Atmos. Chem. Phys. Discuss., 10, C13686–C13691, 2011
www.atmos-chem-phys-discuss.net/10/C13686/2011/

Atmospheric
Chemistry
and Physics
Discussions

Interactive comment on "A Bayesian inversion estimate of N₂O emissions for western and central Europe and the assessment of aggregation errors" *by* R. L. Thompson et al.

Anonymous Referee #2

Received and published: 15 February 2011

The study presents a pixel-based inversion over western Europe for 2007 of N2O fluxes using one single concentration tower. The first part aims at defining the optimal correlation lengths (in time and space), in a similar way as Kaminski et al., 2001, by projecting the aggregation error from the flux space into the observation space. The second part presents the inverse flux estimates, using IER emissions and GEIA as prior fluxes, and compare them to two other flux estimates (EDGAR and the inverse fluxes from Manning et al., 2003). The paper presents the first continental scale inversion of N2O fluxes, but the lack of discussions about the results and the other elements of the inverse system requires major revisions before publication. Considering the very limited number of data available to assess the N2O flux balance over western Europe, this study should include clearly the limitations of the method. The detailed assessment of aggregation errors remains very interesting and well presented, but represents only a part of the total errors that might affect the results.

The general structure of the paper focuses first on the aggregation errors, which is understandable considering the limited number of observations, therefore leading easily to an ill-conditionned system. But the other sources of errors, as transport model errors for example, are poorly documented. The location of the site in a moutain area is highly challenging for transport models, as shown previously in the litterature (e.g. Law et al., 2008). Representation errors might affect the results with systematic errors. The assessment of the different error contributions from boundaries, measurements, and transport, seems somehow arbitrary in section 3.2.1. It also highlights that these components are, individually, about 3 times larger than the aggregation error (P26092-line 6). As a note, the statement at the end of section 3.1.3 seems different than the one for the real case in section 3.2.1 (sum of the squares of the errors about 0.5ppb total compared to 0.1ppb for aggregation errors, but only 0.3ppb vs 0.05 to 0.24ppb in the pseudo data case). The long description of aggregation errors when considering this final ratio is disproportionate as a minor part of the total errors. Following this comment, when the error reduction is presented later, it is more likely that the total observation errors might be under estimated because of the other unexplored components, with potential biases in the transport model errors, or from incorrect boundary conditions. A careful description of the other sources of errors is necessary before considering the error reduction (which combines all of the errors), or at least, a discussion is required after presenting the results.

Concerning the aggregation error calculation, the importance of the correlation lengths in time and space is well presented in the paper. Nevertheless, it remains unclear for the reader that the consistency evaluated with the chi-square test (close to 1 or not) does not imply that the correlation lengths are realistic. This test only ensures the ratio between the observationnal constraint and the number of unknowns (with their associated uncertainties). One could artificially increase the observationnal constraint by adding correlations, and therefore decrease the degree of freedom of the system, as long as the variability in the modeled concentrations still matches the main features in the observations. Flux error correlations have a reality, often hard to measure, but the structures in the errors are a representation of the truth (or should be). This true correlation length might lead to an ill-conditionned inverse system, which means for example that more observation sites are needed. In this study, the definition of the optimal correlation lengths is calculated to fit your assumptions (high frequency data but only one site) to invert fluxes over a large domain. This point has to be discussed in the paper. Considering the degree of freedom of any system, one can always inflate the observational constraint to reach a reasonable balance between observations and unknowns by increasing the error correlation lengths. Several tests could be done to evaluate the importance of the correlation lengths. Considering the shift of source locations, changes in the seasonal cycle amplitude, and the annual mean flux, from sensitivity tests of the solution to the assumptions is a critical aspect of an inverse system.

More generally, the discussion part is mixed with the results, and too limited. Some discussions are part of the results, as your analysis of the aggregation errors, but many other aspects are not considered here. For example, in Table 5, the annual balances of the prior and the posterior are close to each other considering the posterior uncertainties, larger than 50% of the total value. Even EDGAR estimates are within 1 sigma of the posterior estimates. This is not even discussed. In section 3.2.2, called "posterior error", only error reduction is presented, and two general statements are made (more constraint when closer to the site, and too large correlations might lead to over estimated error reduction). This part needs an important improvement, including discussions about the posterior errors, the comparison to previous studies in terms of uncertainties (is this study bringing more reliable estimates than Manning et al for example). As a first study of its kind (for N2O at this scale at least), discussions are more critical than results, because one site only was used, but also because of the low

quality of current prior estimates.

Finally, the correlation between the inverse fluxes and several other variables are part of the results, but many conclusions are clearly preliminary. Many "may" seem to indicate discussions more than results. The table 4 for example, showing only one region correlated with soil moisture, and two out of six with rainfall, is not convincing from a statistical standpoint. Even if the UK is known as a rainy place, the Benelux, western Germany, or western France, have actually about the same amount of precipitation. Why, in these areas, don't you observe similar correlations? Other variables, not intuitively correlated, or not physically related, might show similar results. Even temperature anomalies could be an artefact as you explained at the end of the section 3.2.3. By the way, the temprature anomalies show a clear seasonal cycle, which is somehow surprising. Are you removing the mean temperature at every time step from this value or just an annual mean? This part is clearly a discussion and not results.

Technical comments:

26079: Equation 1: first described by Uliasz 1983 (Uliasz, M.: Application of the Perturbation Theory to the Sensitivity Analysis of an Air Pollution Model, Z. Meteor., 33, 6, 355–362, 1983.)

26084: The correlation length of 30 days seems very arbitrary and long. Using this assumption, an error occuring the 1st of July in your boundary conditions has an impact on the 1st of August (and even after). Does it make sense? Have you tried to decorrelate the boundaries in time, two weeks being long enough.

Figure 2: differences between noisy and initial 7-day pseudo data seem rather small on the figure. The 0.3ppb value compared to daily variations of several ppb could be increased to test the potential of your system.

26087: replace "validity" (e.g. "internal consistency").

section 3.1.2: The description of the impact of the temporal correlation is misleading.

The over (or under) estimation of the constraint is a consequence, but you need to know the truth to conclude in the real case. In your case, you assume a true correlation length, and test the impact of using different values for D. In other terms, you test the impact of wrong correlation lengths in your inverse system. But the optimal correlation length you determine later is valid for your system, maybe not in reality. Explain clearly the missing elements in your study. Could you, for example, bias your retrievals by using large correlation lengths estimated to keep the ratio between observations and unknowns?

26088-20: use "a posteriori" or "posterior"

26089-5: Carouge et al., 2010 (ACP) use the algorithm to determine the impact of aggregation errors.

26091-1 to 16: re-write the paragraph, and explain the figures: why a minimum of chi2 at one day? Is it due to the temporal correlation in the observation errors? Why does it increase again after? You explain why it has a minimum at 168h, but not why it increases again after this point. This figure is interesting, because previous inversions used averaged observations. A good description is needed to understand better the different features, and will bring additional value to the paper.

26092-5: "choice" seems arbitrary. The assumptions are determined, not chosen (or should be, at least). The truth could be far from your "choice".

26093-5 to 8: repetition. In addition, the chi2 tests the consistency, or the ratio between observations and unknowns. A similar chi2 value would be obtained with smaller prior RMS and smaller D (or larger prior RMS and larger D) for example.

Conclusions: the second part of the conclusion has to be re-phrased. Final uncertainties are ignored, results are not discussed (e.g. the correlation for the UK is unique and might be random, the small correction on the annual mean is a direct consequence (i.e. expected) of the limited number of observation sites), and even if the EDGAR invento-

ries are smaller, they are within 1-sigma of your results.

Interactive comment on Atmos. Chem. Phys. Discuss., 10, 26073, 2010.