Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-755-AC2, 2017

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Satellite-derived methane hotspot emission estimates using a fast data-driven method" *by* Michael Buchwitz et al.

Michael Buchwitz et al.

michael.buchwitz@iup.physik.uni-bremen.de

Received and published: 8 February 2017

First of all, we would like to thank the referee for carefully studying our manuscript and for providing critical comments and questions. Below we provide answers to each of these comments and questions. The referee's comments have resulted in a significantly improved version of our manuscript.

Q1: Referee: First, the word "hotspot" in the title is somewhat misleading because the main result of this study is a regional or subregional (relatively large area) estimate of CH4 emissions although pixels (but at coarse resolution) with large enhancements relative to surrounding pixels are identified. I strongly suggest that the authors remove the word "hotspot" from the title because this work essentially estimates emissions for source "regions", for which many studies have already been doing using data from

ground tower sites or aircrafts or remote sensing. The more accurate bottom-up inventories the authors cites (e.g., Jeong et al., 2014) can now identify hotspots with a much finer resolution. At the global scale, the source regions in this study may be considered hotspots, but those areas are really regions or subregions as shown in many previous studies.

Author's reply: A hotspot does not have to be a very small area. It can be a large area, e.g., a country, see Oxford dictionary (https://en.oxforddictionaries.com/definition/hotspot): Definition: "A place of significant activity, danger, or violence." Example sentence: "Madagascar is considered a biodiversity hot spot, an area that is home to great numbers of species and that is under constant assault from human activity". As we apply our method to areas of very different size and to areas emitting large amounts of methane, the term hotspot seems appropriate for this manuscript.

Q2: Referee: Second, the authors try to match their satellite-based XCH4 to another assimilated product. This is disappointing because the value of those satellite products for XCH4 is significantly diminished as they are supposed to be used as independent retrievals of XCH4. The authors need a clear justification for this. Please see the related specific comments below.

Author's reply: In the revised version of the paper we will improve the description of how the methane data product used for "matching" has been generated. This product uses optimized emissions (obtained via assimilation) which are then used to generate the atmospheric methane concentration (via forward simulation). Therefore, the atmospheric concentrations are consistent with the emissions and this is exactly what we need for our purpose. The correctness of the emissions is not relevant for our application but what is relevant is that the link between emissions and concentrations is modelled as good as possible.

Q3: Referee: Third, it looks like that the proposed method ends up with a simple linear

scaling of satellite-derived XCH4 to CAMS, in particular with a single parameter of V, which seems to be estimated as one value for the whole globe (as written it sounds like that; if not please clarify it).

Author's reply: Yes, this understanding is correct. Please see also our detailed answer to the concerns of the other referee. In the revised version of our manuscript we will explain this better and will also add more details on our efforts to use meteorological data to improve on this. We also present an additional investigation using another model, which simulates methane at much higher spatial resolution compared to the used CAMS data set.

Q4: Referee: Also, given the too large uncertainty for individual annual emission estimates, I wonder what value from this study can be added to the scientific community for regional GHG modeling.

Author's reply: The purpose of our method is not to improve regional GHG modelling but to obtain very quickly (rough) methane emission estimates from (large amounts of) satellite data. The results can be used to identify regions where methane emissions are potentially higher than existing emission data bases suggest. We write in the "Summary and conclusions" section: "More detailed assessments likely require the use of much more complex approaches compared to the simple method uses in this study. Nevertheless, simple and fast approaches also have a role to play as they permit to perform quick assessments on possible discrepancies with respect to emission inventories or other data sets and can also be used for plausibility checks for more complex approaches".

Q5: Referee: Page 4, Lines 26 - 28, the sentence needs to be revised because the authors are trying to say two conflicting things in the sentence, making it confusing. Also, I would recommend that the authors be more quantitative instead of saying "agree reasonably well". In terms of data gap, how SCIAMACHY and GOSAT are different, e.g., available data points/pixels at the annual scale?

Author's reply: It is not entirely clear for us why this sentence is confusing. Taking into account the sampling of GOSAT and the fewer number of observations (a factor of 2-3 depending on product) compared to SCIAMACHY we were also surprised about the reasonably good agreement of the emissions as obtained from SCIAMACHY and GOSAT. At present this is a finding based on our end results. We have not aimed at explaining this is in detail in terms of number of observations and required precision, accuracy and sampling as this is a complex topic requiring additional assumptions, e.g., on error correlations, and because we think that this a bit out of scope and not mandatory for our study. Concerning "agree reasonably well": We have added more specific information in the revised version of the manuscript by adding in brackets: "(e.g., in terms of mean value and scatter of the resulting annual emission estimates)". The difference in terms of data gaps is addressed in our manuscript as we show for each investigated target region XCH4 maps for SCIAMACHY and GOSAT in Sect. 4.

Q6: Referee: Page 5, Line 10, I wonder if the authors considered the data scarcity (i.e., small number of data) for the annual averages in terms of uncertainty. For certain pixels, the # of available data would be too small while others have enough for averaging.

Author's reply: We consider this by visual inspection of annual XCH4 maps for each target region (examples are shown in Sect. 4 of our manuscript) and quasi-automatically by varying the size and shape of the surrounding region and by considering the standard deviation of the resulting emissions in our error estimate. We are confident that this is better than explicitly using the number of individual data points for our error estimate as this would require knowledge on error correlations (please note that from previous studies we know that improvement upon averaging will not follow a square root law).

Q7: Referee: Page 5, Line 17, I wouldn't use the term "enhancement" because the surrounding region is not equal to the CH4 "background" region, e.g., the Pacific region for the western US.

Author's reply: In that paragraph we are not using the term "background". "Enhancement" is defined as source region XCH4 minus surrounding region XCH4. If this difference is positive than we have a positive enhancement, i.e., XCH4 is higher over the source region compared to its surrounding area. In this context, it does not matter if the surrounding is equal to a true background or not. This only matters in terms of the accuracy of our method (e.g., additional sources in the surrounding region). In the revised version of our manuscript we have added more information on this accuracy aspect.

Q8: Referee: Page 6, Lines 4 - 9, Looking at Eq. (1), the authors are trying to estimate emissions (flux) for the source region using _XCH4. But _XCH4 is not exactly the local enhancement, but only the relative enhancement to the surrounding region, which itself has some local enhancements. This will lead to underestimation of the emissions for the source region.

Author's reply: Yes, this is true and in the revised version of our manuscript we will highlight this aspect more prominently and provide more details.

Q9: Referee: Page 7, Lines 9 - 10, The authors confirm my point about the underestimation when using Eq. (1). The authors state that "we aim at quantifying the impact of the choice of the surrounding region by varying its size and shape." This makes it very hard to adopt the proposed method in other regions because it involves adjustments of size and shape, likely yielding multiple estimates and subsequently expanding the uncertainty.

Author's reply: Yes, this expands the uncertainty as explained in our paper. As our emission result depends on the chosen surrounding region our uncertainty estimate contains an error term which reflects this. Please note that it is not hard to adopt our method to other regions. For the four source areas discussed in our manuscript we vary the surrounding region using a pre-defined automatic procedure which is the same for all four source regions (see page 10, top).

Q10: Referee: Page 7, Lines 20 - 22, There are two important concerns about the method. First, I expected from the title that the satellite products would provide independent observations as in most of the top-down studies. It is not vey satisfactory to try to match estimates from another product, i.e., CAMS. Also, from what is written here, I find that a single value for V needs a serious justification. Also, I am not convinced why CAMS should provide "true" estimates. Can the CAMS estimates be truly representative of any of the study sites/regions? How well are they compared with the estimates from previous studies for those source regions (maybe the word "true" may not be appropriate here; otherwise needs clarification).

Author's reply: In the revised paper we will show additional results using another model which provides methane simulations at much higher spatial resolution. We also provide more details on why we are using a single value of V. If we apply our method to real satellite data, then the true emissions are not known. However, if we apply our method to simulations the underlying emissions are known. We refer to these emissions as "true emissions", meaning "known emissions". In the revised version of our manuscript we will explain this better.

Q11: Referee: With respect to the optimization of V, this parameter optimization would be the key to this study. However, it seems that there is no explanation or consideration of the errors between the relationship between CAMS and XCH4, which can be defined as: CAMS = f(XCH4, V) + err, where the function f is likely a linear one and err is the irreducible error (e.g., mean 0, normal error). Here for correct estimation of V, we need some independent estimates for err, similar to a linear regression case with errors.

Author's reply: It is true that parameter V is very important as the estimated emissions are directly proportional to it. It is also true that it would be good to have an independent assessment of the error of our estimated emissions. Therefore, we have added in the revised version of our manuscript additional assessment results using another model to compute emission biases for several source regions and we use the results to present more details on the performance of our method.

Q12: Referee: Page 10, Lines 18 - 21, I differ with the authors. The too large uncertainty suggests that the method is not powerful. I would conclude that the only value of the satellite products used in this study is to provide auxiliary information derived from the columnaveraged XCH4 which is linearly scaled to match another model product (rather than independent measurements).

Author's reply: It is not clear for us what is wrong with what we write in lines 18 – 21. Our emission estimates are independent as they are derived from independent satellite retrievals. However, we agree that our large uncertainty limits the power of our method. In this context please see our response given above related to Q4.

Q13. Referee: Page 11, 33-34, Again, the uncertainty is too large. When we think about hotspots, we expect relatively unambiguous isolation of emissions. The papers cited in this work already estimated emissions for the region with much better uncertainty. What policy makers need is identification of hotspots at the level of km scales and emission estimates for those small regions to mitigate sources from them. However, in this study, even the regional annual total yields very large uncertainty. Is there any way to reduce the uncertainty, even at the annual scale?

Author's reply: Please see our response to your concern for Q12 and Q4. Our method is not accurate enough for "policy applications". This would require a much more powerful method. As explained above (and in our manuscript) the main purpose of our fast method is to obtain rough estimates of emissions for source regions of interest using large amounts of satellite data. Via our method, source regions can be identified where emissions are potentially significantly underestimated in emission inventories. These regions can then be studied in detail using more powerful (but also computationally much more demanding) procedures.

Q14: Referee: Table 3. EDGAR v4.2 happens to estimate the same Mt CH4 for both Four Corners and the Central Valley?

Author's reply: We have checked this for both source regions and found that the correct

value for the Central Valley is 0.19, not 0.17. Many thanks for pointing to this! We have corrected Tab. 3.

Q15: Referee: Figure 1. The region needs to be defined more accurately. For example, the region defined as the Central Valley of California in Figure 1 includes Southern California, and is different from that in Table 2.

Author's reply: The purpose of Fig. 1 is to present an overview about the entire globe and to show where the investigated source regions are located and how XCH4 looks like in these areas but also in their surrounding area. It is not the purpose of Fig. 1 to define exactly the source regions. The exact definitions of the source regions is given in Tab. 2.

Q16: Referee: Figure 8 needs some improvements. First, the data points (circles) should match the years on the X-axis label that are represented. Is the "standard deviation" the standard deviation of 7 annual estimates, e.g., for the 2003 - 2009. If this is the case, standard deviation is not very useful. I would be more interested in knowing the overall mean estimate for the multi-year period and the uncertainty about the mean, e.g., during 2003 - 2009. When individual annual estimates have huge uncertainties associated, I don't see the benefit of using standard deviation.

Author's reply: We have improved this figure by changing the annotation of the x-axis (Year -> Time[year]). We use standard deviation as this is a precisely defined quantity in contrast to the computation of the uncertainty about the mean as this would require sufficiently good knowledge of error correlations.

Q17: Referee: Also, the 1-sigma uncertainty in estimated emissions for individual years overlap with the EDGAR estimate, making it hard to statistically evaluate EDGAR. Looking at this at face value, I am not sure if there is any statistical power in the proposed method to say about the regional emission, even at the annual scale.

Author's reply: Please see our response to these aspects (large uncertainty, power of

our method) as given above.

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-755, 2016.