General Comments

More Widespread Discussion of Error Estimates:

We agree that we could have done a better job discussing our results in relation to input and output uncertainties. We went through the text and figures and tried to improve this. Uncertainty is often neglected in inverse model studies, and we feel it's an important diagnostic.

Text:

We noted the estimated uncertainties throughout the sections that discuss the emissions estimates, including the Abstract and Conclusions. Where possible, we tried to interpret our results in the context of the posterior uncertainties and how they change from priors. We also added a short paragraph to the end of section 2.1 about the posterior covariance:

As noted above, the posterior covariance matrix is approximated by using the posterior parameter deviations. Temporal covariance is limited to the period spanned by the assimilation window. Therefore, time aggregated quantities, such as annual uncertainties will likely be overestimates since information about temporal covariations will be limited. Furthermore, as with any inversion, the error covariance matrix ultimately reflects the relative weighting between the model-data mismatch errors and prior emission uncertainties that are specified.

Figures:

Regarding the zonal average figures (8-11), the only error bars are on the observed curves and they are very small. They are calculated using a bootstrap method that quantifies the effect of spatial distribution of observation sites on the global (or zonal) average. We pointed this out in the captions.

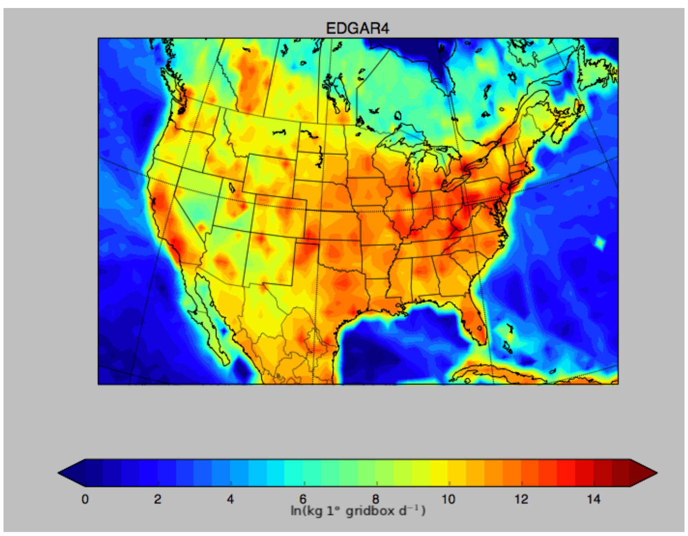
For the histograms (or bar graphs) we felt that adding the error bars for the total emission estimates would further confuse what are already not very easy to interpret plots. We mentioned the average uncertainty estimate in the caption so that readers can compare to the bars. Also, we pointed out that the errors represent the 1-sigma error bounds in each caption.

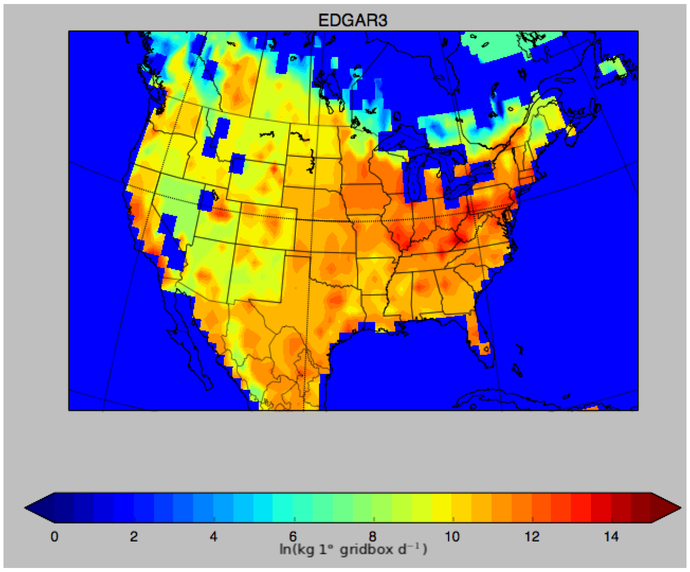
We added error bars to Figure 18 that shows the tropical flux anomalies. While it's true that the error bars are quite large, they are not large enough to disqualify what we say in the text.

Use of Different Versions of EDGAR:

We agree that the underlying distribution of anthropogenic emissions is an important factor in the result regarding apparent increases in fossil fuel emissions for North America. Below we show comparisons of EDGAR 3 and EDGAR4.2 for Jan. 2000 (recall that we use constant emissions). It's clear that the emissions are higher for EDGAR 4.2, especially in the area of interest, and the global total is about 30 Tg/yr higher for EDGAR 4.2.   We will likely explore use of different anthropogenic emissions in future versions of CarbonTracker-CH4, however, our aim in this paper was not to test bottom up emissions inventories, but to see what the effect of keeping emissions constant would be. We were especially interested in the sensitivity of our observing system to quantifying anthropogenic emissions. The lower global total prior we use (along with the fact that it doesn't grow over time) plays a role in the low bias we have globally, but there are also significant uncertainties in natural emissions and the chemical loss.

We added a statement about the possible implication of using low prior anthropogenic emissions to section of the paper that describes North American emissions.





The Shift of Emissions from High Northern Latitudes to the Tropics and Southern Mid-latitudes:

Work on improving the N-S transport in TM5 is ongoing and some possible improvements are being tested. The possibility that TM5 underestimates emissions at high latitudes because it traps emissions at the surface has been a concern, however, comparisons of inversions included in Kirschke et al. (2013) show that results obtained using the TM models are not that different from results obtained using non-TM models.

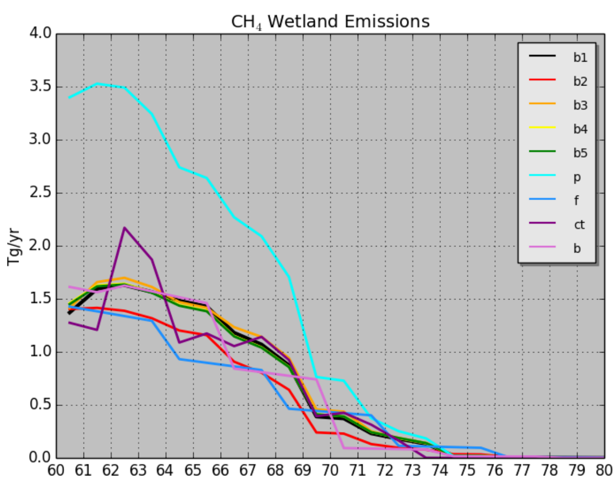


Figure – Comparison of inversions included in Kirschke et al (2013) for high northern latitudes. Note that the TM inversions are violet and purple, while the other inversions use non-TM5 models. (The x-axis is latitude).

ERA-I vs OD Meteo fields:

Early in the development of CarbonTracker (CO_2) a decision was made to use the OD met data. At that time, we did not have ERA-I and other reanalyses did not cover the time span we were interested in. Also, we hoped that we could keep the assimilations very current, although time lags in the availability of observations in practice have still meant that we lag real time by at least 1 year. For example, we don't receive samples from the South Pole during austral winter. Now that ERA-I is available, it will be our first choice for future simulations, however, comparison forward simulations suggest that differences between ERA-I and OD for CH_4 at surface sites is very small, both before and after the change in the vertical levels. Assimilations run with both met data sets for CarbonTracker (CO2) produce virtually indistinguishable results in estimated fluxes (these results may be found on the CarbonTracker (CO2) web site). On the other hand, comparisons with high altitude measurements obtained using the aircore technique suggest that the high-altitude structure of CO2 is represented better with ERA-I, so this product will be a better choice for multi-decadal inversions.

Initialization of the Assimilation:

We neglected to mention how we initialized the assimilation, so we thank the reviewer for pointing this out. We don't believe that the initial conditions propagate very far into the time series of estimated emissions based on synthetic data tests. We added the paragraph below to the end of the first part of section 2:

We initialized the assimilation using an equilibrated distribution produced by another TM5 run that was scaled to match observed zonal average CH_4 mixing ratio for the year 2000. The north-south gradient therefore should represent the observed atmospheric gradient at the surface. Sensitivity runs using synthetic data (not shown) suggest that spin-up effects are restricted to within in the first half year of the assimilation.

As for why the low bias compared with observations occurs, we believe that this is due to prior emissions being too low (indeed this is by design in the case of the anthropogenic emissions) or the loss being too high. In other words, the prior sources and sinks will cause a lower equilibrium value than what the atmosphere would approach if sources and sinks don't vary.

Specific Comments

P2179, line 10: There is no mentioning of the time dimension of the state vector. I presume the 121 refer to a single month?

We added a sentence describing the weekly time step and the assimilation window of 5 weeks.

P2179, line 19: What is meant by satellite observed "hot spots"? Fire counts, burned area, ..?

We clarified this statement to read: “The final terrestrial emission category is biomass burning, which is treated as a separate category due to the existence of strong spatial constraints coming from satellite observations of locations of large fires.”

P2180, line 11: This argument is more often used to justify short assimilation windows. I wonder, however, if there is any evidence of transport model errors accumulating over time. One may argue also that errors representing synoptic scale variations may dissipate on longer time/spatial scales that are better resolved by the course resolution transport model. Much of the observational constraint that inversions make use of come from larger scale mass balances. By reducing the response functions, this signal may end up being aliased to shorter scales. It is difficult to quantify the significance of this, but a more careful formulation seems needed here.

We added a reference to this issue of accumulated transport errors. We tend to agree with the reviewer that transport errors will likely cancel over long temporal and spatial scale, however, this issue has been the subject of heated debate within our group! CarbonTracker was originally developed to treat a dense observational network, however, budget issues ultimately meant far fewer sites that we had originally hoped for. In meantime, computational limitations and changes in input met fields discouraged us from lengthening the assimilation window. Future versions of CarbonTracker-CH4 will have longer assimilation windows, however.

P2183: Since Bergamaschi et al 2007 refers to an inversion, a reference is needed of where natural wetlands emissions come from that where used in that study (or the model that was used to generate them).

We used the wrong reference here – it’s Bergamaschi et al. (2005). We also added more details about this prior: the wetland prior is based on the distribution of Matthews and Fung (1989) and the emission model of Kaplan (2002).

P2187, line 15: Which global model is ‘a global model’?

We obtained the OH fields from Krol et al., and these were produced by a full-chemistry version of TM5 that was adjusted to agree with methyl chloroform. We clarified this in the text.

P2188, line 18: The model resolution of 6x4 degree seems more relevant here than the 1x1 degree of the emission inventories. Besides this, the inversion doesn’t allow changing small-scale emission patterns. It makes me wonder how valid it is to include tall tower measurements in the analysis. An additional error on top of the representation error seems needed here.

We realize that the resolution is an issue (not only for this inversion, but for many others that simulate transport at relatively coarse resolution). The same issue about representing local sources with such a transport model also applies to the use of continental air samples taken using flasks since these are essentially point measurements and not even afternoon averages. Using the background sites is safer, but then the observational constraints in the inversions would be restricted only to the largest scales. We chose the model-data mismatch errors to be large enough to account for uncertainty coming from the transport model and the assumed local distribution of sources based on both forward runs and posterior differences from observations. We find that with the exception of sites that are near strong local sources, we generally get posterior residuals (simulated – observed) that are within our model data mismatch. Our web site shows figures at all sites, but these results are summarized in Table 2 in the 'bias' column.

Some figures are either quite small (7 and 10). This is true also for Figure 9, but that one doesn't seem to provide much information and could probably be left out.

We separated Figure 7 into two separate figures and revised Figure 9.

Technical Comments:

We thank the reviewer for pointing out these problems, and we particularly regret our oversight with the Figure 5 caption! We fixed all of these problems.