Response to Reviewers' comments on

Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty at mesoscale

Thanks to the reviewers for their careful reading of the discussion paper. The reviewers' comments are below in italics, with our responses in normal font.

Reviewer #1

1 General

The paper addresses the relevant issue of quantification of uncertainty in mesoscale modelling and atmospheric transport modelling. It makes a limited contribution which is, though, well defined and the underlying analysis appears to be well designed.

Thank you.

2 Specific remarks

1. A major part of the paper deals with uncertainties of meteorological parameters in the WRF output. This makes sense, but it should also be reflected in the title, which only mentions uncertainty in transport simulations.

While it's true that a major part of the paper covers the uncertainties of the meteorological model output, we did not choose to highlight that in the title. If we were writing a paper about meteorological uncertainty, we would need to include even more about it. There is a very extensive literature on that, and it's not clear that we could add anything to that literature. Instead our purpose is to describe the meteorological uncertainty sufficiently that an interested reader can compare it (however qualitatively) with the uncertainty in the resulting transport.

2. The issue of uncertainty estimation through ensemble techniques in meteorological and transport modelling is vast and it is clear that a piece of work as this here can only cover some limited aspects. Therefore, it is important that the limitations of this study are clear communicated. This would for example include climatic region, season, receptor points compared with observations (e.g., only limited set of surface data, not representing specific features such as topography, coastal, etc.). Furthermore, with respect to the tracer simulation, the spatio-temporal release pattern is important (results could be quite different for point sources). There may be more such aspects. In general, a complete error description has to include also the spatio-temporal covariance patterns – not that I am expecting this study to deliver it, but it deserves to be mentioned.

Text to cover these points has been added in the Introduction: "We also note that the generality of the results is unknown. The region we cover is in the middle of a continent, with only modest terrain, and we only consider six weeks of one season. We use spatially distributed emissions; point sources might produce rather different results."

However, note that in fact we are not comparing the tracer results to surface measurements, but to aircraft data, which has considerable advantages in being away from the surface and covering a large area. On the other hand, it has disadvantages in being deliberately biased toward specific conditions (e.g. avoiding rain) and having limited temporal coverage.

3. Meteorological parameters are evaluated one by one, without looking at the error covariance between them, which obviously could impact the transport. Also vertical motion is not considered.

We attempted to calculate correlation lengths in the meteorological variables and in the tracer fields. It was not clear how to do this in a complete and correct fashion, and the results were not particularly enlightening, so we did not pursue them further. Grid-scale vertical motion is unobservable, so we did not pursue it either.

4. The main results are presented almost exclusively as spreads / standard deviations/ RMSE normalised by the mean value of the quantity under consideration. Even if partly regions with "low values" are left white in the plots, such a ratio is obviously sensitive to this normalising value in the denominator. The authors themselves dismiss a potentially important result (comparison of model error to ensemble spread for CO) as not robust. It would be very useful to show and discuss also absolute values. Scatter plots (double-log), probably the most common form of presenting model-observation or model-model concentration values, should also be included.

In order to keep the paper concise, we have included two plots of absolute values (means and spreads) in the Supplemental material. While useful for checking the reasonableness of the results, these plots are quite difficult to interpret in terms of uncertainty. Scatter plots would not be meaningful because of the lack of background CO in the simulations and the uncertainty of the emissions. Improving the emissions inventory is our long-term goal, but the results in this paper do not depend on perfect emissions.

5. Page 4604, line 20: While backward runs may be used to invert measurements for finding source emissions, neither are backward runs necessary for this purpose, nor is this their sole application.

The text now reads, "…among other uses…"

6. Please avoid the wording "NWP model or reanalysis" – reanalyses are produced by NWP models. What you mean is probably operational NWP model output or reanalyses.

Reworded.

7. In this paper, "FLEXPART" is used for the WRF version of this Lagrangian model, which is a branch-off from the main version running with ECMWF or NCEP global data. I think it would be more clear to use the designation "FLEXPART-WRF".

We have changed the term to "FLEXPART-WRF" throughout where appropriate.

8. At first (p. 4605), mesoscale simulations are defined as resolving features of 10 km in size. Then we are told that simulations were done with 12 km grid spacing. This is a contradiction, as such simulations will resolve only features larger than 24-48 km. One should also note that ECMWF now produces 0.125 degree output – corresponding to the WRF simulations used. If it is still

possible for a revised version to include ECMWF fields into the comparison, it would be very valuable.

The text has been reworded to refer to features "10-100 kilometers in size." However, we note that a simulation on a 12-km grid does contain some information at that scale, since the land surface and physics operate at the grid scale.

Using ECMWF operational output, while appealing, is not practical in our case. That output is not freely available to researchers outside Europe.

9. As runs were reinitialised daily, I presume that no grid nudging was done. Please state that clearly, and also explain the rationale for this approach (what about discontinuities between runs and how they affect transport simulations?).

This text was added in section 2: "No observed data was directly assimilated into WRF, nor were the WRF runs nudged toward any analysis." We don't think an explanation is necessary. Grid nudging is used by some practitioners to prevent long simulations from diverging from the larger-scale solution. A similar effect is achieved by reinitializing the runs each day. We have looked for discontinuities in these runs and in our previous work, and have not found them to be important.

10. Of course, soil moisture is an important parameter. However, one should keep in mind that also soil temperatures could be an issue.

In fact we cycled both soil temperature and moisture, as mentioned several places in the text. However, there is considerable literature on the importance of soil moisture, and little about temperature. We think it is best to leave the text of the third paragraph of section 3 as it is.

11. On page 4614, line 20, rank histograms are introduced through a reference. Please explain this quantity sufficiently well to let readers understand you results without having to look up this reference.

We have added one sentence of explanation. Readers wanting more should really look up the Hamill reference, which presents a number of caveats and cautions that we cannot cover in this paper.

12. I found the meaning of the columns in Table 7 not easy to comprehend. Please try to improve description. Each column should be explained in the caption, referring to it by number or by quoting its column head. Do not use abbreviations such as "std.dev." in the caption. On the other hand, some information such as details about data points can be replaced by links to previous explanation or the text.

We have spelled out the column headings and removed abbreviations.

13. Figure 1: The size ratio between the maps and the colour bar is disproportional. One of the maps, e.g. the station map, should show the mean CO emission field rather than terrain, as the distribution of the emissions is quite relevant for understanding the results.

We have reduced the size of the colorbar and added a panel showing the CO emission field.

14. Figures 2 and 3: Please make sure that the observations are clearly distinguishable from the simulations (e.g. thicker line).

The dark blue line used for the observations is easily distinguished from the other lines, all lines are a reasonable thickness, and the legend makes clear which is the observations.

15. Finally, GMD has a data policy – however, it seems this policy is not applied. At least there should be a statement on data availability.

The WRF simulations and FLEXPART output amount to several terrabytes. We do not have resources to serve that amount of data to others. Interested parties can of course contact us and we will work with them. We have indicated that in the text (beginning of the Conclusions).

Reviewer #2:

In the paper "Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty at mesoscale" by W. Angevine et al. the authors address the impact of transport errors on Lagrangian passive tracer simulations. The authors use six configurations of the WRF model to obtain an ensemble of meteorological situations on the east coast of the USA. They run identical configuration FLEXPART-WRF runs on the resulting wind fields. The output of the Lagrangian model runs is statistically analysed in order to derive estimates of transport uncertainty. The paper is interesting and well written and I recommend publication after a few technical corrections that may make the paper easier to read for certain readers.

Thank you.

Specific/Technical comments:

4608 l7: The turbulence is only horizontal or also vertical?

The Hanna scheme calculates both horizontal and vertical turbulent mixing. This has been clarified in the text.

4608 l15: define age clases

We have added "since emission" to clarify.

4609 l1: Providing the emissions is useful in case someone wants to repeat the experiments.

The emissions are now shown in figure 1. The emissions data can be duplicated given the information in the text, or can be gotten from us directly.

4609 l27: what does "open loop" mean?

This should be clear from the following sentence.

4611 l6: Is there a reference for this uncertainty interpretation?

It is contained, explicitly or implicitly, in most textbook discussions of uncertainty or error analysis. We have included a reference.

4612 l23: For the readers not familiar with US geography, it would be useful to precise where Atlanta is.

We have added text to clarify that Atlanta is in Georgia and give its approximate coordinates.

4612 l28: You mean that because the lower layer is difficult to simulate you are going to focus on a higher layer?

Not exactly. We mean that because the lower layer is more difficult, it is useful to look at a higher layer in addition.

4613 l20: Have you made additional calculations or it is speculation?

We have not made further calculations, but the speculation seems reasonable from the plot shown, and the language ("may be") is sufficiently mitigated.

4614 l15: Define CRN before Figure 7. Maybe mention at intro.

The Climate Reference Network is defined and referred to in conjunction with figure 2, on page 4610.

Table 7: mention CRN before.

See response to previous comment.

Fig 6. made more clear that are relative units.

These are normalized spreads and therefore unitless. This has been clarified in the caption.

Fig 11. clarify: tracer age calculated with trajectories, mixing ratio measured.

A good point; in fact all quantities in figure 11 are simulated, not measured. This is also true in figure 12. It has been clarified in the captions.