Replies to first reviewer comments (Reviewer1)

General comments

Bisht et al. present in their manuscript a new data assimilation system based on the Local Ensemble Transform Kalman Filter (LETKF) method with atmospheric transport described by the atmospheric transport model, MIROC4. The system is applied to the estimation of surface fluxes of methane, using both a network of surface observations and GOSAT satellite retrievals. This study describes the method and tests it using Observing System Simulation Experiments (OSSEs) consisting of performing inversions with synthetic observations and for which the true fluxes are known. A number of sensitivity tests are presented to test the system.

On the whole the methodology is scientifically sound and based on previously published models and algorithms. However, in parts the manuscript is difficult to follow and the text unclear or ambiguous. In particular, I suggest improving the description of the methodology especially regarding the preparation and selection of the pseudo observations (see specific comments). In addition, the results and discussion section could be improved to make it easier to follow.

Based on this, I think the manuscript could become acceptable after minor revisions.

Thank you for reviewing the manuscript and providing us with useful comments and suggestions. We have revised the manuscript to make it easier to read. Our replies are given in black font for your comments in red.

Specific comments

L10: "which is substantially" should be "which has substantially"

L23: suggest removing "absolute" before normalized, since if normalization is done the value is always relative

L32: "much high" should be "much higher"

L35: "that have anthropogenic" should be "that has anthropogenic" (i.e., singular form)

L36, "the global CH4 budget"

Reply: All the above grammatical comments have been incorporated in the revised manuscript.

L37: suggest stating that the given range is for the total of all sources and not to put it in parentheses since it is quite important information

Reply: Incorporated in the revised manuscript.
L38: suggest changing “remaining CH_4 emissions” to “main anthropogenic CH_4 emissions” since you list only anthropogenic ones and not all (e.g., the minor source from incomplete combustion of bio and fossil fuels is not mentioned)

Reply: Incorporated in the revised manuscript.

L43: I think the reaction with Cl radicals actually mostly occurs in the troposphere where Cl is more abundant, see e.g.:

Wuebbles, D., Hayhoe, K. and Kotamarthi, R. (1999), Atmospheric Methane in the Global Environment. In: Atmospheric Methane: Sources, Sinks, and Role in Global Change. (Eds. M. Khalil), Springer-Verlag, New York, NY.

Allan, W., Struthers, H., and Lowe, D. C. (2007), Methane carbon isotope effects caused by atomic chlorine in the marine boundary layer: Global model results compared with Southern Hemisphere measurements, J. Geophys. Res., 112(D4), doi:10.1029/2006JD007369.

Reply: Thank you for the suggestions.
CH_4 loss to Cl takes place in the marine boundary layer (MBL), where sea salt is abundant, but CH_4 is also destroyed in the stratosphere by reaction with Cl (for e.g. Röckmann et al., 2004; McCarthy, 2003). In the modified sentence we didn’t specify troposphere or stratosphere because our model simulations consisted of these in both the layers but we have not explicitly included Cl from sea-salt sources. We rewrite the sentence as follows:

“other loss processes include oxidation by soil, and reactions with O1D and Cl”

Röckmann, T., J. ‐ U. Grooß, and R. Müller (2004), The impact of anthropogenic chlorine emissions, stratospheric ozone change and chemical feedbacks on stratospheric water, Atmos. Chem. Phys., 4, 693–699.

McCarthy, M. C., Boering, K. A., Rice, A. L., Tyler, S. C., Connell, P., and Atlas, E.: Carbon and hydrogen isotopic compositions of stratospheric methane: 2. Two-dimensional model results and implications for kinetic isotope effects, J. Geophys. Res., 108, doi:10.1029/2002JD003183, 2003.

L63: The resolution of the control vector in EnKF methods is strongly limited by the ensemble size, if the number of ensemble members is much smaller than the rank of the error covariance matrix, then this method can give spurious results, see e.g.:

Houtekamer, P. L., & Zhang, F. (2016). Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. Mon. Wea. Rev., 144(12), 4489–4532. http://doi.org/10.1175/MWR-D-15-0440.1.

This limitation is not present in variational methods.

Reply: Thank you for the suggestion. The sentence is modified and the reference is added in the revised manuscript (L66-68).

L70: replace “in the” with “for”, i.e., “for carbon cycle research”

L73: Remove “The” before “assimilation” and change “window” to “windows” since you are not referring to one specific assimilation window, but to them generally.

L74: change “hour” to “hours”

Reply: All the above grammatical comments have been incorporated in the revised manuscript.

L75: The time resolution of the control vector is not the only consideration in the assimilation time window, but the time frame over which the system behaves linearly, and in what time frame the observations respond to the control variables (in this case, determined by atmospheric transport)

Reply: The sentence has been modified in the revised manuscript as follows (L79-80): “The time frame over which the system behaves linearly, and in what time frame the observations respond to the control variables such as, atmospheric transport, as well as observation abundance, must also be taken into consideration.”

L80: change “estimate” to “estimates”

Reply: Corrected in the revised manuscript.

Eq. 1: This equation should be re-written to express x^b and xmean^b (column vectors) as matrices with the same dimensions as X^b (or alternatively for any ith member of the ensemble using the ith column of X^b)

Reply: The equation has been modified in the revised manuscript.

L99: “and is derived” (missing “is”) and change “with” to “using”

Figure 1: Please change “broken line” to “dotted line” as “broken” could also be confused with the dashed line used.

L140: Please spell-out RTPS

L144: Please specify that Eq. 8 is referring to RTPS and not RTPP.

L169: change “accelerates” to “accelerate”

L170: change “observation” to “observations”

L191: change to “applied to the”

L198: change to “initial perturbations are applied”

Reply: All the above grammatical comments have been incorporated in the revised manuscript.

Section 3.3: I don’t see where the locations of the surface network sites are given. It would be helpful to include a figure of these.

Reply: The following sentence has been added to the revised manuscript. In this case, a figure of observation locations would not be useful.

“In the experiment 1, the simulated surface layer CH_4 concentrations at each grid for the entire globe were used as synthetic assimilated observations”

L205: change to “Errors in the estimated fluxes could arise…” I think the authors should also specify that this is in the context of the OSSE. In real-data inversions there are additional sources of potential error, e.g., modelled transport, inappropriate prior or observation uncertainties.

Reply: The sentences have been rewritten as follows (L212-214):

“In real data assimilation, there are additional sources of potential errors, such as, atmospheric transports, and inappropriate prior or observation uncertainties.”

L205: Please clarify if “inflation used” the authors refer to the inflation of the covariance matrix (as described in section 2.1), and if so, is this not coupled to insufficient ensemble size since the inflation is to account for an under dispersive ensemble?

Reply: The “inflation used” is referred as the inflation of the background covariance matrix. It is coupled with insufficient ensemble size. In our study, different inflation methods solve the under-sampling problem differently. To more clearly describe, the relevant sentence has been rewritten as (L210-212):

“Errors in the estimated flux could arise due to the insufficient ensemble size and also the implemented inflation methods to overcome the under-sampling, along with simplified forecast process of emissions. In real data assimilation, there...”

L210-212: I’m not sure what the authors mean by the following:

“We have estimated the CH4 flux for each grid by choosing the observation that influence the grid point using optimal cutoff radius (horizontal covariance localization) of 2200 km and vertical covariance localization of 0.3 in the natural logarithmic pressure (ln P) coordinate.” Could the authors please explain in more detail how observations were selected for assimilation?

In addition could the authors please explain:

“The optimized value of horizontal and vertical localizations…”

The localizations of which variables?

Reply: We explained by adding the following discussion in the revised manuscript (Section 3.3, paragraph 2):

“In this study, the CH4 observations are assimilated by applying the observation error covariance localization (Kotsuki et al., 2020) to reduce the spurious spatial correlation due to smaller ensemble size than the degrees of freedom of the system ($R \leftarrow R \times \exp\left(-\frac{1}{2}\{(d_h/\sigma_h)^2 + (d_v/\sigma_v)^2\}\right)$). Where d_h and d_v denote the horizontal distance (km) and vertical difference (log[Pa]) between the analysis model grid point and observation location. The tunable parameters σ_h and σ_v are the horizontal localization scale (km) and vertical localization scale (log[Pa]), respectively). Using the spatial localization technique, we have estimated the CH4 flux for each grid by choosing the CH4 observation that influence the grid point using optimal cutoff radius (horizontal covariance localization≃3.65$\sigma_{h,v}$; Miyoshi et al., 2007) with horizontal covariance localization (σ_h) of 2200 km and vertical covariance localization (σ_v) of 0.3 in the natural logarithmic pressure (log[Pa]) coordinate. The localization is performed to improve the signal to noise ratio of ensemble-based covariance. Numerous sensitivity experiments have been performed by varying the horizontal and vertical localization length in order to obtain the optimized CH4 flux that best compare with the truth.”

L229-234: I suggest removing the discussion of the assimilation window here and adding the new information to where this is discussed in section 2 (note, the assimilation window is discussed in section 2 (not 2.1).

Reply: Incorporated in the revised manuscript.

L242: XCH4 is not weighted by the prior and averaging kernel, but rather it is a weighting of the prior and the modelled mixing ratios, where the weighting is given by the averaging kernel.

Reply: We modified the equation and the text in the revised manuscript as follows.

$$XCH_4 = XCH_{4(\text{a priori})} + \sum_j h_j\, a_j (CH_{4(\text{ACTM})} - CH_{4(\text{a priori})})_j$$

“Where, XCH_4 is the column-averaged model simulated CH4 concentration. $XCH_{4(\text{a priori})}$ is a priori column-averaged concentration. $CH_{4(\text{ACTM})}$ and $CH_{4(\text{a priori})}$ are the CH_4 profile from ACTM and a priori, respectively. h_j is the pressure weighting function (j is the vertical layer index), and a_j represents averaging kernel matrix for the column retrieval which is the sensitivity of the retrieved total column at the various (‘j’) atmospheric levels.”

L247-251: Similar to my comment above, I think the selection of observations needs further explanation.

Reply: We added the explanation in the revised manuscript as follows (Section 3.4, paragraph 2).
"In this case, the CH4 flux has been estimated for each grid by choosing the CH4 observation with cutoff radius ($\simeq 3.65\ \sigma_{h,v}$) with horizontal covariance localization (σ_h) of 5000 km and vertical covariance localization (σ_v) of 0.35 in the natural logarithmic pressure (log[Pa]) coordinate. The optimal horizontal and vertical covariance localization values are chosen based on trial and error method (those best fits to estimate CH4 flux when compared with truth). A long cutoff radius has been chosen due to sparse observational coverage of GOSAT."

L258: Please change "It could be noticed that the…" to e.g. "Noteworthy is that the…"

L259: change to "15% larger error"

L315: change "discussed" to "discuss" and add "of" before "GOSAT"

Reply: These grammatical comments are incorporated in the revised manuscript.

Section 4.2, L315-329: I find these paragraphs quite confusing. If I understand well, these paragraphs should introduce the sensitivity tests carried-out in this section? If so, please start with the description of these tests, and simply state if the same set-up was used (or not) as Experiment 1.

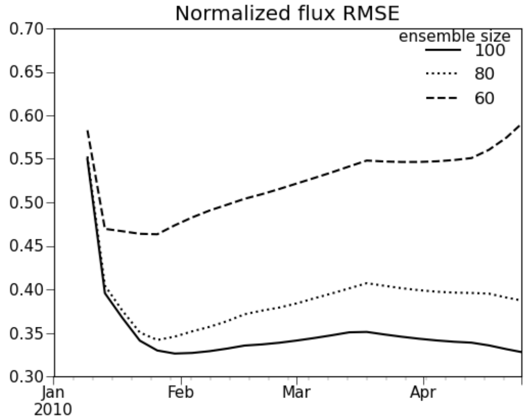
Why were the experiments for assimilation window and ensemble size performed on the satellite observation dataset and not on the "dense surface observation" dataset? Would the results, e.g., for assimilation window, change for surface observations compared to satellite ones?

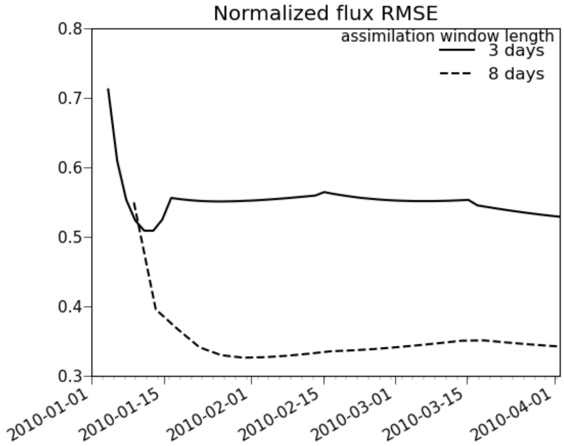
Reply: We modified these paragraphs in our revised manuscript as:

"In this section we discuss the LETKF flux estimation by assimilation of GOSAT synthetic CH4 concentration observations. Figure 6 shows the model simulated mean XCH4 concentration sampled spatiotemporally with GOSAT observations during January and July for the year 2010 (sampling method discussed in Section 3.4). In this case we have shown different LETKF sensitivity experiments such as; LETKF sensitivity to (1) FM, RTPS, adaptive multiplicative inflation (2) assimilation window (3) ensemble size, (4) chi-square test, (5) prior emission uncertainty. In the LETKF sensitivity experiments from 1-4, the initial ensemble spread provided similar way as Experiment 1."

Our Experiment1 is very ideal when we have surface observations at each model grid but in reality, we need to deal with the dataset described in Experiment 2. Therefore, we attempt to demonstrate these sensitivity experiments with more realistic observations. However, we perform the sensitivity experiments with different ensemble sizes and smaller window

lengths (shown below) as asked but found almost similar results as satellite one in our present LETKF setting. We added the following lines at the end of Section 3.3 where we first mentioned about the assimilation window and ensemble size in Experiment1:
"The ensemble size and assimilation window are chosen based on the CH4 flux estimation accuracy calculated by performing sensitivity experiment for ensemble size (60, 80, and 100) and assimilation window (3-days and 8-days), respectively (not shown)."





L359: By "the larger coverage of CH4 observations" in the 8 day assimilation window do the authors mean the greater sensitivity or "footprint" of the observations through the longer computation of atmospheric transport? This should be made clearer.

Reply: For better clarity we modified the paragraph as follows (Section 4.2.2):

"The LETKF data assimilation window length determines the time span of the observations assimilated in each assimilation cycle. We have shown the sensitivity of two assimilation window size configurations; 3 days and 8 days in supporting information Figure S4. Our sensitivity experiments with window size configurations show that 8 days long assimilation window estimates the CH4 flux with better accuracy (~10%) compared to 3 days assimilation window, because more observational information is incorporated into the system with 8 days long assimilation window. This study uses 8 days assimilation window for CH4 LETKF data assimilation."

L375: I think by "relatively diluted flux signal" the authors mean the weaker constraint on surface fluxes provided by satellite observations or the weaker connection of the satellite signal to surface fluxes. I think the term "diluted" is a bit vague.

Reply: We modified the sentence accordingly. Thanks for the suggestion.

Section 4.2.3: I think the chi-square test needs a bit more explanation. For instance, which normal variables are being summed in this test? It would be helpful to write the equation. Also, if I understand correctly what is being tested here, would a value greater or lower than one possibly be also due to an under estimation of the observation error covariance?

Reply: We elaborated the chi-square test in the revised manuscript and introduces equations relevant to this. Assuming that observation error covariance matrix R is known, and the Gaussian assumptions are considered appropriate as in the case of our OSSE, the equation (11) and (12) in our revised manuscript can be used as verification tools for the background error covariance P^b. Nevertheless, it is possible that the observation error covariance is also biased, which could also affect the chi-square estimates. In our OSSE for synthetic dense observation data assimilation we provided constant measurement error of 5 ppb. In our OSSE for synthetic GOSAT observation data assimilation we added the real GOSAT XCH4 retrieval error.

L387: "Our results suggest that, background error covariance matrix is highly underestimated in both RTPS and FM covariance inflation methods (Fig. 7b) and indicates much confidence in the model" – please explain why this gives the authors "much confidence"

Reply: We modified the paragraph as follows for better clarity (L405-410):
"Our results suggest that, background error covariance matrix is highly underestimated in both RTPS and FM covariance inflation methods (Fig. 7b). However, the chi-square values convergence towards 1 is better in the case of RTPS compared to FM covariance inflation method which indicates the improved representation of background errors and then more appropriate data assimilation corrections in the case of the RTPS inflation method."

L395-407: This section is difficult to follow. To start with, by "the flux estimation is extremely sensitive…" do the authors mean that the analysis fluxes are sensitive to the prior uncertainty, and by "provide larger prior uncertainty" do the authors mean to generate the prior error covariance matrix, or do they mean the perturbation to generate the prior fluxes? And in L402, by "the flux estimated error" do the authors mean the error between the analysis and true fluxes, and that this error would be larger when the inflation parameter is calculated grid-wise compared to region-wise?

Reply: Kindly find our clarifications as follows:

"the flux estimation is extremely sensitive…" Yes! the analysis fluxes are sensitive to the prior uncertainty used to perturb prior fluxes.

"provide larger prior uncertainty" mean perturbation to generate the prior fluxes.

"the flux estimated error" mean the error between analysis and true fluxes, this error would be larger when grid-wise initial ensemble spread will be provided.

For better clarity we changed the section as "CH4 LETKF sensitivity to initial ensemble spread" and modify the paragraph as follows (Section 4.2.5):

"In this case, we found that the analysis fluxes are extremely sensitive to the initial ensemble spread if prior fluxes perturbed with more than 5% prior uncertainty. Therefore, we used initial ensemble perturbation with only 2% prior uncertainty. Reducing the initial ensemble spread reduces the CH4 flux estimation sensitivity (>60%). However, it also poses a challenge to mitigate the under-dispersive background error covariance matrix. We performed LETKF data assimilations in this case with RTPS covariance inflation method

(α_{RTPS} = 0.9 optimized value is used here uniformly) with 8-days long assimilation window and 100 ensemble members and calculated the normalized RMSE between analysis and true fluxes (Supporting information Fig. S5). Noteworthy that, the estimated error between analysis and true fluxes (Fig. S5) with this setting (grid-wise initial ensemble spread) is still larger (25%) than the case when region-wise initial ensemble spread provide (Fig. 7a; 100 ensemble size). It suggests that initial ensemble spreads need to be carefully provided that best represents CH4 variability among ensembles to estimate the CH4 flux."

L406: The authors mention that machine learning could be used to determine the spread of the initial ensemble. I think this needs to be explained, i.e., how could machine learning help?

Reply: We removed this statement because machine learning is beyond the scope of this paper.

Replies to second reviewer comments (Reviewer2)

General Comments

Bisht et al present a data assimilation system for local ensemble transform Kalman filter, and evaluate that through OSSEs, particularly testing three covariance inflation methods (fixed multiplicative, relaxation to prior spread, and adaptive multiplicative) and two observing networks (surface dense network and GOSAT satellite network). This manuscript describes several interesting findings. I have three concerns.

1. This OSSE does not account model transport error, which would result in over-optimized solutions.
2. The number of ensemble members is not sufficiently greater than the dimension of the state vectors, which might bias the inversion performance interpretation.
3. Several sections require clarifications, as in the following "Specific Comments".

Thank you for reviewing the manuscript and providing us useful comments and suggestions.

Our reply to Point number 1

We mentioned transport error limitation in our OSSEs at the end of Section 5 of our revised manuscript as follows:
"We have not accounted for the transport error due to meteorological fields in this work (Patra et al. 2011), in case of real observations data assimilation a week-long window may introduce transport errors in CH4 analysis because of nonlinear growth of ensemble perturbations."

Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E., Hess, P., Ito, A., Kawa, S. R., Law, R. M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P. I., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: TransCom model simulations of CH4 and related species: Linking transport, surface flux and chemical loss with CH4 variability in the troposphere and lower stratosphere, Atmos. Chem. Phys., 11, 12813–12837, https://doi.org/10.5194/acp-11-12813-2011, 2011b.

Our reply to Point number 2

In our LETKF data assimilation system the localization approach is used to mitigate spurious correlation due to much smaller ensemble size than the degrees of freedom of the system. In our revised manuscript we added the discussion on localization approach (Section 3.3, Paragraph 2):

"In this study, the CH4 observations are assimilated by applying the observation error covariance localization (Kotsuki et al., 2020) to reduce the spurious spatial correlation due to smaller ensemble size than the degrees of freedom of the system……"

Our reply to specific comments (point number 3): for the specific comments our replies are given in black fonts for your comments in red.

Specific Comments

L43: "Cl in the stratosphere". Suggest including Cl in the troposphere.

Reply: Thank you for the suggestions.

CH4 loss to Cl takes place in the marine boundary layer (MBL), where sea salt is abundant, but CH4 is also destroyed in the stratosphere by reaction with Cl (e.g. Röckmann et al., 2004; McCarthy, 2003). In the modified sentence we didn't specify troposphere or stratosphere because our model simulations consisted of these in both the layers but we have not explicitly included Cl from sea-salt sources. We rewrite it as follows:

"other loss processes include oxidation by soil, and reactions with O1D and Cl"

Röckmann, T., J. ‐ U. Grooß, and R. Müller (2004), The impact of anthro- pogenic chlorine emissions, stratospheric ozone change and chemical feedbacks on stratospheric water, Atmos. Chem. Phys., 4, 693–699.

McCarthy, M. C., Boering, K. A., Rice, A. L., Tyler, S. C., Connell, P., and Atlas, E.: Carbon and hydrogen isotopic compositions of stratospheric methane: 2. Two-dimensional model results and implications for kinetic isotope effects, J. Geophys. Res., 108, doi:10.1029/2002JD003183, 2003.

L122: Typo, "the ensemble forecast of CH4 concentrations"

Reply: Corrected.

L79: "Advanced". Could you please specify what is the advanced aspect of this study, comparing to the previous studies using the same model? Is it the setup of the multi-window optimizing framework, or these inflation methods, or others?

Reply: The advanced aspect of this study includes the use of different inflation methods in our research and the simultaneous estimation of atmospheric concentration and surface fluxes of CH_4.

L188: "by 30%". Unclear if this is uniform bias. According to the later text, the perturbation is not uniform. Could you please specify the way to combine this "30%" with the following regional/grid level perturbation?

Reply: The systematic bias of a prior flux against true flux is assumed to be 30%. Besides, random perturbations equivalent to standard deviation of 6-8 % are added to the a priori flux as the initial ensemble spread. We modify the sentence such as:

"An initial perturbation with standard deviation of approximately 6-8% is applied to the a priori flux as the initial ensemble spread."

L196: "Experiment1". The word is misleading. Confused the readers if these experiments are corresponding to the experiments in section 4.1 and 4.2 (in fact, they are not).

Reply: We attempt to simplify it by modifying the sentences as follows (L202-207):

"This study performs two LETKF data assimilation experiments. In these experiments, we provided initial perturbation on regional basis over land (53 different land regions; Chandra et al., 2021) and at every grid over ocean, no spatial error correlation between grid points is considered among ensemble members. However, in Section 4.2.5, we also discussed the sensitivity of CH_4 data assimilation by providing initial ensemble spread at every grid by considering horizontal spatial error correlation between grid points among ensemble members, with a global mean correlation of 20%."

L196: "regional basis over land" and "every grid over ocean". Please explain why emissions over land and over ocean are perturbed differently.

Reply: We demonstrated in our CH_4 LETKF sensitivity to initial ensemble spread experiment (Section 4.2.5) that, the estimated error between analysis and true fluxes with grid-based initial ensemble spread (both over land and ocean) is significantly larger (25%) than region-wise (region-wise over land and grid-wise over ocean) ensemble spread.

L207: "Only surface layer CH_4 concentrations are used". Both over land and ocean? Please explain if the "dense observation network" include all surface grids or a collection of surface networks. If it is the first case, the word "dense observation network" is confusing.

Reply: It include all surface grids. We replace "dense observation network" to "dense observation data" in our revised manuscript.

L208: "added a constant measurement uncertainty of 5ppb". Please explain the way to add this 5 ppb (uniformly increase/decrease 5 ppb?). Also, typo, space between "5" and "ppb".

Reply: Uniformly increase 5 ppb uncertainty is being added. Typo corrected.

L236: "3.4 Experiment2". In experiment 1, "dense observation formulation", the author added measurement uncertainty of 5 ppb. Please explain why experiment 2 has no observation error, given the fact that satellite observations have larger uncertainties than measurements of surface sites.

Reply: It is already mentioned in the manuscript (L260-262) as follows: "we added the same retrieval (XCH_4) error as GOSAT to the XCH_4 (ACTM simulated) to make the OSSE more realistic and then attempt to estimate the true fluxes."

L406: “Machine learning tools could be used to”. Machine learning comes from nowhere. Please explain why it would help.

Reply: We remove this statement because machine learning is beyond the scope of this paper.