Atmos. Chem. Phys. Discuss., 11, C2019–C2031, 2011
www.atmos-chem-phys-discuss.net/11/C2019/2011/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Optimizing global CO emissions using a four-dimensional variational data assimilation system and surface network observations" *by* P. B. Hooghiemstra et al.

P. B. Hooghiemstra et al.

p.b.hooghiemstra@uu.nl

Received and published: 19 April 2011

We are grateful for the extensive comments and suggestions given by the reviewer. Below we give answers to all comments. We indicated if we changed the text in the manuscript. As indicated in more detail below, major changes have been made to Section 4, where we added a paragraph to elaborate on the observation error. In short, we found that that the main reason for the rejection of the large amount of observational data is the coarse model resolution, in which both high and low observations are poorly reproduced. A detailed study of the model representativeness error showed that it remains extremely challenging to define a model error representation that works well

for all stations.

Furthermore we changed the discussion paragraph on separation of the sources (§4.1). The main conclusion is that the available observations constrain total CO emissions and therefore, the uncertainty reduction in the total emissions is larger than in the individual emission categories. Hence, in the posterior solution, the source categories are negatively correlated. We now more clearly emphasize that this implies a solution space in which the different source categories can hardly be separated.

Since the topic of the current paper is a thorough testing of the 4D-VAR system, we acknowledge current shortcomings and apply improvements in future work.

Answers to specific comments

1) Page 344, lines 10-12: The statement that "the main sink of CO is the reaction with OH, the co-called cleansing agent of the atmosphere" is somewhat redundant with the first sentence, where it is stated that "by reaction with OH, CO influences the oxidizing capacity of the atmosphere."

Changed the statement to: "the main sink of CO is the reaction with OH."

2) Page 347, line 3: The reference for Fisher (1998) here is to an ECMWF seminar. Is this really a valid reference? Also, there is a typo in the date of the reference on page 370. The conjugate gradient approach was introduced decades ago, I would think that there are much more appropriate references for this than Fisher (1998).

We changed the reference "Fisher (1998)" to "Hestenes & Stiefel (1952)" as the first reference to describe the conjugate gradient method.

3) Page 347, lines 11-15: I don't have access to the ECMWF technical report for Fisher and Courtier (1995). How is the Hessian approximated? What is the accuracy of the approximation –i.e. how well does it converge to the true a posteriori covariance? More information would be helpful here since the source uncertainty discussion requires confidence in the Hessian calculation.

The approximation of the Hessian is described in Meirink (2008b). To help the reader in this respect, we added some lines about the convergence (see also the answer to comment 13 below): “In our study, we consider the minimum of the cost function reached when the norm of the gradient of the cost function is reduced by 99%. Typically, less than 30 iterations are needed to achieve this reduction. Although the eigenvalues are not yet converged to 1 by this time, the errors on the scale of a continent seem reasonably converged after a limited number of iterations as shown in Fig. 3.”

4) Page 348, lines 12-15: What is the global, annual mean OH concentration? What is the estimated methyl chloroform lifetime?

The tropospheric, annual mean OH concentration is 1.1 x 106 molecules/cm3. The corresponding methyl chloroform lifetime (total burden/tropospheric loss) is estimated at 4.8 years. These number have been added to the text.

5) Page 350, lines 20-23: The authors claim that their error estimates of 20-48% and 58-72% are realistic for the Western developed world and the developing world, respectively. Do they have a reference for this claim?

The prior grid-scale errors (50% and 250% for the Western developed world and the rest of the world, respectively) are chosen such that the aggregated prior errors for continental-scale regions range from 20 to 100%. We trust the EDGARv3.2 inventory more for the Western developed world and hence these errors are smaller compared to the prior errors for the rest of the world. A similar approach was followed by Stavrakou & Muller (2006), see §3.3 of their paper: “The errors range from 0.35 for sources supposed to be better quantified (i.e., anthropogenic emissions in western Europe, North America, and Oceania) to 0.7 for highly uncertain emission categories (e.g., natural emissions).” In the revised manuscript, we changed the text to: “Therefore, we apply grid-scale errors of 250% of the corresponding grid-scale emission for the developing world (Asia, Africa and South America) and 50% for the Western developed world. With these settings, realistic continental-scale errors are computed for the developing world

(65-75%) and the Western developed world (20-48%) in the range previously used by Stavrakou and Muller (2006)."

6) Page 351, line 3: On what analysis is the 1000 km correlation length scale based? Is there a reference for this?

The correlation length scale of 1000 km is based on the 4D-Var study for methane described in Meirink et al. (2008b). They show that increasing the correlation length increases the region of influence of a measurement; Hence, the aggregation error increases also. However, a smaller correlation length will increase the effective number of variables in the state vector to be optimized. We changed the text to: “For the three emission categories we use a Gaussian spatial correlation length of 1000 km as in Meirink et al. (2008b).”

7) Page 351, lines 3-6: The authors mentioned that they do not expect a pronounced seasonal cycle for anthropogenic emissions. Although this is typically assumed, Petron et al. (GRL, 2004) found that emissions of CO from fossil fuel and biofuel combustion were 30% and 200% greater, respectively, in winter than in summer. Similarly, The a posteriori results of Kopacz et al. (JGR, 2010) show a significant seasonality in CO emissions from North America and Asia. They found that North American emissions were 50% larger in winter than in summer, while Asian emissions were almost a factor of 2 larger in winter than in summer. It would be valuable if the authors were to assess the impact of their assumed temporal correlation on their results. Would a much shorter e-folding timescale significantly change the regional estimates?

The prior inventory (EDGARv3.2) does not provide a seasonal cycle. Therefore, we use a relatively long e-folding temporal correlation length. We present our emission estimates only as yearly totals. The reason for this is that it is already hard to separate the individual sources categories on the yearly time scales. As outlined in the (revised) paper, this is due to the data sparse nature of the station-only inversion. Any seasonal cycle in continental-scale anthropogenic emissions is aliased with seasonal cycles in the NMVOC-CO source and the biomass burning source. This behaviour is also clear from the posterior errors on the yearly emissions. Errors on the monthly emissions are even larger and we found large negative correlations among the subsequent monthly emissions. This indicates that, comparable to the different source categories, also the monthly emissions can not be accurately determined. With the larger data-stream from satellites we hope that this situation will improve.

We changed the text to: “An e-folding temporal correlation length of 9.5 months (0.9 month-to-month correlation) is chosen for anthropogenic emissions. This high month-to-month correlation is used because the prior inventory does not include a seasonal cycle.

8) Page 352, line 1: On what is the assumed 1.5 ppb measurement error based? Is there a reference for this?

This is a good point. Novelli et al. (1998) report a measurement error of about 1-2% for CO mixing ratios < 300 ppb. We selected 1.5 ppb as an error since a typical CO concentration is about 100 ppb. It turns out that using 1.5 ppb might be too conservative as detailed in the discussion in §4.2. We added the reference in the text.

9) Page 352, lines 3-4: A brief explanation of the model error approach used in Bergamaschi et al. (2010) would be helpful for the reader, especially since the Bergamaschi et al. analysis was for CH4 and this paper is focused on CO. A Figure similar to Figure 2 (top panel) of Bergamaschi et al., showing the representativeness error and the overall data uncertainty for a couple of selected stations would be helpful.

The approach used by Bergamaschi et al. (2010) estimates the uncertainty in an observation as follows: the horizontal representation error is approximated using a simple boundary layer model to account for emissions in the box in which the station resides. The vertical representation error is accounted for by the vertical model gradient of modeled CO mixing ratios in the adjacent grid boxes. In the revised manuscript we added Fig. 7 to show the model representativeness error for 3 stations. We also added a

paragraph to the discussion (§4.2) to discuss the observation error in more detail. (see also the answer to comment 10). We changed the text to: "We estimate the model error using the same approach as described in Bergamaschi et al. (2010). First, the impact of local emissions on the simulated CO mixing ratio is accounted for by a simple emission model for observations in the boundary layer. Second, to account for sub-grid variability that can not be resolved, the vertical component of the model error is calculated from the modeled CO mixing ratios in adjacent grid cells. Third, the temporal standard deviation of the modeled CO mixing ratios within a 3 hour window is added to the representation error. With this advanced representation of the model error, we do not account for possible other model uncertainties in vertical transport or the OH field. This will be discussed further in Sections 4 and 5. The model error is usually much larger than the measurement error for stations close to or downwind of emission regions (e.g., Fig. 7). In remote areas in the SH, however, the measurement error is the dominant term in the observational error."

10) Page 352, lines 15-21: I do not understand the justification for throwing out 15-20% of the data as outliers. If those observations represent particular pollution events that the model cannot capture because of the coarse resolution, as the authors claim, then the representativeness errors should account for this. This seems quite arbitrary to me and represents a significant weakness in the analysis. Omitting this much data needs better justification.

We agree with the reviewer that this point indeed requires more explanation. Most of the surface stations used in our inversions represent background sites for which the coarse model simulation is not a severe restriction. But a detailed inspection of the coarse resolution model output shows that it is difficult to simulate pollution peaks that are interspersed with very clean air. Our advanced representation of the model error does work very well for most stations, but underestimates the error on some specific stations. Instead of artificially enhancing the model error for some stations, as is done in e.g. the Carbontracker system (Peters et al. 2010 (Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, Global Change Biology, 16(4), 1317-1337)) we chose to additionally reject outliers from the observational dataset. Another reason for the large amount of rejected points might be too small prior errors on the emissions, resulting in a system that has insufficient freedom to adjust the prior emissions. However, as discussed before, the prior grid-scale errors are chosen to obtain realistic continental-scale emission error estimates. We therefore chose a compromise in which we remove a rather large portion of the data on some stations. We have changed the text in the revised manuscript in the following way:

- in §2.5 (Inversion specifics), we shortly explain the reason for doing the inversions in two cycles.

- in the Discussion (§4.2) we elaborate on the cause of rejection of the data as well as the effect it has on the inferred emission estimates.

11) Table 1: Although removing the outliers results in a better goodness of fit for the a posteriori CO field, removal of the outliers increases the mean a priori bias. For all the stations shown in Table 1, the bias goes from -0.58 ppb to 1.53 ppb. For individual stations such as Alert (ALT), the bias increases from 0.94 ppb to 2.78 ppb. Again, what is the justification for removing the outliers given that doing so means that you are starting the inversion from a more biased a priori state? It would be helpful to see what is the impact of removing the outliers on the inferred sources – i.e. how do the regional source estimates compare in inversion cycles 1 and 2?

We agree with the reviewer that in the second inversion cycle it seems that we start from a more biased prior state. However, by rejecting those observations, the system is capable to fit more observations on other stations: of course for some stations (e.g., ASC, AZR, BMW, CGO etc.) the posterior bias is slightly larger in cycle 2 compared to cycle 1, but much larger reductions in the posterior bias are found for other stations (e.g., BRW, EIC, GMI). To show the effect of removing the outliers on the estimated

emissions, we added the regional emission estimates for the cycle 1 inversion for 2004 in Table 2. The differences between the cycle 1 and cycle 2 emission estimates are discussed in §4.2.

12) Page 353, lines 22-26: Large amounts of satellite data would not necessarily create a situation in which the inversion is not strongly dependent on the a priori. Ultimately, it will depend on the precision of the satellite data. Large amounts of imprecise data are not necessarily better than sparse but precise data.

We agree with the reviewer that “Large amounts of imprecise data are not necessarily better than sparse but precise data.” However, in the future we intend to assimilate both surface network observations and satellite data (with bias correction to account for possible biases in the satellite observations) to further constrain the emissions. First results using MOPITT V4 observations only, indicate that the emissions are better constrained in regions where surface network observations are sparse, compared to the stations-only inversion.

13) Figure 3 and Table 2: With the exception of Asia and Europe, there is little reduction in the uncertainty of the anthropogenic emissions. However, Kasibhatla et al. (GRL, 2002) showed much greater uncertainty reduction for fossil fuel emissions from North America, Europe, and Asia in their inversion analysis of the surface CO data. The authors should comment on why their inversion results are so different from those of Kasibhatla et al. (2002).

The inversion approach used by Kasibhatla et al. (2002) was different from ours as it used the ‘big region approach.’ In this approach it is feasible to compute all matrix inverses (see Eq. 3 and 4) and the posterior error covariance matrix will be exact. In 4D-Var analysis as is pointed out in Section 2, the posterior error covariance matrix is still an approximation and hence, the estimated errors will be an overestimate of the true errors. For the base inversion, we iterated on to a gradient norm reduction factor of 1010. Then the posterior error covariance matrix is converged and error reductions of the same amplitude as given by Kasibhatla et al. (2002) are derived. However, those inversions are very computationally intensive (up to 5 times compared to the current). Therefore, we use a more practical reduction factor of 100. We added text in the manuscript: "It is acknowledged here, that the presented posterior error reductions are much smaller compared to the big region synthesis inversion studies (e.g. Kasibhatla et al. (2002)). However, the differences are mainly explained by the inversion approach used. A synthesis inversion approach optimizes the emissions for a set of big regions. In such a framework, the posterior emission estimates and their errors can be computed by a direct matrix inversion and hence the posterior errors are exact. In the 4D-VAR framework presented here, the cost function is minimized iteratively and considered converged when the norm of the gradient is reduced by a factor 100 (or 99%.) As a special case we continued the iterative process up to a gradient norm reduction factor of 10ˆ10. For this case the approximation of the Hessian of the cost function was converged to the true Hessian. The resulting posterior errors are indeed close to the numbers in Kasibhatla et al. (2002) (not shown). However, a gradient norm reduction factor of 10ˆ10 is not very practical as the computational burden increases up to a factor 5. It should be kept in mind that the error estimates calculated by a 4D-VAR approach always represent an upper limit."

14) Page 358, lines 13-17: The authors claim that the overestimate in May and June is due to an overestimate of the NMVOC source in the a priori, since the inversion particularly reduced the NMVOC in these months. However, as the authors acknowledge on page 355, lines 23-46, the data do not constrain the NMVOC source well. Furthermore, a tight a priori constraint was imposed on the NMVOC source in the inversion. As a result, the NMVOC source should have remained close to the a priori. Also I find it suspicious the bias is confined to just May and June, given that the summertime maximum in the NMVOC source is broad and does not peak as early as May. It seems likely that the large reduction in the NMVOC source reflects the impact of other biases being projected onto the NMVOC source. Indeed, the authors noted on page 361, lines 15-18, that the decrease in the inferred NMVOC source from 2003 to 2004, could be

an artifact of the inversion.

The prior NMVOC-CO is very high in April/May, compared to the other months and the error settings are quite strict. However, the observational part of the cost function can be largely reduced by reducing the NMVOC-CO parameters in May and June, with only small costs in the background part of the cost function (because it is only 1 parameter). We agree with the reviewer that this can result in a projection of other biases on the NMVOC source: the 4D-Var system is clearly not capable to differentiate between the sources, and the decrease in the inferred NMVOC source from 2003 to 2004, is likely an artifact of the inversion. To bring this important message more clearly to the reader, we changed the text of the discussion about this issue in paragraph §4.1.

15) Page 359, lines 14-16: I disagree with the claim that the “comparison with NOAA aircraft profiles showed however that the vertical transport in TM5 is reasonable.” The aircraft data were mainly in the northern hemisphere, whereas the bias with respect to MOPITT is pronounced in the southern hemisphere (at middle and high latitudes). In the absence of more independent aircraft data in the southern hemisphere, over a range of longitudes, one cannot rule out a bias in vertical transport in the model.

We agree with the reviewer and will weaken our claim. New text: “The inversion is capable to improve the comparison with independent observations in the free troposphere over North America.”

16) Page 359, lines 18-19: It is not clear to me how any putative issues with the MOPITT retrievals over deserts is relevant here? Indeed, the bias seems to be larger on land over the deserts, where one would expect the retrievals to be challenging, but the model-MOPITT bias is largest over the oceans in the southern hemisphere. Furthermore, the CO inversion of Jones et al. (ACP, 2009) showed that both TES and MOPITT data resulted in an overestimate of surface CO in the midlatitudes of the southern hemisphere. It is unlikely that the TES and MOPITT data are similarly biased. Can the authors comment on this? My guess is that the model used here as well as the model

used in Jones et al. are biased in their vertical transport.

We agree with the reviewer that the largest discrepancies are found over the remote SH. However, we also discovered an error in the computations. In the revised manuscript we changed Fig. 5. The largest discrepancies are now found over the Southern ocean, but the error also pertain over desert regions, e.g., the Sahara desert. The latter was also observed by de Laat et al. 2010. We further agree with the reviewer that the discrepancy on the SH may be caused by the vertical transport in the model (as also observed by Jones et al. 2009), but the transition from ocean to land at the west coast of Africa remains suspect and might indicate some retrieval issues in that region. Our ongoing studies with inversions using MOPITT data seem to confirm this finding.

17) Page 361, lines 23-26: Are the correlation coefficients of -0.29 and -0.23 statistically significant?

The point here is the following: The assimilated observations only constrain the total CO emissions (that is, the sum of all categories) and therefore the uncertainty in the total CO source is smaller than the uncertainty in the individual components (the source categories). So the inversion results in the 'uncertainty-ellipse' for the relation between the anthropogenic and natural source shown in Fig. 1. The form of the ellipse confirms that the system is capable to constrain total CO emissions, but has difficulties separating them. This is indicated by the negative correlation, shown in the figure by the tilted axes of the ellipse. We emphasize that this correlation does not mean that there is any reason to believe that in reality these emission categories are correlated in any way, it is only the posterior outcome of the current inversion system. Basically, the system is only capable to distinguish sources with different spatial and/or temporal patterns that are specified in the prior emissions/uncertainties. We clarified this in the discussion in §4.2.

18) Page 363, Section 5.1: It would be helpful to show how the regional estimates

respond to the different sensitivity tests. Showing only the global totals in Table 5 is less informative.

We agree with the reviewer and added a table in the revised manuscript (Table 4).

19) Tables 4 and 5: The error for the NMVOC-CO source in Table 5 is 10% for S3 (81 out of 812 Tg CO), whereas in Table 4 the error is listed as 16%. Which is correct?

The prior error on the monthly NMVOC-CO parameter is set to 16%. However, due to temporal correlations, the error aggregated over the whole year is reduced to 10%.

20) Page 364, lines 1-4: I do not understand the discussion here. What do the authors mean when they say that the "in sensitivity study S1 the NMVOC-CO prior error dominates, resulting in less reduction in this source"?

We removed the section on sensitivity studies S1 to S4 since the whole discussion about compensation between the sources is now treated in §4.1.

21) Page 366, line 3: Please change "increase with 75 Tg CO" to "increase by 75 Tg CO

Changed

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 341, 2011.

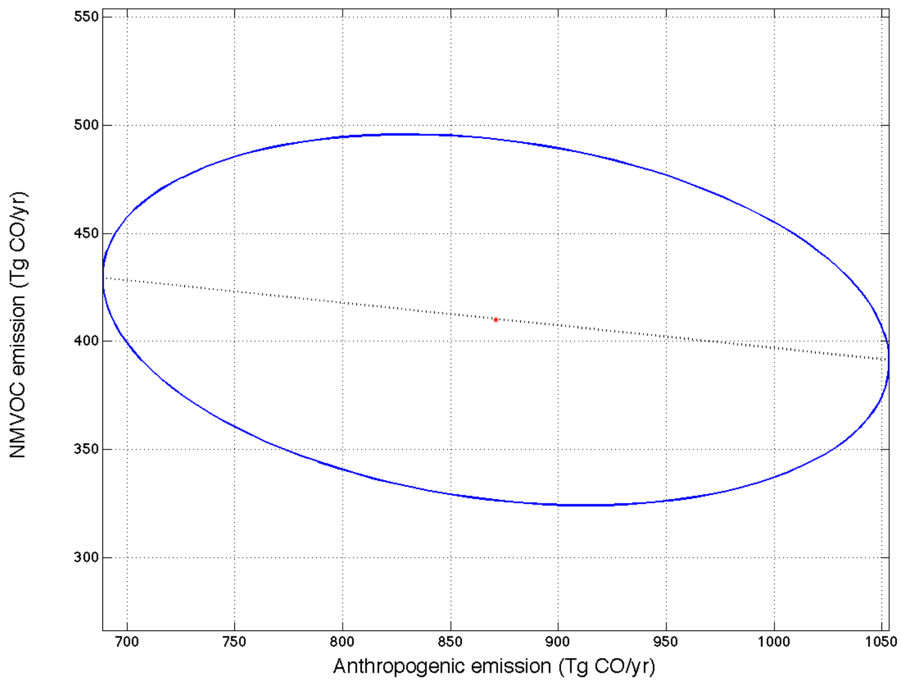

Fig. 1. 95% confidence ellipse for the global total annual emissions for the anthropogenic and NMVOC-CO source for 2004. The center of the ellipse is the posterior emission estimate.