Response to Reviewers

Reviewer #1

General Comments from Reviewer #1

The paper could be shortened without losing important points. Despite already having 13 manuscript figures, in places the text relies heavily on the 15 supplemental figures. Often beginning a discussion point by referencing the supplement that a typical reader will not see. The OA section seems particularly dense given what can be concluded.

The major update in this paper is the emissions, which are characterized throughout the results without much detail in the methods. In the methods, there PM facilities that are revised are mentioned, but no estimate of total changes is provided. Even with a companion paper, I would expect to see a summary. The paragraph on profiles is difficult to follow. Given the two changes, rates and speciation, a summary of changes by relevant species (TOLU, ALKA, AROM, POA) would be helpful. A summary here could help reduce the reliance on supplemental emission figures.

Recommend focusing the paper, removing over-reliance on supplement, removing some of the OA section, and improving clarity in the emission revision methods section.

We appreciate the reviewer's constructive recommendations on how to improve the clarity, length and structure of the manuscript. We have moved the four histograms originally in the Supplementary Information (Figures S1, S4, S8 and S11) into the main paper to improve clarity so that the reader does not have to refer to the SI section to see the histograms. These histograms illustrate important conclusions in the paper, particularly the model improvement for the higher percentile concentration bins. In the TOLU section, we have removed Figure 2, as it can be discussed in a few words. In the OA section, we have removed Figure 8 and its discussion because it is a complex figure and can have multiple interpretations. Removing these two figures and related discussion will help to shorten the manuscript. Table S1, comparing the VOC speciation profiles, has been moved from the SI to the main paper to help clarify the text, now labelled as Table 2.

We have modified the paragraph describing the VOC chemical speciation profiles as follows:

"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 (2) extraction-only facilities (Shell Canada Muskeg/Jackpine, Syncrude Aurora North, and Imperial Oil Kearl). Table 2 shows a comparison of the CEMA plant-specific VOC speciation profiles used in the base case for the two types of OS plants compared with two standard VOC speciation profiles for petrochemical facilities (#9012 "Petroleum Industry – Average", #0316 "Fugitive Emissions, Pipe/Valve Flanges") that were used by SMOKE to speciate more than half of the refinery emissions in the Houston area, the largest petrochemical processing cluster in the U.S. There are significant differences between the base-case OS plant VOC speciation profiles and the two commonly used standard oil refinery standard profiles. The OS integrated extraction and

upgrading plant profiles are higher in long-chain alkenes, toluene, and other aromatics than the standard oil refinery profiles, while the extraction-only OS plant stack profile has the highest long-chain alkane fraction. The two standard oil refinery profiles used in the base-case and revised-case simulations (for speciating U.S. and Canadian refinery emissions) have higher less-reactive species (e.g. propane, acetylene) and higher formaldehyde (profile #9012), than both the CEMA OS plant profiles. Note also that these differences in relative fractions result in substantial differences in the absolute emissions of certain groups of VOCs between the standard profiles for oil refineries and the facility-specific oil sand profiles. For reference, the aircraft-measurement-derived facility-specific VOC speciation profiles used for four OS facilities in the revised-emissions case are presented in Zhang et al. (2018). The aircraft-measurement-derived profiles in Zhang et al. (2018), and used here for the revised case, are composite profiles since they encompass plant, tailing pond and mining emissions. As such, they are not appropriate for comparison with the profiles in Table 2, which are specific to plant emissions."

Table 2. Facility-specific VOC speciation profiles (mass fractions) applied to the surface mining facilities in the Athabasca oil sands region compared to standard speciation profiles for Canadian and U.S. petrochemical oil refineries (in ADOM-II chemical speciation). Data are based on Zhang *et al.* (2018) and references therein. All four profiles are used in the base case simulation.

Species	Shell M/J, Syncrude AN, Imperial Kearl Base-Case Plant Profile (CEMA)	Syncrude ML, Suncor, CNRL Base-Case Plant Profile (CEMA)	CEPS Database Standard Profile #9012 for Oil Refineries in Base Case	SPECIATE Database Standard Profile #0316 for Oil Refineries in Base Case*
EC38 (Propane, Benzene, Acetylene)	0.0	0.0	0.247	0.176
EA3 (Alkane ≥C4)	0.90	0.71	0.623	0.781
EA2 (Alkene ≥C3)	0.007	0.069	0.031	0.002
ETOL (Toluene and other mono-aromatics)	0.001	0.057	0.005	0.008
EARO (Multi-functional aromatics)	0.0003	0.099	0.003	0.003
EHCO (Formaldehyde)	0.00001	0.0003	0.110	0.0

Columns do not add up to unity due to "unaccounted for" or "unassigned" species and/or due to consideration of reactivity weighting for the ADOM-II mechanism.
*Refinery Profile #9012 is a profile from the Canadian Emissions Processing System (Moran, M.D., M.T. Scholtz, C.F. Slama, A. Dorkalam, A. Taylor, N.S. Ting, D. Davies, P.A. Makar, S. Venkatesh, An Overview of CEPS1.0: Version 1.0 of the Canadian Emissions Processing System for Regional-Scale Air Quality Models. In Proc. 7th AWMA Emission Inventory Symp., Research Triangle Park, North Carolina, Air & Waste Management Association, Pittsburgh, Oct. 28-30, 1997.)

As recommended by the reviewer, we have also added a table quantifying the changes to total facility emission rates between the base case and the revised run. This table enables a reader to see the changes in total species emission rates for each facility without needing to refer to the SI to see the 4 emission change maps (Figures S2, S5, S9, S12 in original manuscript). Below is the new table.

Table 1. Facility total emission rates for three lumped organic species and $PM_{2.5}$ calculated with the bottom-up, base case inventory, CEMA facility-specific VOC profiles (labeled Base Case) and the top-down measurement-derived rates (labeled Revised Emission case, scaled to tonnes/year for VOCs or tonnes/Aug&Sept for $PM_{2.5}$). Emission rate increase/decrease of more than ±500 tonnes compared to base case is shown in red/blue.

	Suncor – M/S		Syncrude - ML		Shell – MR/J		CNRL - Horizon	
Species	**Base Case**	**Revised Emission Case**	**Base Case**	**Revised Emission Case**	**Base Case**	**Revised Emission Case**	**Base Case**	**Revised Emission Case**
Mono-Substituted Aromatics (TOLU)	486	1,112	806	1,539	6.8	72	135	393
Multi-Substituted Aromatics (AROM)	1,457	1,569	5,273	1,696	746	88	1,125	500
Long Chain Alkanes (ALKA)	5,636	13,488	12,348	10,022	1,690	14,384	2,651	23,779
Particulate Matter ($PM_{2.5}$)	1251	2537*	1021	3648*	459	2423*	402	1015*

* based on 2-month emission (Aug&Sept) rather than based on annual estimate (Zhang et al., 2018)

We have added the following text to the emissions description in Methods section of paper to discuss this new table:

"Table 1 compares the facility emission rates for four species for the base case and revised-emissions case. The changes are not consistent from species to species and are not uniform across facilities. For example, the ALKA species showed a large decrease for one facility but

increases for the other facilities. Likewise, for a given facility, some species showed increases and some species showed decreases. The SI includes emission difference maps for the oil sand region (absolute and relative differences) showing the spatial distribution of changes. The changes are largest over the surface mines and tailing ponds. The revised $PM_{2.5}$ emissions in Table 1 are derived only for summer months (sum of August and September) due to the uncertainties in extrapolating measurement-based dust emissions to other seasons (Zhang et al., 2018)"

Specific Comments from Reviewer #1

Specific notes: title : recommend signaling the improvements in the title. abstract : has too many methodological details including a reference to an accompanying paper. In 32-33 : not a clearly stated point, esp without having read the paper. In 27-40 : appropriate for an abstract? In 67,69 : use references instead of urls In 121-124 : seems out of place here. In 151 : consider using sections (151 Emissions, 211 Modeling, 225 Observations) In 177-193 : nomenclature is inconsistent and confusing: between paragraphs and between the text and supplemental table. In 179 : missing "(2)" after "and"? In 187 : "other profiles" could include "integrated extraction and upgrading," but I am pretty sure you mean the base-case profiles. In 260-264 : Bias as a function of magnitude is very important considering how your data is being used. It is not uncommon to have a few high points driving the relationship. In addition, the std error in slopes should be reported and used in your analysis. In 276 : since you will apply a similar approach to AROM, you should discuss it here. In 307 : are you referring specifically to the secondary peak? or are there other flyover data? In 339 : Given the inherent uncertainty in this approach, did you perform analyses where you did not subtract peaks from AROM and instead lumped AROM and TOLU? In 341,350,367,395,399,430,550 : These paragraphs begin by introducing patterns that a typical reader should not have to see. This over-reliance on the supplement is distracting and makes your paper hard to read. In 455-457 : this statement is made using complex figures and then made clearly later in Figure 10. Figure 10 succinctly conveys what I believe you were trying to communicate with Figure 8 and 9. Consider removing 8/9 and associated discussion.

1) Title: recommend signaling the improvements in the title.

We have changed the title to emphasize the improvements resulting from the work: "Improved Air Quality Predictions using Measurement-Derived Organic Gaseous and Particle Emissions in a Petrochemical-Dominated Region".

2) Abstract: has too many methodological details including a reference to an accompanying paper.

We have removed some of the methodological detail in the abstract, as well as removing the references since they are already in the Methodology section. The abstract has been reduced from 447 words to 346 words. It now reads as follows:

"This study assesses the impact of revised volatile organic compound (VOC) and organic aerosol (OA) emission estimates in the GEM-MACH (Global Environmental Multiscale–Modelling Air Quality and CHemistry) chemical transport model on air quality predictions for the Athabasca oil sands region in Northern Alberta, Canada. The first emissions dataset that was evaluated (base-case run) makes use of regulatory-reported VOC and particulate matter emissions data for the large oil sands mining facilities. The second emissions dataset (sensitivity run) uses total facility emissions and speciation profiles derived from box-flight aircraft observations around specific facilities. Large increases in some VOC and OA emissions in the revised-emissions data set for four large oil sands mining facilities and decreases in others were found to improve the modeled VOC and OA concentration maxima in facility plumes, 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 locations in North America. A significant increase in the correlation coefficient is reported for the long-chain alkane predictions against observations when using the revised emissions based on aircraft observations. For some facilities, larger long chain alkane emissions resulted in higher secondary organic aerosol production, which improved OA predictions in those plumes. Overall, the use of the revised emissions data resulted in an improvement of the model mean OA bias; however, the decrease in OA correlation coefficient and a remaining negative bias suggests the need for further improvements to model OA emissions and formation processes. The weight of evidence suggests that the top-down emission estimation technique helps to better constrain the fugitive emissions in the oil sands region, which are a challenge to estimate given the size and complexity of the oil sands operations and the number of steps in the process chain from bitumen extraction to refined oil product. This work shows that the top-down emission technique may help to constrain bottom-up emission inventories in other industrial regions of the world with large sources of VOCs and OA**."**

3) Lines 32-33: Not a clearly stated point.

We changed the wording of line 32-33. The new text is as follows: "For some facilities, larger long-chain alkane emissions resulted in higher secondary organic aerosol (SOA) production, which improved OA predictions in those plumes."

4) Lines 27-40: Not appropriate for abstract.

We have removed the statement of future work in the abstract (lines 37-38) to focus and shorten abstract. We discuss IVOCs and future work in the Discussion section of manuscript.

5) Line 67,69: Use reference and not url.

We have removed the URLs and added references for the Canadian and U.S. national pollutant inventories as follows:

Government of Canada, Notice with respect to the substances in the National Pollutant Release Inventory for 2018 and 2019, Canada Gazette Part I, Vol. 152, No. 3, pp. 129-172, ISSN 1494-6076, Ottawa, January 20, 2018.

Office of the Federal Register National Archives and Records Administration, Protection of Environment, Code of Federal Regulations, Title 40, Parts 50 to 51, Special Edition of the Federal Register, U.S. government publishing office, Washington, DC 20402–0001, July 1, 2015.

6) Lines 121-124: Seems out of place here.

The discussion on lines 121-124 in the original manuscript is already in the Methodology section, so it is redundant. We have removed the lines from the Introduction to shorten the manuscript.

7) Line 151: Consider using titles: Emissions, Modeling, Observations.

We have added the subtitles as recommended for 2.1 Emissions, 2.2 Modeling and 2.3 Observations.

8) Line 177-193: Nomenclature is inconsistent and confusing between paragraphs and between the text and supplemental table. Line 179: missing "(2)" after "and".

We have added more descriptive words to clarify the paragraph on chemical speciation profiles (discussed above in addressing general comments). We added the "(2)" text.

9) Line 187: "Other profiles" could include "integrated extraction and upgrading", but I am pretty sure you mean the base-case profiles.

We have modified the text to the VOC speciation profile paragraph, as described above in general comments. The "Other profiles" refers to the standard profiles used to speciate oil refineries in other parts of North America.

10) Line 260-264: Bias as a function of magnitude is very important considering how your data is being used. It is not uncommon to have a few points deriving the relationship. In addition, the std error in slopes should be reported and used in your analysis.

We are deriving a relationship to calculate equivalent observed lumped VOC species to compare to the same model lumped VOC species. We performed a series of linear fits to transform the PTR-MS measured VOC species (e.g. ethyl benzene vs. toluene). We agree that the standard error on slopes can provide the reader more information on robustness of the correlations. We have added this information, $m=0.376\pm0.006$ (for ethyl benzene vs. toluene) and $m=0.0652\pm0.0008$ (for propyl benzene vs toluene). We have also added the y-intercepts and their standard error, $b=0.033\pm0.006$ ppbv (for ethyl benzene vs toluene) and $b=0.0011\pm0.0008$ ppbv (for propyl benzene vs. toluene). In the linear correlation plots, there are not have a few points deriving the entire relationship. The approach of using the chemical speciation from the canister data (e.g. ethyl benzene) and correlating against a fast time tracer in the PTR-MS data (e.g. toluene) is published in Li et al., (2017).

11) Line 276: Since you will apply a similar approach to AROM you should discuss it here.

We would prefer to keep the discussion of TOLU and AROM separate for clarity.

12) Line 307: Are you referring specifically to the secondary peak? Or are there other flyover data?

The secondary peak refers to the second peak "in the figure". We have added the phrase, "in the figure" to help clarify.

13) Line 339: Given the inherent uncertainty in this approach, did you perform analyses where you did not subtract peaks from AROM and instead lumped AROM and TOLU?

We wanted to evaluate the emissions for both TOLU and AROM species since this is how the aromatics are speciated in the GEM-MACH model. We did not evaluate the sum of TOLU+AROM.

14) Line 341, 350, 367, 395, 399, 430 and 550: These paragraphs begin by introducing patterns that a typical reader should not have to see. The over-reliance on the supplement is distracting and makes the paper hard to read.

We have decided to move the supplemental figures with flight patterns as panels in the corresponding time series figures in the main section of manuscript, so the reader does not need to go back and forth to the SI. We have also moved the histograms to the main figures since they justify a conclusion in the paper.

15) Line 455-457: This statement is made using complex figures and then made clearly later in Figure 10. Figure 10 succinctly conveys what I believe you were trying to communicate with Figure 8 and 9. Consider removing 8/9 and associated discussion.

We have removed Figure 8 since it is a complex plot and its interpretation is inconclusive. Figure 9 shows that the model OA bias is improved more for samples that are influenced by petrochemical combustion, as determined by plotting bias against black carbon measurements. This is due to under-predictions in OA from petrochemical-related sources in the base case simulation. The figure also shows that the OA bias is better, but still remains even with the revised emissions, for the samples influenced by petrochemical sources. This is an important conclusion of the paper.

Reviewer #2

A. The study relies heavily on an earlier submission to Atmospheric Chemistry and Physics (Zhang et al., 2018), which describes the emissions used in the model. Publication of this earlier paper seems essential for acceptance of the present manuscript.

This manuscript has been accepted for publication in ACP.

B. In Sections 3.1 through 3.3, the model is compared with the observations of various lumped hydrocarbon species. I would like to see a much clearer discussion of what is actually being evaluated here. Briefly, the aircraft measurements were used to calculate emissions, which are then used as inputs for the model. The output from the model is compared with the measurements again. Not surprisingly, the model with revised emissions, i.e. those driven by the measurements, agrees better with the measurements. The argument can thus be perceived as being circular, but I do believe it is still a useful exercise and also lays the groundwork for Section 3.4 where the organic aerosol is compared between model and measurements. Nevertheless, the paper should describe much more clearly what is being evaluated in this study (for example more detailed atmospheric transport, model resolution, temporal variability in emissions, etc.). Were the data shown in Figures 3-7 used to calculate the emissions? If so, what is learned from this study about the accuracy of the revised emissions? Are the box flights adequate to quantify emissions or is the transport more complex leading to inaccurate emissions estimates? Another option might be to use part of the measurement data to derive emissions and test the model output with these emissions versus another part of the data set. As it is, the paper gives a fairly dry comparison between the measurements and two different models, and does not describe the above subtleties in any detail.

Regarding circularity, it is present but it is very indirect. The goal of the paper is to report how good the model performance can be if we constrain the model with a top-down measurement-based emission data set. Since the two model runs are only different in their emissions, we are isolating the impact of the new emission data set. Of course, the model results still have variability compared to observations due to uncertainties in modeling meteorology and atmospheric dispersion, particularly on the local scale with pollutants still in plumes. We are also evaluating the spatial and temporal disaggregation of the facility-total emission rates, that is, the emissions processing step necessary to connect emissions inventories with AQ model input emissions files. Given all of the uncertainties inherent in the modeling, this study is clear in that we are just isolating the impact of changing the facility-total emissions rates from bottom-up inventory estimates to top-down measurement values.

It is not uncommon for models to make use of emissions derived from measurements (and this is preferred, if the emissions data from inventories are uncertain). For example, Continuous Emissions Monitoring Systems (CEMS) measure the SO_2 and NOx concentrations from sensors on tall stacks in Canada and the United States; the stack concentrations and flow rates are then used to calculate measured-emission rates, which depending on the regulatory environment, may

be reported as hourly emissions or annual totals. It is common for models to use the CEMS emission data. We note that the use of the new emissions in our study did not result in a perfect match with measurements downwind – which shows that either those emissions estimates are not perfect, or (more likely) the remainder of the model processes influencing predicted downwind concentrations are not perfect. We have added the following paragraph to the Discussion section, outlining this issue:

"The use of aircraft observations to both derive emissions data and evaluate the subsequent model evaluations might be taken as circular reasoning. We note first that observation derived emissions are frequently used in modelling (for example, Continuous Emissions Monitoring System concentration observations are used to generate emissions data for large stack emitters), and second, that the emissions are only one component of the overall modelling system. An improvement in the simulated VOC concentrations using observation-based emissions is only guaranteed if the emissions dominate the net model error. While our results show that the new emissions information does improve model performance, the results using that new data are not perfect, indicating other sources of error are contributing to the overall model performance."

C. I found the analysis in Section 3.4 to be quite confusing. Earlier work from this group had shown that low-volatility organic compounds are important to explain the strong SOA formation downwind from the oil sands (Liggio et al., 2016). Therefore, my expectation reading this part of the paper was for the Authors to show better model performance using the improved emissions including for low-volatility organic compounds. However, emissions of these low-volatility organic compounds were not explicitly included in the model and only mentioned as an afterthought in Section 3.4. The conclusion that is conveyed to the reader is that the observed SOA can be better explained using the revised emissions of hydrocarbons and the Authors recommend a better treatment of SOA from monoterpenes and, perhaps, including SOA from low-volatility organic compounds. I find these conclusions to be almost orthogonal to the earlier work published in Nature.

This work is one of several research studies underway at ECCC to try to understand and constrain the potential causes of the observed negative bias in OA formation near the oil sands and elsewhere. Here, we are not saying that all the OA formation can be explained with the revised VOC and POA emissions, rather, that the updates in the emissions of VOC precursors lead to *some* improvements in the overall OA performance, and hence constrain the improvement which might be expected through further work, for example, on the IVOC part of OA formation. The aircraft measurements are mostly from facility box flights, as emissions characterization was the primary goal of the first Oil Sands study. There were only 3 Lagrangian transformation flights out of a total of 21 flights. One of the drawbacks of the 2013 study was that the gaseous IVOC and SVOC species were not explicitly measured by the aircraft – hence the findings in Nature were not confirmed by observations; rather, the modelling carried out for that paper was used to show the levels of IVOC and SVOC required to account for the missing OA mass. A follow-up measurement study, with instrumentation specifically designed to estimate the IVOC and SVOC, both in the gas-phase and particulate phase, has just been completed (final flights on July 5th, 2018). In preparation for the data from that study, and to further close the OA formation

budget, we have asked here "To what extent may the deficit in OA be due to inaccurate emissions and speciation for VOCs and primary particulate matter?" The best approach to simulate relative IVOC contributions to SOA formation is the topic of a research project in progress by a PhD student in Dr. Stroud's group and will use the new 2018 observation data set. Also, while the measurement-derived emissions are missing the gaseous IVOCs and SVOCs, the measurement-derived POA emissions may contain some IVOCs and SVOC species that react quickly in one oxidation step and condense onto particles. This rapid SOA mass produced would be measured by the box flights and, at least partially, accounted for in the updated OA emissions; however labeled here as POA instead of fresh SOA. In this paper, we have tried to minimize this effect by examining the model performance in the "near field" from emission flights close to facilities (assuming 5-km distance from emission source to aircraft box location and 3m/s wind results in 0.5 hr transport time). This is a short time, but not so short that some reactive precursor gases could form SOA. This will be the topic of future box modelling work with the new 2018 measurement-derived gaseous IVOC and SVOC emissions to determine how much of the measurement-derived POA is derived from the fugitive open-pit mining IVOC and SVOC emissions and their rapid particle formation.

In our current work, we see the model is under-predicting the aged background organic aerosol over the boreal forest (outside of plumes when the SO_2 and NOx are low). This is an independent issue from the oil sand emissions. Ongoing work by Dr. Stroud's PhD student is showing that this requires an update to the biogenic SOA yield parameters. In this current work, we do not focus on updating the biogenic VOC emissions, rather just the anthropogenic VOC and OA emissions from the facilities, as this was the manuscript goal.

We have added a Discussion section to the manuscript where these issues and future recommendations are discussed. The Discussion section and Conclusion section are as follows:

4.0 Discussion

The improvement in model PM_1 OA mean 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. The remaining negative bias suggests that other important processes may be missing or under-represented in the model. Three recommendations emerge from recent publications and this current work:

4.1 SOA Formation from Fugitive IVOC Emissions

Recent publications 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*., 2017). In our emissions revision, only a small portion of the IVOCs (dodecane C_{12}) were available in the VOC speciation – these were lumped into the long-chain ALKA lumped species. IVOC species with carbon number $\geq$13 were not measured by the Li *et al.*, (2017) aircraft study of 2013 and thus we do not have revised IVOC emissions included in this work. Furthermore, the ALKA lumped species has an SOA yield more representative of a lower molecular-weight range, and the yield is known to increase with increasing carbon number, so the dodecane contribution would be underestimated as simulated here. Work is currently underway with GEM-MACH to implement a Volatility Basis Set (VBS) approach to SOA

formation. The VBS approach will more adequately represent the intermediate and semi-volatile volatility range and chemical aging of these lower volatility compounds (Robinson *et al*., 2006). Future work will incorporate new IVOC emissions estimates from 2018 box flights around the oil sand facilities and the open-pit mines. This will help remove current uncertainties in the models and likely help improve the negative OA bias in plumes. Implementing the VBS scheme will also enable the PM emissions used here (in both emission data sets) to be distributed into volatility bins.

Also, while the measurement-derived emissions are missing the IVOCs, the measurement-derived POA emissions may contain some gaseous VOCs, IVOCs and SVOC species that react quickly and in one oxidation step yield products that condense onto particles. This rapid SOA mass produced would be measured by the box flights and, at least partially, accounted for in the updated OA emissions; however labeled here as POA instead of fresh SOA. Furthermore, there is the potential for double counting if some of the very reactive gaseous precursors react to form SOA and this is accounted for in the measured POA. In this paper, we have tried to minimize this effect by examining the model performance in the "near field" from emission flights close to facilities. This will be the topic of future box modelling work with the new 2018 measurement-derived IVOC and SVOC emissions to determine how much of the measurement-derived POA is derived from the fugitive open-pit mining IVOC and SVOC emissions and their rapid particle formation.

4.2 Background Organic Aerosol Levels

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, and to the lack of a speciated representation of other biogenic SOA precursors such as sesquiterpenes. Future work will update the biogenic SOA yield coefficients using the VBS approach and recent smog chamber results which account for gas-phase loss of organic species to chamber walls (Ma *et al.*, 2017).

4.3 Spatial Allocation of Emissions

Future field studies should also focus on improving within-facility spatial allocation of PM emissions. 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 truck emission spatial allocation within a facility. The GPS data would also be useful to define the location of freshly excavated open-pit mines within a facility to spatially allocate IVOC emissions.

4) Conclusions

Overall, the weight of evidence suggests that the top-down emission estimation technique applied to the OS surface mining facilities helps to better constrain reported facility-total organic emissions, as shown here by improved model predictions when the revised emissions are employed. We note that emissions from these sources are a challenge to calculate in bottom-up inventories due to the potential for fugitive emissions. For the mono- and multi-substituted aromatics (TOLU and AROM), the revised emission rates from facilities were more fine adjustments, as some facility totals increased and some decreased 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 paraffinic solvents for bitumen extraction (Shell Muskeg/Jackpine and Syncrude Aurora North; Li *et al.*, 2017). Overall, the predictions of alkanes in high concentration plumes improved with the revised emission data set, as shown by the 99th percentile statistic.

For PM_1 organic aerosol, the revised emissions improved the mean bias for predictions; however, a negative bias still exists and the improvement was associated with a decrease in correlation coefficient. The increase in predicted PM_1 OA concentration was largely due to the increase in POA emissions in the revised emissions input files. The POA emissions increased because of a combination of larger measurement-derived PM_1 emissions and the revised ground-observed PM speciation profile having a larger POA fraction. The increase in PM_1 POA emissions were largely allocated spatially to stack locations and this allocation may be a key factor in the degradation of the correlation coefficient, especially if the fine OA actually originates from mine-face fugitive emissions. Future work should focus on improving within-facility spatial allocation of emissions. The remaining negative bias in plumes likely stems from missing IVOC emissions in both the data sets used here, as suggested in Liggio *et al.* (2015). Ongoing field work to measure the IVOC emissions using aircraft box flights is underway in a new 2018 measurement intensive. Upcoming modelling work with GEM-MACH will include the VBS approach to better represent the lower volatility compounds.

Detailed comments:

The Abstract is quite long and even contains two paragraphs. While I do not have specific suggestions, one has to wonder if the content of the paper can be summarized more succinctly.

Along similar lines, I wonder if all five Figures 3-7 are needed to present the comparisons between model and measurements.

Reviewer 1 made similar comments and we have taken both sets of comments to heart. We have shortened the abstract to one paragraph, removing the references to other work and removed some technical description. The paper text is also shorter than the original. We have modified the figures as described in our response to Reviewer 1 (removed 2 figures and moved some SI figures to main section of manuscript for clarity).

Improving 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 on air quality predictions for the Athabasca oil sands region in Northern Alberta, Canada., ~~driven with two different emissions input datasets, using observations from the 2013 Joint Oil Sands Monitoring (JOSM) intensive field study~~. The first emissions dataset that was evaluated (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~~ T~~t~~he second emissions dataset (sensitivity run) uses total facility emissions and speciation profiles derived from ~~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).~~

L~~The l~~arge increases in some VOC and OA emissions in the revised ~~–~~emissions data set for four large oil sands mining facilities and decreases for others were found to improve the modeled VOC and OA concentration maxima in facility 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 significant increase in the correlation coefficient is reported for the long-chain alkane predictions against observations when using the revised emissions based on aircraft observations. For some facilities, larger long chain alkane emissions resulted in higher secondary organic aerosol production, which improved OA predictions in those plumes. ~~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.~~ Overall, t~~T~~he use of the revised emissions data resulted in ~~a large~~ an improvement of the model mean OA bias; however, the decrease in OA correlation coefficient and a remaining negative bias suggests the need for further improvements to model ~~organic aerosol~~ OA 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, t~~The weight of evidence suggests that the top-down emission estimation technique helps to ~~new aircraft-observation-derived organic emissions help to~~ better constrain ~~better~~ the fugitive organic emissions in the oil sands region, which are a challenge to estimate given the size and complexity of the oil sands operations and the number of steps in the process chain from bitumen extraction to refined oil product. ~~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. This work shows that top-down eEmissions estimation techniques, such as those used to construct the inventories in our study, may therefore may help to constrain bottom-up emission inventories in have beneficial impacts when applied to other industrial regions of the world 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(Government of Canada, Canada Gazette, 2018) and the

National Emissions Inventory (NEI) in the United States ~~(https://www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei.html~~(Office of the Federal Register, Protection of Environment, 2015). 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 either used directly,~~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).

The Joint Oil Sands Monitoring (JOSM) program 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 to monitor the potential impacts of pollutant emissions. During JOSM, t~~T~~op-down approaches to estimate emissions based on atmospheric observations provided a unique opportunity to compare with bottom-up calculated emissions for the ~~. 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 (Makar *et al.*, 2018). ~~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).~~ In this model study,

~~Here~~ we make use of both regulatory-inventory-based and aircraft-observation-derived~~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 two different emission data sets ~~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).~~

Formatted: Font: Italic

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.

2.1 Emissions

The Canadian base-case emissions were derived from a hybrid inventory targeting 2013 as the base year, as described by Zhang *et al*. (2018~~7~~). 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.*, 2018~~7~~). 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*. (2018~~7~~). Table 1 compares the facility emission rates for four species for the base case and revised-emissions case. The

changes are not consistent from species to species and are not uniform across facilities. For example, the ALKA species showed a large decrease for one facility and increases for the other facilities. Likewise, at a given facility, some species showed increases and some species showed decreases. The SI section includes figures illustrating the emission difference maps for the oil sand region (absolute and relative difference) showing the spatial distribution of changes. The changes are largest over the surface mines and tailing ponds. The revised $PM_{2.5}$ emissions in Table 1 are derived only for summer months (sum of August and September) due to the uncertainties in extrapolating measurement-based dust emissions to other seasons (Zhang *et al.*, 2018).

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 (2) extraction-only facilities (Shell Canada Muskeg/Jackpine, Syncrude Aurora North, and Imperial Oil Kearl). Table 2~~S1~~ shows a comparison of the CEMA plant-specific VOC speciation profiles used in the base case for the two types of OS plants ~~facilities~~ compared to two standard VOC speciation profiles for petrochemical facilities (#9012 "Petroleum Industry – Average", #0316 "Fugitive Emissions, Pipe/Valve Flanges") that were used by SMOKE ~~for the base case~~ to speciate more than half of the ~~U.S.~~ refinery emissions in the Houston area, the largest petrochemical cluster in the U.S. There are significant differences between the base case OS plant ~~aircraft-observation-based OS~~ VOC speciation profiles and the two commonly used standard oil refinery ~~reference~~ profiles. The OS integrated extraction and upgrading plant ~~facilities~~ profiles are higher in long-chain alkenes, toluene, and other aromatics than the standard~~reference~~ profiles, while the extraction-only plant ~~facilities~~ profile has~~ve~~ the highest

long-chain alkane fraction. The ~~other~~ two standard profiles used for in base case and revised simulation (for speciating U.S. and Canadian refinery emissions) have higher less-reactive species (e.g., propane, acetylene) and formaldehyde (profile #9012), than both the CEMA OS plant -profiles. Note also that these differences in relative fractions result in substantial differences in the absolute emissions of certain groups of VOCs between the standard profiles for oil refineries and the facility-specific oil sand profiles. For reference, the aircraft-measurement-derived facility-specific VOC speciation profiles used for four OS facilities in the revised-emissions case are presented in Zhang *et al.* (2018). The aircraft-measurement-derived profiles in Zhang *et al.* (2018), and used here for the revised case, are composite profiles since they encompass plant, tailing pond and mining emissions. As such, they are not appropriate for comparison with the profiles in Table 2, which are specific to plant emissions.

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.*, 2018~~7~~). In general, significantly higher POA emissions were observed over the open-pit mines for all facilities, except for the Imperial Kearl mine. The~~is~~ impact of the revised POA emissions will be discussed further in Section 3.4.

2.2 Modeling

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.

2.3 Observations

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

Formatted: Font: Italic

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 species~~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 the ~~time and spatially linearly interpolated~~ GEM-MACH model output. The ~~interpolated~~ model grid cell output was extracted along the flight track and interpolated linearly between the two minute model output intervals ~~along the flight track was merged with the observations~~ to create a coincident model and ~~–~~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, y-intercept and correlation coefficient for both these plots (not shown) were as follows: m=0.376±0.006, y=0.0328±0.006, R=0.91 and m=0.0652±0.0008, y-intercept=0.0011±0.0008, 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 2 ~~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 3~~1~~ lists ~~other~~ statistical scores for the TOLU lumped species and the other species considered in this study. 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.

~~Figure S2 illustrates the difference between the 2.5 km TOLU emissions field over the Athabasca OS mining facilities for the revised base case as a spatial map for one selected time~~

~~(Friday , August, 18Z).~~ The largest increases in the TOLU emission, between the revised and base case run, are noted for the Syncrude Mildred Lake facility over the tailing ponds and open pit mine faces. ~~The~~ Table 1 shows the changes ~~shown here reflect net changes~~ on a facility-wide level. Notable ~~actual and relative~~ increases are also calculated for the Suncor Millennium/Steepbank and the Canadian Natural Resources Ltd (CNRL) Horizon facilities.

~~Figure 2 illustrates the correlation between the sensitivity-case TOLU output and the base-case TOLU output.~~ The flights on August 14 and 24 have the largest TOLU mixing ratios for the aircraft study, and both flights (blue and green points) correspond to box flights around the Syncrude Mildred Lake facility. Figure 3a~~S3~~ shows the flight path for August 14, color-coded as a function of the difference between the modelled revised and base-case concentrations. The background is a satellite map image along with the GEM meteorological model wind barbs predicted for that day at 16 UTC. The largest differences in the simulated concentrations (1.8 ppbv) correspond to a location just downwind of the Syncrude Mildred Lake open-pit mine, as expected based on the emission difference map (Figure S1~~2~~) and the southerly wind direction.

Figure 3b shows the time series for a segment of the August 14 flight corresponding to three flight boxes at different heights (green dotted line). The observations are plotted as open circles and the two lines represent the two model results. The model output with the revised VOC emissions clearly captures the main peak of the TOLU concentrations driven by TOLU emissions from the Syncrude Mildred Lake facility. The secondary peaks in the figure (a couple minutes apart from primary peaks) occur~~s~~ from the aircraft flying over the Suncor Millennium/Steepbank facility while on the east side of the flight box pattern. The direct flyover adds uncertainty to the model comparison, as it depends on predicting 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 3~~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 5~~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 3~~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.

The largest decreases in the AROM emission field between the revised and base case emissions ~~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 (refer to Table 1). 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 (refer to Figure S2 for the emission spatial difference map). Figure 6~~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 7a~~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~~Figure 7b 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 8a~~S7~~ shows the ~~box~~ flight track on September 3~~,~~ over and around the ~~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 8b~~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 3~~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 9~~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 3~~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, t~~The revised ALKA emissions are considerably higher for the CNRL Horizon and Shell Muskeg/Jackpine facilities, but have smaller changes for the other

facilities (refer to Table 1), possibly reflecting differences in the processing activities between the facilities.

Figure 10a ~~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 the northeast and notably Fort McMurray (further to the south, not shown) had quite poor air quality. The largest differences in the modelled mixing ratios correspond to observation locations south of CNRL. Positive differences as large as 20 ppbv were simulated for some points. Figure 10b~~7~~ shows the time series for the observations, revised-emissions model results, and base-case model results for the Aug. 26 box flight around the CNRL Horizon facility. A clear improvement in ALKA modelling is observed when using the revised emissions for the plume sampled downwind of the CNRL facility.

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

The other facility that had large increases in ALKA emissions with the revised data was the Shell Muskeg/Jackpine facility (refer to Table 1 and Fig. S3~~9~~). Flight 9 on August 21 was a box flight around the Shell facility. A detailed analysis of this flight showed that for the majority of the data points on this flight, the model run with the base-case emissions showed the best results, except for the three highest measured canister samples, where the model run with the revised

emissions performed better. This likely reflects an uncertainty in the spatial allocation maps used to distribute the emissions with a higher fraction of emissions needed at the point specific locations.

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

The use of aircraft observations to both derive emissions data and evaluate the subsequent model evaluations might be taken as circular reasoning. We note first that observation derived emissions are frequently used in modelling (for example, Continuous Emissions Monitoring System concentration observations are used to generate emissions data for large stack emitters), and second, that the emissions are only one component of the overall modelling system. An improvement in the simulated VOC concentrations using observation-based emissions is only guaranteed if the emissions dominate the net model error. While our results show that the new emissions information does improve model performance, the results using that new data are not perfect, indicating other sources of error are contributing to the overall model performance.

3.4 Organic Aerosol (OA)

Figure 11 ~~S11~~ illustrates the histograms for the organic aerosol observations and model results with base case and revised emissions. A clear improvement is shown in the highest concentration bins (> 15 µg/m^3) with the revised emissions. This can be quantified with the 99th percentile of the data (obs=13.4 µg/m^3, revised=9.3 µg/m^3, base=4.9 µg/m^3. The median

statistics also improved (obs=2.8 µg/m^3, revised=0.84 µg/m^3, base=0.70 µg/m^3). The lower 5th percentile is also significantly under-predicted compared to observations and does not change much between 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 under-prediction in the background OA predicted by the model, which is likely due to too low a level of biogenic SOA formation and aging in both model versions. The importance of widespread biogenic SOA formation from boreal forests has been reported in other work (Slowik *et al*., 2010; Tunved *et al*., 2006).

~~The other~~ Additional statistics are presented in Table 3~~1~~. The mean bias, RMSE and slope all improve for the revised-emissions run, though the correlation coefficient decreases significantly for this run. ~~The decrease in correlation suggests that the improved bias may not always be the result for the right reasons for some points. To investigate this further, this study looked at the model PM_1 OA bias as a function of different observed variables. Figure 8a shows the base-case model bias as a function of observed PM_1 sulfate for individual points. Figure 8a is color-coded as a function of the model %SOA relative to model OA. There is a trend of increasing negative bias in model OA in the base case with increasing observed sulfate aerosol. This suggests that air masses that originate from sources high in SO_2 tend to be under-estimated in simulated organic aerosol. These air masses also tend to have large %SOA in the base model. Figure 8b is the same plot but for the model run using revised emissions. There are a large number of points that are relatively high in observed sulfate (>1 µg/m^3) that change from a negative OA bias to a positive OA bias and these points also shift to being dominantly POA in model composition (low %SOA, blue points). The model output confirms that much of the increase in model OA in the sensitivity run originates from the increase in primary PM emissions and from the increase in the mass fraction of that primary PM assumed to be OA (based on Wang *et al*., 2015). The revised~~

~~emissions simulation has less organic aerosol bias at high sulfate loading suggesting that the improved results are closely linked to model organic emissions co-located with sources of SO_2.~~

To investigate the variability in the OA bias, we plotted the OA bias as a function of different measured variables. Figure 12 ~~9a~~ is a plot of the OA bias as a function of the observed black carbon (BC) aerosol for the base-case and sensitivity runs. The BC is a marker for petrochemical combustion, particularly diesel. For the base-case run, the OA negative bias is observed to increase in magnitude with observed BC. Points with high observed BC correlate well with emissions from the OS open-pit mines (Liggio *et al*., 2017), where the BC is likely emitted from the heavy-hauler trucks. ~~This~~ The locations with the largest OA bias ~~was~~ were also consistent with the locations of mines and the transport wind direction. A review of the OS emission inventories suggest that about 70% of the BC comes from the OS off-road diesel fleet. Including all points, the mean bias improves from -2.8 to -2.4 (see Table 3~~1~~) when using the revised emissions. Figure 12~~9~~b shows a zoomed plot for points with high observed BC (>0.8 µg/m^3). There is a clear improvement in bias for most of these points. The average bias for these high BC points improves from -6.8 µg/m^3 for the base case to -2.6 µg/m^3 for the revised emissions. For emissions processing the increase in PM emissions was assigned to the processing plants (particle bin D<1µm) or the surface mines (particle bin D>1µm). Overall, Figure 12 shows that, while the negative OA bias improves for samples high in BC concentration (i.e. influenced by petrochemical combustion or collocated with petrochemical combustion sources), there still remains an unaccounted for negative OA bias. ~~Note that anthropogenic SOA formation should be minimal for the high BC points close to the emission source, so uncertainties stemming from the model SOA formation mechanism should be reduced. The revised emissions simulation has less~~

~~bias as a function of observed black carbon, suggesting that the improved results are closely linked to model emissions with similar sources to the black carbon~~.

Figure 13 ~~10~~ is a scatterplot of the difference in predicted POA between the revised and base-case emissions runs vs. the difference in predicted total OA. A large fraction of the points fall along the 1:1 line, and hence for these points the difference between the two runs is almost completely due to the increased total primary PM emissions, and increased POA fraction of those emissions, of the revised emissions simulations. The points with largest concentrations along the 1:1 line correspond to flights over the Syncrude Mildred Lake facility on Aug. 16, Aug. 23 and Sept. 3. There is a subset of points, however, that lies below the 1:1 line; these correspond to points with significantly enhanced model SOA between the two runs (Aug. 16 flight over CNRL Horizon and Aug. 21 survey flight over Shell Muskeg/Jackpine).

The focus of the flight on Aug. 21 was ~~an enclosing~~ a box pattern around the Shell Muskeg/Jackpine facility at different altitudes. The approach to this facility, however, also included an overpass of the Syncrude Mildred Lake facility. Figure 14ab ~~S13~~ illustrates the flight path color-coded as a function of POA difference (revised emissions – base case) and SOA difference (revised emissions –base case). The corresponding time series for OA observations, the revised emissions model run, and the base-case emissions model run OA predictions are shown in Figure 14c~~1~~. There is a clear "hot spot" in POA difference in Figure 14~~S13~~a located over the Syncrude Mildred Lake facility. This hot spot corresponds to the first large maxima in the times series in Figure 14c~~1~~ (17:17 UTC). The observations at this time lie between the predictions from the two model simulations, though the overestimate of the revised emissions simulation is closer to the observations than the underestimate of the base-case emissions simulation. The aircraft then entered the box pattern at different altitudes around the Shell

Muskeg/Jackpine facility, and each successive pass around this facility intersected the observed plume on the north-east corner of the flight box (see hot-spot, Figure 14~~S13~~b); the model predicts that the increase in OA is largely due to SOA (as implied by Figure 13~~10~~), and the revised-emissions simulation produces peak OA concentrations that are closer to the observations than the base-case emissions simulation. As is clear from Figure~~.~~ 14c~~1~~, the base-case emissions simulation greatly underestimated the OA relative to observations. In examining the time series, it is also clear that both model simulations are under-estimating the background biogenic OA concentrations, by about 0.5 μg m^{-3}. The height of the peaks relative to background is closer to the sensitivity run peaks than the base-case run peaks.

Figure 15a~~S14~~ shows the difference between revised-emissions and base-case model OA predictions for another case study on September 3, for southerly winds with a box flight over the Syncrude Mildred Lake facility. The flight started and ended with a horizontal zig-zag pattern with overpasses directly over the facility emissions sources. This corresponds to the initial spikes in the model in the time series shown in Figure 15b~~12~~ (8:30 p.m. UTC). Again, the observed height of the peaks lies between the model peak heights for the base-case and revised-emissions simulations. For this flight the background OA concentration is under-predicted by up to 2 $\mu g/m^3$ by the end of the flight. The background air has more measured oxygenated organic aerosol (OOA) (Liggio *et al*., 2015), with an aerosol mass spectra more reflective of laboratory 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 aircraft-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 16~~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 Figure 16~~. S15~~a, we can again see the local maxima in POA difference between the runs over the Syncrude Mildred Lake facility. This corresponds to the peaks in time series at 5:50 p.m. UTC (Figure 16c~~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 Figure 16~~. S15~~b). 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

Figure 16. S15a). Note that Figure 16c . 13 suggests that background OA levels once again seem to be under-estimated in both simulations.

4.0 Discussion

In summary, tThe 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 The remaining negative bias suggests that other important processes may be missing or under-represented in the model. Three recommendations emerge from recent and current work: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.

4.1 SOA Formation from Fugative IVOC Emissions

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, only a small portion of the IVOCs (dodecane C_{12}) were added and lumped placed into the long-chain ALKA lumped species. IVOC species with carbon number ≥13 were not measured by the Li *et al*., (2017) aircraft study and thus we do not have revised IVOC emissions included in this work. Furthermore, tThe ALKA lumped

Formatted: Font: Italic

~~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, so the dodecane contribution would be underestimated. Work is currently underway with GEM-MACH to implement a Volatility Basis Set (VBS) approach to SOA formation. The VBS approach will more adequately represent the intermediate and semi-volatile volatility range and chemical aging of this lower volatility compounds (Robinson *et al.*, 2006). Future work will measure IVOC emissions using box flights around the oil sand facilities and open-pit mines. This will remove current uncertainties in models and help improve the negative bias in plumes. Implementing the VBS scheme will also enable the PM emissions used here (in both data sets) to be distributed into volatility bins.

Also, while the measurement-derived emissions are missing the IVOCs, the measurement-derived POA emissions may contain some gaseous VOCs, IVOCs and SVOC species that react quickly and in one oxidation step yield products that condense onto particles. This rapid SOA mass produced would be measured by the box flights and, at least partially, accounted for in the updated OA emissions; however labeled here as POA instead of fresh SOA. Furthermore, there is the potential for double counting if some of the very reactive gaseous precursors react to form SOA and this is accounted for in the measured POA. In this paper, we have tried to minimize this effect by examining the model performance in the "near field" from emission flights close to facilities. This will be the topic of future box modelling work with the new 2018 measurement-derived IVOC and SVOC emissions to determine how much of the measurement-derived POA is derived from the fugitive open-pit mining IVOC and SVOC emissions and their rapid particle formation.

4.2 Background Organic Aerosol Levels

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. Future work will update the biogenic SOA yield coefficients in the VBS approach using recent smog chamber results which account for gas-phase loss of organic species to chamber walls (Ma *et al.*, 2017).

Formatted: Font: Italic

4.3 Spatial Allocation of Emissions

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 truck~~their~~ emission spatial allocation within a facility. The GPS data would also be useful to define the location of freshly excavated open-pit mines within a facility.

4) Conclusions

Overall, the weight of evidence suggests that the top-down emission estimation technique applied to~~revised aircraft measurement derived organic emissions for~~ the OS surface mining facilities helps to better 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 due to the potential for fugitive emissions. For the mono- and multi-substituted aromatics (TOLU and AROM), 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 paraffinic solvents for bitumen extraction (Shell Muskeg/Jackpine and Syncrude Aurora North; Li *et al*., 2017). Overall, the predictions of alkanes in high concentration plumes improved with the revised emission data set, as shown by the 99th percentile statistic.

Formatted: Font: Italic

For PM_1 organic aerosol, the revised emissions improved the mean bias for predictions; however, a negative bias still exists and the~~is~~ improvement was associated with a decrease in correlation coefficient. The increase in predicted PM_1 OA concentration was largely due to the increase in POA emissions in the revised emissions input files. The POA emissions increased because of a combination of larger ~~aircraft~~ measurement-derived PM_1 emissions and the revised ground-observed PM speciation profile having a larger POA fraction. The increase in PM_1 POA emissions were largely allocated spatially to stack locations and this allocation may be a key factor in the degradation of the correlation coefficient, especially if the fine OA originates from mine-face fugitive emissions. Future work should focus on improving within-facility spatial allocation of emissions. The remaining negative bias in plumes likely stems from missing IVOC

emissions in both the emission data sets used here, as suggested by Liggio *et al.* (2015). Ongoing field work to measure the IVOC emissions using aircraft box flights is underway in a new measurement 2018 intensive. Upcoming modelling work with GEM-MACH will include the VBS approach to better represent lower volatility compounds.

~~A portion of this aircraft-measurement-derived POA emission increase could stem from rapid SOA formation in the interior of the aircraft flight boxes. It was recently discovered that IVOC SOA formation can be important in OA formation downwind of the OS surface mining region and the IVOC emissions came from open-pit mine fugitive emissions. GEM-MACH does not currently include the IVOC emissions from open-pit mines, as the estimation of the IVOC emission rate and SOA aging rate are the subject of ongoing Lagrangian box model studies. Furthermore, the increase in PM_1 POA emissions were largely allocated spatially to stack locations and this allocation may be a key factor in the degradation of the correlation coefficient, especially if the fine OA originates from mine-face fugitive emissions. Future work should focus on improving within-facility spatial allocation of emissions.~~

Acknowledgements

The authors are grateful to all of the participants in the 2013 JOSM intensive field study for their commitment. The authors are also appreciative of the ECCC Pollutant Inventory and Reporting Division (PIRD) and the U.S. EPA for developing, maintaining, and distributing each country's national emission inventories. We also appreciate the efforts of George Marson of ECCC in helping to compile the various emissions inventories from Alberta Environment and Parks, and also CEMA. We also appreciate the analysis of the NAPS VOC measurement group. This study was funded by the Joint Oil Sands Monitoring program and the Climate Change and Air Quality Program.

References

Akingunola, A., P. A. Makar, J. Zhang, A. Darlington, S.M. Li, M. Gordon, M.D. Moran, Q. Zheng, Evaluation of GEM-MACH Air Quality Modelling at 2.5km Resolution Using JOSM 2013 Intensive Campaign: Impact of Continuous Monitoring Emissions Stack Parameters on Model Simulations, accepted in~~submitted to~~ ACP~~D~~, 2018.

Barsanti, K.C., Carlton, A.G., Chung, S.H., Analyzing experimental data and model parameters: Implications for predictions of SOA using chemical transport models, (2013) Atmospheric Chemistry and Physics, 13 (23), pp. 12073-12088.

Cappa, C.D, J. Jimenez, Quantitative estimates of the volatility of ambient organic aerosol, *Atmos. Chem. Physics*, 10, 12, 2010, pp 5409-5424.

Cappa, C.D., Wilson, K.R., Evolution of organic aerosol mass spectra upon heating: Implications for OA phase and partitioning behavior, (2011) *Atmospheric Chemistry and Physics*, 11 (5), pp. 1895-1911.

Chai, T., Kim, H.-C., Lee, P., Tong, D., Pan, L., Tang, Y., Huang, J., McQueen, J., Tsidulko, M., Stajner, I. Evaluation of the United States National Air Quality Forecast Capability experimental real-time predictions in 2010 using Air Quality System ozone and NO2 measurements, (2013) *Geoscientific Model Development*, 6 (5), pp. 1831-1850.

Cho, S., McEachern, P., Morris, R., Shah, T., Johnson, J., Nopmongcol, U., Emission sources sensitivity study for ground-level ozone and PM 2.5 due to oil sands development using air quality modeling system: Part I- model evaluation for current year base case simulation, (2012) *Atmospheric Environment*, 55, pp. 533-541.

Cohan, D.S.; Napelenok, S.L. Air Quality Response Modeling for Decision Support. *Atmosphere*, **2011**, *2*, 407-425.

Côté, J., Desmarais, J.-G., Gravel, S., *et al*., The operational CMC/MRB global environmental multiscale (GEM) model. Part 1: design considerations and formulation, (1998a), *Mon. Wea. Rev*. 126, 1373-1395.

Côté, J., Desmarais, J.-G., Gravel, S., *et al*., The operational CMC-MRB global environment multiscale (GEM) model. Part II: results, (1998b), *Mon. Wea. Rev*. 126, 1397-1418.

Dann, T.F., Wang, D.K., Ambient air benzene concentrations in Canada (1989-1993): Seasonal and day of week variations, trends, and source influences, (1995) Journal of the Air and Waste Management Association, 45 (9), pp. 695-702.

Davies, M., Person, R., Nopmongcol, U., Shah T., Vijayaraghavan, K., Morris, R., and Picard, D.: Lower Athabasca Region Source and Emission Inventory, report prepared by Stantec Consulting Ltd. and ENVIRON International Corporation for Cumulative Environmental Management Association - Air Working Group, http://library.cemaonline.ca/ckan/dataset/0cfaa447-410a-4339-b51f-e64871390efe/resource/fba8a3b0-72df-45ed-bf12-8ca254fdd5b1/download/larsourceandemissionsinventory.pdf, 274 pp., 2012 (last accessed on October 24, 2017).

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.

Government of Canada, Notice with respect to the substances in the National Pollutant Release Inventory for 2018 and 2019, Canada Gazette Part I, Vol. 152, No. 3, pp. 129-172, ISSN 1494-6076, Ottawa, January 20, 2018.

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.

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

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.

Makar, P.A. *et al*., Estimates of Exceedances of Critical Loads for Acidifying Deposition in Alberta and Sasketchwan, accepted in~~submitted to~~ ACP~~D~~, 2018~~7~~.

Mashayekhi, R., Zhao, S., Saeednooran, S., Hakami, A., Ménard, R., Moran, M. D., and Zhang, J.: Emissions Uncertainty Inventory and Modeling Framework: Case Study of Residential Wood Combustion, 15th Annual CMAS Conference, October 24-26, Chapel Hill, NC, https://www.cmascenter.org/conference//2016/slides/mashayekhi_development_emission_2016.pptx, 2016.

McNeill, V.F., Aqueous organic chemistry in the atmosphere: Sources and chemical processing of organic aerosols (2015) Environmental Science and Technology, 49 (3), pp. 1237-1244.

Moran, M.D., Ménard, S., Pavlovic, R., Anselmo, D., Antonopoulos, S., Makar, P.A., Gong, W., Gravel, S., Stroud, C., Zhang, J., Zheng, Q., Robichaud, A., Landry, H., Beaulieu, P.A., Gilbert, S., Chen, J., Kallaur, A., Recent Advances in Canada's National Operational AQ Forecasting System, (2013) NATO Science for Peace and Security Series C: Environmental Security, 137, pp. 215-220.

Office of the Federal Register National Archives and Records Administration, Protection of Environment, Code of Federal Regulations, Title 40, Parts 50 to 51, Special Edition of the Federal Register, U.S. government publishing office, Washington, DC 20402–0001, July 1, 2015.

Pankow, J.F., An absorption model of the gas/aerosol partitioning involved in the formation of secondary organic aerosol, (1994) Atmospheric Environment, 28 (2), pp. 189-193.

Park, S.H., Gong, S.L., Bouchet, V.S., Gong, W., Makar, P.A., Moran, M.D., Stroud, C.A., Zhang, J. Effects of black carbon aging on air quality predictions direct radiative forcing estimation, (2011) Tellus, Series B: Chemical and Physical Meteorology, 63 (5), pp. 1026-1039.

Pouliot, G., Pierce, T., Denier van der Gon, H., Schaap, M., Moran, M., Nopmongcol, U., Comparing emission inventories and model-ready emission datasets between Europe and North America for the AQMEII project, (2012), Atmospheric Environment, 53, pp. 4-14.

Pouliot, G., Denier van der Gon, H.A.C., Kuenen, J., Zhang, J., Moran, M.D., Makar, P.A., Analysis of the emission inventories and model-ready emission datasets of Europe and North America for phase 2 of the AQMEII project (2015) Atmospheric Environment, 115, pp. 345-360.

Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop, A.P., Lane, T.E., Pierce, J.R., Pandis, S.N., Rethinking organic aerosols: Semivolatile emissions and photochemical aging, (2007) Science, 315 (5816), pp. 1259-1262.

Rouleau, M., Egyed, M., Taylor, B., Chen, J., Samaali, M., Davignon, D., Morneau, G., Human health impacts of biodiesel use in on-road heavy duty diesel vehicles in Canada, (2013) Environmental Science and Technology, 47 (22), pp. 13113-13121.

Schultz, M.G., Diehl, T., Brasseur, G.P., Zittel, W., Air Pollution and Climate-Forcing Impacts of a Global Hydrogen Economy, (2003) Science, 302 (5645), pp. 624-627.

Seinfeld, J. H. and Pandis, S. N. (1998). Atmospheric Chemistry and Physics from air pollution to climate change. New York. John Wiley and Sons, Incorporated.

Shrivastava, M., Easter, R.C., Liu, X., Zelenyuk, A., Singh, B., Zhang, K., Ma, P.-L., Chand, D., Ghan, S., Jimenez, J.L., Zhang, Q., Fast, J., Rasch, P.J., Tiitta, P., Global transformation and fate of SOA: Implications of low-volatility SOA and gas-phase fragmentation reactions, (2015) Journal of Geophysical Research, 120 (9), pp. 4169-4195.

Slowik, J.G., Stroud, C., Bottenheim, J.W., Brickell, P.C., Chang, R.Y.-W., Liggio, J., Makar, P.A., Martin, R.V., Moran, M.D., Shantz, N.C., Sjostedt, S.J., Van Donkelaar, A., Vlasenko, A., Wiebe, H.A., Xia, A.G., Zhang, J., Leaitch, W.R., Abbatt, J.P.D., Characterization of a large biogenic secondary organic aerosol event from eastern Canadian forests, (2010) Atmospheric Chemistry and Physics, 10 (6), pp. 2825-2845.

Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M.D., Wyat Appel, K., Bessagnet, B., 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, accepted in~~submitted to~~ ACP~~D~~, 2018~~7~~.

Table 1. Facility total emission rates for three lumped organic species and $PM_{2.5}$ calculated with the bottom-up, base case inventory, CEMA facility-specific VOC profiles (labeled Base Case) and the top-down measurement-derived rates (labeled Revised Emission case, scaled to tonnes/year for VOCs or tonnes/Aug&Sept for $PM_{2.5}$). Emission rate increase/decrease of more than ±500 tonnes compared to base case is shown in red/blue.

	Suncor – M/S		Syncrude - ML		Shell – MR/J		CNRL - Horizon	
Species	**Base Case**	**Revised**	**Base Case**	**Revised**	**Base Case**	**Revised**	**Base Case**	**Revised**
Mono-Substituted Aromatics (TOLU)	486	1112	806	1539	6.8	72	135	393
Multi-Substituted Aromatics (AROM)	1457	1569	5273	1696	746	88	1125	500
Long Chain Alkanes (ALKA)	5636	13488	12348	10022	1690	14384	2651	23779
Particulate Matter ($PM_{2.5}$)	1251	2537*	1021	3648*	459	2423*	402	1015*

* based on 2-month emission (Aug&Sept) rather than based on annual estimate (Zhang *et al.*, 2018)

Formatted Table

Formatted: Subscript

Formatted: Font: Italic

Table 2. Facility-specific VOC speciation profiles (mass fractions) applied to the surface mining facilities in the Athabasca oil sands region compared to standard speciation profiles for Canadian and U.S. petrochemical oil refineries (in ADOM-II chemical speciation). Data are based on Zhang *et al.* (2018) and references therein. All four profiles are used in the base case simulation.

Species	Shell M/J, Syncrude AN, Imperial Kearl Base Case Plant Profile (CEMA)	Syncrude ML, Suncor, CNRL Base Case Plant Profile (CEMA)	CEPS Database Standard Profile #9012 For Oil Refineries in Base Case	SPECIATE Database Standard Profile #0316 For Oil Refineries in Base Case
EC38 (Propane, Benzene, Acetylene)	0.0	0.0	0.247	0.176
EA3 (Alkane ≥C4)	0.90	0.71	0.623	0.781
EA2 (Alkene ≥C3)	0.007	0.069	0.031	0.002
ETOL (Toluene and other mono-aromatics)	0.001	0.057	0.005	0.008
EARO (Multi-functional aromatics)	0.0003	0.099	0.003	0.003
EHCO (Formaldehyde)	0.00001	0.0003	0.110	0.0

Columns do not add up to unity due to "unaccounted for" or "unassigned species" and/or due to consideration of reactivity weighting for the ADOM-II mechanism.

Refinery Profile #9012 is a profile from the Canadian Emissions Processing System (Moran, M.D., M.T. Scholtz, C.F. Slama, A. Dorkalam, A. Taylor, N.S. Ting, D. Davies, P.A. Makar, S. Venkatesh, An Overview of CEPS1.0: Version 1.0 of the Canadian Emissions Processing System for Regional-Scale Air Quality Models. In Proc. 7th AWMA Emission Inventory Symp., Research Triangle Park, North Carolina, Air & Waste Management Association, Pittsburgh, Oct. 28-30, 1997.)

Formatted Table

Table 3~~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.

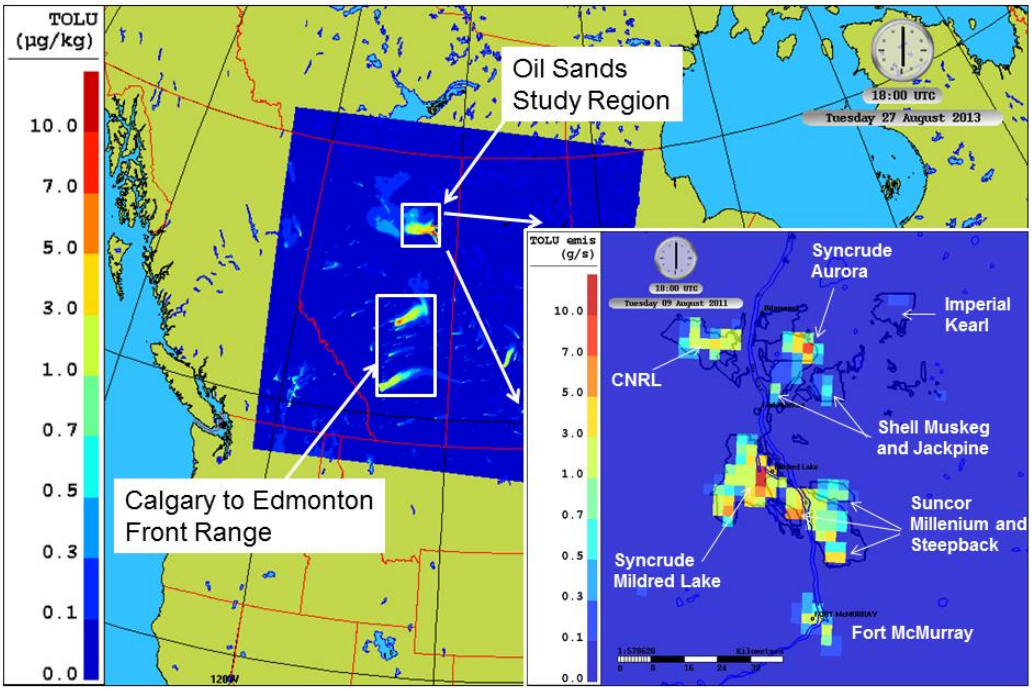


Figure 1. The background image is the nested domain, at 2.5-km grid spacing, covering all of Alberta and Saskatchewan and encompassing the Athabasca Oil Sand study region (white box). The model field shown is for the lumped toluene species (TOLU) mass mixing ratio (µg/kg air). The inserted image on the right is the TOLU emission map (g/s/grid cell) for the Oil Sands study region at the same hour as mixing ratio image on the left. The Oil Sand facility's names are listed in white labels.

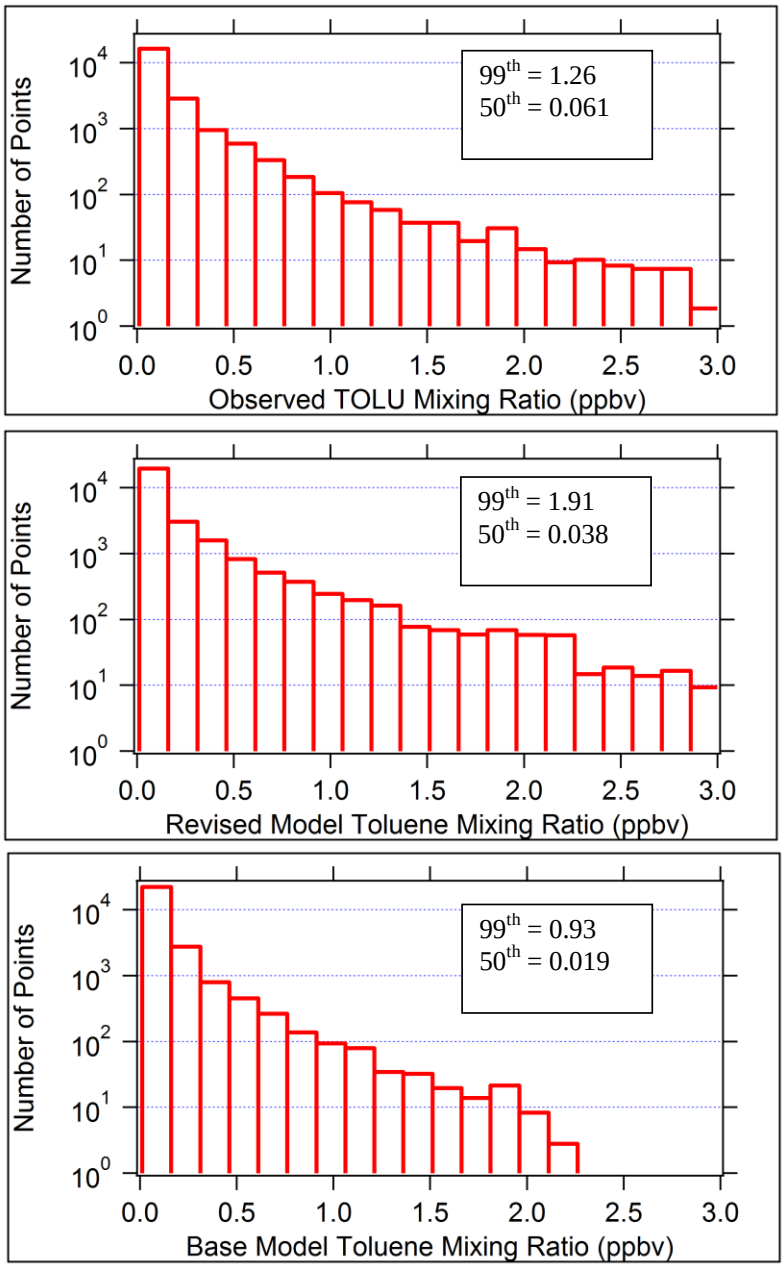


Figure 2abc. Histograms for (a) observed TOLU, (b) revised-emissions TOLU, and (c) base-case-emissions TOLU volume mixing ratios (ppbv). Points correspond to 10-sec averaged aircraft and model data, sorted into 20 bins by volume mixing ratio. The inset boxes show the 50th and 99th percentile values for each histogram.

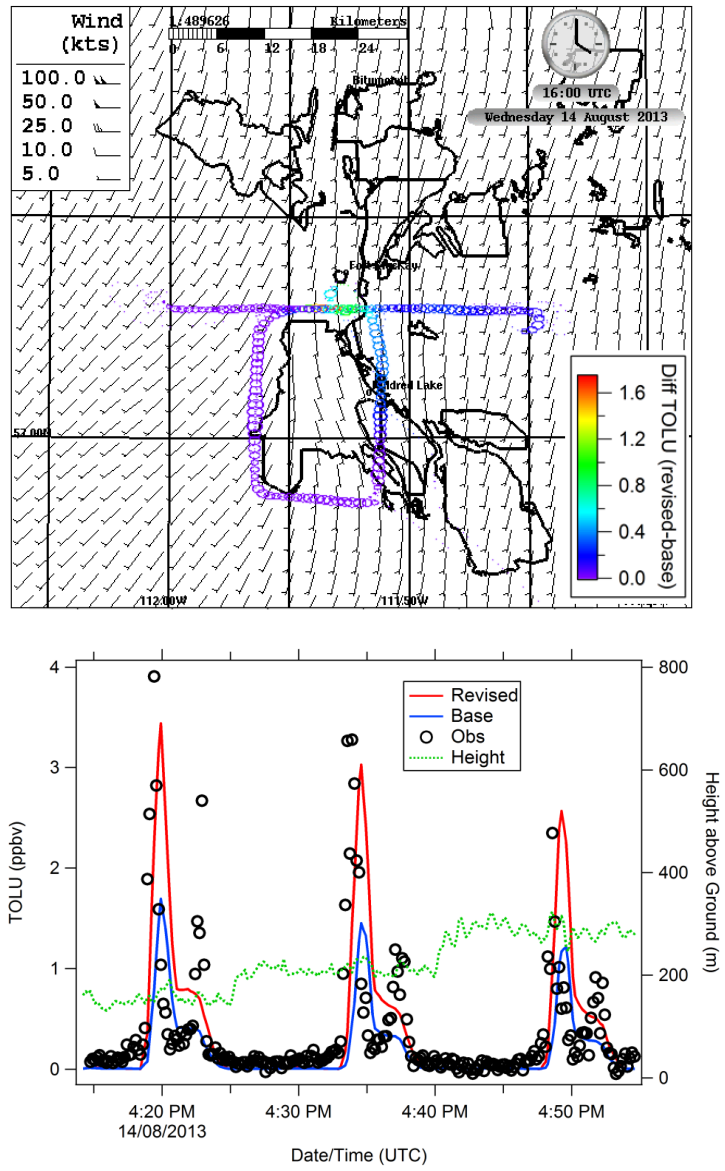


Figure 3ab. Flight track of the aircraft on Aug. 14, 2013 around the Syncrude Mildred Lake facility color-coded by the difference in TOLU volume mixing ratio (pptv) between the revised-emissions and the base-case simulations. The modelled wind barbs at the time of the maximum difference (16 UTC) are included in the background map. Panel b is the time series of observed and model-predicted TOLU volume mixing ratios for the flight on Aug. 14, 2013. The highest magnitude points correspond to a location north of the facility sampled at 3 different altitudes.

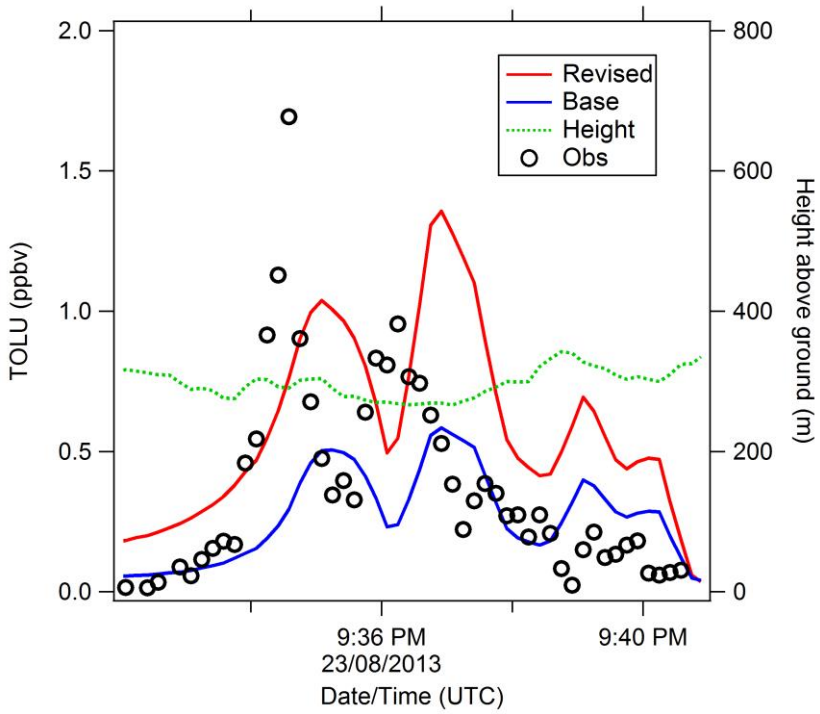


Figure 4. TOLU volume-mixing-ratio time series for a flight on Aug. 23, 2013 over the Suncor Millennium/Steepbank facility, just east of the Athabasca River, on a survey pattern.

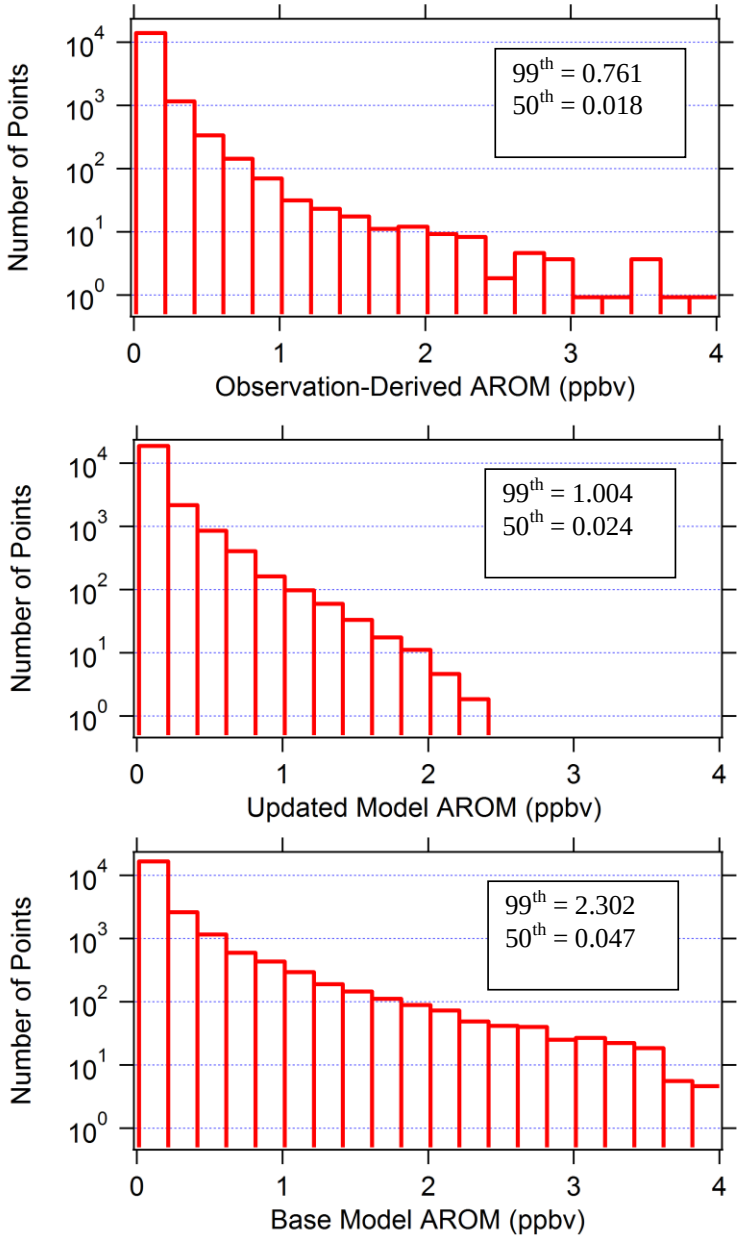


Figure 5. Histograms for (a) observed AROM, (b) revised-emissions AROM, and (c) base model AROM volume mixing ratios (ppbv). Points correspond to 10-sec averaged aircraft and model data, sorted into 20 bins by volume mixing ratio. The inset boxes show the 50^{th} and 99^{th} percentile values for each histogram.

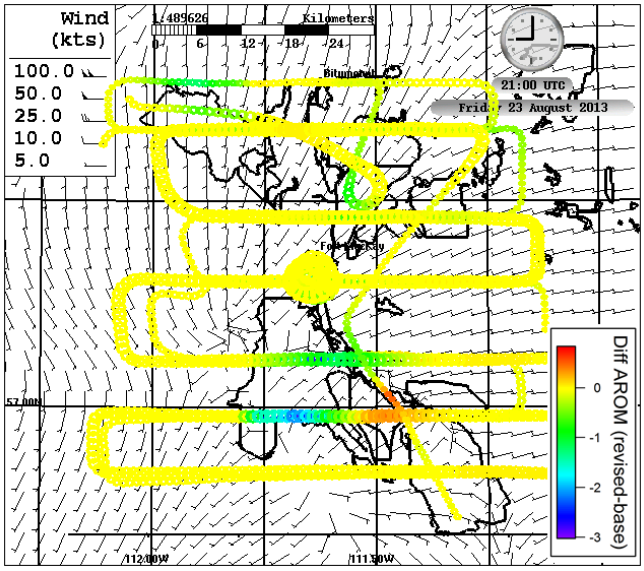


Figure 6. Flight track of the aircraft on Aug. 23, 2013 over all six OS surface mines color-coded by the difference in predicted AROM volume mixing ratio (pptv) between the revised-emissions and the base-case simulations. The modelled wind barbs at the time of the maximum difference (21 UTC) are included in the background map.

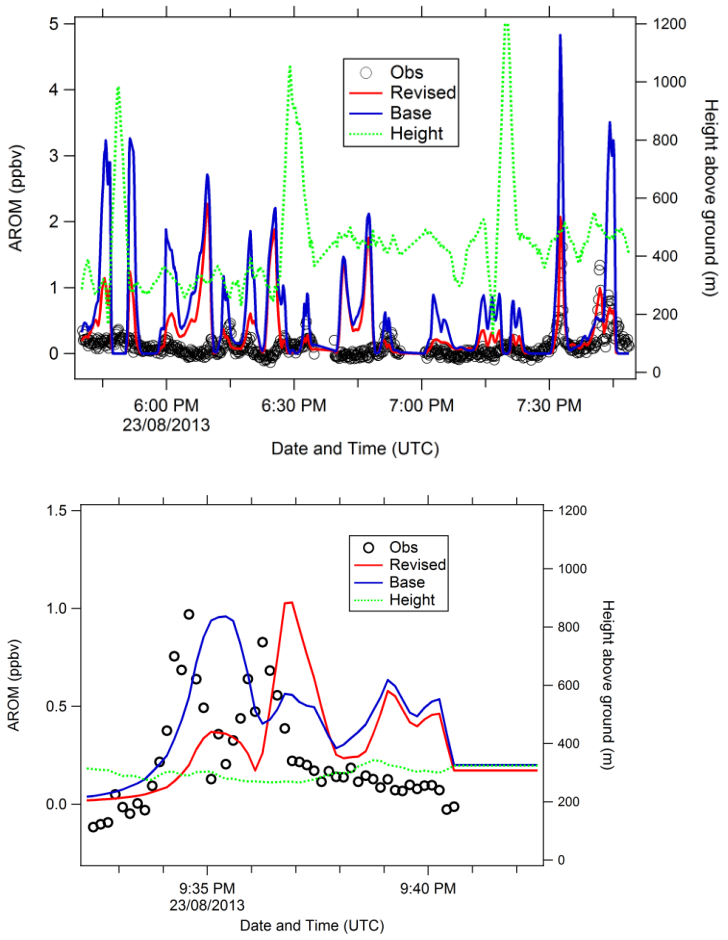


Figure 7ab. Time series of observed and model-predicted AROM volume mixing ratios for the Aug. 23 survey flight. The mixing-ratio peaks in panel A are over the Syncrude Mildred Lake facility (7:30-7:45 p.m. UTC). The 2nd peak in panel B is over the Suncor Millennium/Steepbank facility (9:37 p.m. UTC).

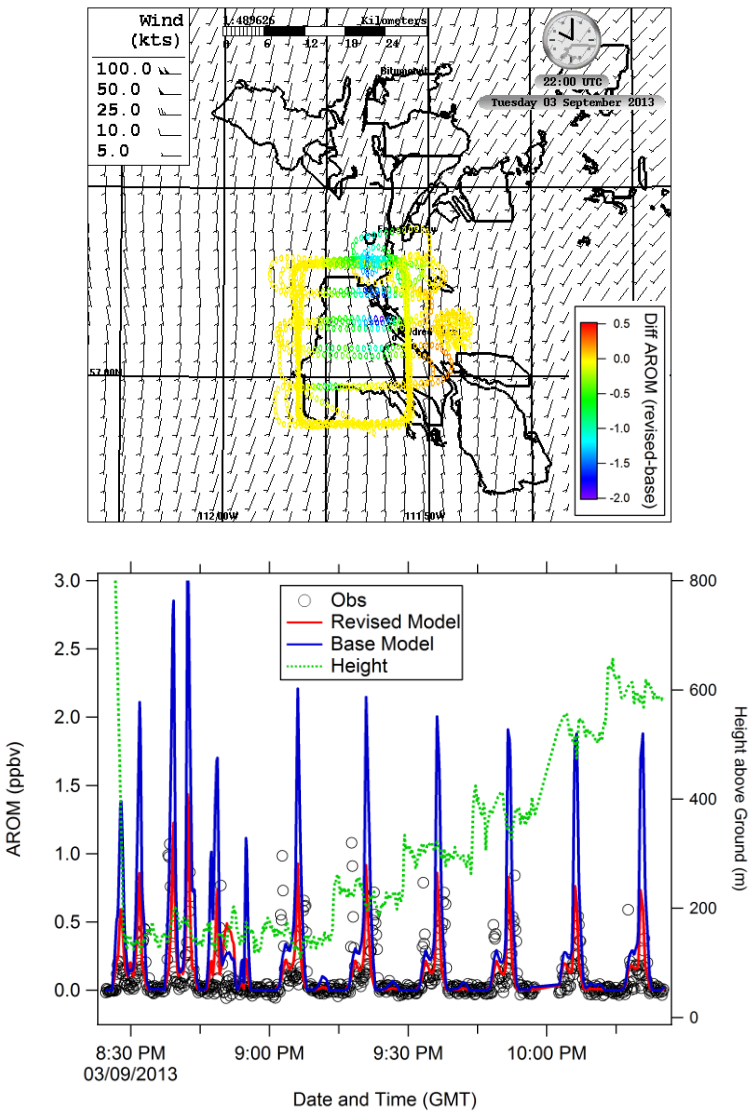


Figure 8ab. The top panel is the flight track of the aircraft on Sept. 3, 2013 over the Syncrude Mildred Lake facility color-coded by the difference in predicted AROM volume mixing ratio (pptv) between the revised-emissions and the base-case simulations. The modelled wind barbs at the time of the maximum difference (22 UTC) are included in the background map. The bottom panel is the AROM volume-mixing-ratio time series for the Sept. 3 flight around and over the Syncrude Mildred Lake facility.

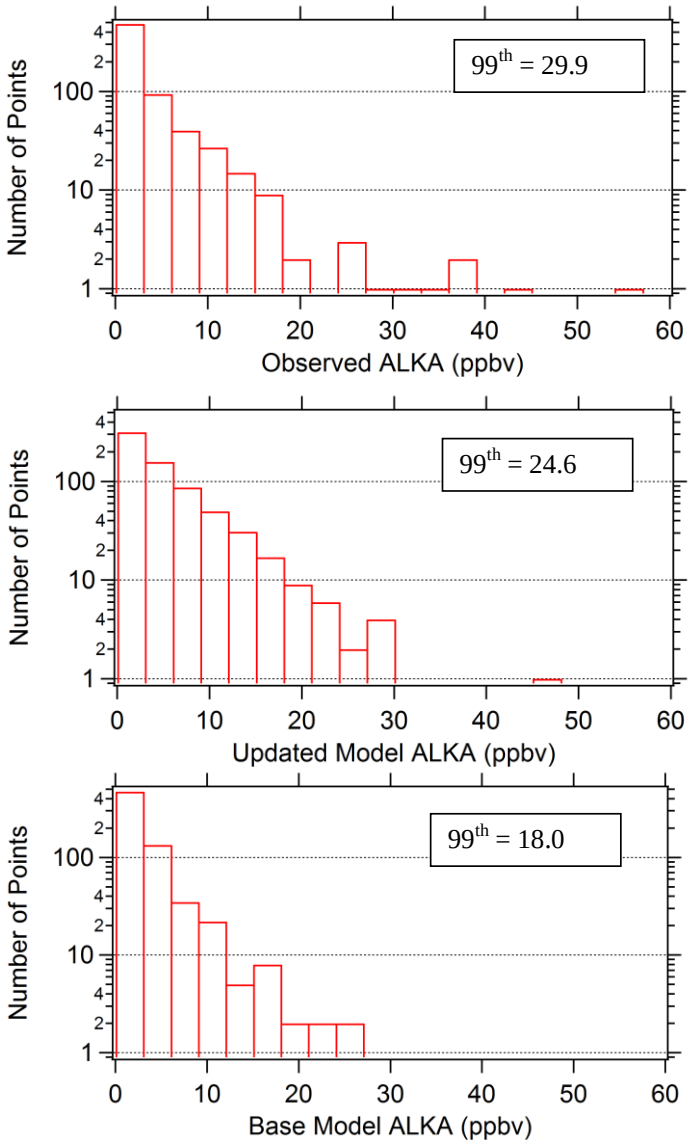


Figure 9. Histograms for (a) observed ALKA, (b) revised-emissions ALKA, and (c) base-case emissions ALKA volume mixing ratios (ppbv). Points correspond to canister grab samples and model data, sorted into 20 bins by mixing ratio. The inset boxes show the 99th percentile value for each histogram.

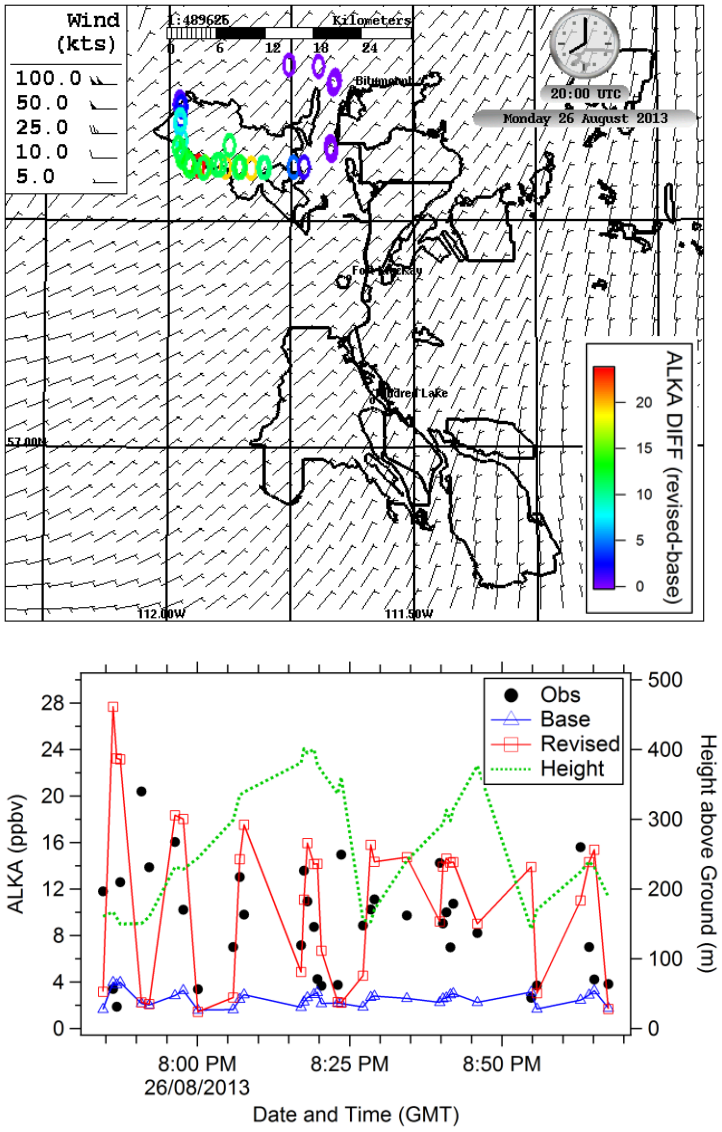


Figure 10ab. Flight track of the aircraft on Aug. 26, 2013 over the CNRL Horizon facility color-coded by the difference in predicted ALKA volume mixing ratio (pptv) between the revised-emissions and the base-case simulations. The modelled wind barbs at the time of the maximum difference (20 UTC) are included in the background map. The second panel is the ALKA volume-mixing-ratio time series for the Aug. 26 flight around the CNRL Horizon facility.

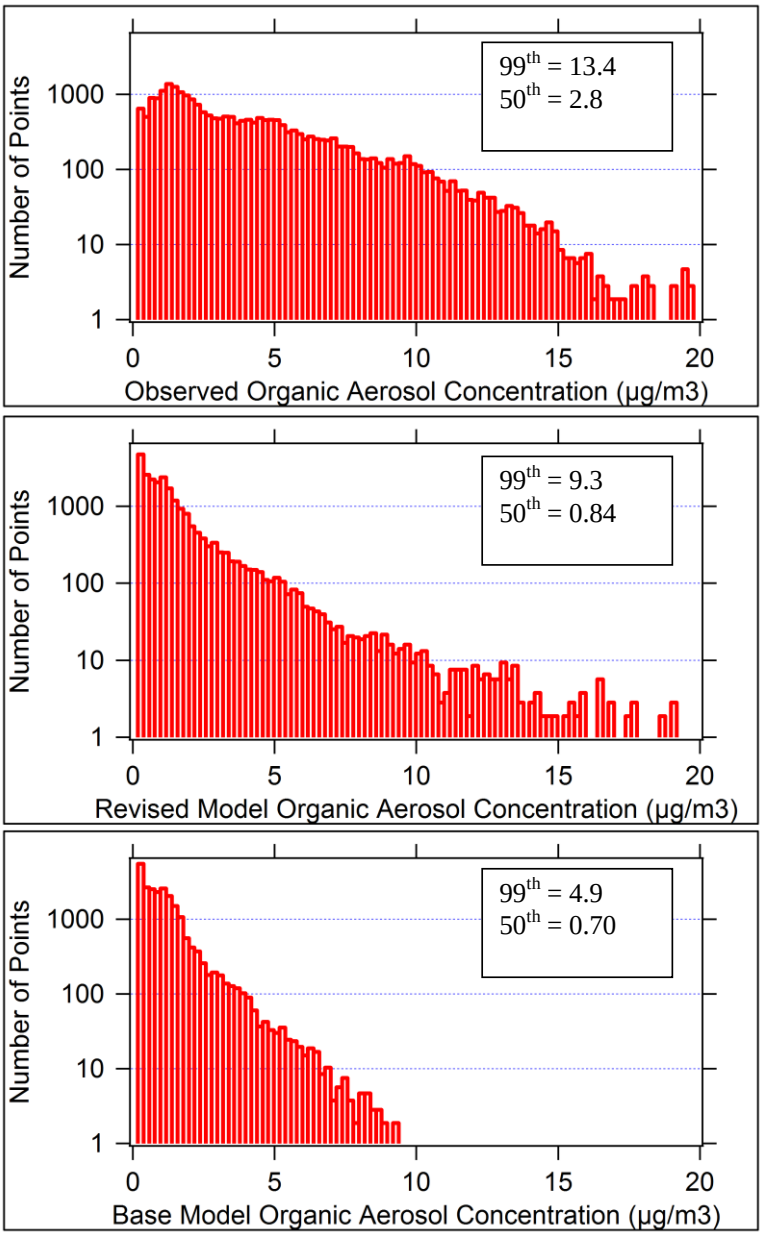


Figure 11. Histograms for (a) observed organic aerosol (OA), (b) revised-emissions OA, and (c) base-case emissions OA concentrations (µg/m^3). Points correspond to 10-sec averaged aircraft and model data. The inset boxes show the 50th and 99th percentile values for each histogram.

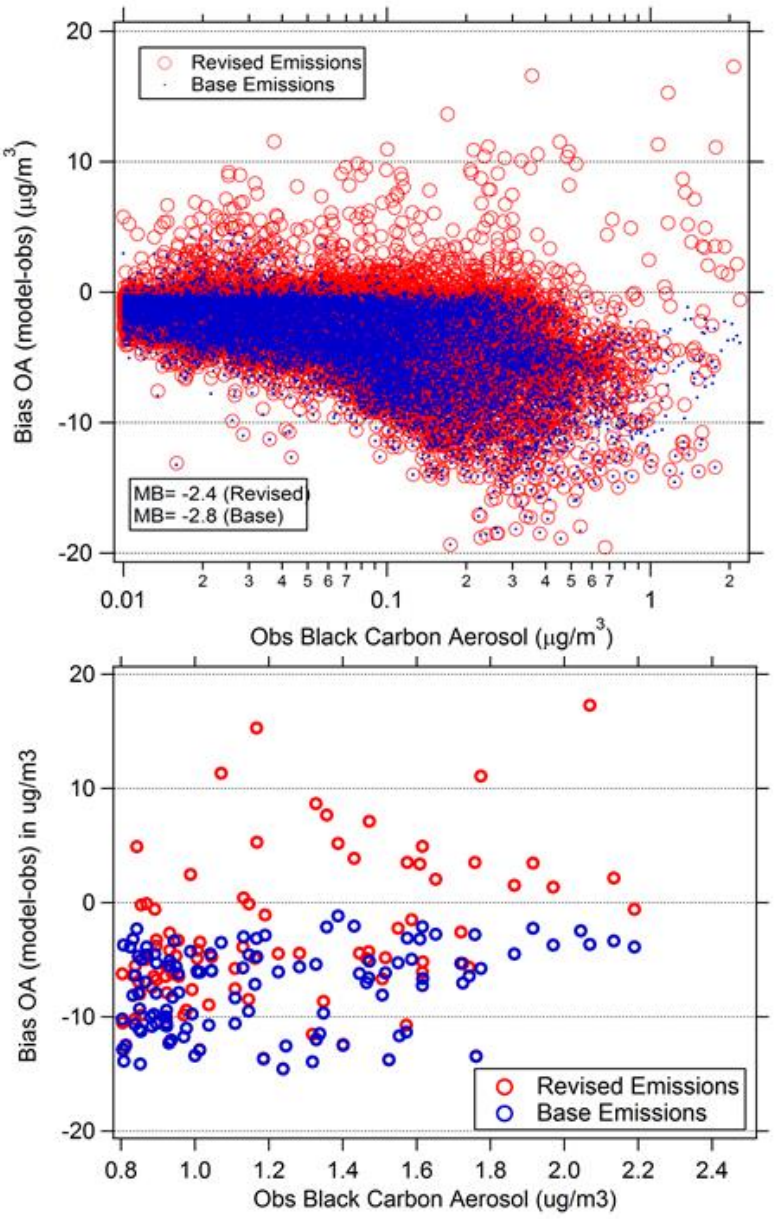


Figure 12ab. Organic aerosol model bias as a function of observed black carbon aerosol. The bottom panel is an enlargement of the upper panel showing only the data points for observed BC>0.8 µg/m^3. The model results for the base-case emissions run are plotted in blue and points in red correspond to the revised-emissions run. The data plotted is for all the aircraft flights.

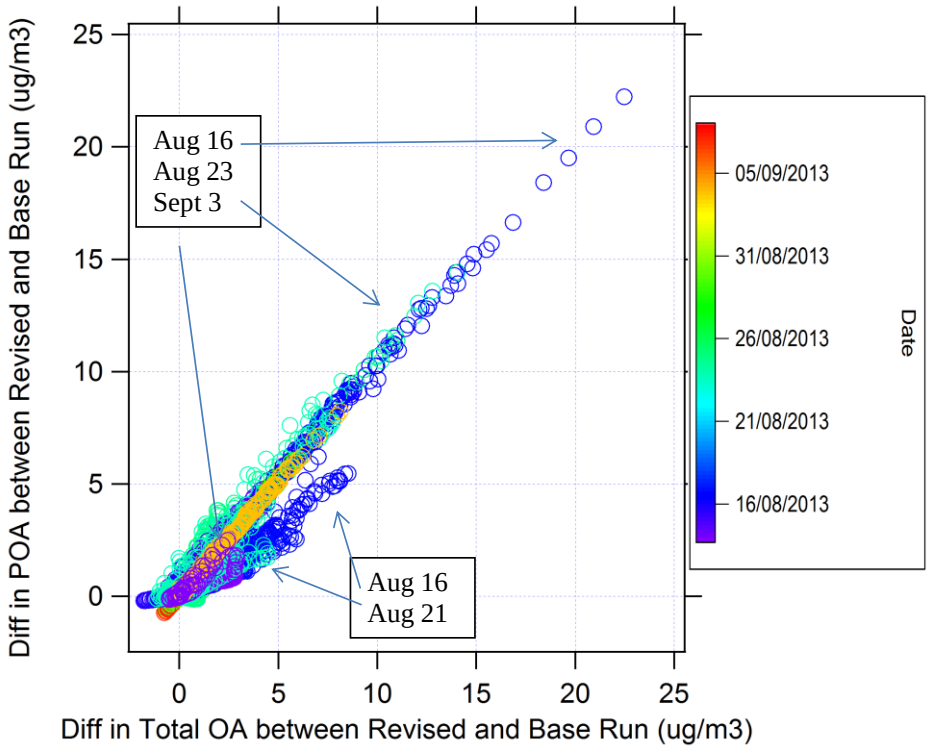


Figure 13. Difference in predicted POA concentrations between revised-emissions and base-case runs plotted as a function of the difference in predicted total OA concentration between the revised-emissions and base-case runs for all flights. Points along the 1:1 line show a difference solely from POA emission changes. Points below the 1:1 line show enhanced SOA formation.

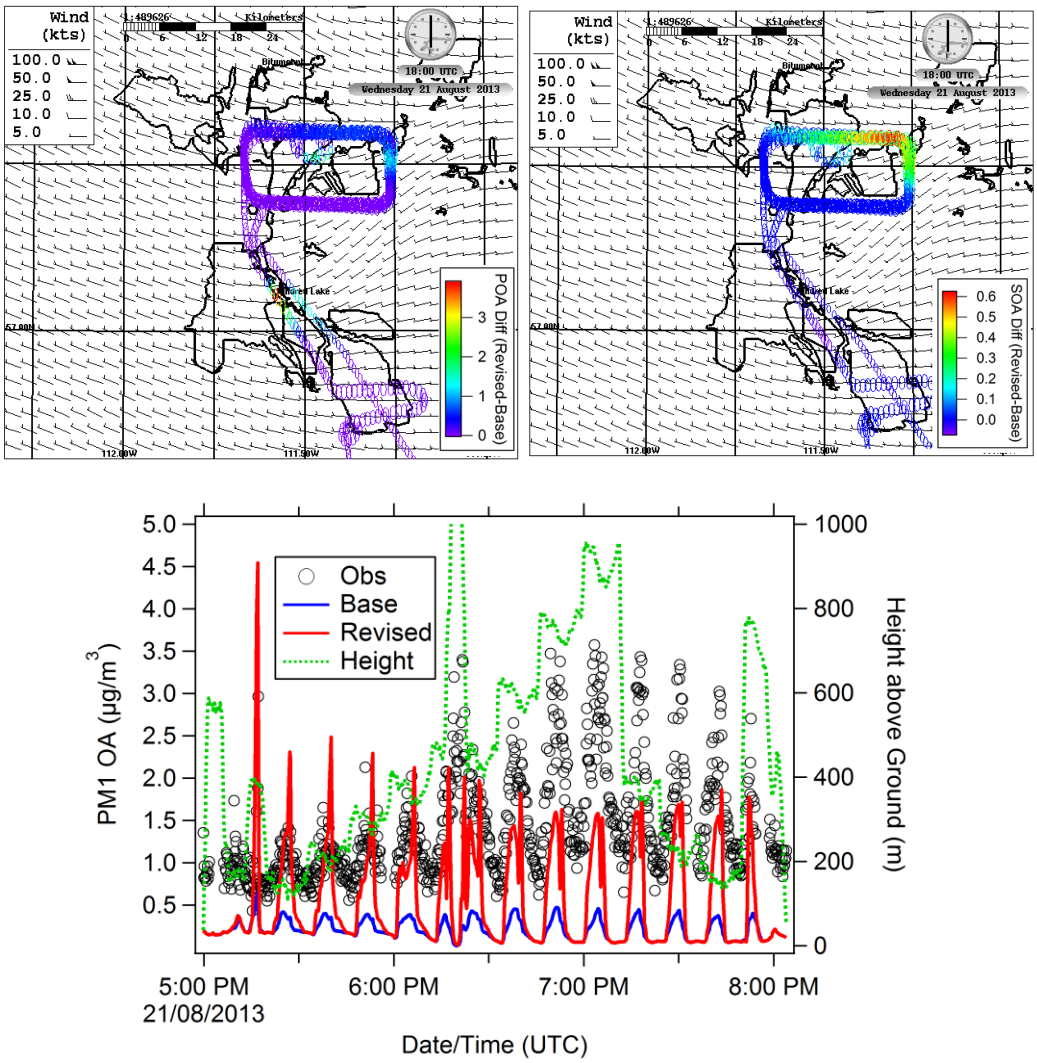


Figure 14abc. Flight track of the aircraft on Aug. 21, 2013 over the Shell Muskeg/Jackpine facility color-coded by the difference in predicted (a) POA concentration (µg/m^3) and (b) SOA concentration between the revised-emissions and the base-case-emissions simulations. The bottom panel is the time series 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.

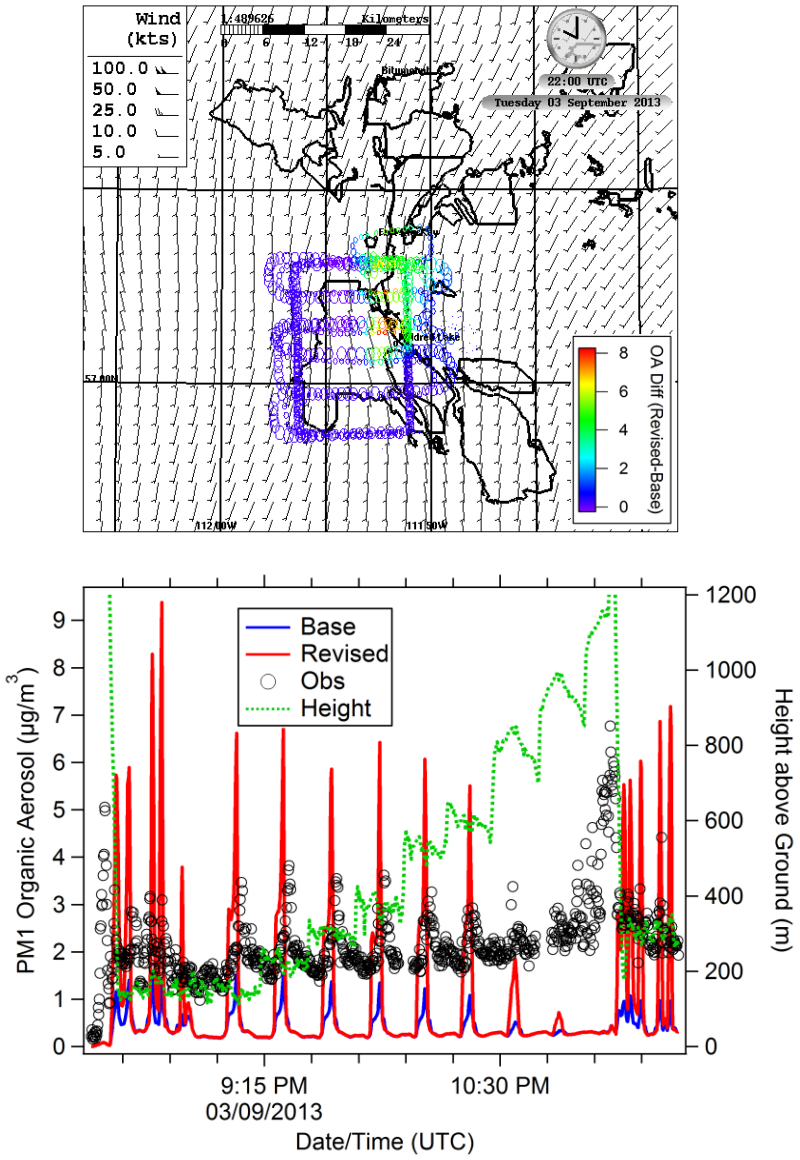


Figure 15ab. Flight track of the aircraft on Sept. 3, 2013 over the Syncrude Mildred Lake facility color-coded by the difference in predicted organic aerosol (OA) concentration (µg/m^3) between the revised-emissions and the base-case-emissions simulations. The bottom panel is the time series for PM_1 organic aerosol concentration (µg/m^3) for the Sept. 3 flight over and around the Syncrude Mildred Lake facility.

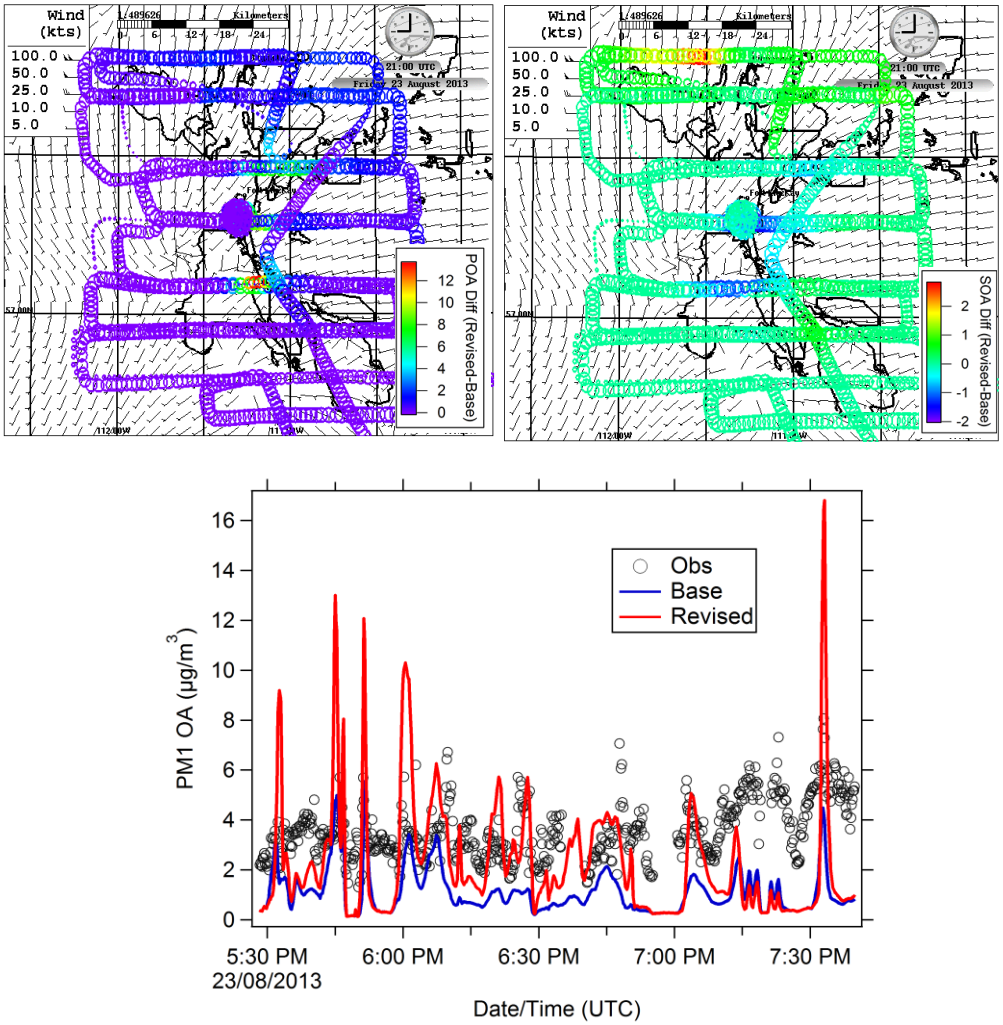


Figure 16abc. Flight track of the aircraft on Aug. 23, 2013 over all six OS surface mines color-coded by the difference in predicted (a) POA concentration (µg/m^3) and (b) SOA concentration between the revised-emissions and the base-case-emissions simulations. The bottom panel is the time series for PM_1 organic aerosol concentration (µg/m^3) for the Aug. 23 survey flight.

Supplementary Information

Improved 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)

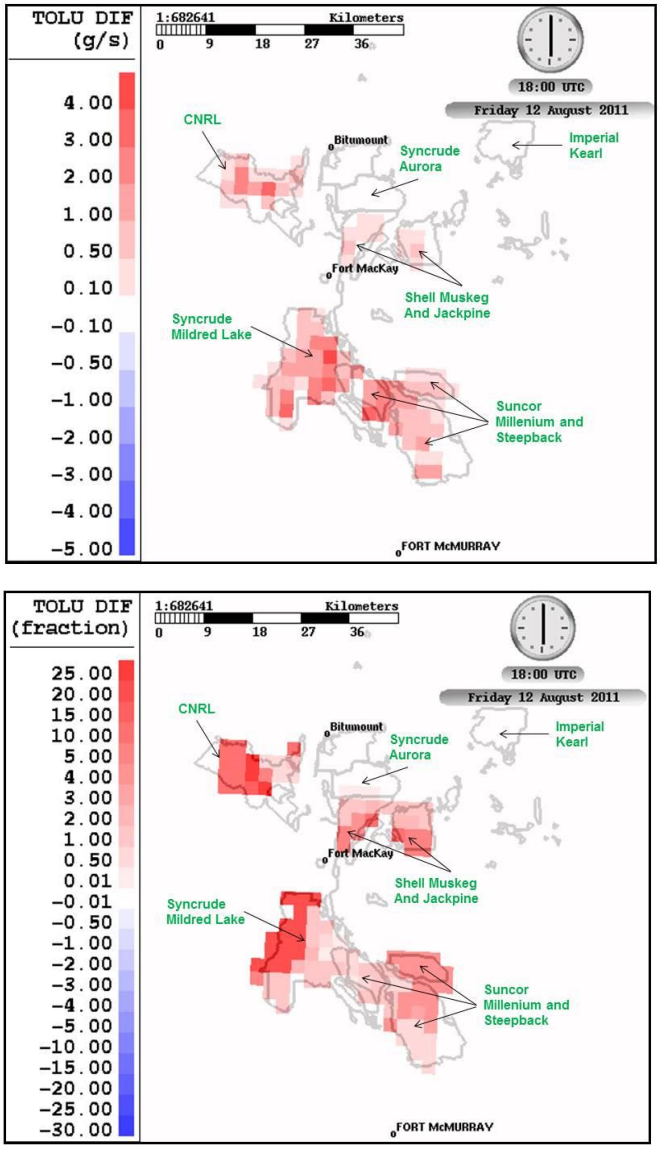
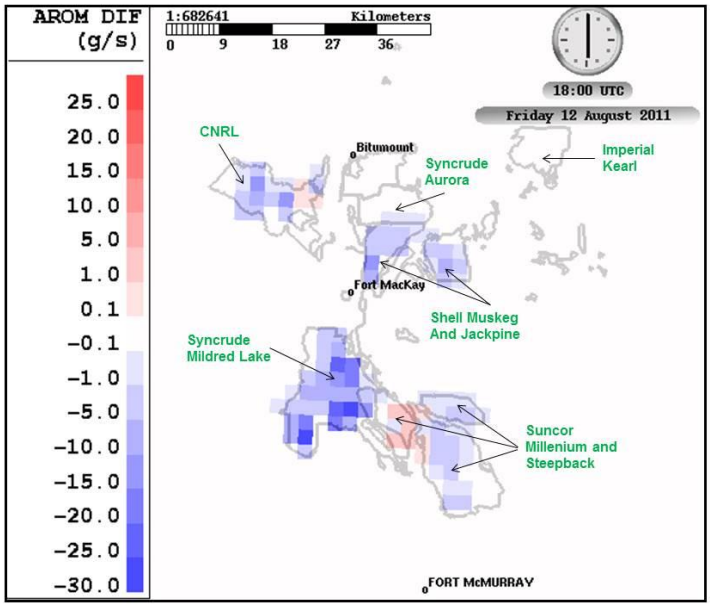


Figure S1. (a) Difference in lumped TOLU emissions (revised-base case) in units of grams/sec for each 2.5-km x 2.5-km grid cell; (b) relative difference calculated as (revised–base)/base.



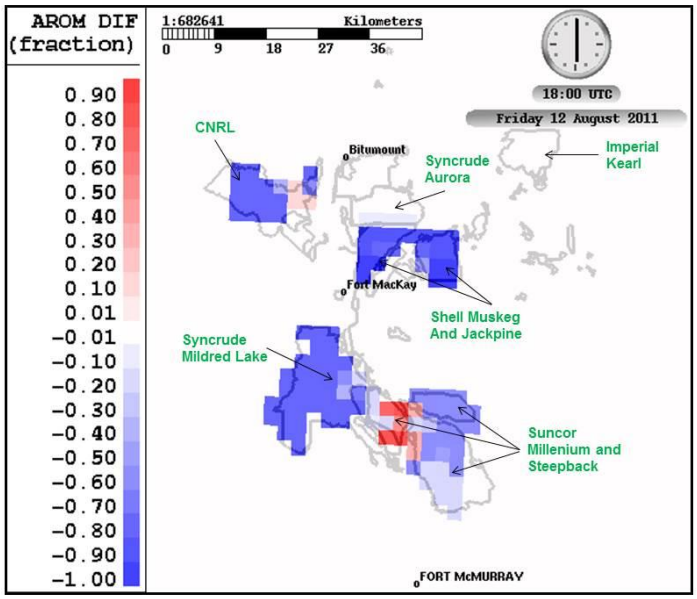
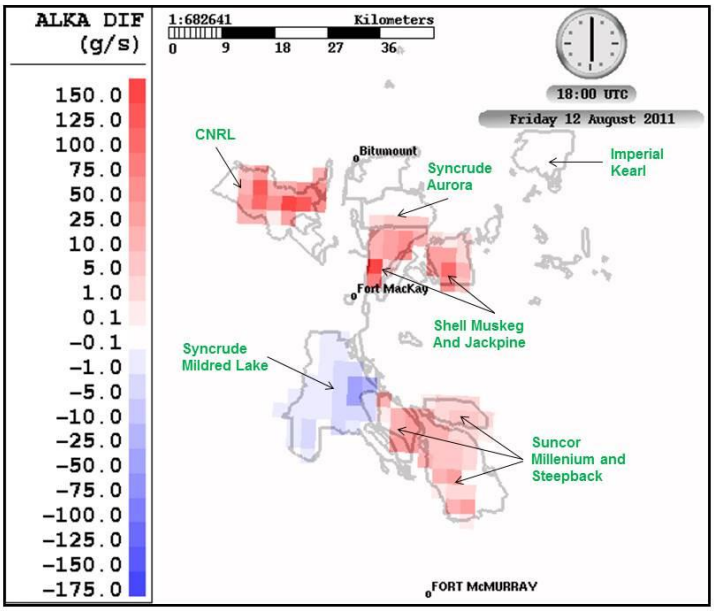


Figure S2. (a) Difference in AROM emissions (revised-base case) in units of grams/sec for a selected date and time; (b) relative difference calculated as (revised–base)/base. Large negative changes are noted over the Syncrude Mildred Lake facility.



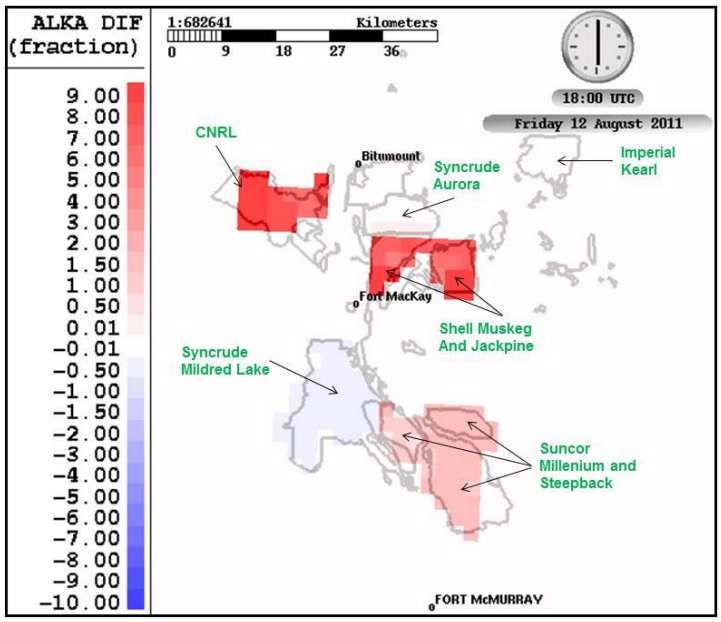
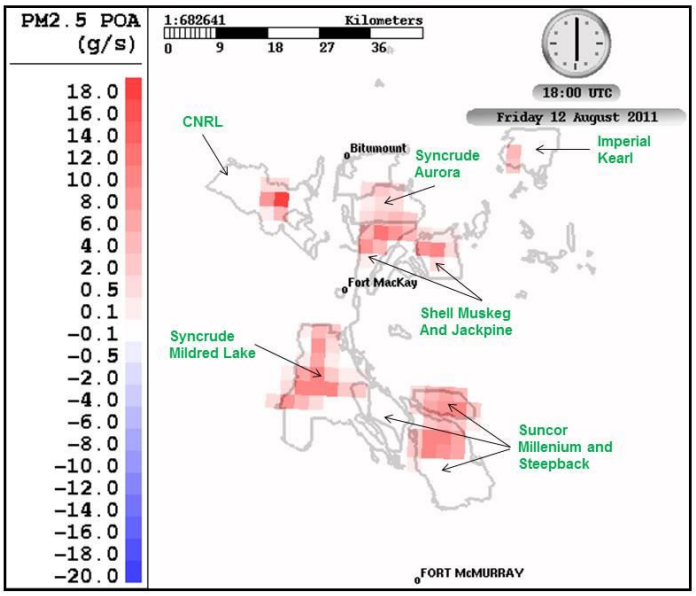


Figure S3. (a) Difference in lumped ALKA emissions (revised-base case) in units of grams/sec for each 2.5-km x 2.5-km grid cell; (b) relative difference calculated as (revised–base)/base.



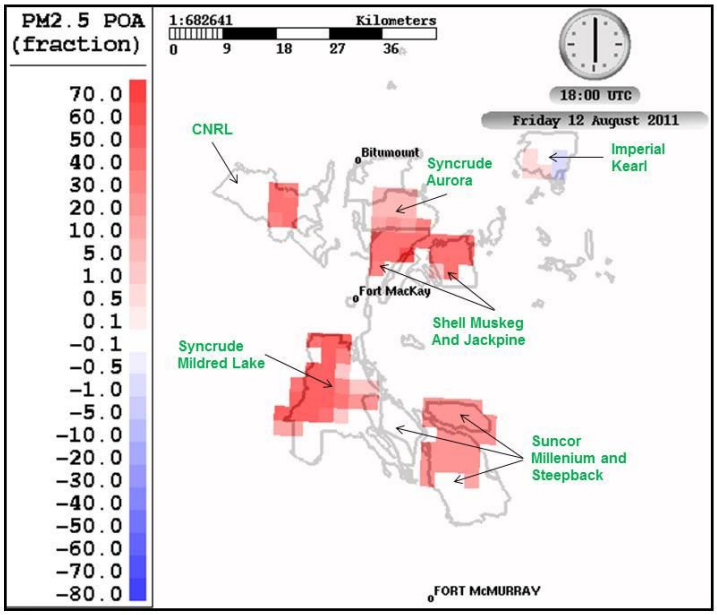


Figure S4. (a) Difference in primary organic aerosol (POA) emissions (revised-base case) in units of grams/sec for each 2.5-km x 2.5-km grid cell; (b) relative difference calculated as (revised–base)/base.