

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 Kearn 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 \geq C4)	0.90	0.71	0.623	0.781
EA2 (Alkene \geq 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.

Species	Suncor – M/S		Syn crude - ML		Shell – MR/J		CNRL - Horizon	
	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, $m=0.376\pm 0.006$ (for ethyl benzene vs. toluene) and $m=0.0652\pm 0.0008$ (for propyl benzene vs toluene). We have also added the y-intercepts and their standard error, $b=0.033\pm 0.006$ ppbv (for ethyl benzene vs toluene) and $b=0.0011\pm 0.0008$ 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 SO₂ and NO_x 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 ECCO 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 SO₂ and NO_x 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 PM₁ 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 PM₁ 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 C₁₂) were available in the VOC speciation – these were lumped into the long-chain ALKA lumped species. IVOC species with carbon number ≥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 PM₁ 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 PM₁ 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 PM₁ emissions and the revised ground-observed PM speciation profile having a larger POA fraction. The increase in PM₁ 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).