

### **Author response to Reviewer #1**

**We would like to thank the reviewers for their useful comments and for their positive assessment of our study.**

#### Reviewer comment:

*Overall, I think the methods and analysis are strong and recommend this paper for publication.*

#### Author response:

**We would like to thank Reviewer #1 for recommending our paper for publication.**

#### Reviewer comment:

*It seems unnecessary to spend so much of the paper discussing the model applied individually to the measurement sites when it is clear that that method does not work as well as analyzing gradient between sites. Other studies have also demonstrated the benefit of using gradients (McKain et al., PNAS, 2015), to the point where many studies start with that method. You should focus on demonstrating that the gradient method is best and then on the results using that method, rather than giving a thorough explanation of a method that does not work well.*

#### Author response:

**We will shorten the discussion on the simulations of the concentrations at individual sites and refer to other publication to move faster to the gradients and to support their use (MacKain et al., 2015, PNAS, Turnbull et al., 2015, JGR etc.). However, we feel that even though there are regional studies analysing gradients instead of the simulation of concentrations at individual sites, the large majority of the “large scale” atmospheric inverse modelling community still uses concentrations at individual sites instead of gradients to constrain their inversions. Among the first inversions at very high resolution for small regions or cities, different strategies are used to remove the “baseline” or “background” conditions, which are often difficult to compare with the use of “gradients” (e.g. Henne et al., 2016). Such an analysis here is useful to promote the use of gradients in the community.**

**Furthermore, analysing time series of concentrations at individual sites helps to connect the analysis of CO<sub>2</sub> and CH<sub>4</sub> wind roses at individual sites (Fig. 2, which provides good initial insights into the signature of local emissions) with the subsequent analysis of the gradients.**

**Finally, even though analysing gradients highly improves the results for CO<sub>2</sub>, this study shows that it is not necessarily the case for CH<sub>4</sub> because the CH<sub>4</sub> emissions are more local.**

**Therefore, we would like to keep a significant section on the analysis of CO<sub>2</sub> and CH<sub>4</sub> at individual sites in our revised manuscript.**

#### **Reference:**

**Henne S, Brunner D, Oney B, Leuenberger M, Eugster W, et al. (2015) Validation of the Swiss methane emission inventory by atmospheric observations and inverse modelling. Atmos Chem Phys Discuss 2015: 35417-35484.**

Reviewer comment:

*Measurement methodology appears to be thorough and designed to attain comparable measurements across the various sites, which is essential.*

*For sites without local sources of CH<sub>4</sub>, does the model do better? If not, why?*

Author response:

The local sources at less than 1 km from the sites cannot be represented correctly in the NAEI inventory, but high emissions at 1 km resolution in this inventory can still be indicative of the probability that such a source is located close to the measurement sites. The inventory indicates that there are significant emissions of CH<sub>4</sub> in the model grid cells in which Poplar and Hackney are located or in the neighbouring grid cells. However, the amplitude of the CH<sub>4</sub> emissions around Poplar and Hackney is moderate and does not correspond to major point sources such as waste processing sites. The NAEI inventory does not indicate significant CH<sub>4</sub> emissions within 5 km of the Teddington or Detling. Therefore, even though the urban sites are more likely to be influenced by local CH<sub>4</sub> sources (such as gas leakages from the gas distribution network) than the suburban and rural sites, none of the sites should have a major CH<sub>4</sub> point source, such as landfills or farms, in their vicinity (at a distance smaller than 5km).

Table 1 shows that Teddington and Detling exhibit lower model–data discrepancies than the two urban sites, which suggests that the model would do better for sites with less CH<sub>4</sub> emissions in their vicinity. However, as explained in Section 3.4, the use of constant boundary conditions for CH<sub>4</sub> is a major cause of large model–data discrepancies applying to all sites, and whose amplitude is larger than that of the discrepancies due to emissions in the vicinity of the sites. This explains why the discrepancies are not substantially higher at urban sites than at the sub-urban and rural sites.

The text will be amended to include this analysis.

Reviewer comment:

*Conclusions: What tests could you propose in order to be assured that other sites (perhaps at higher altitudes, etc.) be useful for inversion analysis and improving upon bottom-up inventories?*

Author response:

Conducting such measurements and analysing the skill of the model to represent them, such as in our study, would be the natural way to test this. Various alternative approaches exist to determine which type of signal/observation bears information about large scale fluxes and would be well represented by the km-resolution models presently used for the atmospheric inversions. Such approaches include the analysis of the CO<sub>2</sub>/CH<sub>4</sub> atmospheric variability at very high resolution using a high resolution transport model, mobile measurements, or a very dense array of measurements in a small area. We will briefly discuss these in the conclusions section of our revised manuscript (see below).

Reviewer comment:

*You vaguely state that the large model-data misfits mean that your network is not up to that task, but could be more specific about how you came to that conclusion.*

Author response:

We will better clarify that the analyses demonstrate that the CO<sub>2</sub> signal measured at Hackney and Poplar is highly impacted by local sources, which cannot be represented with the 2 km resolution model. This high impact applies to both the short-term variability and to the mean concentrations (i.e. over long timescales). Therefore, we can hardly expect state of the art inversion approaches based on the 2 km resolution model to have sufficient skills to filter the signal of the city scale emissions from that of the local emissions without subgrid scale analysis such as those discussed in the answer to the previous comment.

Regarding CH<sub>4</sub>, the discussion is different (see the answer below).

Reviewer comment:

*What would be necessary to achieve an adequate network,*

Author response:

The analysis of Bréon et al. 2015 and the subsequent studies of city scale inverse modeling at LSCE indicate that CO<sub>2</sub> measurements at levels higher than 15 magl, and located in suburban areas at opposite edges of the urban area, can be used for city scale CO<sub>2</sub> inversion when assimilating cross-city upwind–downwind gradients. Exploiting CO<sub>2</sub> measurements at more than 15 magl in the core of the urban area could remain a challenge as shown by the analysis of Bréon et al. 2015 for the measurements at the top of the Eiffel Tower in Paris. This challenge may be addressed using networks with different types of measurements (e.g. integrated column measurements), averaging data from sufficiently dense sampling to get information about the spatial scales relevant to the model, or using local (for each site) very high resolution model simulations to help detect under which conditions the large scale signal vs. local signals can be filtered from the measurements. Following Reviewer 2's suggestions, these ideas will be listed in the conclusion. Still, these are prospective ideas that need to be tested and evaluated.

These ideas could also apply for monitoring CH<sub>4</sub>. However, the situation can sometimes be very different for CH<sub>4</sub>. McKain 2015, PNAS could conduct a city scale assessment of the emissions of Boston, but this likely relies on the fact that the fugitive CH<sub>4</sub> emissions from the gas distribution network are high in large cities in the US (Philipps et al. 2013). However, Lowry 2001 diagnosed that the gas distribution in London generates less than 20% of the total emissions, which are dominated by waste treatment in this city. The CH<sub>4</sub> emissions from the gas distribution network in other European cities such as Paris and Rotterdam seem to be very low (results from the CH<sub>4</sub> mobile campaigns in the frame of the Carbocount-city project). Therefore, for many cities, including London, the major component of the CH<sub>4</sub> emissions originates from specific sites that are generally located outside the central urban area (e.g. landfills, waste water treatment plants, gas compression sites). Consequently, the city scale approach is not systematically adapted to city CH<sub>4</sub> emissions and local approaches (such as mobile measurements around the sites and local models) would often be more suitable.

The new manuscript will better discuss these points.

References:

McKain KK, Down A, Raciti SM, Budney J, Hutyra LR, et al. (2015) Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. *Proceedings of the National Academy of Sciences of the United States of America* 112: 1941-1946.

Phillips NG, Ackley R, Crosson ER, Down A, Hutyra LR, et al. (2013) Mapping urban pipeline leaks: Methane leaks across Boston. *Environmental Pollution* 173: 1-4.

Lowry D, Holmes CW, Rata ND, O'Brien P, Nisbet EG (2001) London methane emissions: Use of diurnal changes in concentration and  $\delta^{13}\text{C}$  to identify urban sources and verify inventories. *Journal of Geophysical Research* 106: 7427.

Reviewer comment:

*and how would you verify that the network is good enough?*

Author response:

**See above the answer to the beginning of the same reviewer comment.**

Reviewer comment:

*Specific Comments:*

*P. 8, Ln. 6: Is the Picarro air stream dried? If not, I question the 0.021 ppm uncertainty in CO<sub>2</sub> using the Rella correction. The Rella correction has an uncertainty of >0.1 ppm at water levels greater than 1%, and I have found in lab tests that a water correction specific to each Picarro instrument is necessary to achieve 0.1 ppm accuracy in undried air streams.*

Author response:

**Indeed, recent laboratory measurements indicate larger uncertainties associated with the water vapor correction for the CRDS/Picarro analyzers. To our knowledge the most exhaustive study of this effect was conducted at the ICOS Metrology Laboratory and presented at the recent WMO GGMT Meeting in San-Diego (Laurent et al., 2015). This study evaluated the water vapor correction applied to 14 G2401 instruments. For all instruments but one, the uncertainties at a water vapor content of 1.5% are within +/- 0.05 ppm. The outlier instrument shows a bias of 0.12 ppm. Similar tests for CH<sub>4</sub> showed an uncertainty of +/- 1 ppb for all instruments. We propose to change the uncertainties associated to the water vapor correction according to this study.**

**Reference: Laurent O. et al., ICOS ATC Metrology Lab: metrological performance assessment of GHG analyzers, 18th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases, and Related Measurement Techniques (GGMT-2015), La Jolla, California, September 13-17, 2015 [http://www.wmo.int/pages/prog/arep/gaw/documents/GGMT2015\\_A6\\_LAURENT.pdf](http://www.wmo.int/pages/prog/arep/gaw/documents/GGMT2015_A6_LAURENT.pdf)**

Reviewer comment:

*P. 10, Ln 28: For summer, the biosphere is very important to the CO<sub>2</sub> flux. It would be nice to have a few more sentences describing the biosphere model, including how emissions in the city are treated (are they non-zero?)*

Author response:

**We will add a comment on this to the revised manuscript. Our natural CO<sub>2</sub> flux estimate should provide a poor representation of the role of the ecosystems within the city, given that the C-TESEL model producing the simulations we use is run at ~15 km resolution. It does not have a specific implementation of the urban ecosystems.**

Reviewer comment:

*P. 13, Ln. 25: Specify "bottom-up emission inventory" for clarity*

Author response:

**It will be done.**

Reviewer comment:

*P. 14, Ln 25: You describe the modeled mixing layer height a 13% lower than that measured with the lidar. In our experience, the agreement between model and measurement varies significantly day to day and month to month – if that is true for your data it would be useful to state that, and to indicate that the 13% is an average*

**Author response:**

**We will clarify it and we will add further details of the variability of the model-data MLH misfits.**

Reviewer comment:

*P. 15, Section 3.3: How would you expect these wind errors to impact the modelled concentration? How much error would you expect them to introduce and in what direction?*

**Author response:**

**It is highly difficult to translate an error in the wind into an error in terms of concentrations since it strongly depends on the emissions and their spatial distribution (and thus on the uncertainties in the emissions and their spatial distribution in the model) around a given site.**

**It also depends on whether the wind error is transitory or whether it is responsible for errors in the long-range transport from remote areas to the local site, in which case it could raise errors in the signature of the remote fluxes.**

**All these considerations together prevent us from proposing a typical error in the modelled concentrations for a typical wind error.**

**However, we can state that, in general, for urban sites, if the wind speed is too low then the concentrations will be too high in the model since lower wind speeds increase the signature of the high city emissions.**

**This will be discussed in the revised version of the manuscript.**

Reviewer comment:

*P. 16, Ln 12: “We have also excluded data from 29th August and 23rd to 24th September since the model simulated very large GHG peaks on these days which do not occur in the data.” Why does the model produce these large GHG peaks? Can you use that to gain insight into the model?*

**Author response:**

**We believe that these peaks were produced by the combination of low mixing height and of zonal wind direction, which dramatically reduced the model horizontal numerical diffusion to unrealistically low values.**

**We avoided entering into such a qualitative and uncertain discussion in the paper. At the most, it reveals some artefacts of the numerical recipes of the models.**

Reviewer comment:

*P. 16, Section 3.4: What strikes me in Figure 4 is that the modelled CO<sub>2</sub> is often very similar to the background CO<sub>2</sub>, and you don't address that at all.*

**Author response:**

**As discussed in the text, this is revealing of the role of the boundaries that often dominates in this variability. See also the answer below.**

**Reviewer comment:**

*Could you give some explanation of why that is and what it says about the model that you have virtually no emissions added from the boundary?*

**Author response:**

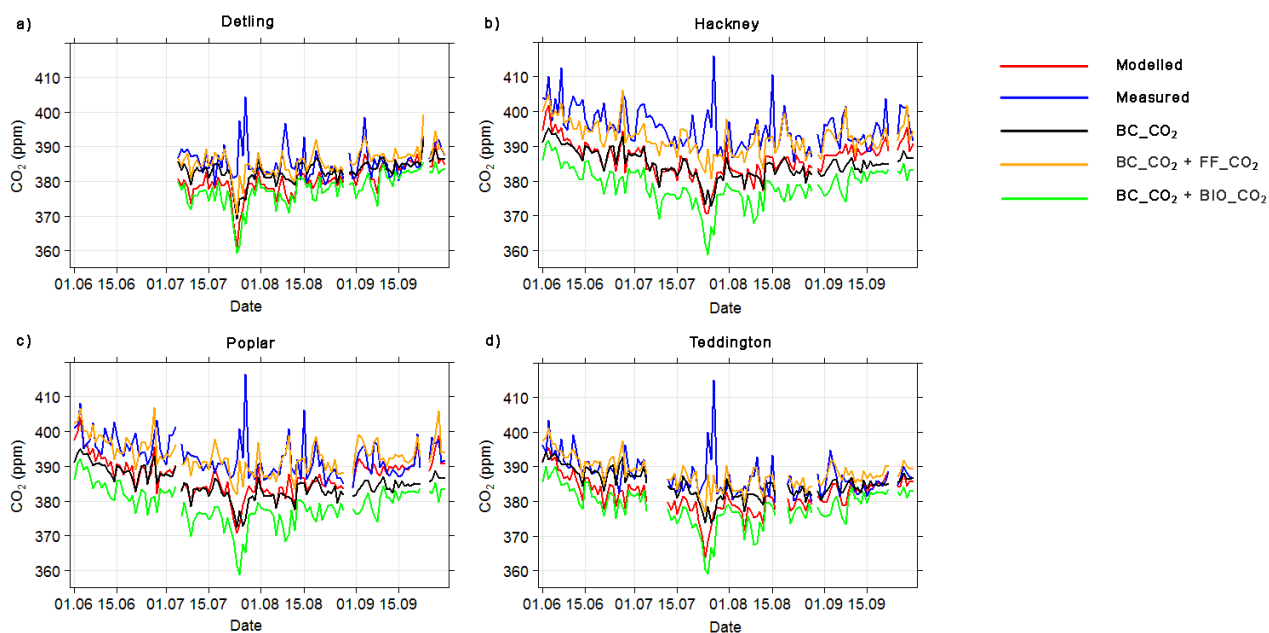
**Actually, when looking at Figure 4, it appears that at DET and TED the total CO<sub>2</sub> is significantly lower than the CO<sub>2</sub> from the boundary due to the natural fluxes in Southern England. The emissions from London are high enough to then shift the total CO<sub>2</sub> back to the boundary level at POP and HAC. This will be discussed in the new manuscript.**

**Reviewer comment:**

*It would also be useful if you included separate lines for biosphere and anthropogenic emissions so we could see if in fact there is an impact of anthropogenic emissions, but they are being negated by the biosphere.*

**Author response:**

**We now plot separately the signature of the anthropogenic and biospheric fluxes added to the boundary CO<sub>2</sub>.**



**This plot confirms that the signature of the urban emissions balances that of the natural fluxes in Southern England for the urban sites, except at the end of the simulation period (in September) when it exceeds it.**

**Reviewer comment:**

*We have actually seen a pattern similar to this in a WRF-STILT model of Boston emissions, and found that it was an artifact of using the model in the city, which we are working to fix.*

**Author response:**

In our study, we do not see it as an artefact, but just as an indication of the similarity of the impacts of the natural fluxes in Southern England and of the emissions in London. When looking at the time series in detail, we find that the discrepancies are significant (especially in September when they become large) and the similarity only applies to the typical amplitude of both impacts.

**Reviewer comment:**

*P. 17, Section 3.5: How is it that you see so little enhancement in CO<sub>2</sub> when modeling the sites individually, but so much greater of an enhancement when modeling the difference between 2 sites?*

**Author response:**

See the answer to the previous comment. Furthermore, Figure 4 shows a clear enhancement from DET or TED to POP and HAC since at DET and TED, the total CO<sub>2</sub> is significantly below the CO<sub>2</sub> from the boundaries, while at POP and HAC it is at the level of the CO<sub>2</sub> from the boundaries. Again, all this discussion will be included in the new manuscript.

**Reviewer comment:**

*P. 20, Ln 9: How many data points are included when you filter for wind speed? Are there enough points for reliable statistics?*

**Author response:**

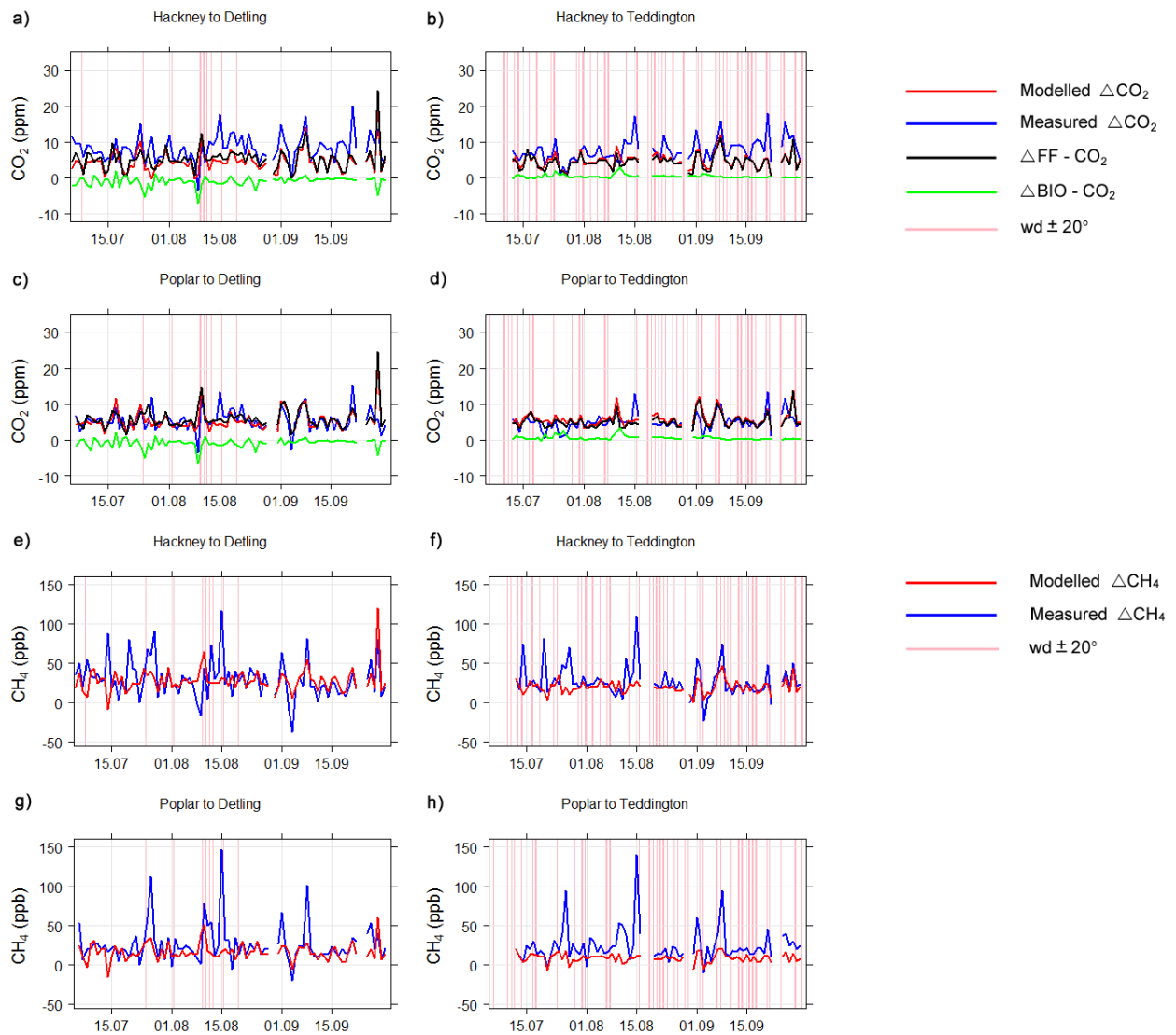
Yes, 18% of HAC–TED and 16% of POP–TED gradients were within this filtered dataset, which corresponds to 101 and 93 observations, respectively. The text will be amended to provide this information.

**Reviewer comment:**

*P. 20, Section 3.6: Could you show a time series of model and observations for the wind filtered data? Or instead you could you markers or shading to show which portions of the time series in Figure 6 were used.*

**Author response:**

Shading has been added to indicate which are the data that are selected according to the wind direction when using this filtering approach.



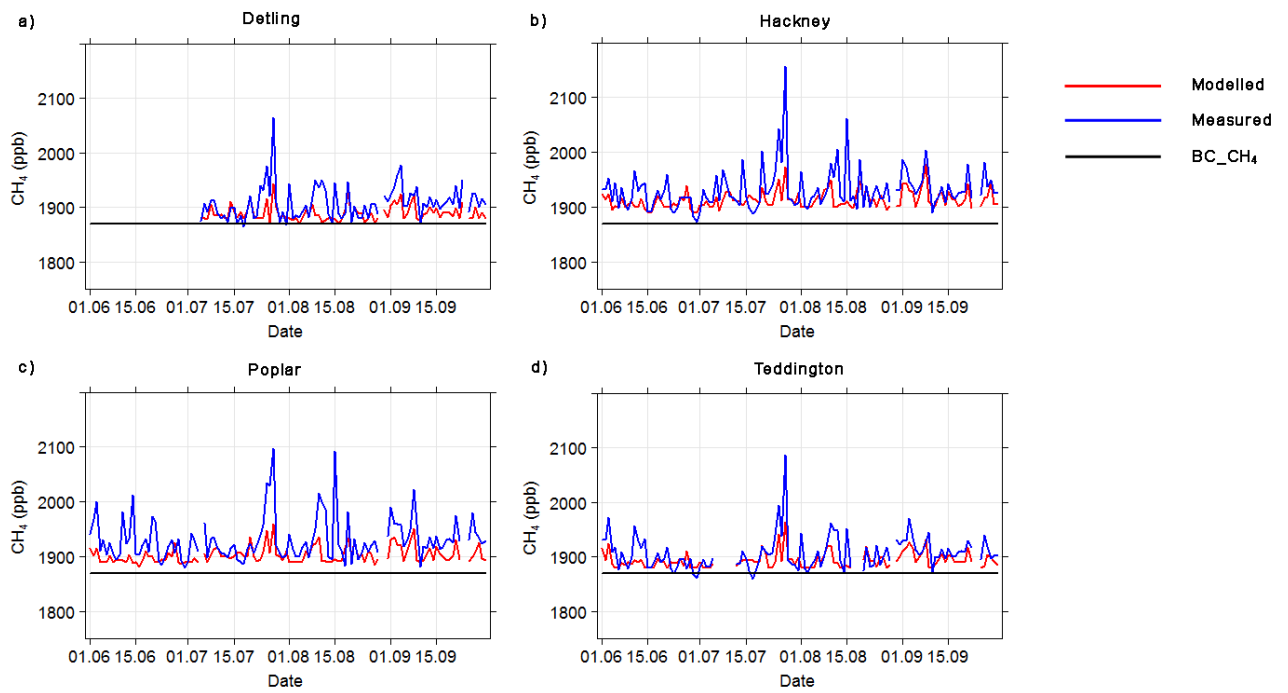
Reviewer comment:

Figure 5: It would be useful to show the background concentration (even if it is constant).

**Author response:**

**The background concentration is added to this figure.**





Reviewer comment:

*Figure 6 e,f: It is hard to make sense of this. I would rather see separate plots as for CO<sub>2</sub>.*

Author response:

**The CH<sub>4</sub> data has now been split into separate plots (see above).**