs), with the subsequent use of VOC speciation profiles (splitting factors) to determine the relative contribution of the individual VOCs to the total VOC emissions. Uncertainties in the availability and assignment of appropriate VOC speciation profiles, spatial and temporal allocation factors (Mashayekhi *et al*., 2016), and/or unaccounted-for emitting activities, result in the need to evaluate the impact of these assumptions through the comparison of CTM predictions with ambient observations.

The Athabasca region of northeastern Alberta, Canada has one of the largest reserves of oil sands (OS) in the world. The OS deposits are composed of bitumen, minerals, sand and clay. Oil sand near the surface is mined by open-pit mining techniques. The oil sand is then transported by heavy hauler trucks to crushers, followed by the addition of hot water to make the oil sand flow through pipelines to a bitumen extraction facility. Here, the bitumen is separated from the sand and clay by the use of organic solvents. The product is then upgraded on-site to crude oil or transported to a remote upgrader facility. Volatile organic compounds from the bitumen have the potential to escape into the atmosphere as fugitive emissions during the mining, extraction, processing, or tailing discharge steps. The complexity and vast size of the oil sands operations make generating pollutant emission input files for CTMs a challenge (Cho *et al*., 2012; ECCC & AEP, 2016).

Organic compounds in the atmosphere are oxidized over time and, in the presence of sufficient levels of oxides of nitrogen, are important precursors to ozone formation (Seinfeld and Pandis, 1998). VOCs and semi-volatile organic compounds (SVOCs) are also precursors to secondary organic aerosol (SOA) formation (Griffin *et al*., 1999; Kanakidou *et al*., 2005; Robinson *et al*., 2007; Kroll and Seinfeld, 2008; Slowik *et al*., 2010; Stroud *et al*., 2011; Gentner *et al*., 2017). If the organic compounds have sufficiently low saturation vapor pressures, then upon release into the atmosphere they remain particle-bound and are classified as primary organic aerosol (POA). Many specific organic compounds can also be toxic to human health and require explicit reporting in emission inventories (Stroud *et al*., 2016).

Top-down approaches to estimate emissions based on atmospheric observations provide a unique opportunity to compare with bottom-up calculated emissions. One such approach has recently been applied for Athabasca OS facilities in Alberta, Canada (Gordon *et al*, 2015; Li *et*

al., 2017). The mass-balance approach that was used is based on using box-shaped aircraft flight patterns around a facility and measuring pollutant concentrations and meteorological variables (wind speed and direction, air density). In this approach, the difference in pollutant mass fluxes entering and leaving the box is used to determine the total facility-wide emission rate, subject to assumptions such as minimal losses due to chemical oxidation between the emissions location and the nearby aircraft observations. This emission estimate can then be compared with the reported bottom-up emission total.

The Joint Oil Sands Monitoring Program (JOSM) was developed by the federal government of Canada and the Alberta provincial government with input and consultation from the local indigenous population and industry stakeholder groups. Environment and Climate Change Canada (ECCC)'s chemical transport model, GEM-MACH (Global Environmental Multi-scale-Modelling Air quality and CHemistry) is being used in JOSM to assess the impact of current emissions and future emission changes on local air quality and downwind regional-scale acid deposition. Evaluations of the model performance in different configurations and with respect to other pollutants may be found elsewhere in this special issue (Makar *et al*., 2017; Akingunola *et al.*, 2017).

Here we make use of both regulatory-inventory-based and aircraft-observation-based emissions data for VOCs and primary particulate emissions for six large OS mining facilities as inputs to GEM-MACH in order to assess the impact of these sources of information on model predictions of VOC concentrations and organic aerosol (OA) formation. The base-case inventory, which was derived from regulatory reporting, and updates for point sources, spatial and temporal allocation, and measured top-down, facility-total aircraft-measurement-based

emission rates are described in detail for VOCs in Li *et al* (2017) and for particulate matter (PM) in Wang *et al*. (2015) and Zhang *et al*. (2017).

2 Methods

The GEM-MACH model uses the ECCC operational weather forecast model (GEM) as the core operator for dynamics and microphysical processes (Côté *et al*., 1998a,b; Girard *et al*., 2014). GEM-MACH is an "on-line" CTM - the chemistry, vertical diffusion, and pollutant deposition routines exist as a set of subroutines contained and called from within GEM's meteorological physics package (Moran *et al*., 2010, Makar *et al*., 2015a,b). The gas-phase chemistry scheme is based on the ADOM-II mechanism, originally developed for continental boundary-layer oxidant formation. The VOC lumped species used in GEM-MACH are described in Stroud *et al*. (2008). The focus here is on evaluating volatile aromatic and alkane species of anthropogenic origin. The aerosol size distribution is described by a 12-bin sectional approach based on the Canadian Aerosol Module (CAM) (Gong *et al*., 2003; Park *et al*., 2011). The SOA scheme is based on a two-product fit to smog chamber data using the SOA yield equations derived from gas/particle partitioning theory (Pankow 1994; Griffin *et al*., 1999; Barsanti et al., 2013). In the GEM-MACH model's current SOA formation algorithms, after initial particle formation, the organic compounds in the particle phase are assumed to be converted rapidly to non-volatile mass, as observed by recent studies (Cappa and Jimenez, 2010; Cappa *et al*., 2011; Lopez-Hilfiker *et al*., 2016) and recommended by modelling studies (Shrivastava *et al*., 2015). However, other recent observation studies suggest that SOA 'chemical aging' over hours to days is quite complex, and involves further gas-phase oxidation and fragmentation reactions (Jimenez *et al*., 2009; Donahue *et al*., 2014), as well as potential particle-phase oxidation and oligomer reactions (McNeill *et al*., 2015). The particle oligomer reactions are rapid, often acid-catalyzed,

and can result in conversion to non-volatile mass (Liggio *et al*., 2005; Kroll *et al*., 2005). We discuss below the evidence from this work on the likelihood that these additional missing processes are still impacting our model organic aerosol bias.

The Canadian base-case emissions were derived from a hybrid inventory targeting 2013 as the base year, as described by Zhang *et al*. (2017). This base year was chosen to align with the JOSM intensive field study period, which provided the observations for the model/observation comparisons that follow. Canadian emissions for industrial facilities, including the Athabasca OS mining facilities, were obtained from the 2013 NPRI. The U.S. base-case emissions were obtained from the 2011 U.S. NEI Version 1 (Eyth *et al*., 2013).

These base-case, bottom-up emissions inventories were processed with the SMOKE emissions processing tool (https://www.cmascenter.org/smoke), which includes three major steps corresponding to spatial allocation, temporal allocation, and chemical speciation (for NOx, VOC, and PM). The base-case VOC speciation profiles used by SMOKE for the OS surface mining facilities were obtained from the CEMA (Cumulative Environmental Management Association) inventory (Davies *et al*., 2012; Zhang *et al*., 2015).

For the sensitivity run, speciated VOC emissions from the base case for four OS mining facilities (Suncor Millenium/Steepbank, Syncrude Mildred Lake, Shell Canada Muskeg/Jackpine, and CNRL Horizon) were revised by replacing them with the top-down emission rates estimated by Li *et al*. (2017) while primary PM emissions were revised for six oil sand facilities (Suncor Millenium/Steepbank, Syncrude Mildred Lake, Shell Canada Muskeg/Jackpine, CNRL Horizon, Syncrude Aurora North, and Imperial Oil Kearl) (Zhang *et al*., 2017). The VOC and PM chemical speciation profiles used for these facilities were also revised using the aircraft-observed VOC speciation (Li *et al*., 2017) and ground-based PM filter

analysis (Wang *et al*., 2015), respectively. The set of emissions input files making use of these revisions is hereafter referred to as the "revised emissions", while the original emissions input files without these changes is referred to as the "base-case emissions". A detailed description of the development of the emission inventory and emissions processing steps to create the model-ready files (hourly gridded emission fields for the same domain and grid spacing as the model) for the base case and revised version are described in Zhang *et al*. (2017).

Depending on whether bitumen extracted from the oil sand is upgraded on site or not, the OS mining facilities can be classified into two broad types: (1) integrated extraction and upgrading facilities (Suncor Millenium/Steepbank, Syncrude Mildred Lake, and CNRL Horizon) and extraction-only facilities (Shell Canada Muskeg/Jackpine, Syncrude Aurora North, and Imperial Oil Kearl). Table S1 shows a comparison of the VOC speciation profiles used for the two types of OS facilities compared to two speciation profiles that were used for the base case to speciate more than half of the U.S. refinery emissions in the Houston area. There are significant differences between the aircraft-observation-based OS VOC speciation profiles and the two commonly used refinery reference profiles. The OS integrated extraction and upgrading facilities are higher in long-chain alkenes, toluene, and other aromatics than the reference profiles, while the extraction-only facilities have the highest long-chain alkane fraction. The other profiles have higher less-reactive species (e.g., propane, acetylene) and formaldehyde, than the OS profiles.

The primary PM emissions from the OS facilities originate largely from off-road heavy-duty diesel trucks, plant stack emissions, and fugitive and wind-blown dust. The 2009/10 CEMA inventory was used to specify the tail-pipe emissions from the off-road mining fleet and the 2013 NPRI inventory was used for fugitive road-dust emissions. The base-case inventory did not include wind-blown dust. For the revised inventory, the PM size distribution was measured

during the 2013 field study for all six facilities and these data were used to constrain the revised PM emission input data set. Note that the PM emissions estimates based on the aircraft-measured aerosol data included the contribution of wind-blown dust emissions. The aircraft-based PM emissions were re-binned for the 12 GEM-MACH PM size bins. The first eight size bins correspond to mass up to diameter 2.56 µm. Interestingly, the aircraft measured a much higher fraction of particulate mass in bin 8 (bounded by diameters 1.28 and 2.56 µm) compared to the mass fraction in bin 8 from the area-source PM size-distribution profiles used by SMOKE in processing the base-case emissions. In addition, a PM chemical speciation profile specific to OS fugitive dust emissions was created from an analysis of deposited dust collected from surfaces in the OS region (Wang *et al*., 2015); this speciation profile replaced the standard fugitive dust profile for unpaved roads from the U.S. EPA SPECIATE v4.3 database (https://www.epa.gov/air-emissions-modeling/speciate-version-45-through-40) in the revised emissions processing. The resulting organic carbon fraction in the observation-derived PM speciation profile was higher than that of the base-case emissions by about a factor of 3 (Zhang *et al.*, 2017). In general, significantly higher POA emissions were observed over the open-pit mines for all facilities, except for the Imperial Kearl mine. This will be discussed further in Section 3.4.

The GEM-MACH model was run in a nested configuration with an outer domain covering the continental U.S. and Canada and an inner domain covering Alberta and Saskatchewan. The continental-scale GEM-MACH model (10-km resolution) and the Canada-wide GEM weather model (2.5-km resolution) were run first. These provided the chemical and meteorological lateral boundary conditions, respectively, for the high-resolution GEM-MACH 2.5-km resolution run, which has a domain covering the provinces of Alberta and Saskatchewan (Figure 1). The

two models providing boundary conditions were run on a 30-hour cycle, of which the first six hours were spin-up and discarded, while the remaining 24 hours provided boundary conditions for the 2.5-km GEM-MACH simulation. The initial conditions subsequent to the starting model simulation for each overlapping 24-hour 2.5-km GEM-MACH simulation came from the end of the previous 2.5-km GEM-MACH simulation. This strategy was used to allow the two boundary condition simulations to make use of assimilated meteorological analyses. The sequence of model simulations was started for August 10, 2013 and run until September 7 to cover the 2013 JOSM intensive field study period.

The NRC (National Research Council) Convair two-engine turboprop aircraft was used to collect air-quality observations during the JOSM 2013 intensive field study. The aircraft was equipped with a suite of instruments to measure air quality over 22 flights (see Li et al., 2017, Fig. S1). Most of the flight hours focused on "box" flight paths; these took the aircraft around the periphery of facilities at different heights, with the goal of deriving facility-wide emission rates by using observations of chemical concentrations and winds to estimate the mass of pollutants entering and leaving the box enclosures. Coupled with a mass-conserving flux model (Gordon *et al*., 2015), these aircraft data were used to estimate emissions from the encircled facilities.

VOC and PM observations were collected by the instrumented research aircraft using different technologies. A proton-transfer-reaction mass spectrometer (PTR-MS) was used to measure a select number of VOCs at high temporal resolution (1-sec) (Li *et al*., 2017). An aerosol mass spectrometer (AMS) was used to measure PM_1 mass and non-refractory chemical composition (Liggio *et al*., 2016). A Single Particle Soot Photometer (SP2) was used to measure refractory black carbon aerosol (Liggio *et al*., 2016). Black carbon is used in our analysis, as a proxy for transport from open-pit mine-face sources. A number of canisters were filled with

ambient air on each flight and returned to the lab for GC-FID and GC-MS analysis of VOCs (Li *et al*., 2017). The canister VOC analysis measured 154 different C_2 to C_{12} hydrocarbons (Dann and Wang, 1995). Each flight typically filled ~30 stainless-steel canisters. The resulting observation data were compared to the model output generated as described above. The 2.5-km GEM-MACH runs used a 120-s chemistry time step; 120-s model output values were linearly interpolated in time and space to the aircraft observation locations; all comparisons which follow make use of the resulting model/observation data pairs for the two simulations.

3 Results and Discussion

We present our evaluation results for four VOC classes: mono-substituted aromatics in section 3.1; multi-substituted aromatics in section 3.2; long-chain alkane species in section 3.3; and organic aerosols in section 3.4.

3.1 Toluene and other Mono-Substituted Aromatics (TOLU)

The aircraft PTR-MS measurement data set was averaged to 10-sec intervals for comparison to time and spatially linearly interpolated GEM-MACH model output. The interpolated model output along the flight track was merged with the observations to create a coincident model–measurement time series. The model lumped TOLU species includes toluene and other mono-substituted aromatics with the two most important additional species being ethyl-benzene and propyl-benzene. Therefore, we must derive an equivalent observed lumped TOLU species for a comparison. We used all of the canister VOC data from the field study to create ethyl-benzene vs. toluene and propyl-benzene vs. toluene scatterplots. The corresponding slope and correlation coefficient for both these plots (not shown) were as follows: m=0.376, R=0.91 and m=0.0652, R=0.90, respectively. Thus, we derived an observed TOLU equal to the PTR-MS C7 aromatic

multiplied by the factor 1.4412 (sum of m=1.0 C7+0.376 C7+0.0652 C7). This new observation-derived TOLU was used in the statistical comparison with model output TOLU, which follows.

Histograms of mixing ratio were created using the observed TOLU, the revised-emissions model output, and the base-case model output. Figure S1 illustrates the histograms using 20 mixing-ratio bins and an increment of 0.2 ppbv per bin. It is clear that there are more high values (>2 ppbv) produced by the sensitivity model run with revised emissions compared to the base-case model run. The number of observations in the highest value bins lies between the results from the revised and base-case versions. This can be quantified by using the 99% percentile statistic (obs=1.258 ppbv, revised=1.906 ppbv, base=0.934 ppbv). The 99% percentile means that 99% of the data points are lower than the value. The median concentration of the observations (0.061 ppbv) is higher than both the revised (0.038 ppbv) and base-case model (0.019 ppbv) simulated values, but is closer to the revised version. Table 1 lists other statistical scores for the TOLU lumped species. The mean bias goes from a negative value with the base-case run to a positive value with the revised emissions. There is little difference in the correlation coefficient for the model vs. observation scatterplot between the base-case and sensitivity run. The changes to the VOC emissions for the revised-emissions run affected their total mass and speciation, and the observations were made sufficiently close to the sources that there was little time for oxidation. The main sources for VOCs are the processing plants, tailing ponds, mine faces, and off-road vehicles and their spatial allocation (from CEMA, 2010) did not change significantly between the two model-emission versions. The main differences in the model time series between the two simulations are thus in magnitude of concentrations, and hence relatively invariant correlation coefficients might be expected.

286 Figure S2 illustrates the difference between the 2.5-km TOLU emissions field over the
287 Athabasca OS mining facilities for the revised-base case as a spatial map for one selected time
288 (Friday , August, 18Z). The largest increases are noted for the Syncrude Mildred Lake facility
289 over the tailing ponds and open pit mine faces. The changes shown here reflect net changes on a
290 facility-wide level. Notable actual and relative increases are also calculated for the Suncor
291 Millennium/Steepbank and the Canadian Natural Resources Ltd (CNRL) Horizon facilities.
292 Figure 2 illustrates the correlation between the sensitivity-case TOLU output and the base-case
293 TOLU output. The flights on August 14 and 24 have the largest TOLU mixing ratios, and both
294 flights (blue and green points) correspond to box flights around the Syncrude Mildred Lake
295 facility. Figure S3 shows the flight path for August 14, color-coded as a function of the
296 difference between the modelled revised and base-case concentrations. The background is a
297 satellite map image along with the GEM meteorological model wind barbs predicted for that day
298 at 16 UTC. The largest differences in the simulated concentrations (1.8 ppbv) correspond to a
299 location just downwind of the Syncrude Mildred Lake open-pit mine, as expected based on the
300 emission difference map (Figure S2) and the southerly wind direction.
301 Figure 3 shows the time series for a segment of the August 14 flight corresponding to three
302 flight boxes at different heights (green dotted line). The observations are plotted as open circles
303 and the two lines represent the two model results. The model output with the revised VOC
304 emissions clearly captures the main peak of the TOLU concentrations driven by TOLU
305 emissions from the Syncrude Mildred Lake facility. The secondary peak occurs from the aircraft
306 flying over the Suncor Millennium/Steepbank facility while on the east side of the flight box
307 pattern. The direct flyover adds uncertainty to the model comparison, as it depends on predicting
308 accurately the early-stage vertical mixing of the plume from the Suncor facility.

Figure 4 is a time-series segment for August 23 corresponding to a fly-over of the Syncrude Mildred Lake (earlier peak in time) and Suncor Millennium/Steepbank (later peak in time) facilities at a constant altitude of 300-magl. Winds were light on this day with variable swings in direction. A double-peak pattern is observed in both the model and observations with a 1-min time shift needed to align the peaks. For this fly-over, the magnitude of the peaks is better represented with the revised emission model version. We note also that the one-minute lag time of the model peaks illustrates the difficulties in prediction of plume location at high resolution; this corresponds to an error in the forecast position of the plume of 6 km, or 2.4 of the model's grid-cells, given the aircraft's typical flight speed of 100 m s^{-1}. Small errors in wind direction, the potential for point sources located near grid-cell boundaries to effectively be re-located to the grid centroids, as well as directional errors in the forecasted winds, can contribute to these offsets between observed and simulated concentration peaks.

For the TOLU lumped species, the overall statistical scores change from a small negative bias to a small positive bias with little change in correlation coefficient, which may be controlled in part by the meteorological model accuracy as noted above. However, for the locations where the absolute difference in emissions has changed the most (see Fig. S2), the model time series for the revised emissions shows improvement and this is reflected in the improvement in the slope statistics in Table 1. The larger deviation between model and observed RMSE for the revised emissions likely reflects the error in positioning of the plumes inherent in both simulations – in the revised emissions simulation, the positioning error noted above likely contributes to the increased RMSE value.

3.2 Multi-Substituted Aromatics (AROM)

The model lumped AROM species includes all multi-substituted aromatics, with the most important species being the xylene isomers and trimethylbenzene isomers. These two species match with the PTR-MS C8 and C9 aromatic fragments, respectively. However, the observed C8 aromatic also includes ethyl-benzene and the C9 aromatic also includes propyl-benzene, which are lumped with TOLU in the model VOC speciation. Thus, we need to subtract these unwanted species from the totals used to compare to the model lumped AROM species. To do this, we use their correlation slopes with PTR-MS C7 aromatic from Section 3.1. The new observation-derived AROM was calculated from the PTR-MS measurements as follows: C8 + C9 – 0.376 C7 – 0.0652 C7.

Figure S4 shows the histograms for the lumped AROM species for 10-sec averaged points along all the flight tracks. The base model has a large number of high value points (> 2ppbv), many more than the model simulations with the revised emissions, and also more than the observations. This can be quantified by using the 99% percentile (obs=0.7607, revised=1.004, base case=2.302). The median value for the observations is 0.0182 ppbv, smaller than both the model versions (revised=0.0236 ppbv, base case=0.0466 ppbv), but closer to the model driven by the revised emissions. Table 1 lists other statistical scores for the AROM lumped species. The mean bias and RMSE are smaller for the revised emissions run compared to the base case. However, there is a small degradation in the correlation coefficient with the sensitivity run.

Figure S5 displays the difference between the 2.5-km AROM emissions field over the Athabasca OS mining facilities for the revised emissions and base-case emissions. The largest decreases are again over the Syncrude Mildred Lake facility. There were also notable decreases over the CNRL Horizon and Shell Muskeg/Jackpine facilities, but positive changes in emissions were noted over the Suncor Millennium/Steepbank facility. Figure S6 shows the flight track for

the August 23 survey flight, which flew over all the facilities. The background map shows model winds were light and variable on this afternoon. The flight track is color-coded as a function of the difference between AROM from the sensitivity-base case. Consistent with the emission changes, negative difference in ppbv were modelled over Syncrude, Shell, and CNRL and positive differences in ppbv over Suncor.

Figure 5 shows the time series for a segment of the August 23 survey flight over Syncrude Mildred Lake and Suncor Millennium/Steepbank. The largest maxima are for times over Syncrude (7:31Z) and, while both runs show an over-prediction in plumes, the sensitivity run predictions are closer to observations. Panel B is the time series for a short segment later in the flight for locations over the Syncrude (earlier peak in time) and Suncor (later peak in time) facilities. For the Suncor maximum, the sensitivity run with revised emissions has a better prediction for the magnitude of the mixing ratio change.

Figure S7 shows the box flight track on September 3, which was focused on quantifying emissions from Syncrude Mildred Lake facility. The flight path also included some turns over the Suncor Millennium/Steepbank facility. Similar to the August 23 flight, there are negative differences in the predicted AROM mixing ratio between the sensitivity and base runs over Syncrude and positive differences over Suncor. The decreases in mixing ratio are as large as 2 ppbv. Figure 6 is the time series for a segment of the September 3 flight. The observed mixing ratios are closer to the predictions from the revised-emissions model run compared to the base-case run.

In general, from the overall statistics and the case studies, the revised-emissions model run output for lumped AROM species compares better to observations than the base-case emissions

run, reducing the mean bias by a factor of three, and the RMSE by a factor of two (Table 1), and giving a better overall performance for the histograms of AROM concentration.

3.3 Long-Chain Alkanes (ALKA)

The long-chain alkanes (C_4 to C_{12}) were sampled by filling canisters with ambient air on-board the aircraft. Figure S8 presents the histogram for the long-chain alkanes. The mixing ratios are divided into 20 bins each with a width of 3 ppbv. From the observed histogram, there is a wide range to the mixing ratios with a small number of very large concentrations, but also the first bin (0 to 3 ppbv) has a high percentage of the points. The model gas-phase mechanism represents all higher –carbon-number alkanes by a single lumped species, with chemical and physical properties derived from C_4 to C_8 alkanes. The base-case run calculates lower ALKA mixing ratios than the model version using revised emissions. The model using revised emissions is much better at reproducing the higher concentration points, particularly above 12 ppbv. This is quantified by the 99% percentile of the data sets (obs=29.9, base=18.0 revised=24.6). Other statistics for the lumped ALKA species are shown in Table 1. The mean bias went from a small negative value to +1.98 ppbv. The slope decreased by a small value, but the y-intercept increased, which also increased the RMSE for the run with the revised emissions. The correlation coefficient improved significantly for the model run with revised emissions.

Figure S9 shows the difference in emission rate between the revised emissions and the base-case emissions. Interestingly, the revised emissions are considerably higher for the CNRL Horizon and Shell Muskeg/Jackpine facilities, but have smaller changes for the other facilities, possibly reflecting differences in the processing activities between the facilities.

Figure S10 shows the differences between the two model predictions for ALKA at the observation canister sample locations, for the flight on August 26. On this day, winds were from

401 the northeast and notably Fort McMurray (further to the south, not shown) had quite poor air
402 quality. The largest differences in the modelled mixing ratios correspond to observation locations
403 south of CNRL. Positive differences as large as 20 ppbv were simulated for some points. Figure
404 7 shows the time series for the observations, revised-emissions model results, and base-case
405 model results for the Aug. 26 box flight around the CNRL Horizon facility. A clear improvement
406 in ALKA modelling is observed when using the revised emissions for the plume sampled
407 downwind of the CNRL facility.

408 There were two other box flights around the CNRL Horizon facility. The flight on August 20
409 also showed an improvement in ALKA predictions when using the new emission data set.
410 Winds were from the west on this day. The flight on Sept. 2 showed the opposite trend, with
411 more of an over-prediction with the revised emissions. Winds were from the north on this day.
412 The background ALKA on this flight was predicted to be higher for the sensitivity run; however,
413 the differences in mixing ratio between background and plume were still over-predicted with the
414 revised emissions and under-predicted with the base emissions.

415 The other facility that had large increases in ALKA emissions with the revised data was the
416 Shell Muskeg/Jackpine facility (Fig. S9). Flight 9 on August 21 was a box flight around the Shell
417 facility. A detailed analysis of this flight showed that for the majority of the data points on this
418 flight, the model run with the base-case emissions showed the best results, except for the three
419 highest measured canister samples, where the model run with the revised emissions performed
420 better. This likely reflects an uncertainty in the spatial allocation maps used to distribute the
421 emissions with a higher fraction of emissions needed at the point specific locations.
422

423 Overall, the ALKA statistics show mixed improvements associated with the revised emissions
424 using the entire data set. The correlation coefficient does improve significantly. The large
425 increases in ALKA emissions for the CNRL facility did improve the model maxima for time
426 series downwind. The analysis suggests further improvement in spatial allocation for the Shell
427 facility may be needed. The higher ALKA mixing ratios also feeds back to higher SOA
428 formation downwind of these facilities, as discussed below.

429 **3.4 Organic Aerosol (OA)**

430 Figure S11 illustrates the histograms for the organic aerosol observations and model results
431 with base case and revised emissions. A clear improvement is shown in the highest concentration
432 bins (> 15 μg/m^3) with the revised emissions. This can be quantified with the 99th percentile of
433 the data (obs=13.4 μg/m^3, revised=9.3 μg/m^3, base=4.9 μg/m^3. The median statistics also
434 improved (obs=2.8 μg/m^3, revised=0.84 μg/m^3, base=0.70 μg/m^3). The lower 5th percentile is
435 also significantly under-predicted compared to observations and does not change much between
436 the two model runs (obs=0.49 μg/m^3, revised=0.036 μg/m^3, base=0.035 μg/m^3). This reflects an
437 under-prediction in the background OA predicted by the model, which is likely due to too low a
438 level of biogenic SOA formation and aging in both model versions. The importance of
439 widespread biogenic SOA formation from boreal forests has been reported in other work (Slowik
440 *et al.*, 2010; Tunved *et al.*, 2006).

441 The other statistics are presented in Table 1. The mean bias, RMSE and slope all improve for
442 the revised-emissions run, though the correlation coefficient decreases significantly for this run.
443 The decrease in correlation suggests that the improved bias may not always be the result for the
444 right reasons for some points. To investigate this further, this study looked at the model PM_1 OA
445 bias as a function of different observed variables. Figure 8a shows the base-case model bias as a

446 function of observed PM_1 sulfate for individual points. Figure 8a is color-coded as a function of
447 the model %SOA relative to model OA. There is a trend of increasing negative bias in model OA
448 in the base case with increasing observed sulfate aerosol. This suggests that air masses that
449 originate from sources high in SO_2 tend to be under-estimated in simulated organic aerosol.
450 These air masses also tend to have large %SOA in the base model. Figure 8b is the same plot but
451 for the model run using revised emissions. There are a large number of points that are relatively
452 high in observed sulfate (>1 µg/m^3) that change from a negative OA bias to a positive OA bias
453 and these points also shift to being dominantly POA in model composition (low %SOA, blue
454 points). The model output confirms that much of the increase in model OA in the sensitivity run
455 originates from the increase in primary PM emissions and from the increase in the mass fraction
456 of that primary PM assumed to be OA (based on Wang *et al.*, 2015). The revised emissions
457 simulation has less organic aerosol bias at high sulfate loading suggesting that the improved
458 results are closely linked to model organic emissions co-located with sources of SO_2.
459 Figure 9a is a plot of the OA bias as a function of the observed black carbon (BC) aerosol for
460 the base-case and sensitivity runs. For the base-case run, the OA negative bias is observed to
461 increase in magnitude with observed BC. Points with high observed BC correlate well with
462 emissions from the OS open-pit mines (Liggio *et al.*, 2017), where the BC is likely emitted from
463 the heavy-hauler trucks. This was also consistent with location of mines and the transport wind
464 direction. A review of the OS emission inventories suggest that about 70% of the BC comes
465 from the OS off-road diesel fleet. Including all points, the mean bias improves from -2.8 to -2.4
466 (see Table 1) when using the revised emissions. Figure 9b shows a zoomed plot for points with
467 high observed BC (>0.8 µg/m^3). There is a clear improvement in bias for most of these points.
468 The average bias for these high BC points improves from -6.8 µg/m^3 for the base case to -2.6

469 µg/m^3 for the revised emissions. For emissions processing the increase in PM emissions was
470 assigned to the processing plants (particle bin D<1µm) or the surface mines (particle bin
471 D>1µm). Note that anthropogenic SOA formation should be minimal for the high BC points
472 close to the emission source, so uncertainties stemming from the model SOA formation
473 mechanism should be reduced. The revised emissions simulation has less bias as a function of
474 observed black carbon, suggesting that the improved results are closely linked to model
475 emissions with similar sources to the black carbon.
476 Figure 10 is a scatterplot of the difference in predicted POA between the revised and base-
477 case emissions runs vs. the difference in predicted total OA. A large fraction of the points fall
478 along the 1:1 line, and hence for these points the difference between the two runs is almost
479 completely due to the increased total primary PM emissions, and increased POA fraction of
480 those emissions, of the revised emissions simulations. The points with largest concentrations
481 along the 1:1 line correspond to flights over the Syncrude Mildred Lake facility on Aug. 16,
482 Aug. 23 and Sept. 3. There is a subset of points, however, that lies below the 1:1 line; these
483 correspond to points with significantly enhanced model SOA between the two runs (Aug. 16
484 flight over CNRL Horizon and Aug. 21 survey flight over Shell Muskeg/Jackpine).
485 The focus of the flight on Aug. 21 was an enclosing box pattern around the Shell
486 Muskeg/Jackpine facility at different altitudes. The approach to this facility, however, also
487 included an overpass of the Syncrude Mildred Lake facility. Figure S13 illustrates the flight path
488 color-coded as a function of POA difference (revised emissions – base case) and SOA difference
489 (revised emissions –base case). The corresponding time series for OA observations, the revised
490 emissions model run, and the base-case emissions model run OA predictions are shown in Figure
491 11. There is a clear “hot spot” in POA difference in Figure S13a located over the Syncrude

492 Mildred Lake facility. This hot spot corresponds to the first large maxima in the times series in
493 Figure 11 (17:17 UTC). The observations at this time lie between the predictions from the two
494 model simulations, though the overestimate of the revised emissions simulation is closer to the
495 observations than the underestimate of the base-case emissions simulation. The aircraft then
496 entered the box pattern at different altitudes around the Shell Muskeg/Jackpine facility, and each
497 successive pass around this facility intersected the observed plume on the north-east corner of the
498 flight box (see hot-spot, Figure S13b); the model predicts that the increase in OA is largely due
499 to SOA (as implied by Figure 10), and the revised-emissions simulation produces peak OA
500 concentrations that are closer to the observations than the base-case emissions simulation. As is
501 clear from Fig. 11, the base-case emissions simulation greatly underestimated the OA relative to
502 observations. In examining the time series, it is also clear that both model simulations are under-
503 estimating the background biogenic OA concentrations, by about 0.5 μg m^{-3}. The height of the
504 peaks relative to background is closer to the sensitivity run peaks than the base-case run peaks.
505 Figure S14 shows the difference between revised-emissions and base-case model OA
506 predictions for another case study, for southerly winds with a box flight over the Syncrude
507 Mildred Lake facility. The flight started and ended with a horizontal zig-zag pattern with
508 overpasses directly over the facility emissions sources. This corresponds to the initial spikes in
509 the model in the time series shown in Figure 12 (8:30 p.m. UTC). Again, the observed height of
510 the peaks lies between the model peak heights for the base-case and revised-emissions
511 simulations. For this flight the background OA concentration is under-predicted by up to 2 μg/m^3
512 by the end of the flight. The background air has more measured oxygenated organic aerosol
513 (OOA) (Liggio *et al.*, 2015), with an aerosol mass spectra more reflective of laboratory
514 monoterpene SOA (Han *et al.*, 2017). During the box pattern, the peak heights in the

observations more closely match the base model peaks. The PM_1 emission rates derived from the five box flights around Syncrude Mildred Lake did show more variability than for the other facilities. The average of five measurement-derived PM_1 emission rates was used to revise the PM_1 emissions for Syncrude in the revised emissions data used by the model. Interestingly, the largest observed OA value was measured in the spiral into the free troposphere near the end of the flight. There is no corresponding peak in the model at this time. The model peaks again only after the flight path has dropped into the boundary layer. Note that there was no corresponding increase in acetonitrile observed in the free troposphere so the source of the elevated OA is not likely from biomass burning, but may represent long-range transport from other sources.

The last case study is for the survey flight on August 23. Figure S15 shows the corresponding flight path color-coded by POA difference (revised - base case emissions; panel a) and SOA difference (revised – base-case emissions; panel b). From Fig. S15a, we can again see the local maxima in POA difference between runs over the Syncrude Mildred Lake facility. This corresponds to the peaks in time series at 5:50 p.m. UTC (Figure 13). The observed peaks are closer in magnitude to the base-case model peaks at this time. The peak at 7:40 p.m. UTC corresponds to another time later in flight over the same location. The peak in SOA difference at 6:20 p.m. UTC is downwind of the CNRL Horizon facility (red points in Fig. S15b). The observations show a more broadly spread-out peak at this time than is predicted by the model, perhaps indicating a greater degree of turbulence or wind variability in the observations than predicted by the model. Both modelled and observed meteorology had light wind speeds with a high degree of variability in direction on this day. The variability in the observed winds at the local Mildred Lake weather station was large on this afternoon with hourly-averaged wind directions of 40˚, 290˚, 180˚, 20˚, 40˚ from 12-16 UTC and wind speeds all less than 6 km/hr.

These light, variable winds result in a more dispersed nature of the observed organic aerosol. The peak in observations at 6:25 p.m. UTC is represented well by the revised model. This corresponds to a location over Shell Muskeg/Jackpine (light blue points in Fig. S15a). Note that Fig. 13 suggests that background OA levels once again seem to be under-estimated in both simulations.

In summary, the improvement in model PM_1 OA bias due to the use of the revised emissions is encouraging; however the decrease in correlation coefficient suggests that the spatial allocation of PM_1 emissions may need further refinement or other important processes may be missing or under-represented in the model. The under-prediction in background OA was a general finding from the study; the cause is believed to be due to underestimated biogenic SOA, due to the lumping of biogenic monoterpene emissions into the anthropogenic ALKE model species in the model's gas-phase mechanism, and the lack of speciated representation of other biogenic SOA precursors such as sesquiterpenes. Biogenic SOA mass yield stoichiometric coefficients, based on more recent chamber experiments that consider vapor wall loss, should also be used for future modeling. Future aircraft observations should include a biogenic emission closure flight, where the aircraft flies a box pattern over a boreal forest location where the tree speciation is uniform and observed and modeled surface emission fluxes are then compared.

Recent publications also suggest that fugitive intermediate volatile organic (IVOC) emissions from the OS open-pit mines are needed to represent SOA formation downwind of the OS region (Liggio *et al*., 2016). In our emissions revision, a portion of the IVOCs (C_{12}) were placed in the long-chain ALKA lumped species. The ALKA lumped model VOC species has an SOA yield more representative of a lower molecular-weight range, and the yield is known to increase with increasing carbon number. Work is currently underway with GEM-MACH to implement a

Volatility Basis Set (VBS). The VBS approach will more adequately represent the intermediate and semi-volatile volatility range and chemical aging (Robinson *et al*., 2006). Future field studies should also focus on improving within-facility spatial allocation. For example, within-facility data such as the GPS location of the mining trucks would be helpful to derive their activity diurnal profiles and to improve their emission spatial allocation within a facility.

4) Conclusions

Overall, the weight of evidence suggests that the revised aircraft-measurement-derived organic emissions for the OS surface mining facilities help to constrain reported facility-total organic emissions for fugitive sources, as shown by improved model results when the revised emissions are employed. We note that emissions from these sources are a challenge to calculate in bottom-up inventories. For the mono- and multi-substituted aromatics, the emission rates from facilities were more fine adjustments, as some facility totals went up and some went down and the overall biases compared to observations improved for AROM but degraded for TOLU. However, the model's ability to predict very high aromatic concentrations in plumes improved with the revised emissions, as shown by the 99th percentile statistic and the case studies.

For the long-chain ALKA species, the revised emissions may have over-corrected, on average, as shown by the increase in mean bias for the entire aircraft data set. However, the correlation coefficient did improve significantly for the long-chain alkane predictions, suggesting the combination of alkane emission increases for some facilities and decreases for others helped to improve the spatial distribution of ALKA emissions. The results for some facilities suggest that further improvement could be achieved by putting more emissions at extraction processing plant locations (i.e., adjusting within-facility spatial allocation). Interestingly, the alkane emission increases, derived from aircraft data, were associated with the facilities that use

585 paraffinic solvents for bitumen extraction (Shell Muskeg/Jackpine and Syncrude Aurora North;
586 Li et al., 2017). Overall, the predictions of alkanes in high concentration plumes improved with
587 the revised emission data set, as shown by the 99th percentile statistic.

588 For PM_1 organic aerosol, the revised emissions improved the mean bias for predictions;
589 however, this improvement was associated with a decrease in correlation coefficient. The
590 increase in predicted PM_1 OA concentration was largely due to the increase in POA emissions in
591 the revised emissions input files. The POA emissions increased because of a combination of
592 larger aircraft-measurement-derived PM_1 emissions and the revised ground-observed PM
593 speciation profile having a larger POA fraction. A portion of this aircraft-measurement-derived
594 POA emission increase could stem from rapid SOA formation in the interior of the aircraft flight
595 boxes. It was recently discovered that IVOC SOA formation can be important in OA formation
596 downwind of the OS surface mining region and the IVOC emissions came from open-pit mine
597 fugitive emissions. GEM-MACH does not currently include the IVOC emissions from open-pit
598 mines, as the estimation of the IVOC emission rate and SOA aging rate are the subject of
599 ongoing Lagrangian box model studies. Furthermore, the increase in PM_1 POA emissions were
600 largely allocated spatially to stack locations and this allocation may be a key factor in the
601 degradation of the correlation coefficient, especially if the fine OA originates from mine-face
602 fugitive emissions. Future work should focus on improving within-facility spatial allocation of
603 emissions.

604
605 **Acknowledgements**
606
607 The authors are grateful to all of the participants in the 2013 JOSM intensive field study for their
608 commitment. The authors are also appreciative of the ECCC Pollutant Inventory and Reporting
609 Division (PIRD) and the U.S. EPA for developing, maintaining, and distributing each country's

610 national emission inventories. We also appreciate the efforts of George Marson of ECCC in
611 helping to compile the various emissions inventories from Alberta Environment and Parks, and
612 also CEMA. We also appreciate the analysis of the NAPS VOC measurement group. This study
613 was funded by the Joint Oil Sands Monitoring program and the Climate Change and Air Quality
614 Program.

615 **References**
616
617 Akingunola, A., P. A. Makar, J. Zhang, A. Darlington, S.M. Li, M. Gordon, M.D. Moran, Q. Zheng, Evaluation of
618 GEM-MACH Air Quality Modelling at 2.5km Resolution Using JOSM 2013 Intensive Campaign: Impact of
619 Continuous Monitoring Emissions Stack Parameters on Model Simulations, submitted to ACPD.
620
621 Barsanti, K.C., Carlton, A.G., Chung, S.H., Analyzing experimental data and model parameters: Implications for
622 predictions of SOA using chemical transport models, (2013) Atmospheric Chemistry and Physics, 13 (23), pp.
623 12073-12088.
624
625 Cappa, C.D, J. Jimenez, Quantitative estimates of the volatility of ambient organic aerosol, *Atmos. Chem. Physics*,
626 10, 12, 2010, pp 5409-5424.
627
628 Cappa, C.D., Wilson, K.R., Evolution of organic aerosol mass spectra upon heating: Implications for OA phase and
629 partitioning behavior, (2011) *Atmospheric Chemistry and Physics*, 11 (5), pp. 1895-1911.
630
631 Chai, T., Kim, H.-C., Lee, P., Tong, D., Pan, L., Tang, Y., Huang, J., McQueen, J., Tsidulko, M., Stajner, I.
632 Evaluation of the United States National Air Quality Forecast Capability experimental real-time predictions in 2010
633 using Air Quality System ozone and NO2 measurements, (2013) *Geoscientific Model Development*, 6 (5), pp. 1831-
634 1850.
635
636 Cho, S., McEachern, P., Morris, R., Shah, T., Johnson, J., Nopmongcol, U., Emission sources sensitivity study for
637 ground-level ozone and PM 2.5 due to oil sands development using air quality modeling system: Part I- model
638 evaluation for current year base case simulation, (2012) *Atmospheric Environment*, 55, pp. 533-541.
639
640 Cohan, D.S.; Napelenok, S.L. Air Quality Response Modeling for Decision Support. *Atmosphere*, **2011**, *2*, 407-425.
641
642 Côté, J., Desmarais, J.-G., Gravel, S., *et al*., The operational CMC/MRB global environmental multiscale (GEM)
643 model. Part 1: design considerations and formulation, (1998a), *Mon. Wea. Rev*. 126, 1373-1395.
644
645 Côté, J., Desmarais, J.-G., Gravel, S., *et al*., The operational CMC-MRB global environment multiscale (GEM)
646 model. Part II: results, (1998b), *Mon. Wea. Rev*. 126, 1397-1418.
647
648 Dann, T.F., Wang, D.K., Ambient air benzene concentrations in Canada (1989-1993): Seasonal and day of week
649 variations, trends, and source influences, (1995) Journal of the Air and Waste Management Association, 45 (9), pp.
650 695-702.
651
652 Davies, M., Person, R., Nopmongcol, U., Shah T., Vijayaraghavan, K., Morris, R., and Picard, D.: Lower Athabasca
653 Region Source and Emission Inventory, report prepared by Stantec Consulting Ltd. and ENVIRON International
654 Corporation for Cumulative Environmental Management Association - Air Working Group,
655 http://library.cemaonline.ca/ckan/dataset/0cfaa447-410a-4339-b51f-e64871390efe/resource/fba8a3b0-72df-45ed-
656 bf12-8ca254fdd5b1/download/larsourceandemissionsinventory.pdf, 274 pp., 2012 (last accessed on October 24,
657 2017).
658

Dickson, R.J., Oliver, W.R., Emissions models for regional air quality studies, (1991), *Environ. Sci. Technol*. 25:1533-1535.

Donahue, N.M., A.L. Robinson, E.R. Trump, I. Riipinen, J.H. Kroll, Volatility and aging of atmospheric organic aerosol, Topics in Current Chemistry, Volume 339, 2014, Pages 97-144.

Environment and Climate Change Canada & Alberta Environment and Parks: Joint Oil Sands Monitoring Program Emissions Inventory Compilation Report, http://aep.alberta.ca/air/reports-data/documents/JOSM-EmissionsInventoryReport-Jun2016.pdf, 146 pp, 2016.

Eyth, A., Mason, R., and Zubrow, A.: Development and Status of EPA's 2011 Modeling Platform, 12th CMAS Conference, 28-30 Oct., Chapel Hill, North Carolina, https://www.cmascenter.org/conference//2013/slides/eyth_development_status_2013.pptx, 2013.

Gentner, D.R., Jathar, S.H., Gordon, T.D., Bahreini, R., Day, D.A., El Haddad, I., Hayes, P.L., Pieber, S.M., Platt, S.M., De Gouw, J., Goldstein, A.H., Harley, R.A., Jimenez, J.L., Prévôt, A.S.H., Robinson, A.L., Review of Urban Secondary Organic Aerosol Formation from Gasoline and Diesel Motor Vehicle Emissions (2017) Environmental Science and Technology, 51 (3), pp. 1074-1093.

Girard, C., Plante, A., Desgagné, M., Mctaggart-Cowan, R., Côté, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P.A., Zadra, A., Staggered vertical discretization of the canadian environmental multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type (2014) Monthly Weather Review, 142, pp. 1183-1196.

Gong, S.L., Barrie, L.A., Blanchet, J.-P., von Salzen, K., Lohmann, U., Lesins, G., Spacek, L., Zhang, L.M., Girard, E., Lin, H., Leaitch, R., Leighton, H., Chylek, P., Huang, P., Canadian Aerosol Module: A size-segregated simulation of atmospheric aerosol processes for climate and air quality models 1. Module development (2003) Journal of Geophysical Research D: Atmospheres, 108, pp. AAC 3-1 AAC 3-16.

Gordon, M., Li, S.-M., Staebler, R., Darlington, A., Hayden, K., O'Brien, J., and M. Wolde: 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. doi:10.5194/amt-8-3745-2015, 2015.

Griffin, R.J., Cocker III, D.R., Flagan, R.C., Seinfeld, J.H., Organic aerosol formation from the oxidation of biogenic hydrocarbons, (1999) Journal of Geophysical Research Atmospheres, 104 (D3), art. no. 1998JD100049, pp. 3555-3567.

Han, Y., Stroud, C.A., Liggio, J., Li, S.-M., The effect of particle acidity on secondary organic aerosol formation from α-pinene photooxidation under atmospherically relevant conditions, (2016) Atmospheric Chemistry and Physics, 16 (21), pp. 13929-13944.

Houyoux, M.R., Vukovich, J.M., Coats, Jr., C.J., Wheeler, N.J.M., Kasibhatla, P.S., Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project, (2000), J. Geophys. Res., 105, 9079-9090, 1999JD900975.

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., Dunlea, E.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., Worsnop, D.R., Evolution of organic aerosols in the atmosphere, (2009) Science, 326 (5959), pp. 1525-1529.

Kanakidou, M., Seinfeld, J.H., Pandis, S.N., Barnes, I., Dentener, F.J., Facchini, M.C., Van Dingenen, R., Ervens, B., Nenes, A., Nielsen, C.J., Swietlicki, E., Putaud, J.P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G.K., Winterhalter, R., Myhre, C.E.L., Tsigaridis, K., Vignati, E., Stephanou, E.G., Wilson, J., Organic aerosol and global climate modelling: A review, (2005) Atmospheric Chemistry and Physics, 5 (4), pp. 1053-1123.

Kelly, J., Makar, P.A., Plummer, D. Projections of mid-century summer air-quality for North America: effects of changes in climate and precursor emissions, (2012) Atmospheric Chemistry and Physics, 12, pp. 5367-5390.

Kroll, J.H., Ng, N.L., Murphy, S.M., Varutbangkul, V., Flagan, R.C., Seinfeld, J.H., Chamber studies of secondary organic aerosol growth by reactive uptake of simple carbonyl compounds, (2005) Journal of Geophysical Research Atmospheres, 110 (23), art. no. D23207, pp. 1-10.

Kroll, J.H., Seinfeld, J.H., Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere, (2008) Atmospheric Environment, 42 (16), pp. 3593-3624.

Lee, P.; Ngan, F. Coupling of Important Physical Processes in the Planetary Boundary Layer between Meteorological and Chemistry Models for Regional to Continental Scale Air Quality Forecasting: An Overview. *Atmosphere* **2011**, *2*, 464-483.

Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., The contribution of outdoor air pollution sources to premature mortality on a global scale, (2015) Nature, 525 (7569), pp. 367-371.

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., Wentzell, J.J.B., Differences between measured and reported volatile organic compound emissions from oil sands facilities in Alberta, Canada, (2017) Proceedings of the National Academy of Sciences of the United States of America, 114 (19), pp. E3756-E3765.

Liggio, J., Li, S.-M., McLaren, R., Reactive uptake of glyoxal by particulate matter, (2005) Journal of Geophysical Research D: Atmospheres, 110 (10), pp. 1-13.

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.R., Lu, G., Staebler, R.M., Han, Y., Tokarek, T.W., Osthoff, H.D., Makar, P.A., Zhang, J., Plata, D.L., Gentner, D.R., Oil sands operations as a large source of secondary organic aerosols, (2016) Nature, 534 (7605), pp. 91-94.

Liggio, J., C.A. Stroud, J. Wentzell, *et al*., Quantifying the primary emissions and photochemical formation of isocyanic acid downwind of Oil Sands operations, Environ. Sci. Technol.,.2017, Dec 6. doi: 10.1021/acs.est.7b04346.

Lopez-Hilfiker, F.D., *et al*., Molecular composition and volatility of organic aerosol in the Southeastern U.S. : Implications for IEPOX Derived SOA, Environ. Sci. Technol., 50, 5, 2016, pp 2200-2209.

Makar, P.A., Gong, W., Milbrandt, J., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., Im, U., Balzarini, A., Baró, R., Bianconi, R., Cheung, P., Forkel, R., Gravel, S., Hirtl, M., Honzak, L., Hou, A., Jiménez-Guerrero, P., Langer, M., Moran, M.D., Pabla, B., Pérez, J.L., Pirovano, G., San José, R., Tuccella, P., Werhahn, J., Zhang, J., Galmarini, S. Feedbacks between air pollution and weather, Part 1: Effects on weather, (2015a) Atmospheric Environment, 115, pp. 442-469.

Makar, P.A., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., Milbrandt, J., Im, U., Balzarini, A., Baró, R., Bianconi, R., Cheung, P., Forkel, R., Gravel, S., Hirtl, M., Honzak, L., Hou, A., Jiménez-Guerrero, P., Langer, M., Moran, M.D., Pabla, B., Pérez, J.L., Pirovano, G., San José, R., Tuccella, P., Werhahn, J., Zhang, J., Galmarini, S. Feedbacks between air pollution and weather, part 2: Effects on chemistry
(2015b) Atmospheric Environment, 115, pp. 499-526.

769 Makar, P.A. *et al*., Estimates of Exceedances of Critical Loads for Acidifying Deposition in Alberta and
770 Sasketchwan, submitted to ACPD, 2017.
771
772 Mashayekhi, R., Zhao, S., Saeednooran, S., Hakami, A., Ménard, R., Moran, M. D., and Zhang, J.: Emissions
773 Uncertainty Inventory and Modeling Framework: Case Study of Residential Wood Combustion, 15th Annual
774 CMAS Conference, October 24-26, Chapel Hill, NC,
775 https://www.cmascenter.org/conference//2016/slides/mashayekhi_development_emission_2016.pptx, 2016.
776
777 McNeill, V.F., Aqueous organic chemistry in the atmosphere: Sources and chemical processing of organic aerosols
778 (2015) Environmental Science and Technology, 49 (3), pp. 1237-1244.
779
780 Moran, M.D., Ménard, S., Pavlovic, R., Anselmo, D., Antonopoulos, S., Makar, P.A., Gong, W., Gravel, S., Stroud,
781 C., Zhang, J., Zheng, Q., Robichaud, A., Landry, H., Beaulieu, P.A., Gilbert, S., Chen, J., Kallaur, A., Recent
782 Advances in Canada's National Operational AQ Forecasting System, (2013) NATO Science for Peace and Security
783 Series C: Environmental Security, 137, pp. 215-220.
784
785 Pankow, J.F., An absorption model of the gas/aerosol partitioning involved in the formation of secondary organic
786 aerosol, (1994) Atmospheric Environment, 28 (2), pp. 189-193.
787
788 Park, S.H., Gong, S.L., Bouchet, V.S., Gong, W., Makar, P.A., Moran, M.D., Stroud, C.A., Zhang, J.
789 Effects of black carbon aging on air quality predictions direct radiative forcing estimation, (2011) Tellus, Series B:
790 Chemical and Physical Meteorology, 63 (5), pp. 1026-1039.
791
792 Pouliot, G., Pierce, T., Denier van der Gon, H., Schaap, M., Moran, M., Nopmongcol, U., Comparing emission
793 inventories and model-ready emission datasets between Europe and North America for the AQMEII project, (2012),
794 Atmospheric Environment, 53, pp. 4-14.
795
796 Pouliot, G., Denier van der Gon, H.A.C., Kuenen, J., Zhang, J., Moran, M.D., Makar, P.A., Analysis of the emission
797 inventories and model-ready emission datasets of Europe and North America for phase 2 of the AQMEII project
798 (2015) Atmospheric Environment, 115, pp. 345-360.
799
800 Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop, A.P., Lane, T.E.,
801 Pierce, J.R., Pandis, S.N., Rethinking organic aerosols: Semivolatile emissions and photochemical aging, (2007)
802 Science, 315 (5816), pp. 1259-1262.
803
804 Rouleau, M., Egyed, M., Taylor, B., Chen, J., Samaali, M., Davignon, D., Morneau, G., Human health impacts of
805 biodiesel use in on-road heavy duty diesel vehicles in Canada, (2013) Environmental Science and Technology, 47
806 (22), pp. 13113-13121.
807
808 Schultz, M.G., Diehl, T., Brasseur, G.P., Zittel, W., Air Pollution and Climate-Forcing Impacts of a Global
809 Hydrogen Economy, (2003) Science, 302 (5645), pp. 624-627.
810
811 Seinfeld, J. H. and Pandis, S. N. (1998). Atmospheric Chemistry and Physics from air pollution to climate change.
812 New York. John Wiley and Sons, Incorporated.
813
814 Shrivastava, M., Easter, R.C., Liu, X., Zelenyuk, A., Singh, B., Zhang, K., Ma, P.-L., Chand, D., Ghan, S., Jimenez,
815 J.L., Zhang, Q., Fast, J., Rasch, P.J., Tiitta, P., Global transformation and fate of SOA: Implications of low-volatility
816 SOA and gas-phase fragmentation reactions, (2015) Journal of Geophysical Research, 120 (9), pp. 4169-4195.
817
818 Slowik, J.G., Stroud, C., Bottenheim, J.W., Brickell, P.C., Chang, R.Y.-W., Liggio, J., Makar, P.A., Martin, R.V.,
819 Moran, M.D., Shantz, N.C., Sjostedt, S.J., Van Donkelaar, A., Vlasenko, A., Wiebe, H.A., Xia, A.G., Zhang, J.,
820 Leaitch, W.R., Abbatt, J.P.D., Characterization of a large biogenic secondary organic aerosol event from eastern
821 Canadian forests, (2010) Atmospheric Chemistry and Physics, 10 (6), pp. 2825-2845.
822
823 Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M.D., Wyat Appel, K., Bessagnet, B.,
824 Brandt, J., Christensen, J.H., Chemel, C., Coll, I., Ferreira, J., Forkel, R., Francis, X.V., Grell, G., Grossi, P.,

Hansen, A.B., Miranda, A.I., Nopmongcol, U., Prank, M., Sartelet, K.N., Schaap, M., Silver, J.D., Sokhi, R.S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S.T., Galmarini, S., Operational model evaluation for particulate matter in Europe and North America in the context of AQMEII, (2012) Atmospheric Environment, 53, pp. 75-92.

Stroud, C. A., Morneau, G., Makar, P. A., Moran, M. D., Gong, W., Pabla, B., Zhang, J., Bouchet, V. S., Fox, D., Venkatesh, S., Wang, D., and Dann, T.: OH-reactivity of volatile organic compounds at urban and rural sites across Canada: Evaluation of air quality model predictions using speciated VOC measurements. Atmos. Environ., **42**, 7746-7756, 2008.

Stroud, C.A., Makar, P.A., Moran, M.D., Gong, W., Gong, S., Zhang, J., Hayden, K., Mihele, C., Brook, J.R., Abbatt, J.P.D., Slowik, J.G., Impact of model grid spacing on regional- and urban- scale air quality predictions of organic aerosol, (2011) Atmospheric Chemistry and Physics, 11 (7), pp. 3107-3118.

Stroud, C.A., Zaganescu, C., Chen, J., McLinden, C.A., Zhang, J., Wang, D., Toxic volatile organic air pollutants across Canada: multi-year concentration trends, regional air quality modelling and source apportionment, (2016) Journal of Atmospheric Chemistry, 73 (2), pp. 137-164.

Tunved, P., Hansson, H.-C., Kerminen, V.-M., Ström, J., Maso, M.D., Lihavainen, H., Viisanen, Y., Aalto, P. P., Komppula, M. and Kulmala, M.: High natural aerosol loading over boreal forests,, Science, 312, 261–263, doi:10.1126/science.1123052, 2006

Wang, X., Chow, J. C., Kohl, S. D., Percy, K. E., Legge, A. H., and Watson, J. G.: Characterization of PM2.5 and PM10 fugitive dust source profiles in the Athabasca Oil Sands Region, Journal of the Air & Waste Management Association, 65:12, 1421-1433, DOI: 10.1080/10962247.2015.1100693, 2015.

Zhang, J., Zheng, Q., Moran, M. D., Makar, P. A., Akingunola, A., Li, S.-M., Marson, G., Gordon, M., Melick, R., and Cho, S.: Emissions preparation for high-resolution air quality modelling over the Athabasca oil sands region of Alberta, Canada. 21st Intern. Emissions Inventory Conference, 13-17 April, San Diego, http://www.epa.gov/ttn/chief/conference/ei21/session1/zhang_emissions.pdf, 18 pp, 2015.

Zhang, J. *et al*., Emissions preparation and analysis for Multiscale Air Quality Modelling over the Athabasca Oil Sands Region of Alberta, Canada, submitted to ACPD, 2017.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

Atmospheric Chemistry and Physics Discussions

Table 1. Statistical scores from the model simulations with revised and base case emissions; all statistics are relative to observations.

Lumped Species	Simulation	Mean Bias (ppbv)	RMSE (ppbv)	Slope	Y-intercept (ppbv)	Correlation Coefficient, R
TOLU	Base Case	**-0.041**	**0.277**	0.217	**0.063**	**0.32**
	Revised Emissions	0.049	0.386	**0.426**	0.125	0.31
AROM	Base Case	0.152	0.435	**0.957**	0.154	**0.41**
	Revised Emissions	**0.044**	**0.227**	0.383	**0.083**	0.37
ALKA	Base Case	**-0.123**	**5.556**	**0.378**	**2.028**	0.24
	Revised Emissions	1.98	6.403	0.335	4.097	**0.34**
OA	Base Case	-2.79	3.866	0.186	**0.252**	**0.59**
	Revised Emissions	**-2.37**	**3.632**	**0.292**	0.273	0.49

RMSE is the root mean square error. Y-intercept corresponds to the model intercept of a model vs observation correlation plot. Mean bias is the model-observation mean score. The better score for a given pair of statistics is shown in **bold-face** font.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

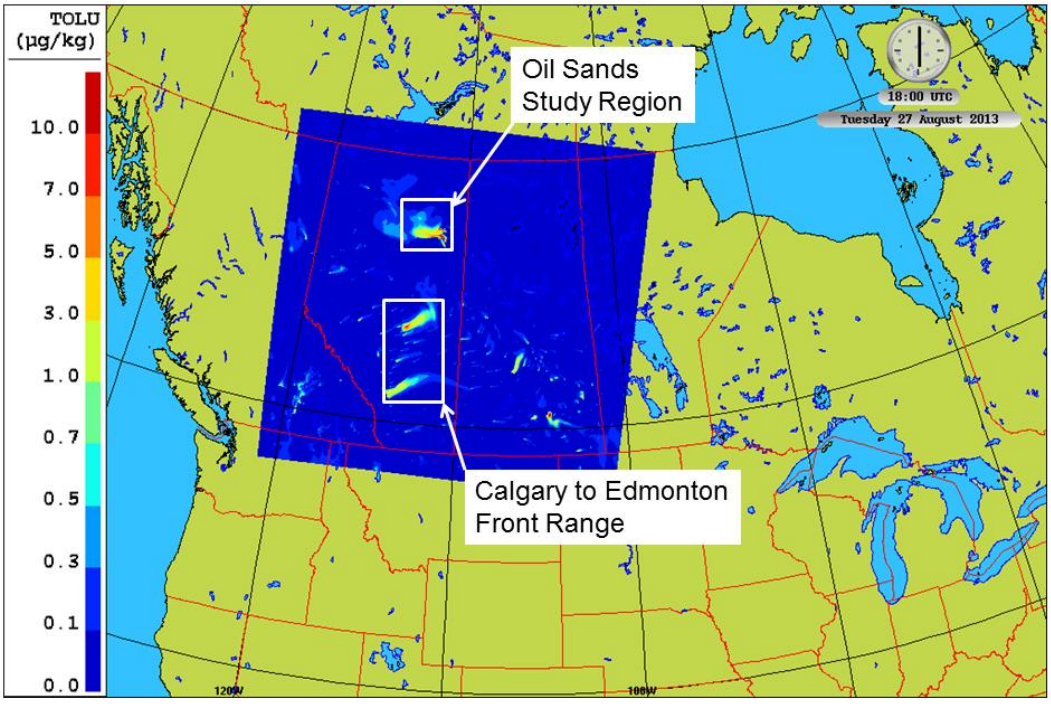


Figure 1. Nested domain at 2.5-km grid spacing (blue field) encompassing the Athabasca Oil Sands study region (white box). An outer North American domain (not shown) provides the lateral boundary conditions for this nested 2.5-km grid spaced domain. The model field shown is for the lumped toluene species (TOLU) mass mixing ratio (µg/kg air).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

Atmospheric Chemistry and Physics Discussions

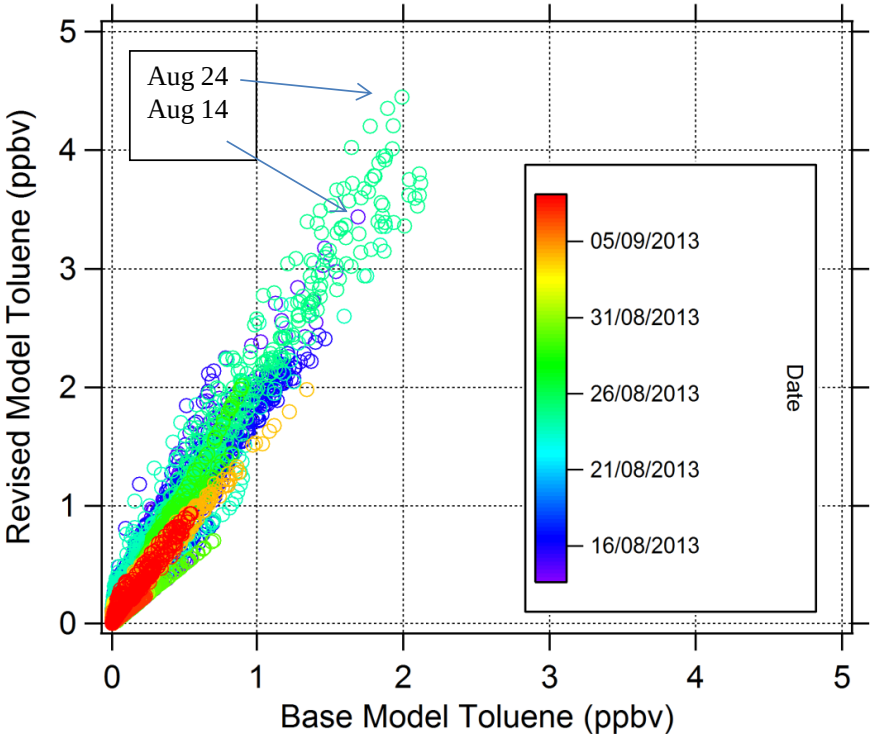


Figure 2. Revised-emissions TOLU vs. base-case TOLU, color-coded by date, for the data extracted along all the Convair flight tracks.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

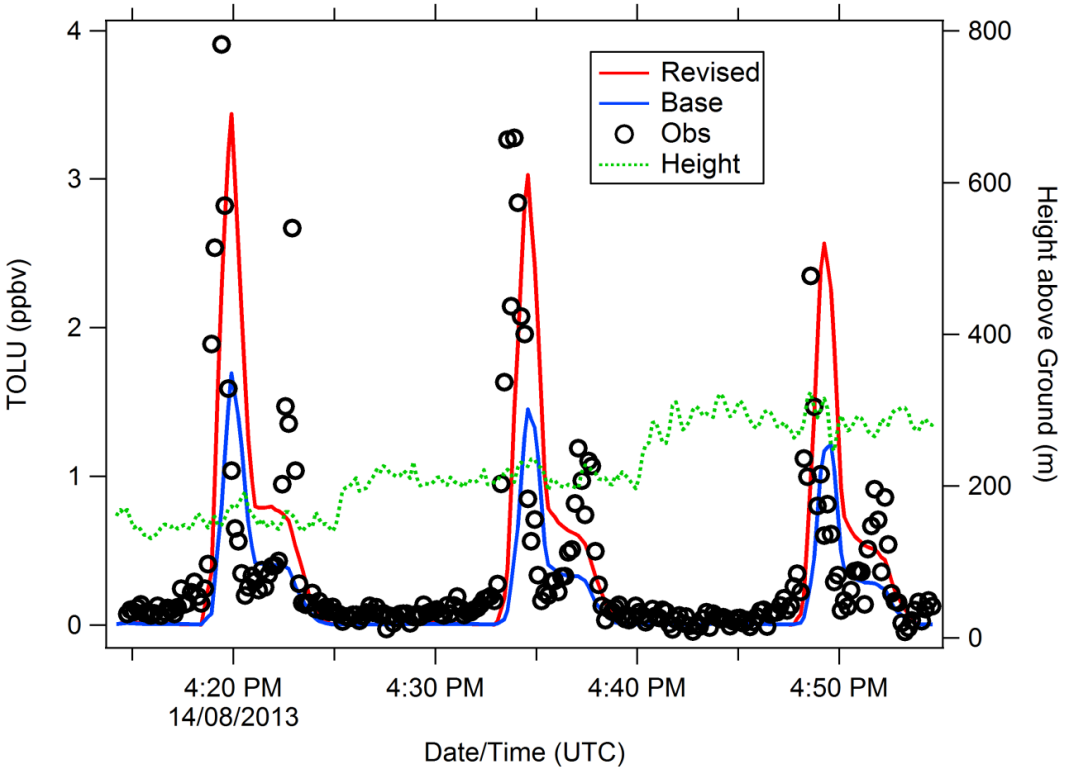


Figure 3. Time series of observed and model-predicted TOLU volume mixing ratios for a flight on Aug. 14, 2013 following a box pattern around the Syncrude Mildred Lake facility. The highest magnitude points correspond to a location north of the facility sampled at 3 different altitudes.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

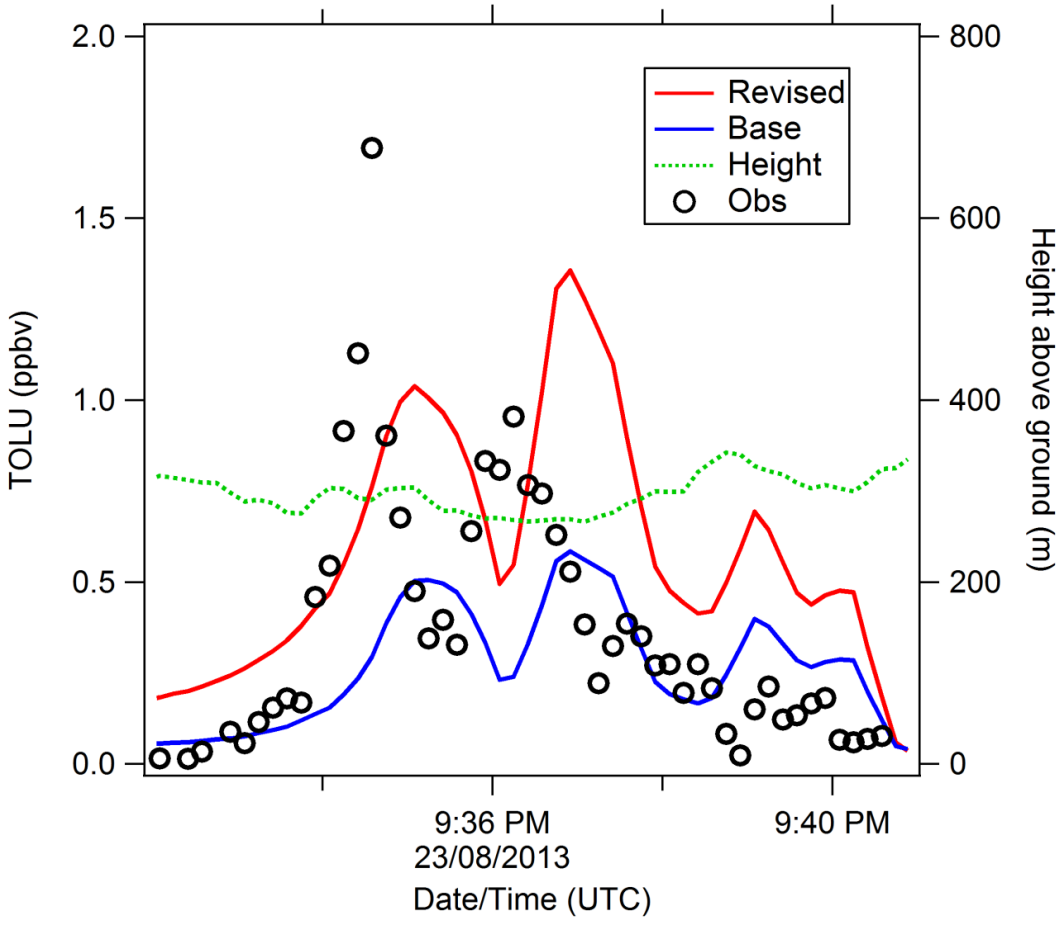
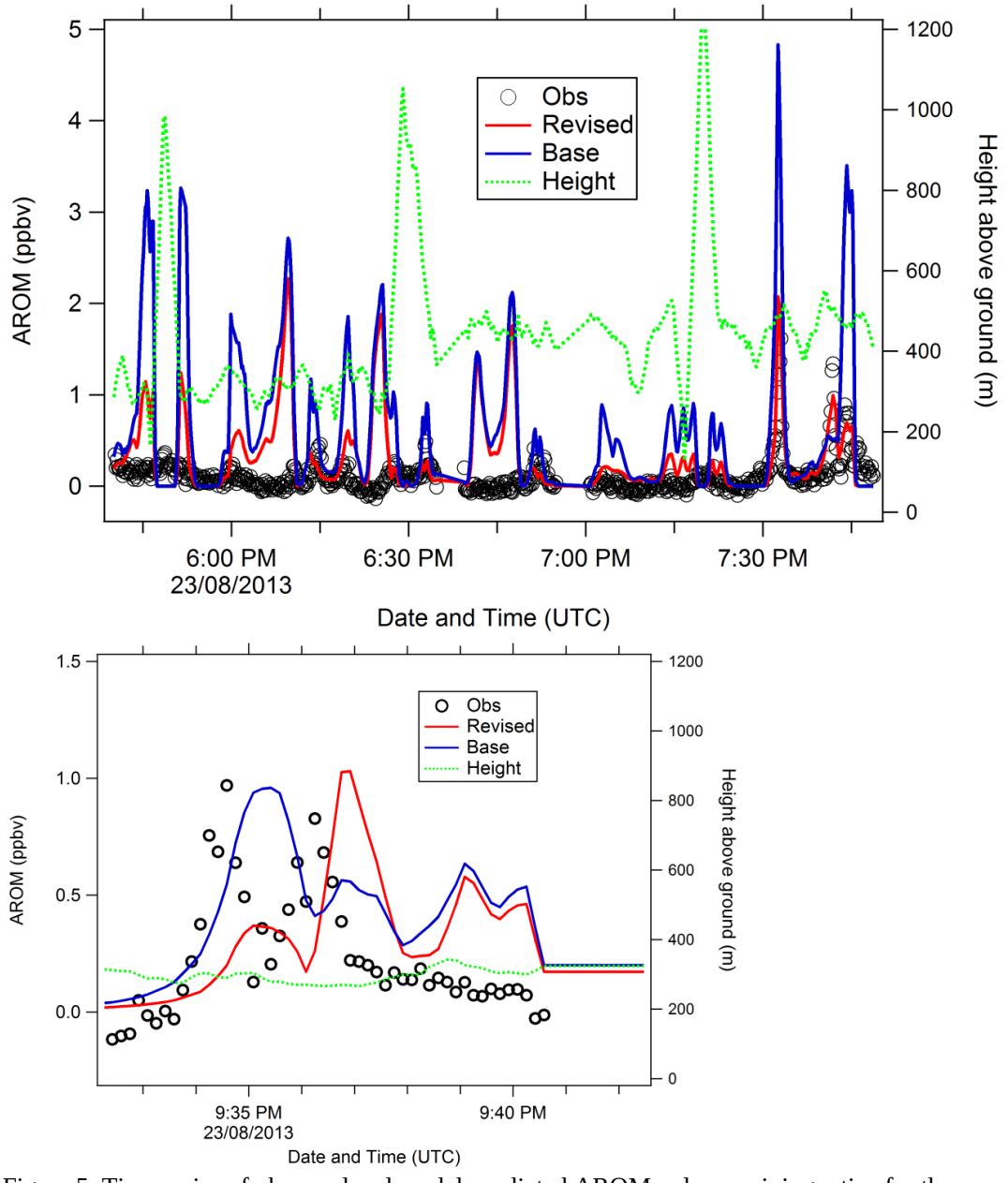


Figure 4. TOLU volume-mixing-ratio time series for a flight on Aug. 23, 2013 over the Suncor Millennium/Steepbank facility, just east of the Athabasca River, on a survey pattern.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

Atmospheric Chemistry and Physics Discussions



962

963

964 Figure 5. Time series of observed and model-predicted AROM volume mixing ratios for the
965 Aug. 23 survey flight. The mixing-ratio peaks in panel A are over the Syncrude Mildred Lake
966 facility (7:30-7:45 p.m. UTC). The 2nd peak in panel B is over the Suncor Millennium/Steepbank
967 facility (9:37 p.m. UTC).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

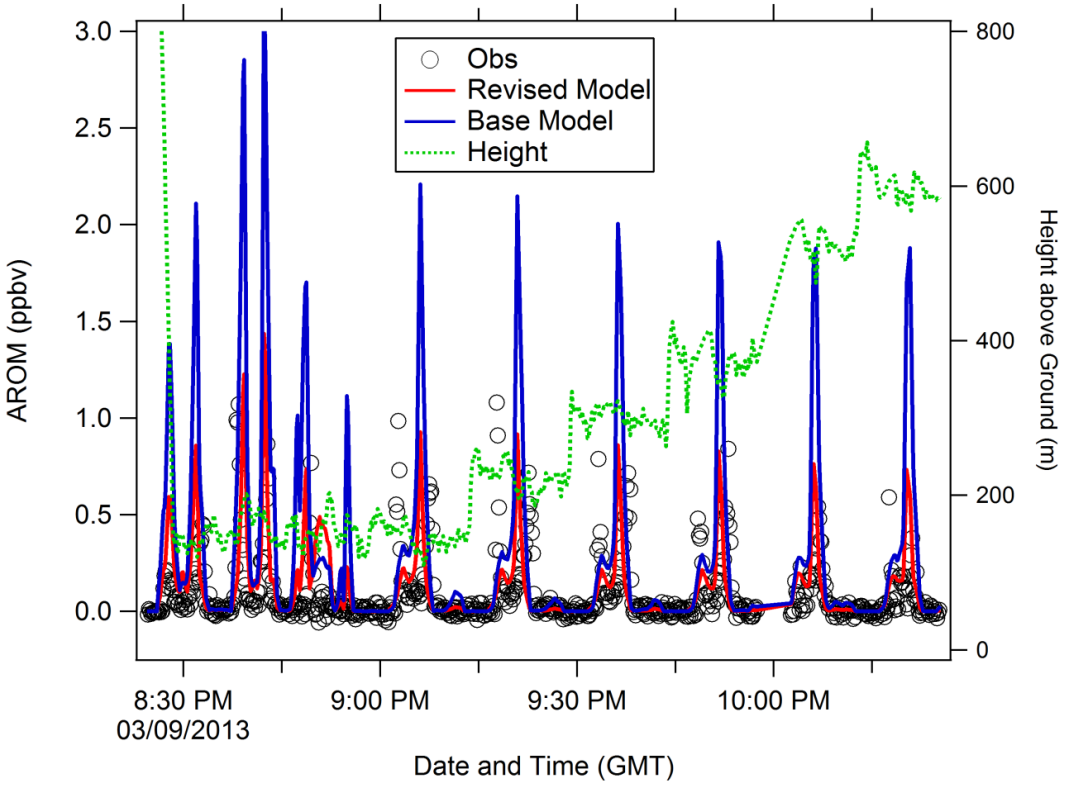


Figure 6. AROM volume-mixing-ratio time series for the Sept. 3 flight around and over the Syncrude Mildred Lake facility.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

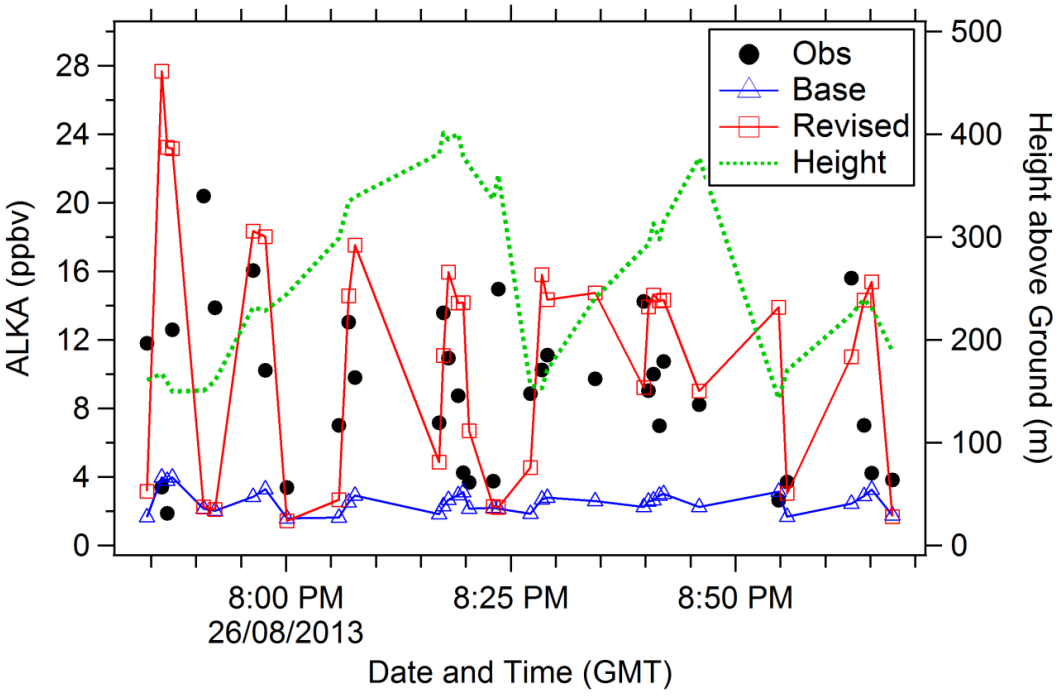


Figure 7. ALKA volume-mixing-ratio time series for the Aug. 26 flight around the CNRL Horizon facility.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

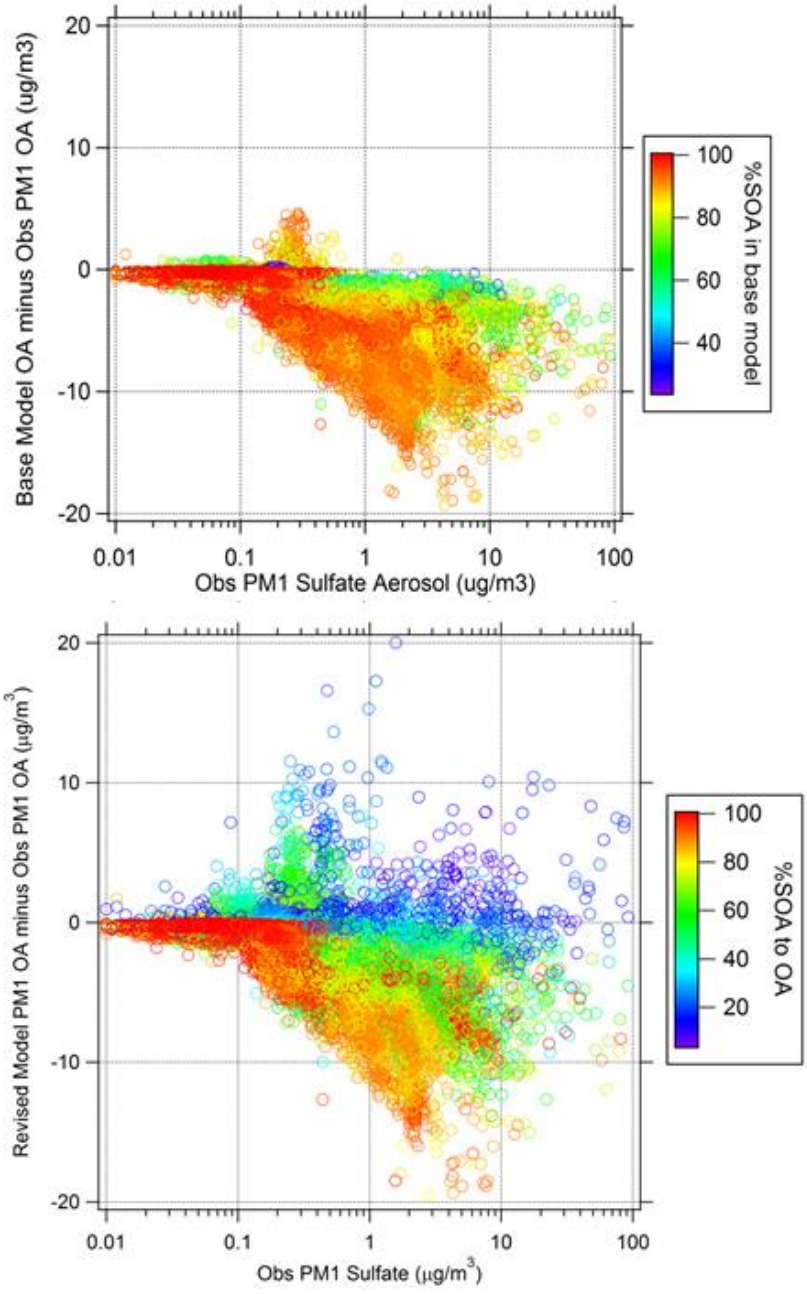


Figure 8. Model bias for individual organic aerosol data points plotted as a function of observed PM_1 sulfate aerosol: (a) base-case simulation; (b) revised-emissions simulation. The data plotted is for all the aircraft flights.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

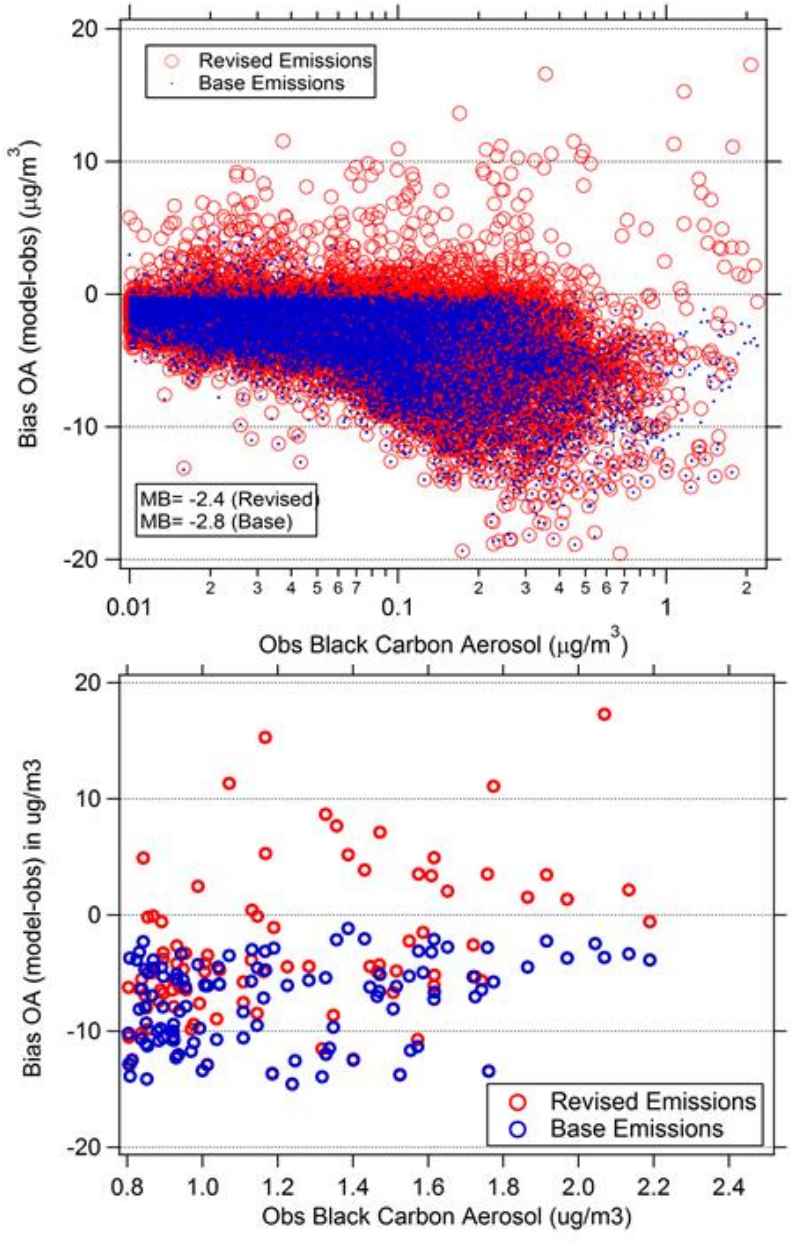


Figure 9. Organic aerosol model bias as a function of observed black carbon aerosol. The bottom panel is an enlargement of the upper panel showing only the data points for observed BC>0.8 µg/m^3. The model results for the base-case emissions run are plotted in blue and points in red correspond to the revised-emissions run. The data plotted is for all the aircraft flights.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

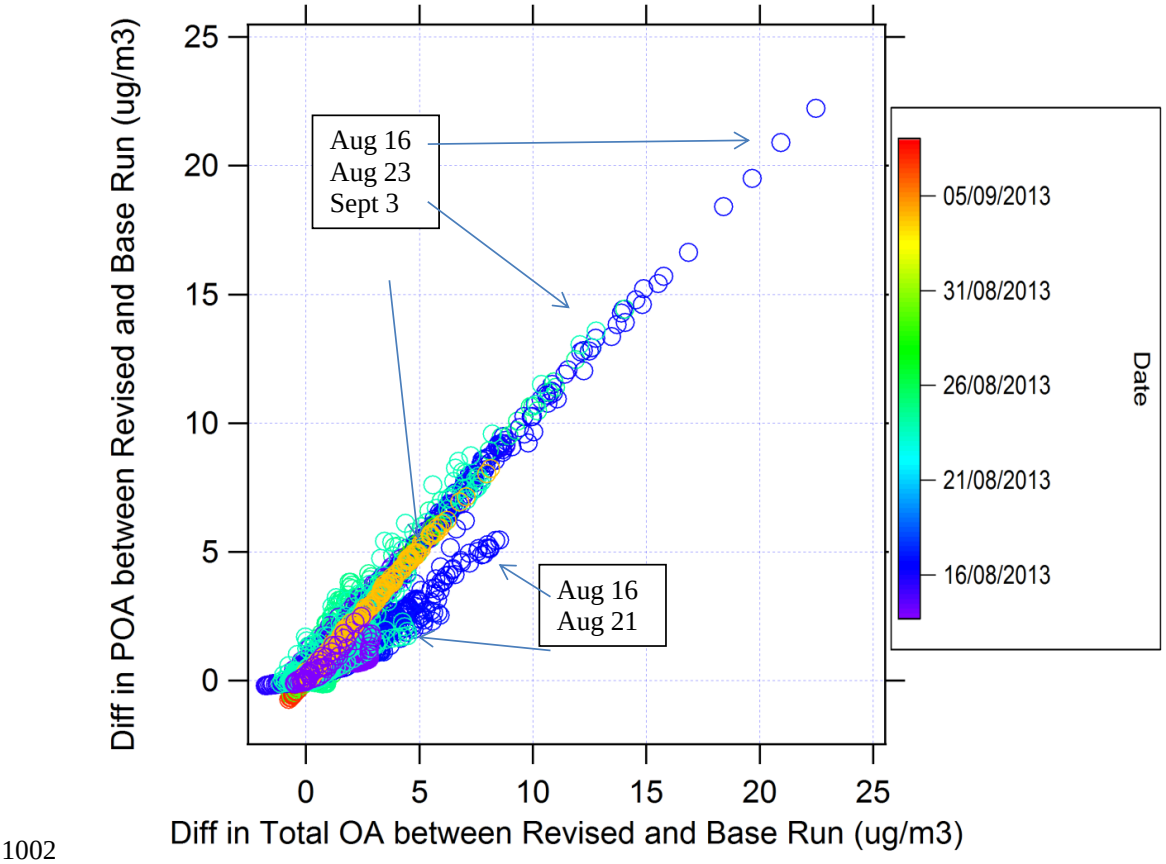


Figure 10. Difference in predicted POA concentrations between revised-emissions and base-case runs plotted as a function of the difference in predicted total OA concentration between the revised-emissions and base-case runs for all flights. Points along the 1:1 line show a difference solely from POA emission changes. Points below the 1:1 line show enhanced SOA formation.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

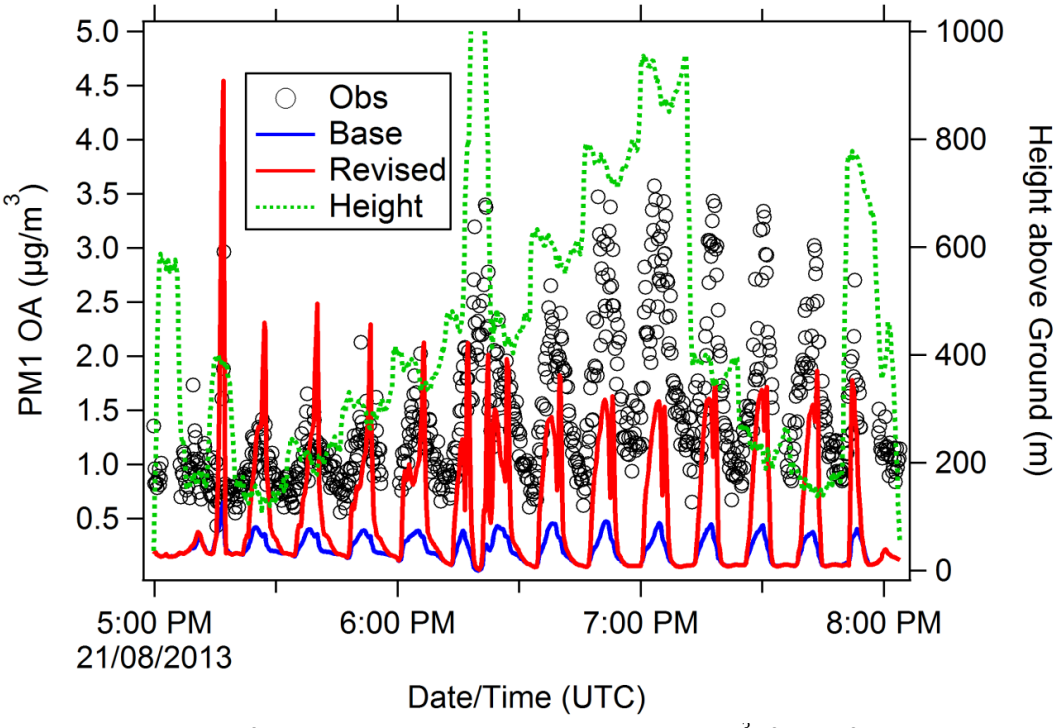


Figure 11. Time series for PM_1 organic aerosol concentration (µg/m^3) for the flight on August 21 crossing over the Syncrude Mildred Lake facility and then circling around the Shell Muskeg/Jackpine facility.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-93
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 4 April 2018

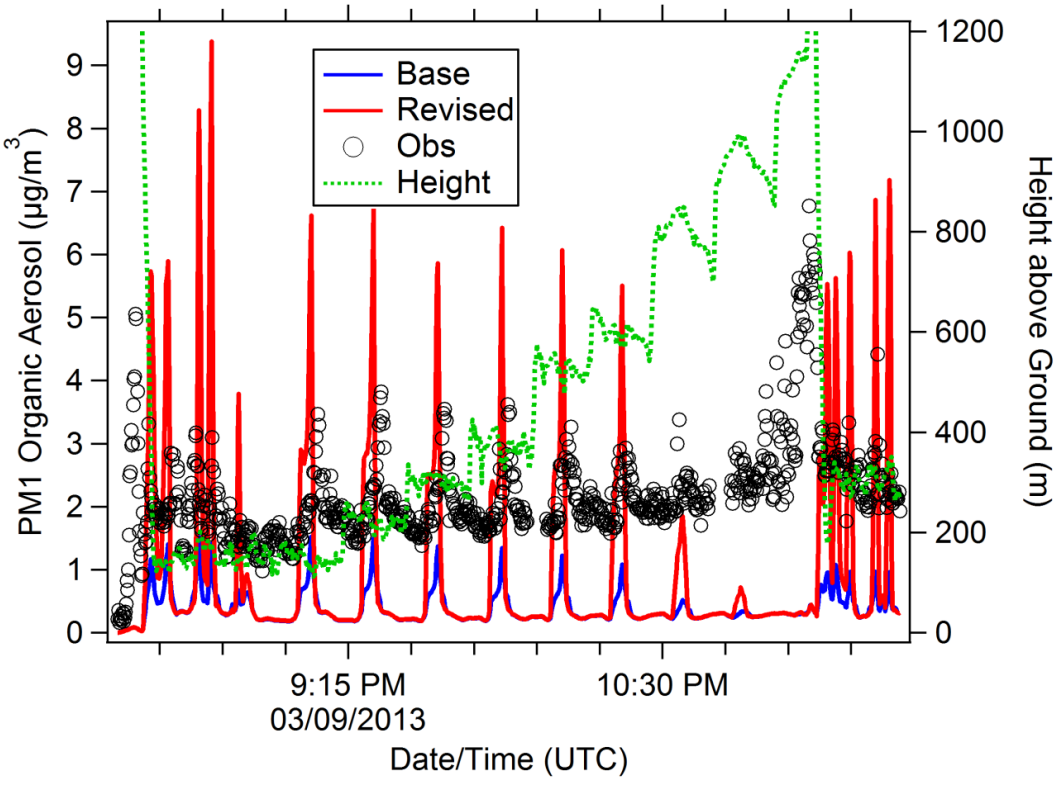


Figure 12. Time series for PM_1 organic aerosol concentration ($\mu g/m^3$) for the Sept. 3 flight over and around the Syncrude Mildred Lake facility.

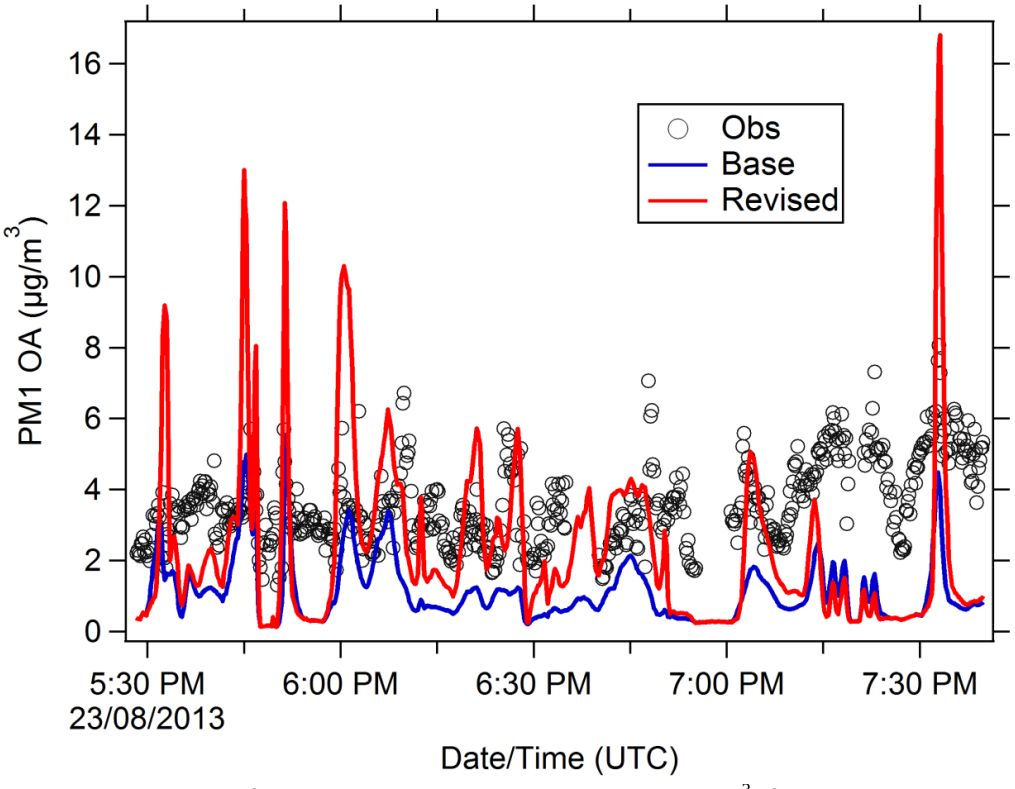


Figure 13. Time series for PM_1 organic aerosol concentration (µg/m^3) for the Aug. 23 survey flight.