Response to anonymous referee #1

We would like to thanks the referees for their thorough comments.

General comments:

It is known that the short-term regional scale variability in total column CH4 is dominated by local sources is well as the dynamics of weather systems. The model system that has been described takes the first into account, but not the second. I mean, it represents the weather but not the corresponding regional-scale patterns in XCH4. I was looking for ways in which the boundary conditions of CH4 where taken into account, but didn't find any. As a result, the fluxes that are derived may well account for unaccounted variations at the domain boundary.
In the inversion, we have represented the boundary conditions by selecting the observations of X_{CH4} dry air mole fractions unaffected by the local sources in our domain. These observations correspond to ideal wind conditions for each day assuming that over such a small domain, no spatial gradient in X_{CH4} mole fractions are expected and therefore one observation location is sufficient. The wind direction is changing but the minimum value is consistent over the two days at about 1.83-1.832ppm, except for a brief decrease to 1.825ppm on January 16th (for about 20 minutes). On January 15th, different instruments measure this value depending on the location relative to the wind direction. Two main reasons motivated our approach. First, we considered that the relative magnitude of the feedlot emissions is several orders of magnitude larger than any other sources in the area. The Chino feedlot is one of the largest sources of CH_4 in the Los Angeles basin, which greatly simplifies the problem of the background conditions (i.e. the relative contributions from other sources are reduced). Nearby sources could still be significant if they are located at a very short distance from the sensors. But drive-around data and upwind EM27/SUN observations did not suggest any prevalent sources around the feedlot, as illustrated by the various EM27/SUN sensors measuring the same background at different times on January 15th. Second, if we were to consider a more regional approach to the problem, we would need a very accurate representation of CH_4 mixing ratios (to few ppb) and therefore an accurate mapping of CH_4 sources within several kilometers around the feedlot in order to simulate the signals from other contributors. This approach is likely to produce larger errors due to missing sources and incorrect magnitudes in the inventories as well as errors in the atmospheric transport, which could all introduce significant biases in the background mixing ratios (much larger than 2ppb). We would also need to define the inflow of CH_4 from outside Los Angeles which would require another set of data outside the city limits. For these reasons, we decided to use an observation-based approach to avoid the complications due to model and inventory errors. We have modified the text to discuss the problem and explain more precisely our approach.

In Chen et al. an upwind-downwind approach was taken, which might indeed work under ideal conditions, with a well defined wind direction. As can be seen in Figure 4, this logic falls apart for low wind speeds. It would have been instructive to plot the zero line in this figure.
We agree and zero lines have been added to each plot.

Unlike the 24th of January, the gradient with respect to Harvard 2 is going everywhere, most notably on the 15th. Therefore I don't understand the role of Harvard 2 is a background on days with a low wind speed and doubt that it can be used that way. The authors may argue that variations in the background may be less important on days with lower wind speed, but this could en should then be demonstrated.

We needed to have the same background instrument for all 3 days of measurements to test the mass balance approach. The Harvard2 X_{CH4} mean and standard deviation are the lowest of all the observations (Figure 3); which is why we have chosen this instrument as a measure of background conditions. We added a sentence on this limitation in the text.

Before the WRF-LEF model is used to fit the FTS data it should be demonstrated that it has a reasonable skill in simulating the observed variability. The fit in S2 doesn't look great (what are the R2 values?), which makes me wonder about the comparison with the prior model. In this context it would be very useful to compare the use of WRF-LES with the use of WRF alone. What do we gain using LES? The results may not look great, but we'd learn about where the remaining problems are.

We have added a paragraph and a reference to a paper that will be published in the coming weeks (Gaudet et al., 2017). This study presents the ability of the Large Eddy Simulation mode to represent plume structures over short distances (less than 10km) compared to the mesoscale mode. Here, we have no simulation of the CH_4 concentrations from coarse-resolution grids. Only 111-m concentrations were simulated.

In line 370 it is mentioned that the modelled wind speed has an error of 'only' 1 m/s increasing further in the PBL. However, on the low wind-speed days this is a very substantial fraction of the total wind speed (according to Figure 1). What is the sensitivity of the emission estimation to errors in the modelling of wind speed and direction? I'd like again to stress that it is important that we learn about what is critical in this approach. The size of the errors in wind speed and direction warrant closer inspection and a test of the potential impact.

The impact of transport errors on inverse estimates is an important topic that has been discussed in previous studies at lower resolutions (e.g. Lauvaux and Davis, 2014). In our study, it remains difficult to address this problem without the use of an ensemble of simulations. Instead, we decided to improve the transport using the WRF-FDDA system. We refer to another study (Deng et al., 2017) that will be published in the coming weeks. In this paper, we performed multiple simulations using WRF-FDDA and propagated the impact of the assimilation into the inverse fluxes. We showed that transport errors were significantly reduced when assimilating vertical profiles of meteorological observations. Here, our estimates of model errors, based on the chi2 normalized distance, suggest that transport errors are relatively small (less than 3ppb) allowing us to produce an inverse estimate for the area. We added a paragraph about model errors.

It is clear to me why the mass balance approach is not used on the 15th and 16th of January. However, it is not clear why the WRF-LES method is not used on the 24th. This would provide a good opportunity to make a direct comparison between the two methods. It may not be the best case to demonstrate the added value of WRF LES, but as a general consistency check to include this comparison is nevertheless necessary.

We decided to avoid January 24th as the wind direction remains constant during the entire day. The inversion uses primarily the information from various wind conditions to constrain the distribution of the CH_4 sources, similar to a triangulation approach. The results for January 15th and 16th revealed that sources are highly variables across the domain. Therefore, the lack of constraints on the spatial distribution (i.e. no change in wind direction) would have been limiting on January 24th. We explain our choice in the section 4.3.1.

Figures 4 and 5 confirm my worry about the emission pattern that is shown in Figure 12 and I wonder why it doesn't receive more attention. The pattern correlates with the configuration of the measurement 'network'. For an inversion the easiest way to fit the data is to modify the emissions in the same grid box as where the measurements are. It may either be that the measurements are very locally influenced, in which case they are not really representative of the domain that is being optimized. Otherwise the inversion may be trying to fit uncertain variables that are not part of the state vector, such as boundary conditions (see my earlier comment) or the emission distribution within 2x2km2 regions. With in situ measurements at LANL as high as 30 ppm CH_4, this may well be an important factor. To me this sounds certainly like a more relevant factor to mention than the emission of a 1 cow per year power plant.
Figures S4 and S5 provide additional details to understand the attribution of signals in the state vector. Overall, the inversion is using the triangulation of information from various wind directions, which is why the two-day inversion is likely to produce the best estimate over the period. In general, with 2 sites, the flux corrections are applied to the nearby pixels. But this result is not valid when using 3 or 4 sensors. For example, on January 15th, two sensors (i.e. LANL and Caltech) show no correction in their vicinity whereas major changes are located near the other two sensors. The optimization is consistent across configurations of 3 or 4 sensors, detecting a major source in the North-East and the South-West, but not near the other two sensors. On January 16th, the fluxes remain unchanged near the Caltech instrument which demonstrates that the system is doing better than simply attributing the corrections to the nearly pixels. We provide additional comments based on Figure S4.

It was not clear to me why the simulated annealing approach was chosen. Since the inversion problem is linear, I don't see why its solution should be any different from the Bayesian method. If I understand well, the method was introduced to deal with the difficulty to define the B matrix. But how does simulated annealing solve this problem? Is it just efficiency at which different options for the B matrix can be tried out? If the greens functions are available, then I wonder how simulated annealing could be faster than a Bayesian inversion. This should be explained better.
We have described the Simulated Annealing (SA) in a separate paragraph. The main reason to perform the inversion in two steps is that SA has no prior information which means that if a pixel is not observed, the system will produce an infinite number of solutions. In addition, SA is producing scores (i.e. mismatches between observations and model concentrations) but no optimal solution is guaranteed. SA simply scans the space of solution (i.e. state space) without converging, whereas the Bayesian inversion will provide the optimal analytical solution considering our prescribed prior errors. In theory, if the optimal solution was found by the SA, then the two methods should agree, but there is no guarantee of such agreement.

<u>Specific comments:</u>

Line 98: 'on a high-wind data'. What does this mean? Why is Chen 2016 not included in Table 1?
This has be changed to 'recorded on favorable meteorological conditions (e. g. on a high-wind with constant direction)'. Chen et al. 2016 estimation has been added to Table 1.

Line 117: In this sub section I am missing numbers for the estimated accuracy of the mobile column measurements. How realistic are the fits to the data derived later on in the light of these uncertainties?

We have added a line about accuracy of these mobile column measurements: "Using Allan analysis, it has been found out that the precision of the differential column measurements ranges between 0.1-0.2 ppb with 10 min averaging time (Chen et al., 2016)".

Line 196: 'one way nested grids': S1 should provide further information about nesting. For CH4 the logical on-way nesting is from the small to the large domain. For meteorology, however, the reverse seems true. Some further explanation is needed.

"one-way" refers to the meteorology only. CH_4 concentrations are simulated only in the 111-m grid. Considering the low risk that CH_4 molecules would re-circulate over the area after having been advected out of the small domain, we ignored the coupling between the grids for CH4. This problem is more significant at larger scales when circulation of air masses is more circular around pressure centers or due to the terrain/surfaces. We provide additional details in S1.

Line 207: To avoid confusion about optimizing methane fluxes and optimizing meteorology it should be stated more clearly that 'data assimilation' is referring to the latter here.

This has been clarified. We have added "to optimize meteorological fields" after "Data assimilation".

Line 255: What does the random draw refer to: the starting point of simulated annealing, the first guess, a random modification of the prior uncertainty? Does this modify the actual cost function that is optimized?

Simulated Annealing is a random walk which implies that random draws of the prior emissions are made at every iteration. There is no optimization of a cost function in Simulated Annealing, only a score calculated for every possible solution. We clarified the sentence.

Line 257: 'Mean absolute error' This approach favors the setup that gives the largest freedom to the prior fluxes. This need not be the best solution, or maybe I do not understand what is meant here (see the previous point).

The impact of using the Mean Absolute Error is minor. We could have used the squared distances (like in a typical least-square regression analysis) or the Mean Error which would account for source compensation.

Line 290: 'C' io 'SC'?

This has been changed

Line 373: 'More relevant to this study' It is unclear why this should be the case. The difference between the two metrics represents a cancellation of wind speeds in calculating the mean. Such errors could still affect the estimated emissions, e.g. when winds with cancelling errors come from different directions.

The sentence refers to the fact that the first error is computed for the Free Atmosphere (Free Troposphere) whereas our local enhancements are located in the PBL. We clarified the sentence.

Line 458: It is clear that the measurements in Figure 11 represent ruminant emissions. However, these samples are not representative of the air masses that are sampled with the FTS instruments. Therefore this result does not refer to the origin of the enhancements in total column CH4 discussed earlier. Confusion should be avoided on this point.
We have added 'measured by the Picarro instrument' to clarify this point.

Line 579: 'The main advantage ...' This statement should be supported by evidence or a reference to other work. We have added a reference (Wunch et al. 2011).

Figure 1: The wind roses do not add up to 100%. For clarity, we have not shown in this Figure wind data corresponding to null wind speeds, which is why the roses do not add up to 100%.

Figure 5: Please repeat that anomalies are relative to Harvard 2. We have added this in the caption.

Figure 6: It should be made clear that these errors refer to errors in WRF-LES simulated wind speed and direction. We have added 'Errors in WRF-LES simulated wind speed and direction:' at the beginning of the legend.

Figure 7: Mention the minimum WMO threshold value. We added the value in the caption.

S1: How about the top boundary in the LES part of the domain? The molecules of CH_4 are not well-mixed over the PBL. For this reason, we avoided the use of PBL height as a relevant criterion in our calculation.

<u>Technical corrections:</u>

All of them have been changed

Line 288: 'V is' io 'Vis'

Line 337: area of Chino in m2

Line 353: 'has' io 'have weaknesses'

Figure 8: Information on axes parameters and units should be given in the figures themselves instead of the caption. Please also add labels to the rows (as is done for columns)

Figure 9: Figure axes and legend without labels.

Title S1: 'models' io 'modes'

Figures S4 and S5: Text is too small to read.