Dear Editor,

We thank the referee for their valuable comments and have replied to each question/comment below. Our responses are shown in green text and any changes made to the manuscript are italicized. The line numbers are in reference to the revised manuscript (unless stated otherwise).

Thank you for considering this manuscript for publication in ACP.

Response to Referee #2

This paper presents an estimate of dry deposition fluxes using aircraft observations.

Such "regional" estimate provide a very useful constraint that can be used to improve the representation of dry deposition in global models with implications for both air quality and ecosystems. The observation-based estimate is compared with the deposition velocity calculated by the GEM-Mach model (TON and TOS) and by a suite of dry deposition algorithms (TOS only). The authors conclude that Vd is significantly underestimated and show that revisions to the representation of Ra, Rb, and Rc could reduce the model bias.

The study is interesting and fits very well within ACP.

However, I do have concerns regarding the robustness of some of some of the results (especially for NOy) and I am unable to recommend this study for publication in ACP without significant clarification.

Comments

1) line 111. What is the sensitivity to organic nitrogen?

If the sensitivity is low, how does it affect the conclusions of the study?

The NO_y measurement is sensitive to organic nitrogen. Conversion is expected to be near 100% (e.g. Williams et al., 1998). This has been added to the text at Lines 121-123 with references.

2) line 191. Is this also a minor loss for NOy?

The upward flux is now also calculated for NO_y. Text was included at Line 462-464: '*For N, the upward flux was estimated to be ~570 g* km^{-2} hr^{-1}*, so although a larger flux than S, it is about factor of 18 lower than the TON fluxes derived from observations*.'

3) eq (2). If I am not mistaken, the authors assume that X_SO2 = - X_pSO4, if so this should be made clear. I would also suggest to write equation (4) as

Delta T_TOS = Delta T_SO2 + Delta T_pSO2 = -D_SO2 - D_pSO4.

Yes, $X_{SO2} = X_{pSO4}$. This has now been explicitly indicated at Line 252-254. We have chosen to keep equation 4 as is, but for increased clarity, as per the Referee's suggestion, we have added Equation 5 at

Line 255: $\Delta T_{TOS} = \Delta T_{SO_2} + \Delta T_{pSO_4} = -D_{TOS} - D_{pSO4}$ (Note the referee likely meant Delta T_pSO4 rather than Delta T_pSO2.)

4) The authors mention that the region is very dusty. This suggests that some SO2 (and much HNO3) could react on dust. Since coarse SO4/NO3 are not measured by the AMS, such flux could be mistakenly counted as dry deposition. The authors need to clarify how this is accounted for.

This comment is addressed as per Referee #1's similar comment and repeated here.

Yes, indeed the AMS only measures submicron aerosols (~60nm to ~1µm) and that pSO_4 can be present in both fine- and coarse-mode particles. The authors did identify this issue and identified a method to account for the coarse-mode fraction. This was done by using a ratio of measured PM_1/PM_{20} from aircraft particle instruments to adjust the AMS pSO_4 concentrations. This explanation was provided in the Supplementary in Section S4, but it now moved to the main manuscript in the Methods section for increased clarity of how this was accounted for.

The text is as follows: "*Since the AMS measures only particle mass < 1 µm (PM_1) in diameter, the mass of SO_4 formed through OH oxidation was scaled to account for all particle sizes that H_2SO_4 vapor could potentially condense on. The scaling factor was determined using the surface area ratio of PM_1/PM_{20} from the aircraft particle measurements. PM_1 measurements were from the UHSAS and PM_{20} were from the FSSP300. As the ratio did not vary significantly in the plumes, one single value was used between each set of screens; in F19 the ratio between screens ranged from 0.6 to 0.8, in F20 the ratio ranged from 0.8 to 0.9, and in F7 the ratio ranged from 0.7 to 0.9 (Liggio et al., 2016).*"

5) line 275 and line 453

It would be helpful to summarize the differences between the different dry deposition algorithms listed here (Table 1 of Wu et al. (2018), for instance).

Without such information, it is very difficult to understand the impact and validity of the changes in Ra and Rb recommended by the authors in the GEM-MACH model.

We agree a summary (including the Wu et al 2018 reference) would be useful and is now provided in Section 2.5, Lines 330-339:
"*The five deposition algorithms considered are denoted ZHANG, NOAH-GEM, C5DRY, WESLEY and GEM-MACH and are compared in Wu et al. (2018) (except the algorithm in GEM-MACH). The five algorithms all use a big-leaf approach for calculating V_d i.e. V_d is based on the resistance-analogy approach for calculating dry deposition velocity where V_d is the reciprocal sum of three resistance terms R_a, R_b and R_c. Although the approach is similar, the formulations of R_a, R_b and R_c between the algorithms are substantially different (Table 1 in Wu et al., 2018). Results from Wu et al (2018) suggest that the differences in R_a+R_b between different models would cause a difference in their V_d values on the order of 10-30% for most chemical species (including SO_2 and NO_2), although the differences can be much larger for species with near-zero R_c such as HNO_3.*"

6) line 351. Deposition velocities vary a lot across the different members of the NOy family.

Differences in NO emissions between model and observations (Tables 1 and 2) could lead to biases in the ratio of NO to NO2 or the conversion rate of NOy to HNO3, which would impact the simulated Vd(NOy).

Careful evaluation of the O3 and NOy simulation are needed to support the authors' conclusions regarding Vd(NOy).

Responding similarly to a comment from Referee #1, we expect that the deposition rates of different N species do vary. Using the GEM-MACH model, Makar et al. (2018), describe the relative contributions of different TOS and TON species towards total S and N deposition in previous modelling estimates in the Athabasca oil sands region. TON was dominated by dry NO_2 (g) deposition flux close to the sources (>70% of total N close to the sources), and dry HNO_3(g) deposition with increasing distance from the sources (remaining always < 30% of total N, with other sources of TON having minor contributions (< 10%). Using the observations here, we calculate the deposition rates for TON. Although TON encompasses a range of different N species with expected differences in their deposition rates, it was not possible to derive their individual deposition rates separate from their chemical formation/losses from our observations. This prevents us from being able to confirm these relative contributions, or determine measurement-based estimates of the deposition velocities of the components of TON.

For increased clarity, text has been added at Lines 260-262: "*Although TON encompasses a range of different N species with expected differences in their deposition rates, it was not possible to quantitatively separate their chemical formation/losses from their deposition rates with this method*."

For further discussion, we also add text at Lines 505-513 '*Using the observations, it was not possible to derive individual TON deposition rates separate from their chemical formation/losses. In previous modelling work, Makar et al. (2018), use the GEM-MACH model and describe the relative contributions of different TOS and TON species towards total S and N deposition in the AOSR. TON was dominated by dry NO_2 (g) deposition fluxes close to the sources (>70% of total N close to the sources), and dry HNO_3 (g) deposition increases with increasing distance from the sources (remaining always < 30% of total N), and other sources of TON having minor contributions to deposition (< 10%). Although TON encompasses a range of different N species with expected differences in their deposition rates, comparisons of $V_{d\text{-}TON}$ with the model show, nevertheless, that overall large differences do exist*.'

With regard to the comment on differences between the measurements and model amongst the TON (NO_y) species (e.g conversion to HNO_3; we think the reviewer intended to say '….the conversion rate of NO_2 to HNO_3') - such differences do indeed limit the comparisons of $V_{d\text{-}TON}$ to a bulk comparison. Differences in the partitioning of TON species between the model and measurements may influence the comparison of V_{dTON}, but the comparison in a bulk sense is still useful as the results indeed show that such a difference ie factor of 2 exists. We do have a statement at Line 576-577 that emphasizes the need for TON deposition velocities to be further investigated: '*Yet, for other algorithms and for **TON** compounds, the model low-biases in V_d remain to be investigated*.'

7) I assume that changing Ra (and Rb?) will not only impact the removal of chemical

species but also the heat, moisture, and momentum fluxes in GEM-Mach. Could the authors discuss the magnitude of these changes?

Yes, this is true. Ra parameterizes turbulent transport in the boundary layer, which occurs through turbulent eddies and behaves similarly (but not identically) for heat, moisture and momentum. It becomes more complicated for Rb since molecular and thermal diffusion processes take over from turbulent processes. Using the Zhang and Wu parameterizations rather than those currently implemented in GEM-MACH would decrease the Ra and Rb for the momentum, heat and moisture fluxes as well, but still remain within the range of what is expected based on published parameterizations. A similar shift of probability distribution functions in terms of relative magnitude as shown in Fig. 5b) would be expected for these fluxes as well. No changes made in the manuscript.

Minor comments:

1) line 20

"Dry deposition fluxes decreased exponentially with distance" This statement is unclear. Distance from where?

Agreed! We have clarified by modifying the statement at Line 21 as such: '*Dry deposition fluxes decreased exponentially with distance from the Athabasca oil sands sources, located in northern Alberta, resulting in lifetimes of 2.2-26 hours*.'

2) line 250

GEM-MACH has not been introduced yet. Delete or define earlier.

GEM-MACH was actually introduced, defined (and referenced) earlier at Line 227 (now Line 277). However, for additional clarity, we added text at Line 308 so the sentence that the reviewer commented on now reads: '*The measurement-derived* $\boldsymbol{V_d}$ *are compared with those from the air quality model GEM-MACH which uses inferential methods*.'

3) line 294

Introduce notation F7 as flight 7 (F7)

Although the flight notation was already defined earlier (in Section 2.1 Lagrangian Flight Design (current Line 91), for clarity, the notation is now repeated at Section 3.1, first line (Line 355; was Line 291).

4) fig. 1 Flight 20 shows two plumes for TON but 1 plume for TOS. Could the authors comment on this difference?

This comment is addressed as per Referee #1's similar comment and repeated here.

We have added further clarity regarding the sources of TON and TOS. Added text at Lines 355-367 ‘*Three aircraft flights, Flights 7 (F7), 19 (F19) and 20 (F20) were conducted in Lagrangian patterns where the same plume emitted from oil sands activities was repeatedly sampled for a 4-5 hour period and up to 107-135 km downwind of the AOSR. The first screen captured the main emissions from the oil sands operations with no additional anthropogenic sources at subsequent screens downwind. The main sources of nitrogen oxides were from exhaust emissions from off-road vehicles used in open pit mining activities and sulfur and nitrogen oxides from the elevated facility stack emissions associated with the desulfurization of raw bitumen (Zhang et al., 2018). As depicted in Figure 1, F7 and F19 captured a plume that contained both sulfur and nitrogen oxides. The westerly wind direction and orientation of the aircraft tracks on F20 resulted in the measurement of two distinct plumes; one plume exhibited increased levels of sulfur and nitrogen oxides mainly from the facility stacks, and the other plume contained elevated levels of nitrogen oxides, mainly from the open pit mining activities, and no* SO_2.’