Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-182-RC2, 2016

Geoscientific Model Development Discussions

Interactive comment on "Development of CarbonTracker Europe-CH_4 – Part 2: global methane emission estimates and their evaluation for 2000–2012" *by* Aki Tsuruta et al.

Anonymous Referee #2

Received and published: 6 October 2016

Overview:

This paper presents an evaluation and interpretation of CarbonTracker Europe-CH4 (CTE-CH4). CTE-CH4 estimates global methane emissions by assimilating in-situ observations of methane mole fractions. It estimates scaling factors to prior emissions on the regional scale (order of 1000 km). It covers the time period 2000-2012, which includes a period of fairly constant global mean methane concentrations (2000-2006) and a period of growing global mean methane concentrations (2007-2012).

The model evaluation includes a comparison of the model with two different state spaces (sizes and shapes of scaling factor regions), as well as two different vertical mixing schemes. The work reported good agreement between predictions made by

the posterior model and independent data from aircraft profiles, TCCON, and GOSAT - with the best agreement found using a vertical mixing scheme from Gregory et al., 2000 and a state space with fewer elements. However, there were biases in the comparison to some of the independent data, as well as in the model residuals.

The model interpretation focused on trends in emissions from each of the optimized regions and the sensitivity of the estimates to the choice of state space and vertical mixing scheme. The authors attribute the renewed growth of global mean methane concentrations to increased anthropogenic emissions in the South American temperate region, Asian temperate region.

I think that there is some scientific potential in this work, but the article requires major revisions in order to realize that potential. I found the presentation to be below publication quality, and I had a few moderate but material concerns about the scientific quality.

I had to work far too hard to understand the work and its importance, which I attribute to a combination of poor communication of the message, a generally low quality of writing, and flaws in the figures.

My scientific concerns are:

1. that the model assumed a fixed lifetime for CH4 even though the authors explicitly acknowledge that this assumption is unlikely to hold. It is very important that the authors qualify any reported results from this model with this assumption.

2. that the inversion violates the assumptions that form its foundation in a way that likely aliased biases in the posterior emissions estimates.

3. that the prior and model-data mismatch uncertainty estimates appear to be arbitrary, and that no tests (e.g., reduced chi-squared statistic) were given to demonstrate that they accurately reflect the actual uncertainty distribution.

4. that the model evaluation examined only the maximum a posteriori estimate of the

inverse model and did not give an assessment of the uncertainty estimates (similar to point 3). This paper evaluates a model that generates a statistical distribution as output – that distribution should be evaluated in its entirety.

I think that point 1 could be dealt with by adjusting the language of the paper and more fully acknowledging the shortcomings, and that points 2, 3, and 4 could be dealt with by revising the uncertainty characterization and model evaluation analysis.

Detailed Comments:

Scientific Significance:

This article presents an evaluation and interpretation of an inverse modeling system designed to attribute the renewed growth in the global mean methane concentration since 2007. This is an important problem, and the literature on the subject is quickly evolving.

This work makes an incremental advancement by testing CTE-CH4. Due to the scientific concerns I raise in the Overview and Scientific Quality sections of this review, I would not believe that the findings of this work uncover the cause of the renewed growth with scientific rigor. However, if CTE-CH4 is further developed it might be a useful tool in contributing to the solution of the problem.

Scientific Quality:

I have reservations about the scientific quality of this work, but most of them could be addressed by refining the language to be more forthright about the limitations of the model and by revising the uncertainty characterization and model evaluation metrics.

My issues are as follows:

1. The authors assumed a lifetime of CH4 that varied seasonally, based on a one-year climatology of hydroxyl concentrations. They assumed that the CH4 lifetime had no inter-annual variability. This assumption is not safe, and the authors acknowledge it -

citing Rigby et al., 2008, Ghosh et al., 2015, and Dalsoren et al., 2016 (lines 21–31 on page 5). The atmospheric lifetime of CH4 likely changed during the study period, and error in the lifetime of CH4 would be directly aliased onto the retrieved emissions.

Additional evidence for changing OH concentrations and therefore changing atmospheric lifetimes of CH4 can be found in:

- McNorton J et al. (2016) Role of OH variability in the stalling of the global atmospheric ch4 growth rate from 1999 to 2006. Atmospheric Chemistry and Physics 16(12):7943–7956.

- Montzka SA et al. (2011) Small interannual variability of global atmospheric hydroxyl. Science 331(6013):67–9.

- Prather MJ, Holmes CD, Hsu J (2012) Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry. Geophysical Research Letters 39.âĂÍ

- Rigby M et al. (2013) Re-evaluation of the lifetimes of the major CFCs and CH3Cl3 using atmospheric trends. Atmospheric Chemistry and Physics 13(5):2691–2702.

One would not expect the authors to implement a changing atmospheric lifetime at this point, but the issue should be treated very seriously and highlighted as a future research need. The caveat of a fixed methane lifetime should appear in the abstract.

2. The inversion setup violates one of the fundamental assumptions from which it is derived in a material way that leads me to doubt the validity of the conclusions. An inversion of this sort assumes that the error in the prior is a second-order (a.k.a. weak-sense) stationary Gaussian random process with zero mean.

The authors use the EDGAR 4.2 FT2010 emissions inventory as a prior anthropogenic emissions field. This is a high-resolution (0.1x0.1 degree) inventory that is known to be (and demonstrated in the paper to be) biased in its spatial distribution over a broad spectrum of scales and also biased in its temporal trend.

They set the prior error variance for the total emissions from any region to 0.8. The assignment is arbitrary and very likely too high. The authors effectively eliminate the bias by over-estimating the random error in the prior. Still, the posterior estimate is at the extreme of the error bounds, and so the bias in the prior affects the posterior estimate. This is visible in Fig. 3, where biases in the latitudinal distribution and seasonal cycles are visible in all posteriors.

Additionally, the scaling factors are resolved at spatial scales of thousands of kilometers. The error in the prior varies at scales much smaller than this – producing a severe representation error. The problem is therefore likely under-parameterized (equivalent to having covariance lengths that are too long, or regions too large), and so adequate scaling factors cannot be derived that permit unbiased residuals at individual sites. As a result, the authors find strong biases in the residuals, and even throw out some sites.

An example of such sites are given in the paper, and the authors remove them:

"Strong negative bias as found in Bukit Koto Tabang, Indonesia (BKT) (-25 to -27 ppb) and Mt. Kenya, Kenya (MKN) (-18 to -23 ppb), such that the posterior mole fractions were especially low during June-October. This suggests that the measurements at those latitudes are not representative of large regions optimized in the model."

To solve this problem, the authors would need to perform the inversions at high resolution using covariance length scales constrained as part of an objective error characterization.

3. There is no objective justification given for the uncertainty used in the study. The authors state: "Variance of the scaling factors was set to 0.8 for all regions, except for 'Ice' region (Fig. 2)" (sic).

It is important that the error statistics of the combined prior and model-data mismatch error match the error statistics of the actual mismatch between the prior model and the observations.

For more information on this problem, please see:

Michalak AM et al. (2005) Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions. J. Geosphys. Res. 110:D24107.

4. The model evaluation examined only the maximum a posteriori estimate of the inverse model and did not give an assessment of the uncertainty estimates (similar to point 3). This paper evaluates a model that generates a statistical distribution as output – that distribution should be evaluated in its entirety.

The analysis of residuals and independent data should test whether the posterior model-data mismatches follow the expected distributions.

Scientific Reproducibility:

This work is reproducible – in fact there have already been other studies that have conducted similar experiments and come to similar conclusions.

Presentation Quality:

This work is very poorly presented. My reservations about the presentation fall into three broad categories: 1) messaging, 2) writing, and 3) figures.

1. Messaging

This work is presented as model evaluation and interpretation. The paper goes into great detail about the variations in every region, giving their trends, comparisons to other regions, and comparisons to other papers. The work needs to be boiled down to a set of key messages. My understanding is that the main messages are those described in the Overview section of this review.

The body of the paper needs to be focused on providing the scientific justification for the given messages.

2. Writing

The paper requires extensive revision by an English language editor. Problems include:

- incorrectly cased letters (e.g., "south America" should be "South America"). - inconsistent tenses and active vs. passive voice (e.g., in the abstract, line 29 "We use three configurations. . .", then line 32 "The posterior estimates were evaluated. . ."). - broken sentences (e.g., page 3, line 22 "To estimate biospheric emissions, information from an underlying ecosystem distribution map is useful, which defines the location of the sources and can help distribute larger regions over which the atmospheric signals integrate."). - truisms (e.g., page 9, line 20 "The growth rate (GR) of atmospheric methane mole fractions showed that the posterior estimates are closer to the observations than the prior, as expected"). - paragraphs that are incredibly long and rambling (e.g., page 14, line 7 to page 15, line 6; page 16, line 13 to page 17, line 5; and page 18, lines 5 - 30).

An exhaustive list of problematic sentences would be too long for this review.

3. Figures

The figures in this paper have a number of issues. Points a-c absolutely must be addressed in order for the paper to be publishable.

a) Figs. 9, 10, 11, 14, S4, and S5 include data and/or error bars that run off of the figure.

b) Many of the figures include error bars but do not specify their meaning (1 standard deviation? 95% credible interval?)

c) Figs. 4, 5, 6, 7, 8, 9, 12, 13, S1, S2, S3, and S6 are not colorblind safe.

d) Figs. 6 and 7 are difficult to read because of closely placed points.

e) Fig. 4 (top panel) should include observations.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-182, 2016.