# Response to Referee

## General comments

The publication “Towards accurate methane point-source quantification from high-resolution 2D plume imagery” by Jongaramrungrang et al. deals with the important aspect of inverting atmospheric concentration gradients to fluxes or emission rates of not well-constraint CH4 point sources. In order to invert these observations, wind information in the measured area is needed and equally important as the concentration measurements themselves. However, as outlined in the introduction of the manuscript, acquiring wind information, preferably, simultaneously and at an adequate spatial and temporal resolution can be a challenging task. Therefore, the authors propose a new method, which solely relies on spatially resolved imaging data from airborne remote sensing instruments and high resolution large eddy simulations (LES)to estimate the prevailing wind speed during the time of the overflight.The wind speed is then derived from the shape of the observed plume and used during the inversion process to compute a flux of the investigated sources.The descripted method is a novel and promising approach to quantify CH4 point source emissions from aircraft but also potentially from space without having to rely on real wind measurements. The manuscript fits well in the scope of AMT and I recommend publication after some modifications along the line of the comments below. In general, the manuscript is well written. The method is described in a comprehensible way, however, some additional information concerning the figures would improve their readability (see also specific comments). Additionally, the authors should add some more information regarding the stated errors and their propagation to the predicted fluxes(see also specific comments).In my opinion, the manuscript could be strengthened by adding more extensive comparisons and applications to real data. So far, most parts of the approach were developed based on (theoretical) model (LES) studies, whereas only few real observations were used to support the novel approach. Therefore, I would recommend to expand on the already given examples in the manuscript:(1) controlled release experiment and(2) analyses of overflights shown in Figure 1 (see also specific comments).

We thank the reviewer for the constructive comments and appreciate the thoughtful review.

## Specific comments

P2, L12:Consider to also add publications regarding the HyTES instrument(e.g., Hulley et al., 2016) or other imaging instruments, which also have CH4 point sources successfully detected, e.g., MAKO (Tratt et al., 2014).

We added Hulley et al. 2016 accordingly to cite publication for the HyTES instrument.

P2, L24f:I would suggest to either only focus on remote sensing studies (by satellite and aircraft)and remove Conley et al., 2016,or to better distinguish between remote sensing and in-situ studies. If in-situ studies are included I would also recommend to add, e.g., Cambaliza et al., 2015,Gordon et al., 2015, and Lavoieet al., 2015.,which have also performed extensive analyses regarding airborne in-situ observations and resulting fluxes.

We would like to mention other studies done by an airborne in-situ approach using a mass balance calculation based on the enhancement downwind of the source. We added the citations from Conley et al., 2016; Jacob et al., 2016, Cambaliza et al., 2015,Gordon et al., 2015, and Lavoieet al., 2015.

P3, L6:"... under various background wind speeds and surface heat fluxes.": Later in the manuscript (P6, L14), it is stated that only one value for the heat fluxes was used("The surface sensible and latent heat fluxes are 400 and 40 W/m$^2$.").

We have conducted sensitivity analysis for other values of latent and surface heat fluxes and added an additional small section on this to show that the changes in sensible and latent heat fluxes do not significantly impact the results and are less crucial than wind speed.

P3, L11:"... and presence of ancillary information on the actual wind speed(Section 5.3).": Based on this statement, I would expect to see a comparison between the derived wind speed based on the new method and actual wind observations or reanalysis data in Section 5.3., however, I was not able to find a paragraph referring to actual wind speed observations or reanalysis data.

We added the result of the predicted wind compared to the value from controlled release experiment to section 5.3 P10 L30. The geostrophic wind speed is predicted to be 3.3 +- 1.2 m/s, compared to the surface sonic wind at the source measured at 1.9 m/s. This is consistent given that geostrophic wind is typically higher than the surface wind speed.

P4, L1:Why not also showing a real example observation of HyTES. Maybe, there is one available for the same scene as shown in Figure 1. If this is the wrong place, because in Figure 1, the authors intend to show the high variability of the plume structure during multiple overflights, I would recommend, if available, to add a comparison of an AVIRIS-NG and HyTES observation after Figure 5, where simulated plumes observed by the two instruments are compared.

We will add a scene from HyTES at the same location within approximately the same time for this comparison. However, this will be restricted to Clutter Matched Filter retrievals, which are similar but nor identical to the Optimal Estimation based retrieval for which the averaging kernels were computed.

P4, L1ff:This would also be a good place to add some more information regarding the applied retrieval algorithms.As Frankenberg et al., 2016, are already cited multiple times and the plume shown in Figure 1 (bottom right) appears to be similar to the one inFigure S1 in Frankenberg et al., 2016, I assume that either the matched filter technique or the IMAP-DOAS method, described in that publication, has been used to retrieve the columns shown in Figure 1 or the ones from the controlled release experiment at the end of Section 5.3used for a flux inversion.Similar for the HyTES algorithm if real observations are added.

For AVIRIS-NG scenes illustrated in Figure 1 the linearized matched filter technique was used. For HyTES, Clutter Matched Filter Approach (CMF) was used.

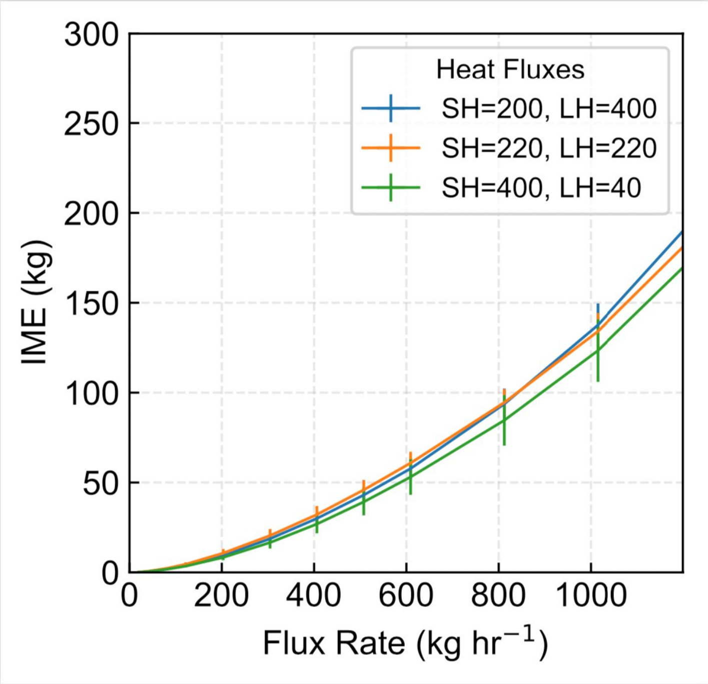
P4, L26:What is the basis of the chosen threshold of 500 ppm-m? As stated in (P3, L30), it is connected to the measurement precision of the AVIRIS-NG instrument and is thus a property of the instrument (Why does it then differ from the value of 200 ppm-m applied in Frankenberg et al., 2016, Section 'IME' in 'Materials and Methods'?). Furthermore, what is the reasoning in using the same threshold also for HyTES simulations, e.g., as shown in Figure 5?

The 500 ppm-m has been chosen as a somewhat conservative threshold, close to what we expect as single-measurement noise for typical scenes In Frankenberg et al, a lower threshold was used but data were also smoothed and plume segmentation algorithms applied. As for HyTES, we used the same threshold just to exemplify the differences due to averaging kernels only, as opposed to thresholds. We will make this clear in the revised version. For now, we don't have a unique quantitative detection threshold for HyTES, which would strongly depend on surface temperature and other variables. Thus, it is mostly for illustration at the moment.

P5, L5:Please add some explanation for the 'sum' and the parameter 'm' in Equation 2to the text.
In the Gaussian plume model, we assume equilibrium to find the concentration of plume in 3-dimensional space. For the vertical direction, when we assume an inversion height at Z_i, the model can assume reflective boundary condition. The parameter m multiplied by Z-i indicates the height that the reflection occurs and the summation over this parameter m shows the total concentration at each height within 0 to Z_i.

P6, L10-15:What is a reasonable range of the proposed values (initial inversion height, potential temperature,specific humidity, surface sensible and latent heat flux) for initializing the LES model? Are these educated guesses or are they based on actual field or reanalysis data? Given these ranges, the authors could also verify their assumption "... our method ... should not be significantly impacted ..." in the conclusions (P12, L6) by performing various LES runs using different initial values. I would agree with the authors that the column-integrated enhancements are not significantly influenced by the surface heat fluxes if the threshold were 0 ppm-m. However, I am curious if and how Figure 6 and 7 might change under various initial conditions, or whether they are well within the error bars.

These values are based on typical field campaign data. To show how the result from our method applies to the field of different conditions, we have added additional LES runs with different combination of sensible and latent heat fluxes (SH and LH) in two more cases: SH = LH, LH > SH to compared with the typical condition we use in this paper (SH > LH)



We found that the relationship between observed IME, flux rate and wind speed under new conditions (orange and blue lines) lie within 1-SD error from our original condition (green line). This shows that the uncertainties associated with the change in these conditions will not significantly impact our method and are captured well with the range of errors we have analysed.

Based on these result, we added a small section on this discussion.

P8, L27:Do the authors only mean Frankenberg et al., 2016, by "... has been ignored in previous studies." or are there further ones?

[Previous studies in this case refer to both Frankenberg et al., 2016 and Varon et al,. 2018 both of which assume flux is linearly proportional to IME as mentioned in the introduction.]

P9, L33:Could the authors add the fitted polynomial also to Figure 10?Why do the authors use a fifth-degree polynomial? Is there a physical relationship, which relates wind speed and plume angular width by a fifth-degree polynomial or is it just the 'best' fit?Could the relationship in Figure 10 also be explained by an exponential curve? Assumption: The relation in Figure 10 is determined by averaging over an ensemble of LES realizations. If this is done, I expect the resulting plume to be approximatelyGaussian as seen in Figure 4, e-f.

We added the fitted polynomial to Figure 10 and its caption, and based on some experiments we found the fifth degree polynomial to fit the relationship well without overfitting.

Although the resulting plume could be similar to Gaussian, because of the threshold value to mask out the plume, this could make the width of the plume not necessarily relate to the flux rates by exponential relationship. We will add the analysis for the exponential fit.

P10, L5-13:Basically, in this paragraph the authors summarize their developed method. I have some questions/comments to the used example.a) The error bars (shaded area) shown in Figure 11:Are they related to the errors shown in Figure 6 or to Figure 10? b) Have the flux rates, shown in Figure11, been corrected for the missing IME (as indicated in Figure 7)? c) How is the error (1-SD of the plume angular width) shown in Figure 10 related/translated to the wind speed error, which then linearly propagates to the estimated flux?

a) The error bars from Figure 11 are directly related to Figure 6. For a given IME, we found the possible range of fluxes for a given value of wind speed from the relationship from Figure 6.

b) Yes, the flux rate in Figure 11 is based on the relationship from Figure 6, in which the IME is the IME that is observed on the scene after applying the threshold.

c) Based on Figure 10, at a given angular width measured from the plume, we can predict the wind speed from the middle line. The associated uncertainties of the wind speed are approximated by the possible range within 1-SD area. We assume that by projecting a value of plume width onto the corresponding range of wind within 1-SD shading area, we obtain uncertainties for predicted wind speed that approximately represent 1-SD error for the wind speed as well.

We modified the text in this paragraph accordingly in the manuscript.

P10, L14-17: Could the authors be more precise in terms of the given "average percentage error" of 30%? I assume it is the mean value of the vertical error bars shown in Figure 12.However, as in reality not only entire fields/regions of CH4 sources (as then investigated in P10, L18-25)are investigated but also single plumes, which are typically observed only once, an interesting measure would also be the average of the absolute differences (also in percent) of the predicted flux rates and the corresponding prescribed flux rates.This would give an idea of the magnitude of the bias one can expect from the method.I assume that the observed bias in the flux for the controlled release experiment of ~32% lies within this computed theoretical value.

The average percentage error here is meant to be the average of the percentage differences between the predicted value and the actual value in Figure 12. Each point in Figure 12 is showing prediction for a single plume measurement, not for the entire field.
To make this clearer, we have accordingly revised the sentence to be “The average of the percentage differences (in absolute terms) between the predicted value and the actual value for single point source predictions is approximately 30%,...”

P10, L25: What do the authors exactly mean by “mean percentage of error” ? Is that the average of all differences between actual and predicted flux OR is it a similar error as computed for Figure 12 in P10, L14-17? If the authors refer to the latter one, I would suggest to also compute the average of all differences between actual and predicted flux (not the average of the absolute differences as in the previous comment). For example, a positive or negative value would then quantify an over-or underestimation caused by the method on average. The same exercise can be done for Figure 12 because it appears (as for Figure 13) that more predicted fluxes lie above the 1-to-1 line then below especially for larger fluxes.

We computed both the mean of absolute differences from all these aggregates which is 5.1% with a standard deviation of 3.9%, and the average of all differences (negative and positive) which results in 2.9% with the standard deviation of 5.9%. We added this result to this section.

P10, L26ff: The possibility to compare the novel approach, which is mostly based on ‘theoretical’ models,to real data is a huge strength of the publication. Therefore,I would recommend to expand this part of the publication. Some starting points are already given.First of all, the authors could add some more information regarding the already analyzed controlled release experiment allowing for a better judgement by the reader.Useful information would be (a)a figure showing the overflight and the retrieved plume and CH4 column enhancements(similar to Figure 1), (b)the fitted wind speed, which is then used to invert the IME to a flux, and (c)as the observation is based on a controlled release experiment, do the authors have access to real wind observations on-site or at least to meteorological reanalysis data, which can then be compared to the fitted wind to test its plausibility?Additionally, the authors nicely show multiple overflights of one source within ~25 minutes by the AVIRIS-NG instrument in Figure 1. It would be an interesting opportunity to apply the developed method to the four overflights shown in that figure and discuss the resulting fitted winds and inverted fluxes.

We have added more details of the fitted wind speed and the scene of the controlled release experiment. The results for applying this method to multiple overflights is taking longer to run. We will add this result to our revised version of the manuscript.

P10, L29:I assume the given estimated emission of 118 kg/hr is already corrected by the potentially missed IME as indicated in Figure 7, right? Additionally, could the authors elaborate on what error sources are included in the error estimate (of 30kg/hr) of the predict flux.
Yes, it has been corrected. In our method, we apply the observed IME to the relationship in Figure 6 which has the relationship between observed IME, wind speed and flux rates. The geostrophic wind speed is predicted to be 3.3 +- 1.2 m/. The error from the predicted wind speed results on the error of 30 kg/hr  has been propagated from the error in wind speed prediction. We have added this details in this section.

P15, Figure 2:Please clarify whether altitude on z-axis is given in meters above sea level or above ground level. Additionally, please harmonize the minimum altitudes of the computed averaging kernels, either to 0 m or to a specific surface elevation. Consider also adding the aircraft altitude(s) which the examples CAKs are valid / have been computed for.
The altitude on z-axis is given above ground level. In the Thermal case (HyTES) the flight altitude is an important factor for the CAK. In Figure 2, the CAK of HyTES was computed for an altitude of about 3 km. For the shortwave range, however, the CAK of AVIRIS-NG is not impacted significantly by flight altitude. In this paper, we focus our analysis on the results from AVIRIS-NG scenario.

P16, Figure 3:Consider adding the true IME (idealized threshold of 0 ppm-m) to the caption.Additionally, consider adding labels for the stability classes to the caption so that the reader sees immediately their meaning without looking up the relevant information in the cited publications, e.g.,A = very unstable; B = moderately unstable; …

The idealized IME when threshold is 0 ppm-m has been added accordingly. Without any threshold, the IME actually depends on the box size as enhancements approach the domain boundaries.  . In this case, we calculated the IME that would have been observed within the box as shown in the figure if the threshold were to be 0. We have added the meaning for each label as A=very unstable, B=unstable, C=slightly unstable. We have added this description to the caption of Figure 3 for more clarity.

P17, Figure 4:Consider adding the true IME (idealized threshold of 0 ppm-m) to the caption (as suggested forFigure 3),and the variance of IME(of the 60 individual snapshots)to each plot in the right column for the three cases of wind speed so that the reader can assess the statement from (P7, L10f).

Based on the values of IME across snapshots, we computed the variance of IME, and have added these values to the right column panels. We have also added the idealized IME similar to what was described above in P16, Figure 3.

P18, Figure 5:Consider adding the true IME (idealized threshold of 0 ppm-m) to the caption (as suggested forFigure 3)
We have also added the idealized IME similar to what was described above in P16, Figure 3.

Figure 3-5:For clarification: The wind shown in Figure 3 (Gaussian plume model) is not directly comparable to the wind shown in Figure 4 and 5 because the latter one is the geostrophic wind, whereas the former one is the wind at plume level(s), correct?
Yes, that is correct. We added this clarification to the Figure captions.

P19, Figure 6: Which threshold was used for Figure 6, 500 ppm-m? Consider adding this information to the caption.
Yes, we have used the threshold of 500 ppm-m in the analysis of Figure 6. We have added this information to the caption of Figure 6.

P20, Figure 9:Please add meaning of vertical bars to caption.
The vertical bars represent the standard deviation of the normalized IME at a given angle across all snapshots. We added this detail to the caption of Figure 9.

P21, Figure 10:Please add meaning of shaded area also to caption.
The shaded area represents one standard deviation from the mean plume angular width for each wind speed. The standard deviation is computed across different values of flux rates and snapshots. We have added the meaning of the shaded area to the caption of Figure 10.

P21, Figure 11:Please add meaning of shaded area also to caption.
The shaded area represents one standard deviation for the flux rate at a particular value of IME for each wind speed. The 1-SD for the flux rate is approximated by the possible range of flux rates resulted from an observed IME and wind speed in Figure 6. I added the meaning of the shaded area to the caption of Figure 11.

P22, Figure 13:Consider using a density plot for better visualization of the data cloud.
Because we would like to illustrate the comparison between predicted and actual values at each point, showing the data cloud could be more helpful in our case.

Technical corrections

P1, L18:“... Large Eddy Simulation...” “... Large Eddy Simulations”
P1, L29:“... large geographical area ...” “... large geographical areas ...”
P2, L13:“... at a resolution of 3-m ...” “... at a resolution of $3x3m^2$ ...” or “... at a resolution of 3 m ...”
P2, L20:“... the retrievals measure the fine ...” “... the instrument observes the fine ...”
P3, L26:“... approximately 15 minutes revisit time.” “... approximately 10 minutes revisit time.” (compare to Figure 1)
P5, L27:“... the plumes structure.” “... the plume’s structure.”
P6, L3:LES is already defined in(P3, L5)
P10, L11:“... 1-SD error bars in the plot.” “... 1-SD error bars are shown in the plot.”
P10, L22:“... large enough represent ...” “... large enough to represent ...”
We corrected all of these technical corrections accordingly.