Reply by the authors to Referee #2'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 #2 (RC2)

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 a response 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.

The manuscript presented by Luhar and co-workers presents an analysis of methane emissions from a region in Queensland, Australia, that contains a mix of different source processes of which coal seam gas production is the one mostly targeted and discussed in the study. Overall the study used valid and up to date methods. The manuscript is well structured and easy to follow. Quantifying uncertain methane emissions on the regional scale by in-situ observations and atmospheric inversion techniques is an important task supporting emission reductions and as such the study deserves publication. However, the authors should include some additional discussion of how their results may be used in the future by gas companies and/or authorities. I recommend the manuscript for publication after a number of minor issues (as listed below) are addressed/clarified by the authors.

Response: Thank you for your comments.

Changes in manuscript: Regarding some additional discussion of how the results may be used in the future by gas companies and/or authorities, we have included the following text at the end in Conclusions (lines 904-913):

"The methods developed in this study could be used to improve the monitoring and management of greenhouse gas and other air emissions from the onshore gas industry, including that in the Surat Basin. They provide independent information to industry and communities living in gas development regions on one of the main environmental impacts potentially arising from onshore gas developments. Improved quantification of methane emissions on the regional scale is an important step in emissions reductions from the onshore gas sector and possibly other industries. The present top-down method is particularly suited to distributed emissions with potentially unknown locations across a large geological gas reservoir and gas production infrastructure. If monitoring is deployed before gas exploration and production begins then a baseline would be established from which emissions from the industry might be detected. Ongoing top-down quantification, with monitoring stations located close to where emissions appear and with source-specific information from tracers could provide the information necessary to validate emissions from the gas industry to support greenhouse gas inventories."

Minor comments

Page 1, Line 2: Why is the term 'efficient' used in the title? What is efficient about this inversion approach? Further explain or omit from title.

Response: We have decided to omit the term 'efficient' from the title. The reason for its use was the application of the MCMC sampling method and the backward plume approach which make computations very efficient. However, we admit that this is not the first time these approaches have been used in inverse modelling in general.

Changes in manuscript: 'efficient' omitted in the title.

P2, L58: Given the involved uncertainties in transport and inverse modeling, 'verification' may be a too strong term. Validation is often the preferred terminology.

Response: Point taken. 'Verification' replaced by 'validation'.

Changes in manuscript: As above.

Figure 1: A zoom into the study region including the location of the observational sites would be useful. This would also help to understand any orographic features of the domain.

Response: Point taken. We include an orographic map (Figure 1b) and also a Google Earth map showing the surface characteristics (Figure 1c) of the study domain. The Ironbark and Burncluith monitoring sites and the three biggest towns in the area are also shown.

Changes in manuscript: As above.

P4, L90: Were the inlets mounted on small towers or on rooftops? Please briefly mention even if described elsewhere.

Response: Inlets were mounted on masts.

Changes in manuscript: At line 103, we say '…with inlets placed on masts at a height of 10 m'.

Bottom-up inventory: Which emission processes were separated for the agricultural sources? Enteric fermentation, manure handling, etc.? The information in the supplement is very brief and I was not able to obtain the cited report by Katestone. Since this is the dominating emission source in the area, it would be good to give a few more details and also to briefly discuss the uncertainties in these estimates.

Response: We have now included the full Katestone report "*Surat Basin Methane Inventory 2015 – Summary Report*" in the Supplement S6 (it was prepared for us, i.e. CSIRO, by Katestone). It provides a comprehensive detail as to how the bottom-up inventory was constructed (largely by Lisa Smith of Katestone, who is a co-author on the present paper), including agricultural sources and uncertainties.

Changes in manuscript: As above.

P6, L145: What was the number of cattle in the feedlots? How do the emission factors per livestock unit compare between feedlots and free range? How were emissions from animal waste treated in the two cases?

Response: We now give the Katestone report in the Supplement S6 which provides this information.

Changes in manuscript: As above.

P7, L164f: What is this rough estimate based on? It seems to be rather large considering that the main source is cattle and per livestock emission factors are more certain than 50 %. Is the livestock number that uncertain?

Response: Yes, this was a very rough estimate, and we do not have any solid justification for it. Therefore, we have decided to delete it and modify the paragraph. The Katestone report that we now provide in the Supplement S6 provides more information about the bottom-up emissions.

Changes in manuscript: We have deleted this sentence.

Figure 3: What is the reasoning about showing these specific towns? Is there any larger population in the area?

Response: These are only given as reference points. We think that not all town locations are necessary. We now only present the locations of the three biggest towns, i.e. Dalby, Roma and Chinchilla (population 12700, 6850 and 6600, respectively), in the region.

Changes in manuscript: The above is stated in the Figure 1 and Figure 3 captions.

Section 4.3: The analysis in the supplement is quite useful. How does the wind rose comparison look for the filtered observation data. Does it improve? What is the mean bias for the filtered data? Next to wind speeds, mixing layer heights are critical when doing regional scale transport modeling and emission inversions. How is the mixing layer height treated in TAPM? Is there any way of comparing mixing layer heights for the target area and period or are their previous evaluations available for the model?

Response: In the Supplement S4, we now present a wind rose comparison for the filtered data (Figure S4) and provide the corresponding model performance statistics for meteorology (Table S1). With the filtering, the mean wind speed is predicted slightly worse, but the wind components are predicted better, which implies that there is an improvement in the estimation of wind direction with filtering.

The mean bias for the unfiltered and filtered data is now reported in Table S1.

Regarding mixing height, because TAPM is a fully prognostic, coupled meteorological and dispersion model, the predicted three-dimensional meteorological and turbulence fields are used directly by the dispersion component to predict concentrations. Therefore, there is no explicit use of mixing height as a parameter and the atmospheric mixing is taken care of by the predicted turbulence fields. Some of the model parameters that represent turbulence (and hence mixing) include friction velocity (mechanical turbulence) and surface heat flux (buoyancy-generated turbulence) have previously been evaluated in some of the studies cited (e.g. Luhar and Hurley, 2003; Hurley and Luhar, 2009; Luhar and Hurley, 2012; and Luhar et al., 2014)

In the Supplement S4, the link https://scholar.google.com.au/scholar?oi=bibs&hl=en&cites=13876071272134760358 to TAPM citation database provides additional references for TAPM application and evaluation.

Changes in manuscript: The Supplement S4 is modified with new Figure S4 and Table S1 included.

We provide some additional information about the meteorological component of the model in Section 4.1, which also details how turbulence is calculated (lines 193-204).

Modified paragraph:

"The model has previously been applied to a variety of flow, turbulence and dispersion problems at various scales, such as those reported by Luhar and Hurley (2003), Luhar et al. (2008), Hurley and Luhar (2009), Luhar and Hurley (2012), Luhar et al. (2014), Matthaios et al. (2017), and Luhar et al. (2020), which include model evaluation studies."

P12, L249f: Another important source of uncertainty is that of representativeness of the point measurement for the model grid cell (5x5 km). What are the observations compared to? Simulated values interpolated to the location of observation or grid cell containing the observation site? Are there any important sources in the closer vicinity of the sites (<10 km)?

Response: The hourly-averaged model predictions on the innermost grid domain were extracted at the lowest model level (10 m) at the grid point nearest to each of the monitoring sites for comparison with the observations. This is now stated in the text.

We agree that the model's representation of point measurements by grid-cell averaged values is another source of uncertainty, and it is now stated in the text.

The location of the two measurement stations was based on criteria given in Section 2, first paragraph, to "optimise the size and frequency of detection of methane emissions from the broader CSG source region without being unduly impacted by individual sources in the proximity of the measurement sites". (Other practical considerations are noted in the reference (Day et al., 2015), namely access, power, security, landowner assistance and possible future developments that would impact the site.) The sites were selected to avoid potential large, sustained methane sources within 10-20 km or even small sources within about a kilometre of the measurement inlet. Surveys of maps and by vehicle involving mobile methane monitoring of the area around the site identified few such sources. Small sources that were closer to the inlets (mainly Burncluith) were identified and their signals filtered from the data as described in Section 2. As a result, we expect that the hourly-averaged filtered data (Section 2) are as representative as possible of the atmospheric methane concentration across the 5 × 5 km grid cell containing the observation site, and can be directly compared to the model simulations.

Changes in manuscript: As above is summarise in the first para of Section 4.4, and it is mentioned that this is another possible source of differences between the observations and model predictions.

P13, L282: Not immediately clear what top 5 % refers to. How do these top 5 % simulated events compare to the observations? Are these also the highest observed concentrations?

Response: These are the highest 5% of the modelled concentrations, i.e. all the values above the 95^{th} percentile. The idea here was to determine the dominant source types that contribute to highest modelled concentrations. A comparison of the modelled and observed concentrations has already been made in Figures 5 and 6, and it is clear that the highest concentrations are generally underestimated by the model at both sites (more so at Ironbark).

Changes in manuscript: The sentence is modified to '… the highest 5% of the modelled hourly-averaged methane concentrations (i.e. all the concentrations above the 95th percentile)'.

P19, L405ff: So if I understand this correctly, the source receptor relationship for a time t is constructed from output of c* at different times according to the value of tr at individual grid points. First, I am wondering if this could be illustrated for an example case where one would show the field c* for a given time and then the reconstructed source receptor relationship for the same time. Second, it seems that there will remain some form of smearing out of the transport history in time. How much does this conflict with filtering data by time of day instead of using the complete data set. Also what was the rational of using hourly data in this case instead of working with longer aggregation times for which the effect should be smaller?

Response: The Referee is correct. We now explain it a bit better in the text and also present an illustrative example case in new Figure 7 that shows the field *c** for a given time and the reconstructed source receptor relationship for the same time (lines 440-461).

Occasionally, there may be some remains of smeared out transport history in time, but generally the intensity and the frequency of this is very small.

We do not think our method of reconstructing the hourly source-receptor relationship would conflict with the filtering of the data. This relationship is continuous with time, and its value at a particular hour would match the data points at that hour.

The rationale of using the hourly-averaged data rather (for which the effect of transport history would be smaller) was to maximise on the available information to constrain the inversions better. We could use longer averages, but that would have reduced the number of concentration data. Also, the wind direction variation inherent in the hourly data aids in better 'triangulation' of sources; the degree of this variation is progressively reduced as the averaging times are made longer. However, one could use longer aggregation times to see what difference that makes, but we have not attempted that. (Lines 493-497).

Changes in manuscript: As above, and new Figure 7 (a, b, c).

P20, L436: What about the sub-grid variability of these sources? Is it kept for the transport simulation and a factor for the larger grid boxes optimised or is the emission flux constant within the large grid boxes. What about the different source categories? Are they treated separately as was done for the forward simulation? Not clear from this description, later on it becomes clear that only total emissions are optimised.

Response: For the purposes of inferring emission rates using the inverse modelling, 11 × 11 source grid points are considered within the study domain. No sub-grid variability of these emission rates is considered. Given the limitation as to the type and amount of concentration observations we have for inversion, the inverse methodology used does not distinguish between different source categories. This is mainly because the concentration of methane alone was monitored and not tracers specific to methane source types. Therefore, there are no separate sources categories in the inferred emissions, unlike what was done for the forward simulation - only total emissions are optimised.

Changes in manuscript: This is clarified in the text (lines 500-507).

P21, L460: Does high probability mean small uncertainty of the posterior? That would be surprising when starting from larger prior uncertainties.

Response: The sentence is not correct and has been deleted.

Changes in manuscript: As above.

P21, L461: Above, it was speculated that the uncertainty of the bottom-up approach was 50 %. Here it is suggested that 0.5 % should be used in an inversion. That seems to be a contradiction. Please elaborate on the small sigma_p. Also, is sigma_p the uncertainty of the total emissions in the inversion grid or that of individual grid cells?

Response: We are not confident about the previously speculated uncertainty of 50% in the bottom-up approach, and have, therefore, deleted the sentence.

Following this comment and another comment below by the Referee, we revised Section 6 on inversion using the 'synthetic' concentration data considerably, with new model runs using an increased prior uncertainty (5% and 10%) and only considering times when the valid (or filtered) observations were actually available. This is more realistic, and the results now provide a better guidance to inversion using the real data.

sigma_p is the uncertainty (standard deviation) in the prior of the individual source and is specified as % relative to the prior mean value (there are 11 × 11 sources considered for the emission inference).

Changes in manuscript: Section 6 revised, with clarification in the text.

Section 6.1: Usually, one would add random or auto-correlated noise to the synthetic observations as a test up to which degree of uncertainty the inversion can obtain useful information. Was this not done here at all?

Response: We did not add any random or auto-correlated noise to the synthetic observations. But we performed new synthetic runs with the same 3.5 ppb uncertainty in the synthetic concentrations as that in the concentration observations for real inversions.

Changes in manuscript: The synthetic inversion Section 6 modified (lines 514-586) with modified and new plots (Figures 9–11).

Page 21, L464: How would the results change if only the synthetic observations were used at times when valid (filtered) observations were actually available? The latter was a considerably larger number of observations, so it is not clear how the results presented in this section can be propagated to the inversion with the more limited data set.

Response: This is a valid point. Following the Referee comment, we have revised the section on "Inversion using the 'synthetic' concentration data" considerably. We now present inversion results by using the synthetic observations only for times when the valid (or filtered) observations were actually available. The uncertainty in the synthetic concentrations is now taken to be the same (i.e. 3.5 ppb) as that in the concentration observations for real inversions. This now provides a better propagation of the results presented in this section to the next section on inversion using the real observations.

Changes in manuscript: The synthetic inversion Section 6 modified (lines 514-586) with modified and new plots (Figures 9–11).

Section 6.2: Next to the posterior emissions it would be good to show simulated time series (synthetic obs, prior, posterior) and some performance stats in order to get a feeling for the inversion performance. This is done later on with additional forward simulations, but it should also be done with the concentrations directly obtained from the source receptor relationships and the coarse resolution emission setup as used in the inversion. Something to add to the supplement.

Response: Good point, but because this is a case of synthetic time series, we thought that rather than presenting the simulated time series it would be better to actually compare the inferred emissions with the bottom-up inventory emissions that were used to simulate the 'synthetic' concentrations (which in turn were used in the inversion). We have done this exercise and presented the results in Section 6.2 along with some performance statistics (i.e. linear least-squares fits and correlation coefficient). These new results also lead to a better linkage of this section on synthetic inversion to the next section on real inversion.

Changes in manuscript: As above (lines 535–586). New figures 10 and 11. Modified Figure 9.

P22, L491: I don't like the terminology "no prior". There is a prior! Why not call the case "uniform" prior, which would describe the used PDF.

Response: We now call it non-informative uniform prior in the text.

Changes in manuscript: As above.

Section 7.1.3: Again it would be useful to see simulation performance for the three uncertainty levels.

Response: Simulation performance for the three uncertainty levels is now given in new Table 1. We thought it was more appropriate to give it in Section 7.2 on validation than in Section 7.1.3.

Changes in manuscript: As above.

P23, L535: So if the best estimate results from using a Gaussian prior distribution, I wonder why an MCMC approach was used at all. Wouldn't it be much more efficient to use the analytical solution of the Bayesian theorem for Gaussian PDFs in this case?

Response: The Referee is correct with the Gaussian prior distribution. However, our idea was to formulate our inverse modelling tool with MCMC so that it is more generally applicable than just for the Gaussian PDFs—something that could be useful for future applications that we may consider.

Changes in manuscript: None.

Figure 11b: Why not show the relative posterior uncertainty? Couldn't this be more directly compared to sigma_p?

Response: Point taken. The plot has been replaced by the relative (%) posterior uncertainty and the corresponding text modified accordingly.

Changes in manuscript: As above.

P25, L551: Not clear which grid point this is referring to. Why is it relevant?

Response: This is grid point (11, 4), which corresponds to a relatively strong coal mine source in the bottom-up inventory (Figure 3d).

Changes in manuscript: Change made in the text.

P26, L577: Give information on which case 3 inversion is used here (sigma_p=?).

Response: This is Case 3 with 3% prior uncertainty relative to the mean.

Changes in manuscript: Change made in the text.

P28, L611ff: This argument could also be supported by comparing the emissions from the non-CSG sub-domain. Do they differ significantly between bottom-up and posterior? If so, what are the possible reasons?

Response: Emissions from the non-CSG subdomain are now compared (lines 739-742), new plot 17c and Table 2 and the discussion.

Changes in manuscript: As above.

P29, L629: Which sigma_p level?

Response: This is Case 3 with 3% prior uncertainty relative to the mean.

Changes in manuscript: Change made in the text.

Figure 14: Include uncertainties. That would allow judging of how well 3-monthly emissions are constraint and if there is a real difference with time. Other studies have shown seasonality in agricultural emissions. Could this be a possibility here as well? Or does it have to do with a seasonality in the source receptor relationships?

Response: Uncertainties are now included. There is also an additional plot (Figure 17c) for the 3-monthly variation of the inferred emissions for the non-CSG area (which is dominated by grazing cattle emissions as per the bottom-up inventory). In this plot, we also present a 3-monthly climatological average (1992 – current 2020) of rainfall at the Dalby airport, located next to the town of Dalby, within the study domain. There is a good correlation ($r = 0.79$) between the non-CSG area methane emissions and the rainfall, suggesting that the 3-monthly emission variation could possibly be explained in terms of the seasonality in agricultural and wetland emissions influenced by rainfall.

Another potential contributor to the temporal variability in the inferred emissions is the seasonality of the winds in the area which influence the source-receptor relationships. We have not explored this possibility here.

Changes in manuscript: As above. Figure 17c included. Lines 772-783.

Technical comments

P1,L21: 'identical TO' ...

Response: Correction made.

Figure 5: Add explanation of dashed line to figure caption.

Response: Point taken.

Changes in manuscript: We say '…and the dashed line is the 1:1 line (i.e. perfect agreement)'.

Figure 8: It seems to be more logical to start with the bottom-up emissions on the left (8a) and show the posterior on the right (8b).

Response: Point taken.

Changes in manuscript: The plots have been swapped and the figure caption and text modified accordingly.