Interactive comment on "Bayesian inverse modeling of the atmospheric transport and emissions of a controlled tracer release from a nuclear power plant" *by* Donald D. Lucas et al.

Donald D. Lucas et al.

ddlucas@alum.mit.edu

General response to Referee #1

We highly appreciate the feedback from Referee #1. The referee recommended a few minor technical corrections and commented on three issues. The issues are related to 1) the *mse* and *corr* metrics used to optimize model and measurement differences, 2) the apparent detrimental effects of nudging and data assimilation on the simulations, and 3) the information on modeling errors obtained by plugging in the known source values (location, time, and amount). We address these issues in further detail below

and will revise the manuscript for clarity and content based on the reviewer feedback. Again, we thank the referee for the informative comments.

Minor comments

1. *mse and corr metrics*
   The reviewer asks about the advantage of fitting the mean-squared-error and correlation metrics instead of the actual concentrations for the optimization. The advantage is mainly one of statistical convenience. We could, in principle, fit statistical models to the actual concentrations (e.g., using gradient boosting), and then use the statistical models to minimize concentration differences with observations. Fitting the concentrations is more challenging, however, because the statistical model becomes functional and the size of the problem is larger. Referring to Eqs. 7 and 8 in the manuscript, our vectors and covariance matrix are only two dimensional (*mse* and *corr*). The terms in these equations would have more dimensions if we directly modeled the concentrations (1,148 without applying a dimensional reduction technique like principle components analysis). Further, we would have to estimate spatial and temporal correlations in the covariance matrix. We have collaborated with statisticians to tackle this more challenging statistical problem (Francom et al., 2016). This effort has shown that the additional complexity of modeling the concentrations can pay off in the form of tighter parameter estimates.

2. *Nudging effects*
   Like the reviewer, we find it counterintuitive that increased levels of meteorological nudging appear to degrade the agreement with concentration observations. Given the relatively small likelihood differences between the nudging options though, this result is not highly robust and not one of our major conclusions. Nonetheless, we also surmise that tradeoffs between nudging, variability, and

physics parameterizations in complex terrain may be playing a role. At higher levels of nudging, there is less variability, and therefore less opportunity for parameterizations to compensate for terrain-related errors. At low levels of nudging, there is more variability and a greater possibility to find better matches with observations. In future work it would be interesting to further diagnose nudging in complex terrain.

3. *Model error from known source*
   The reviewer recommends adding results using the known source parameters to help quantify model error. Both reviewers made the same excellent suggestion, so we expanded Sect. 5.1 in the revised manuscript to accommodate these results. A new scatterplot in the section shows the mean squared error and correlation for the 162 WRF runs with the known source parameters. These variations are not as large as they are across the full 40,000 member ensemble, though they are still significant. We also show that the known source variations are explained mainly by differences in the reanalysis data sets and initialization times, which agrees with our previous variance analysis in Sect. 5.3. Other factors, including nudging levels, do not play a very large role.

Specific comments

In addition to the above comments, we will also revise the manuscript by

- providing a better describing of Latin hypercube sampling,

- and improving figure captions (e.g. Fig. 3).