Response to re-review of Anonymous Referee #1 and Co-Editor Decision letter for ACP manuscript No.: acp-2017-273 [Peter Bergamaschi et al., Inverse modelling of European CH₄ emissions during 2006-2012 using different inverse models and reassessed atmospheric observations]

5 This document describes the detailed point-by-point response to re-review of Anonymous Referee #1 and Co-Editor Decision letter by Jens-Uwe Grooß (13 Nov 2017)

(1) *comments from Referee / Co-Editor [in italics]*

(2) **author's response [bold]**

(3) **author's changes in manuscript [bold blue].**

10

1. Response to re-review of Anonymous Referee #1

I do not find the changes between this version and the original manuscript to be compelling enough to change my
15 *original stance on this manuscript. It is not clear to this reviewer that this manuscript contributes much to the current literature.*

We do not agree with this statement, since we had provided a point-by-point reply addressing all points raised by all 3 referees and revised significantly the manuscript.
20 **We believe that our paper makes a significant contribution to the literature in the field of inverse modelling of European / regional CH₄ emissions and includes many new elements compared to the previous [*Bergamaschi et al., ACP,* 2015] paper:**

(1) use of the new quality-controlled and harmonized InGOS CH₄ in-situ data set including 18 stations. To our knowledge this is the largest (and most consistent) data set of European atmospheric measurements used
25 **in any study of top-down estimates of European CH₄ emissions so far.**

(2) extension of time series (now 7-years period 2006-2012, compared to 2-years period in [*Bergamaschi et al.*, 2015], including a short analysis of CH₄ trends.

(3) detailed analysis of potential contribution of natural sources

(4) comprehensive validation of model results using an extended set of aircraft observations, providing for
30 **the first time quantitative estimates of potential biases in derived regional emissions.**

Point #1 ("I find the wetland hypothesis wholly unconvincing"): The authors have added one sentence to the abstract and one sentence to the conclusions.

We did not only include the additional sentences in the abstract and the conclusions, which clearly emphasize the uncertainties of the wetland emissions, especially regarding their spatio-temporal emission distribution (furthermore these uncertainties are discussed in sections 4.1 (already in the original discussion paper)). Following the suggestion of this referee we had included an analysis of the seasonality of derived CH₄ emissions in scenario S3 (which is not using any detailed a priori inventory nor any a priori seasonal cycle; see section 4.1 and new Figure 5S, which confirms that the derived emissions are driven by the observations and not by the a priori inventories). Furthermore, we had addressed the further specific point raised by this referee on this topic, especially the potential impact of the background (see section 4.2, new Figure 14S).

Point #2 ("the methods description is poor, making it hard to gain any insight from the different inversions"): The methods description is still poor. There is still just a single paragraph in the main text describing the inversions.

As explained in our response to the reviewers, the inverse modelling system are described in the supplementary material (SM), section 1 "Atmospheric models" (summarizing the main elements of each system; this description extends over more than 4.5 pages). Furthermore, all seven inverse models are described comprehensively in separate specific papers (see references in the SM).
Since for most models used in this study only smaller updates were applied (compared to previously published applications), we think that is more appropriate to keep the summarizing description of the individual models in the SM.
However, following the request of this referee we had updated the description in the SM (e.g. the applied a priori probability density functions (pdf's) for the individual models, and the assumed uncertainties for the observations (including estimates of model errors). Furthermore, we had included the applied boundary conditions (background) in Table 3 and the information about the optimization of the background in the revised main paper.
Following also the request of the co-editor, we included further model details in the main paper (section 3.2 / table 3), as detailed below (see under "2. Response to Co-Editor Decision letter by Jens-Uwe Grooß (13 Nov 2017)")

It is left to the reader to guess at why the inverse models obtain, in some cases, radically different emissions. Figure 2 is a good example of this. The emissions from COMET look totally different from the others (e.g., why is there a source in Northern Poland that isn't in the prior or any other model?). Figure 2 seems to only be mentioned a single time in the manuscript (in the first paragraph of Section 4.1).

A significant part of the 'visual' differences in the spatial patterns between different models is related to different spatial resolutions (and, as explained in the text, for NAME related to the averaging of emissions at larger distances from the observations). Integrated over larger areas (e.g. whole EU-28), the models show a remarkable consistency (apart from the generally lower CH₄ emissions derived by NAME). Differences on smaller spatial scales are probably partly due to differences in model transport and different weighting of

2

the observations (i.e. different assumptions of model-data mismatch errors), but may reflect to some extent some noise of the inverse modelling systems.

In order to address this specific new comment of the reviewer, we have added the following short additional paragraph in the manuscript (in section 4.1):

5   **"Apart from this specific feature of the NAME model, also some further differences in the spatial patterns derived by the different models are apparent. One example are the relatively high emissions derived by the COMET model in north-western Poland / north-eastern Germany. Such differences on smaller spatial scales are probably partly due to differences in model transport and different weighting of the observations (i.e. different assumptions of model-data mismatch errors), but may reflect to some extent also some noise of the**

10   **inverse modelling systems."**

In the author's response they state: "But this [understanding the differences between top-down emissions] is actually not the goal of this study (and would require further specific modelling experiments). The objective of this study is to use the model ensemble to provide more realistic overall uncertainty estimates (from the range of the

15   *inverse models) and to evaluate the model performance by validation against independent observations." Was this not already done in the 2015 paper by many of the same authors (Bergamaschi et al., "Top-down estimates of European CH4 and N2O emissions based on four different inverse models", ACP, 2017)? The authors have just added in a few more models and a few more years of data.*

20   **We assume that the referee refers here to our [*Bergamaschi et al.*, ACP, 2015] (and not 2017) paper. As outlined above we think that the current paper provides many new elements and provides a significantly extended and improved analysis of European CH$_4$ emissions. What the referee calls "just added... a few more years" is a significant extension of the time period (covering now the 7-years period 2006-2012, while in [*Bergamaschi et al.*, 2015] only the 2 years 2006 and 2007 were analyzed).**

25

In general, the conclusions drawn do not seem to be in-line with the analysis. For example, the conclusions of this manuscript state (final paragraph): (2) transport models need to be further improved, including their spatial resolution and in particular the simulation of vertical mixing, aren't some of these models finer-resolution then others and include different treatments of vertical mixing? Do these models actually perform better?

30

We had presented detailed conclusions in-line with our analysis in section "5. Conclusions". The referee refers here to our more general conclusions at the end (which are rather recommendations to further improve top-down estimates in the future). In our study the models with higher spatial resolution do not perform better than TM5-4DVAR with resolution of 1x1 degrees (see short discussion at beginning of section

35   **4.2). Nevertheless we consider the further development of high-resolution models essential to further improve the top-down estimates (e.g. [*Henne et al.,* 2016]).**

Point #3 ("it's not clear to this reviewer that their "novel" approach to estimate bias is actually an advancement"): Regarding the 3rd point, their "novel" approach to estimate bias is not particularly useful for estimating biases (as the authors claim). The difference between simulated and measured enhancements is the term that defines the model-data mismatch in the cost-function. As I showed in my previous review, $c_{obs} - c_{mod} \equiv \Delta c_{obs} - \Delta c_{mod}$. Following

5   *this, if there were no difference between the simulated and measured enhancement then the inverse model would not deviate from the prior. Stated another way, if $\Delta c_{obs} = \Delta c_{mod}$ then the top-down emissions would be equal to the prior.*

As explained in our response to the reviewers we do not agree with the statement of the referee. The term

10   **mentioned by the reviewer is indeed part of the cost function, which the inverse modelling systems aim to minimize. However, this applies to the observations that are actually used (assimilated) in the inversions, while we are analyzing here independent observations that were <u>not</u> used in the inversion - which is a common method to validate inverse models. The concept of using independent observations for validation of inverse models is described in detail in the recent review paper by *Michalak et al.,* [2017] and has been**

15   **applied in many studies (e.g. [*Alexe et al.*, 2015; *Monteil et al.*, 2013; *Houweling et al.*, 2014; *Bergamaschi et al.,* 2013]). The rationale behind this approach is to analyze, how well the inverse model perform in areas which are less constrained by the observations. As mentioned in our initial reply to the reviewer, and explained in section 4.2 of the manuscript, especially the validation of the vertical profiles (against independent aircraft profiles) is very important, since the inverse models assimilate only surface observation.**

20   **Therefore, potential errors in the vertical mixing of the models can introduce significant biases in the derived emission.**

Commonly, however, such comparisons against independent observations are performed to diagnose only qualitatively, if the inverse models have biases. The novel aspect of our method is that we provide for the first time quantitative estimates of the derived regional emissions. We had included in the revised version

25   **also an analysis of the correlation between the derived relative bias (rb_{BL}) and regional model emissions around the aircraft profile sites, which confirms that rb_{BL} can be used to diagnose biases in the regional model emissions (new Figure 16S).**

Since we consider Point #3 as absolutely invalid, no changes have been made in the revised version regarding this point.

30

2. Response to Co-Editor Decision letter by Jens-Uwe Grooß (13 Nov 2017)

1. It is not clear to me, why one model that seems to be always at the low end of the deduced CH4 emissions: the Lagrangian NAME model. Is that because of the Lagrangian formulation of the model? Or because the lack of observation stations?

35   **The low emissions of the NAME model are likely - at least partly - related to the positive bias in the background CH_4 used for NAME (as discussed in section 4.2 and shown in the figures 14S (for the aircraft data), since the regional models invert the difference between the observations and the assumed background.**

In fact, also at most continental atmospheric monitoring stations (which are applied in the inversion), the background used for NAME (and TM3-STILT) is significantly higher (~10 ppb) compared to the TM5-4DVAR background (Figure 15S).

The potential impact of the significant positive bias of the background in NAME (and TM3-STILT) on the derived CH$_4$ emissions is also mentioned in the conclusions of the revised version.

It would be certainly useful to perform additional tests with the NAME model using the TM5 baselines in order to quantify the impact of the baseline on the emissions. Unfortunately, however, the NAME team was not yet able to perform this additional test, since they had no further resources after the end of the InGOS project.

10

We don't see any obvious reason why the Lagrangian models should yield lower emissions compared to Eulerian models (however, to our knowledge, this has not been investigated in any study in a systematic way). In fact, also STILT is a Lagrangian model and yields similar emissions as the global / regional Eulerian models.

15 **Regarding the observations: All models use the same observational data set (however, the details, how the observations are used in the inversions, differ, in particular the assumed model-data mismatch error, and hence the weighting of the individual observations in the different models).**

Short general description of the applied the assumed model-data mismatch errors has been included in section 3.2 (see below)

20

*2. Can you understand that the annual cycle of wetland emissions is reduced in Northern Europe with respect to the WETCHIMP study? and opposite than the annual cycle increases in the three other parts? Would you say that the WETCHIMP study is incorrect in that respect***?**

25 **We can only speculate about the potential reasons. E.g. the CH$_4$ emissions in the wetland models are highly sensitive to the assumed assumed temperature dependence (Q$_{10}$ values), but also on water table and soil properties (in particular, soil organic carbon content). We note that for Northern Europe the seasonality in the a posteriori emissions derived by TM5-CTE are actually very similar compared to the WETCHIMP mean / median. Nevertheless, it is obvious that the other models derive smaller seasonal cycles in Northern**
30 **Europe. For the other regions in Europe the derived seasonality is still broadly within the minimum-maximum range of WETCHIMP (even though the seasonality in the mean/median of the WETCHIMP is clearly smaller).**

We would not say that WETCHIMP is 'incorrect', but clearly the uncertainties are very large, as reflected in the very different spatio-temporal emission patterns of the different individual WETCHIMP inventories.

35

To Point #2 of the review:

Although the model descriptions are updated in the supplement, I would likely ask you to add some more information to section 3.2. /table 2, such that in the paper the model diversity would be better understandable

without reading the supplement. That should be a few more sentences about the methods (Lagrangian/Eulerian)
underlying data (if different) resolution etc.

We included further details in section 3.2 / table 3 of the main paper

- short description of the applied meteorological fields in the text

5 **- short description of the applied inverse modelling technique (added both in the text and table 3)**

- type of model (Eulerian / Lagragian) has now been added also in table 3 (was already in the text)

- short general description of the applied uncertainties of the observations (including the measurement and model uncertainties)

(model resolutions were already described in the text and included in table 3)

10

technical corrections:

line 11 4.3 (2.3-8.2) Tg CH4 yr -1

'Tg' has been added

line 26 globally averaged tropospheric CH4 mole fraction

15 **'tropospheric' has been added**

(although this is not exactly the average over the entire troposphere - but with the explanation following in the text ('global average from marine surface sites') it should be clear.)

20

References

Alexe, M., et al., Inverse modelling of CH_4 emissions for 2010-2011 using different satellite retrieval products from GOSAT and SCIAMACHY, Atmos. Chem. Phys., 15, 113–133, doi: 10.5194/acp-15-113-2015, 2015.

Bergamaschi, P., et al., Atmospheric CH_4 in the first decade of the 21st century: Inverse modeling analysis using
25 SCIAMACHY satellite retrievals and NOAA surface measurements, J. Geophys. Res., 118, 7350–7369, doi: 10.1002/jgrd.50480, 2013.

Bergamaschi, P., et al., Top-down estimates of European CH_4 and N_2O emissions based on four different inverse models, Atmos. Chem. Phys., 15, 715-736, doi: 10.5194/acp-15-715-2015, 2015.

Henne, S., D. Brunner, B. Oney, M. Leuenberger, W. Eugster, I. Bamberger, F. Meinhardt, M. Steinbacher, and L.
30 Emmenegger, Validation of the Swiss methane emission inventory by atmospheric observations and inverse modelling, Atmos Chem Phys, 16(6), 3683-3710, doi: 10.5194/acp-16-3683-2016, 2016.

Houweling, S., et al., A multi-year methane inversion using SCIAMACHY, accounting for systematic errors using TCCON measurements, Atmos. Chem. Phys., 14, 3991–4012, doi: 10.5194/acp-14-3991-2014, 2014.

Michalak, A. M., Randazzo, N. A., and Chevallier, F.: Diagnostic methods for atmospheric inversions of long-lived
35 greenhouse gases, Atmos. Chem. Phys., 17, 7405-7421, https://doi.org/10.5194/acp-17-7405-2017, 2017.

Monteil, G., et al., Comparison of CH_4 inversions based on 15 months of GOSAT and SCIAMACHY observations, J. Geophys. Res. Atmos., 118, doi:10.1002/2013JD019760, 2013.