Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., referee comment RC2
https://doi.org/10.5194/acp-2020-1129-RC2, 2021

Comment on acp-2020-1129

Anonymous Referee #1

Referee comment on "Quantification of uncertainties in the assessment of an atmospheric release source applied to the autumn 2017 ^{106}Ru event" by Joffrey Dumont Le Brazidec et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-1129-RC2, 2021

Review of the article: "Quantification of thr modelling uncertainties in atmospheric release source assessment and application to the reconstruction of the autumn 2017 Ruthenium 106 source" by Dumont Le Brazidec et al.

The manuscripts presents an evaluation of the impact on assumptions surrounding statistical model selection on posterior estimates of source properties of (unknown) radiological releases. This manuscript will be informative to researchers and operational users. I have tried to avoid repetition if the previously posted comment. One major drawback of the manuscript is that much of the motivation is based on a straw-man argument, i.e. the original Gaussian set-up is designed in the manuscript such that it will fail. Below are a number of suggestions for revisions to improve the manuscript, followed by technical comments.

- Paragraph starting line 110: This is a straw-man argument. The assumption from the start is that the error is larger for higher measurements. This may not always be valid, e.g. an incorrectly dispersed wide plume with a very high concentration. It may be that the error is much larger for the smaller measurements than the larger measurements. There are other approaches to improve the validity of Gaussian (or any) likelihoods through transformation of variables. For example, using a non-linear forward model. Caveats, justification and 'typical errors' needs explaining, preferably at the start of section 2.
- Much of the arguments surround having independent and identically distributed (iid) model-measurement error in the covariance. Many of the arguments throughout can be countered by the use of non-iid covariances, e.g. tR, where the diagonal of R is the measurement value and t is a scaling – equivalent to having e.g. a 10% model-measurement error. This needs further discussions and better justification for the arguments proposed (e.g. non-negativity).
- Section 3.2.3: It would be useful for many readers to provide a brief conceptual introduction to MCMC methods (i.e. asymptotically exact methods not reliant on closed form solutions or conjugacy).
- Section 3.4: This section needs expanding considerably. It is a paper with lot of content. At current, the summary provides an overview of the approach of the paper

but no summary of the finding. The summary should summarise the results and findings to adequately inform the lazy reader on the paper's content.

- I understand the following would require a lot of extra work, and so do not mandate it for publication. It would however, much improve the paper. Seeing as an ensemble is used, it would seem sensible to me to use a simulated dataset (a simulation using an ensemble member) to draw conclusions from the various experiment. It is not perfect, but would be useful to have a 'truth'.
- It would be useful in the analysis and plot to also show the original case of a Gaussian likelihood. This is needed to prove the worth of using a non-Gaussian likelihood.

Technical comments:

Title: "modelling uncertainties" is ambiguous as it can refer to a statistical model or a transport model. The title also is not grammatically correct. A suggested improvement is "Quantification of uncertainties in atmospheric release source assessment applied to the autumn 2017 Ruthenium 106 source".

Abstract, line 5: "improve on these distributions" is vague. 'Better quantify' or 'improve estimates of these distributions' would be better.

Line 7: A space is not needed before a colon in English.

Line 8: Clarify 'model errors' (I assume transport errors?)

Line 10: 'several suited distributions for the errors are advised' the passive voice reads as though you are advising the distributions. Better "we suggest several suitable distributions for the errors" or "are suggested" if sticking with the passive.

Line 17: sources or a source; remove 'many'; 'Therefore' doesn't follow – delete.

Line 31: 'finds it origin in' to 'originates from'

Line 38: This is an oversimplification of the weighting strategy. See, for example, importance resampling or Ensemble Kalman filter methods.

Line 46: Why index vector x elements with x1, x2 and then ln(q)? Perhaps x3=ln(q) would be clearer.

Line 47: R is better described as the covariance matrix containing the model-measurement errors.

Line 52: Introduce for the reader what the prior is (i.e. the probability distribution of prior knowledge before considering data).

Line 57: reconstructed posterior distribution

Line 60: ‘transformation’ is better than ‘parameterisation’

Line 61: ‘are the results’ would be better as ‘are the results of a simulation’

Line 62: ‘and are therefore depending’ to ‘and depend on’

Line 70: This isn’t an expansion.

Line 103: A cost function is a non-probabilistic concept and so better to refer to as simply the negative log-likelihood.

Equation 4: There should be no divide by 2 in the first term.

Line 112: A space not full stop is needed between units

Line 115: Capital G on Gaussian.

Line 120: Unless there has been a transform (e.g. ln(y)). Square bracket is facing the wrong way.

Line 123: Space between units.

Line 156: ‘Large multiple’

Line 178: ‘as this paper’

Line 278: What are the upper/lower bounds of the uniform distribution?

Line 280: What are the shape parameters of the log-gamma distribution?

Line 348: ‘Harmful’ is an incorrect choice of work here. You can just delete it;

Line 348-350: This sentence isn’t correct. Observations don’t have a high likelihood. Please rephrase.

Line 351: Change ‘totally legitimate’ to ‘valid’

Line 353-355: This sentence does not make sense. I’m unsure of its meaning, please revise.

Line 368: 4c and 5a

Equation A1 and A2: Second term is incorrect, not divide by 2 but multiplied by 2.