Atmos. Chem. Phys. Discuss., 10, C14710–C14717, 2011
www.atmos-chem-phys-discuss.net/10/C14710/2011/

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.

R. L. Thompson et al.

rona.thompson@lsce.ipsl.fr

Received and published: 25 March 2011

We thank the reviewer for his/her thoughtful and constructive comments.

General comments

"The location of the site in a mountain area is highly challenging for transport models, as shown previously in the literature (e.g. Law et al., 2008). Representation errors might affect the results with systematic errors."

While it is true that atmospheric transport at mountain sites is generally poorly represented by global models (as discussed in Law et al. 2008), the Lagrangian model

STILT is able to reproduce the observed atmospheric variability of a number of biospheric tracers reasonably well for the mountain site, Ochsenkopf, as was shown in Thompson et al. 2009. Thus, the model representation error for this site is considerably smaller than that found for a global model. Furthermore, Fig.6a, shows that even with the prior flux model, the observed variability at Ochsenkopf is captured fairly well (R2=0.44). However, we certainly agree that this error is still important and an insufficient estimate of this error could cause systematic errors in the retrieved fluxes. We estimated the model representation error, σtran, as the mean of the 3-day standard deviations of the simulated N2O mixing ratio at Ochsenkopf. The reason for this is that transport models are generally not able to represent synoptic scale variability well, which typically has time spans of circa 3-days. The obtained value of 0.3 ppb also reflects the prior-model – observation differences (which additionally contain an error component due to the flux errors) quite well. The boundary error, σbnd, has a component due to errors in transport as well as a component due to errors in the concentration fields used. We estimated σbnd as the standard deviation of the boundary contribution to the mixing ratio observed at Ochsenkopf. Conversely, the measurement error, σmeas, is relatively well constrained by the measurements of target gas and by the comparison of the in-situ measurements with flasks (see Thompson et al. 2009). A description of each of these errors has been added to the manuscript on p.26092.

"The assessment of the different error contributions from boundaries, measurements, and transport, seems somehow about 3 times larger than the aggregation error"

Unfortunately there was an error in the text of the manuscript (p.26092, l3). The aggregation error estimate for the 7-day resolved fluxes is in fact 0.2 ppb, and thus is comparable with our estimates for the other errors. We apologise for this mistake.

"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)."

As we mention above, there was an error in the value of the aggregation error (p.26092, l3), it is in fact 0.2 ppb but the quadratic sum of the errors, 0.5 ppb is the correct value.

For the pseudo-data tests, the quadratic sum of errors was in fact 0.2 ppb, (the quadratic sum of 0.1 ppb for each of the transport, boundary, and measurements errors). With hindsight, it would have been perhaps better to have used a larger error for the pseudo-data tests reflecting that of the real data, however, the choice of error does not affect the pseudo-data test results. Also the aggregation errors (p26090, l21) given for the flux resolutions of 7 to 30 days, were actually the variances ($\sigma 2$). Thus the errors (σ) are 0.22, 0.27, 0.32 and 0.49 for 7, 10, 14, and 30 days, respectively. We apologise for not making this clearer in the original manuscript and have modified the text accordingly.

"The long description of aggregation errors when considering this final ratio is disproportionate as a minor part of the total errors."

Again, there was an error on p.26092 for the value of the aggregation error. The actual error (0.2 ppb) is in fact more comparable to the other errors. We decided to focus on this error rather than the other error sources for the following reasons: the measurement error is relatively well constrained, as mentioned above, as it can be assessed by the comparison with independent measurements, e.g. flask samples analyzed independently and by measurements of target gases. Therefore, although the measurement error is large, it does not warrant as much examination as the other errors. On the other hand, the transport and boundary errors are less well constrained and, therefore, do warrant detailed examination. We have included a description of how these were determined (p.26092) and although these estimates were based on rather simple assumptions, the magnitude of these errors reflects the prior-model – observation differences (which include additionally the errors in the fluxes). There is still much work to be done in developing methods to better determine these errors but a detailed exploration of this is outside the scope of this paper. We chose instead to focus on the aggregation errors, which we found to be a non-negligible part of the total

error in the observation space, as mentioned above.

"…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."

The assumed error in the observation space is likely an overestimate rather than an underestimate based on the comparison of the assumed error (0.5 ppb) and the prior-model – observation differences (again see Fig. 6a). Furthermore, the chi-squared value of 1.17 indicates to some extent an appropriate choice of errors. Therefore, we do not think that this biases the calculated error reductions to any significant extent.

"…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 observational constraint and the number of unknowns (with their associated uncertainties)."

The chi-square value is the ratio of the actual model-observation and model-flux errors and their prescribed errors and, therefore, has an ideal value of 1.0. As the reviewer correctly notes, this ratio also reflects the choice of correlation scale length but a value of close to one does not necessarily mean that the correlation scale length is correct. It is not possible to determine the true values of the correlation scale lengths from these synthetic tests. Instead, one can only gauge the sensitivity of the results to the choice of correlation scale length. This is mentioned in the manuscript (p.26089).

"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."

The correlation scale lengths of the prior flux errors are independent of the frequency of the observations and the number of sites. We chose a value of 300 km for the spatial correlation scale length on the basis of the scale of weather systems and land-cover, and a value of 30 days for the temporal correlation scale length so that errors in seasonal fluxes are uncorrelated. This description has been added to the text (p. 26083).

"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"

Yes, it is true that the annual posterior emissions are very close to those of the prior. We agree that this point was missing in the discussion and we have now included a section (section 3.2.4) about the annual mean fluxes and the country totals.

"In section 3.2.2, called "posterior error", only error reduction is presented... This part needs an important improvement, including discussions about the posterior errors, the comparison to previous studies in terms of uncertainties. . ."

This section has been expanded to include discussion of the posterior errors (section 3.2.4).

"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 artifact as you explained at the end of the section 3.2.3. By the way, the temperature 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?"

We compared the regional mean posterior flux with the regional anomalies in tempera-

ture, soil moisture and rainfall at each 7-day time-step, thus the results for each region were based on 50 data points, N=50 (it was not quite a full year). The results were tested for significance at the 95% confidence level. These data are shown in Fig. 8. The anomalies were the differences from the annual mean and were normalized. We chose to test for correlations with these 3 parameters (temperature, soil moisture and rainfall) as these are strongly implicated in the strength of N_2O emission although other physical parameters also play an important role e.g. soil type, amount of nitrogen substrate and oxygenation (e.g. Freibauer and Kaltschmitt, 2003). The amount of N_2O emission is dependent on all these (plus other) factors. However, integrated over large regions and in time (7-day intervals), we found that temperature variability dominated over the other 2 tested parameters. The differences between regions for the result of correlation with rainfall and soil moisture depend strongly on other factors not considered in this analysis, such as soil type, oxygen and nitrate substrate availability, as well as on land management. We consider that it is outside the scope of this manuscript to address all these influences.

Technical comments

p26079: Eq 1. This equation was shown in Lin et al. JGR 2003, who refer to the book by Uliasz et Pielke, Computer Techniques in Environmental Studies, 1990. We have included this reference instead.

p26084: The variability in the boundary has only a small influence on the variability observed at Ochsenkopf (where the variability is mostly determined by processes occurring within the domain). Therefore, we do not expect the result to be sensitive to the correlation scale length of the boundary condition. Also, given that the boundary mixing ratios have a resolution of 2 weeks, and that missing values were estimated via our gap-filling procedure (SSA) we do expect some correlation between the errors in these values, so we do not consider a correlation of 0.37 over 30 days to be a misrepresentation of the correlation.

Fig.2: An additional figure (Fig. 2b) has been added to show a 'zoom' of the different timeseries', which shows the differences between them more clearly.

p26087: 'validity' has been changed to 'internal consistency'

section 3.1.2: the reviewer correctly notes that our tests for the correlation lengths test only the influence of choosing a wrong correlation length, however, do not help to determine the true length in the real data inversion. We have made this point clear (in section 3.1.2). The determination of the true flux error correlation lengths is an ongoing area of research and goes beyond the scope of this study.

p26089, l5: We are not quite sure what the reviewer is referring to here. If the reviewer means the Kaminski algorithm for estimating the aggregation error, Carouge et al. ACP, 2010 (parts 1 & 2) do not use this algorithm. Carouge et al. investigate the effects on correlation and normalized SD (w.r.t the true fluxes) of aggregating the posterior fluxes in time and space but do not investigate the influence of the spatial or temporal resolution in the inversion itself i.e. the aggregation error.

p26091, l1 to 16: averaging the observations means also an averaging of the surface influence function, so the surface influence becomes more spread-out with longer averaging intervals. Therefore, by averaging the observations the ability to resolve the fluxes in different pixels is reduced. This is why the RMSE is the smallest at the highest observation frequency (this is the case with perfect transport but of course may not hold when the transport errors are large). The secondary RMSE minimum at 168 hours, corresponding to the flux resolution in the model, and results from the effect averaging the observation has on reducing the influence of aggregation error in the fluxes. The chi-square value is determined by 3 factors: the posterior-prior flux difference (fpost-fprior), the error in the observation space (Ffpost-y), and the magnitude of the prescribed errors for both of these. The posterior-prior flux difference is largest for the 3-hourly observations and decreases with averaging interval length, whereas the opposite is true for the observation errors. However, the shape of the chi-square curve

is not quite the balance of these two, as it also depends on the estimate of errors. The minimum at 24 hours is likely due to the prior assumption of the temporal correlation scale length in the observation error covariance matrix, while the minima at 168 hours is likely because the prior error estimate is close to the actual error, since there is no aggregation error at this time. We have now re-written the section 3.1.4.

p26092, l5: this has been changed.

p26093, l5-8: these lines do not seem to correspond with the reviewer's comment, so it is not sure what was meant.

Conclusions: we agree with the reviewer's comment and the conclusions section has been modified accordingly.

Please also note the supplement to this comment:
http://www.atmos-chem-phys-discuss.net/10/C14710/2011/acpd-10-C14710-2011-supplement.pdf

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