Reply by the authors to Referee #1's comments on
"Quantifying methane emissions from Queensland's coal seam gas producing Surat Basin using inventory data and an efficient regional Bayesian inversion" (#acp-2020-337)

Anonymous Referee #1 (RC1)

We are grateful to the Referee for taking the time to read our manuscript and making a number of valuable comments. In the following, we provide our responses to these comments (the Referee's comments are shown in blue). The locations of the changes made refer to those in the non-tracked version of the revised manuscript.

This paper employs a dataset of quasi-continuous measurements over an 18-month period from two monitoring stations in the middle of a region characterized by a mix of largely anthropogenic methane sources to optimize gridded methane emission inventory estimates. It aims to scale inventory emission estimates for individual grid boxes with a focus on the coal seam gas (CSG) industry. Given the current lack of atmospheric data to inform CSG methane emissions in Australia and elsewhere, this paper is a useful addition to the literature to help researchers improve their methods to quantify emissions from this source. The analysis is very detailed, the paper is well written, and the tables and figures are well presented.

Response: Thank you for your comments.

However, I have two major comments/questions that may be important for the bottom-line implications of the study:

1. The background methane mole fraction estimation (Supplementary S3) requires some more discussion. As Figure 3 shows, both monitoring stations are surrounded by known methane sources that are being quantified here. The monitoring stations do not measure the background air entering the spatial domain for which the emissions are being quantified here (hence background estimation). Filtering peaks during the early afternoon may exclude the largest point sources, but not necessarily the area sources that are clearly shown to exist in Figure 3. Does this estimation method create a high bias for the background levels, and in extension a low bias for the posterior emissions (especially from distributed sources like CSG wells)? Could this explain why all inverse setups produce smaller posterior total emissions than the prior despite the acknowledgment in the paper that the inventory may miss some sources (so the inventory itself may be underestimated)? Note that the opposite is true when looking only at the CSG sub-domain, which is situated largely between both monitoring stations (thus the sources in the CSG sub-domain affect estimated background values to a lesser extent), which appears to underscore this conundrum. It is also noteworthy that such underestimation may be masked also in the q-q plots comparing observed and modeled concentrations because a potentially underestimated prior and overestimated background would compensate each other.

Response: The reviewer has a valid point. Specification of background in a regional model is tricky. Ideally, this requires methane measurements at many locations around the perimeter of the study domain or modelling methane at much larger scale (preferably global), with all sources, sinks and chemical processes accounted for, which could then provide concentration boundary conditions needed for the regional modelling. Notwithstanding the difficulty in carrying out such a major computational task, there are modelling difficulties and uncertainties associated with emissions, representation of processes, model resolution issues etc. There could be other ways to calculate

background too, such as satellite data and model-data assimilation. Nevertheless, we believe that for the hourly-averaged, ground-level background concentrations needed in regional modelling study like ours, in-situ observations near the ground are still a better means to derive the background provided there are sufficient number of monitors sited at favourable locations than using a larger scale model.

In our case, we are limited by only two monitors (i.e. Ironbark and Burncluith) within a relatively large study domain. This reflects the operational and budget constraints of this project and is likely typical of many others. We calculated the hourly background using a methodology described in the Supplement S3 that utilised methane concentration measurements from the two monitors. It assumes that under vigorous atmospheric mixing conditions in the daytime, the measured concentrations within study domain represent methane levels both within and outside the domain boundaries, so that the measured concentrations can be taken to represent the background under such conditions. Figure 4 in the paper shows how the derived background defines the baseline for the methane measurements, which we have treated as the real background.

Because the background concentration is calculated from the measurements within the source region under study, there is a possibility that it represents an upper limit on the magnitude of the background, meaning that the real background is potentially lower than what we have used (as alluded to by the referee).

To examine the sensitivity of the emission inference to the background methane, we have done an additional inversion using an alternate background time series and this is described in detail in the new Supplement S5. The alternate background was constructed using our original background methane and marine baseline methane measurements from the Cape Grim Baseline Air Pollution Station (https://capegrim.csiro.au), located on the north-west tip of Tasmania (40.7ºS, 144.7ºE) (see the Supplement S5). The measurements from the Station were filtered for the marine baseline air (in southern mid latitudes), and the baseline methane thus represents concentration levels without the direct influence of the continental sources. As shown in Figure 1 below, the alternate background falls between the Surat Basin background as used in our study and the Cape Grim baseline (i.e. between the two bounds), and is, on average, lower than the previously used Surat background by 2.8 ppb. (On average, the Cape Grim marine baseline was 8.4 ppb lower than the original Surat background used).

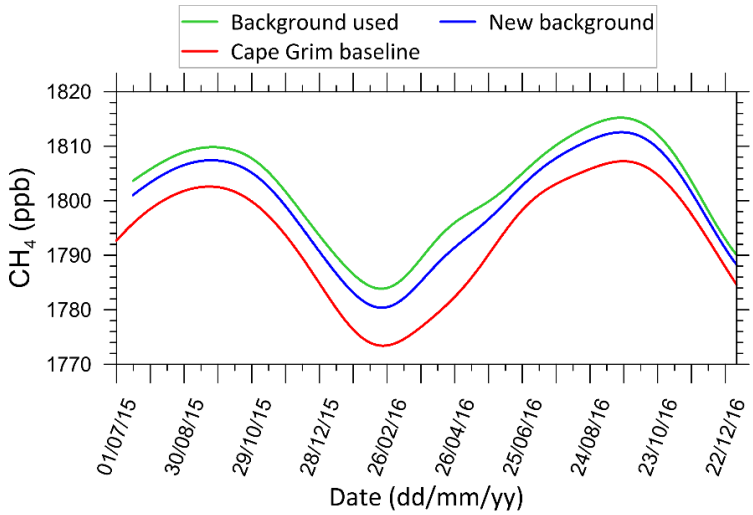

Figure 1. The average hourly background CH_4 concentration (ppbv) time series (green line) as used in the present paper. The hourly-averaged Cape Grim baseline methane is shown as a red line. Blue line is the alternate background.

The inversion results in Table 1 below show that compared to the inferred emissions obtained using the original background methane the alternate background gives total emissions that are 6.8% higher, while the increase is smaller at 3.9% in the CSG subdomain and larger at 8.5% in the non-CSG region. The overall increase is expected because the increase in the measured concentrations by 2.8 ppb as a result of the use of the alternate background needs to be accounted for by the inversion by enhancing the amount of inferred emissions.

We also find that the amount of increase in the inferred emissions with the alternate background is almost uniformly spread through the study domain relative to the total emission, and that there are no significant spatial distributional shifts in the inferred emissions with the two background choices. This means that if these enhanced emissions are used in a forward model simulation, they would lift the modelled concentrations throughout the region by a very similar amount (likely by 2.8 ppb).

Table 1: Inferred emissions ($\times 10^6$ kg yr^{-1}) obtained using the original methane background variation used in the paper (Case 3c in the paper, with the bottom-up inventory as a Gaussian prior with 3% uncertainty relative to the mean) and those obtained using the alternate methane background variation. The values in the parentheses are % change over the original inferred emissions.

Methane background	Total	CSG subdomain	Non-CSG subdomain
Original background (as used in the paper)	165.8	63.6	102.2
Alternate background	177.0 (+6.8%)	66.1 (+3.9%)	110.9 (+8.5%)

The above analysis demonstrates that there is an increase in the amount of inferred emissions with the alternate background and that this increase is smaller in the CSG subdomain relative to the original inferred emission.

Changes in manuscript: The new Supplement S5 given with full details of the above calculation, and the results are also summarised in Section 7.5 of the revised paper.

2. How are the higher-end modelled methane concentrations (but low occurrence, potentially not due to the infrequent emission, but rather due to their being point sources with fewer opportunities to be sampled) weighted against the overall average methane (but high occurrence) in the inversion model framework? Is this objectively weighted in the model (and if so, how), or is it a model design choice?

Response: In our inversions, the hourly-averaged methane measurements obtained during July 2015–December 2016 are combined in one Bayesian calculation to derive time invariant top-down emissions on an 11 × 11 source grid within the domain. Our inverse model framework is, in principle, able to discriminate between a source with a high emission rate but with infrequent impact at a sampling point and a source with a low emission rate but with frequent impact at the sampling point. This is because the concentration observations at the sampling point would reflect representative signals from these two types of sources, and this information when used in the source-receptor relationship would optimise the source emission rates accordingly such that they

best describe the concentration observations. In practice, however, the success in discriminating sources depends on the quality and quantity of available concentration observations, their spatial coverage, and on the number of source parameters that need to be quantified. This is where the specification of the prior plays a very important role because the information available (through concentration observations) may not be adequate to estimate the source parameters properly. This is demonstrated in our study.

Therefore, essentially, the only source weighting in our inverse framework is through the specification of the prior, and there is no other source weighting included/needed in the model apart from what is implicit through the Bayesian approach.

Changes in manuscript: We do not think that there is any change needed in the paper and hope that the above clarification is satisfactory.

Below is a list of detailed comments that may help clarify arguments and language, and correct potential errors.

Main article:

1. Ln 39: For balance, there's an ongoing discussion about the contrasting evidence (contemporary local measurements vs. ice-core 14C data) regarding the magnitude of the fraction of natural geologic seepage: https://www.elementascience.org/articles/10.1525/elementa.383/

Response: We have included two references to the bottom-up global estimates of natural geologic seepage.

Changes in manuscript:

We have modified the original wording to:

"However, a study using measurements of carbon-14 in methane recently showed that nearly all methane from fossil sources is anthropogenic, contrasting with the bottom-up estimates of significant natural geologic seepage (Etiope et al., 2019; Etiope and Schwietze, 2019), and that fossil fuel methane emissions may be underestimated by up to 40% (Hmiel et al., 2020)."

References:

Etiope, G, Ciotoli, G, Schwietzke, S and Schoell, M. 2019. Gridded maps of geological methane emissions and their isotopic signature. Earth Syst Sci Data 11: 1–22. DOI: 10.5194/essd-11-1-2019

Etiope, G. and Schwietzke, S., 2019. Global geological methane emissions: an update of top-down and bottom-up estimates. Elem Sci Anth, 7(1), p.47. DOI: http://doi.org/10.1525/elementa.383

2. Ln 58: "independent": I suggest "atmospherically based" instead since inverse estimates are by definition not completely independent of the prior/inventory.

Response: Point taken.

Changes in manuscript: Modification made.

3. Ln 71: Through this or any top-down approach? Would be valuable to mention if other top-down approaches have been used in Australia in the past.

Response: Point taken.

Changes in manuscript: We have added the following text:

"To our knowledge, this study is the first in Australia to quantify regional scale CH_4 emissions through a top-down approach employing transport modelling and concentration measurements, although studies at other spatial scales with broadly similar approaches have been reported, e.g. by Luhar et al. (2014) and Feitz et al. (2018) for single point sources at local scale and by Wang and Bentley (2002) at continental scale with Australian methane emissions divided into eight source regions."

References:

Luhar et al. (2014) and Feitz et al. (2018) already cited in the paper.

Wang, Y. P., and S. T. Bentley, S. T.: Development of a spatially explicit inventory of methane emissions from Australia and its verification using atmospheric concentration data, Atmospheric Environment, 36, 4965–4975, https://doi.org/10.1016/S1352-2310(02)00589-7, 2002.

4. Ln 158: Would re-phrase that the two operators account for 1.5% of CSG production activity in the region, not emissions (which would be difficult to establish with any accuracy).

Response: Point taken.

Changes in manuscript: The sentence is changed to "…but it was established that these two operators, with a total of 256 wells, only accounted for about 1.5% of the CSG activities that may be related to emissions."

5. Ln 189ff: Spatial resolution of 2.5º × 2.5º means (roughly) 250 x 250 km2. How, then, is it possible to apply it at 5 x 5 km2? Regarding the meaning of the 6 hour availability of met re-analyses, does it means that the temporal resolution is 6 hours?

Response: There is some misunderstanding here. The spatial resolution of 2.5º × 2.5º corresponds to the synoptic-scale fields of the horizontal wind components, temperature and moisture that are required as input boundary conditions for the outermost domain of TAPM. These fields given at 6-hourly intervals were sourced from the U.S. NCEP (National Centers for Environmental Prediction) reanalysis database. The TAPM model outputs hourly-averaged fields of meteorology and concentration at a specified horizontal resolution, which in the present application was 5 km × 5km.

Changes in manuscript: The above has been made clearer in the 2nd last paragraph of Section 4.1 of the revised paper (lines 212-220). Some more details of the model are given in the 2nd paragraph of this Section (lines 193-204).

6. Ln 291: I assume you're referring to the bottom-up emission inventory?

Response: Yes. Thanks for pointing that out. Correction made.

Changes in manuscript: As above.

7. Ln 666: Arguably Figure 14b cannot be used to support the trend in the CSG activity data. According to Ln 609, only 4% of the sub-domain emissions are due to CSG wells (and unclear whether the same processing facilities would emit more given more throughput), so any increase in well count may hardly be detectable by the monitoring stations. Thus, the insight here seems to be not that measurements aren't supporting the CSG increase, but that the existing monitoring setup is likely unable to detect.

Response: We have modified the text to improve clarity. A curve for the number of wells is also included in Figure 16 (Figure 19 in the revised paper).

Changes in manuscript: The paragraph revised as follows (lines 810-818):

"However, Figure 19 (*which is old Figure 16*) also shows that there is a downward trend in the amount of flared/vented gas. Considering, based on the bottom-up inventory in Section 3, that venting (from processing) is the biggest contributor (88%) followed by flaring (8%) (from both processing and production) to the total CSG methane emissions, it is plausible that despite the increase in the CSG development in the area the CSG-related methane emissions have not increased, and that they may have even gone down. The temporal variation of the inferred emissions in Figure 17b (*which is old Figure 14b*) for the CSG dominated area also does not indicate any consistent increase in emissions from 2015 to 2016. Thus, the 33% higher top-down emission estimate from the CSG area compared to the inventory estimate cannot be explained in terms of the growth in the CSG production from 2015 to 2016 and is possibly related to underestimated or missing emissions in the inventory. This also implies that the emissions from CSG may be more closely related to practices in the industry than to the amount of CSG produced."

Supplementary:

1. Ln 99: Emissions of methane due to incomplete combustion of CSG

Response: Change made.

Changes in manuscript: As above.

2. Ln 100ff: Why are methane GWPs used for methane emissions from incomplete combustion, fugitives, and coal extraction? It sounds like the underlying EFs are given in CO2e, which seems illogical.

Response: The calculation methods used to estimate methane emissions from CSG activities are consistent with the Australian National Greenhouse and Energy Reporting (NGER) program. We now attach the Katestone report "*Surat Basin Methane Inventory 2015 – Summary Report*" in the Supplement S6 of the paper, which explains in full detail how these emissions were calculated.

Changes in manuscript: As above.