Response to Editor's Comments

Editor's Comments (responses in blue text):

This is very interesting study demonstrating the challenges of simulating the interaction between meteorology and plume emission conditions and should be of considerable interest to air quality modellers.

Thank you to the author's for their detailed responses and the revised manuscript, which is acceptable for publication after technical corrections.

A couple of areas that the authors may consider in their revisions:

(1) In one of the sensitivity analysis the results showed that: "Nearly half (48%) of the predicted plume rise values are less than half the observed values and a large fraction (34%) of the predicted plume rise values are more than double the observed values." - This seems like a rather large dichotomy, which could potentially be explained with available information leading to some insight. The authors should consider exploring this result and briefly sharing what may be the reason(s).

We have looked into the data in more detail and it appears the dichotomy is related to the stack sources and their different effluent exit velocities. This is an interested result that and the following text is added at line 902 (at the end of Section 4.4.4)..

"The high fraction of underpredicted plume rise (48%) and underpredicted plume rise (35%) using the combined buoyancy/momentum formula of Eq. 20 warrants extra investigation. Of the 83 plume to stack matches used in this analysis, 40 are underpredicted (ratio < 0.5) and 29 are overpredicted (ratio > 2). Of the 40 which are underpredicted, 34 are Suncor stacks. Of the 29 that are overpredicted, 22 are Syncrude stacks. All 4 plume-to-stack matches with CNRL stacks are underpredicted. Hence there is a very strong correlation with stack location. This is consistent with the results discussed in Section 4.4.1, since the Syncrude stacks have high effluent exit velocities (e.g. Table 1), the Suncor stacks have low to moderate effluent exit velocities, and the CNRL stacks have moderate exit velocities. Combining the buoyancy and momentum with Eq. 20 appears to overestimate the influence of momentum, while simultaneously underestimating the influence of buoyancy."

(2) While a strength of this study is the use of observed/measured inputs in the plume rise formulations, which tended to show an underestimate of plume rise and little skill in predicting rise for individual plumes, it could be useful to readers to also see plume rise results for these cases based upon using the modeled (GEM-MACH) inputs. Perhaps this information is in the identified companion paper, however, it would be of value to briefly present these results in this paper so that all the information is in one place. Could an additional line be included in Figure 3 and/or the statistics in Table 3? If these results are added then some text would be important to describe the result. However, ideally, further linking to the companion paper would avoid the need for too much additional revision in this paper to explain what was done.

This information does indeed appear in the companion paper. We have added a brief summary of some of the main results of that paper in a new section at the end of the Discussion at line 914. (just before the Conclusions), as well as directing the reader to that paper for the details:

"**4.5 The influence of stack-location-specific meteorological data – Companion Paper**

Our focus within this work was the use of the available measurement data as a proxy for the meteorological conditions at the stack locations themselves. However, significant differences could be seen in the data between the different measurement platform locations (see Table 2). In subsequent work in our companion paper (Akingunola *et al.* 2018, this issue), high resolution meteorological model forecast simulations for the region were carried out. These suggested the presence of significant spatial heterogeneity in the meteorological parameters used to drive both the Briggs parameterization and the layered method. Predicted meteorological parameters at the meteorological measurement platform locations were substantially different from those at stack locations. When tested using the model-predicted at-stack meteorological values, and NPRI stack emissions data, the Briggs parameterization and the layered approach resulted in very different plume rise behaviour. Predicted surface SO_2 concentration performance was substantially improved across all metrics when the layered approach was used, and aircraft SO_2 comparisons improved for all metrics aside from bias. For the predicted plume heights, the slope of the model observation line was -0.16 for the Briggs parameterization, and 0.97 for the layered approach, with the former under-predicting, and the latter over-predicting the aircraft-observation-estimated plume height. The reader is directed to Akingunola *et al.* (2018) for a discussion of these issues, which suggests that accuracy of estimates of the driving meteorological parameters at the stack locations has a controlling influence on the performance of the layered approach, and with the layered approach recommended for future development. "

Non-public comments to the Author:

Some minor technical issues I found. A careful proof-read of the final manuscript would be helpful as I have not gone through this thoroughly.

All the minor corrections listed below have been incorporated and the issue has been proof-read again.

L144: "The atmosphere is considered .." not "the plume is .."

L196: "found" not "hound"

Table 1 caption: "the flight" not "flight"

Line 489: Consider saying "Non-stationarity" instead of "Stationarity" or say: "The assumption of stationarity..."

Line 513: the variable 's' is used for two different things. This is a bit confusing.

Fig 4 caption: "data" not "date"

L723: Seems more appropriate to say "estimated" not "calculated"

Great work!