Dear Referee,
dear Editor,

we would like to thank the referee for the detailed revision of our manuscript. In the following we respond also in great detail. The referee's comment/text and our response/text can be distinguished by blue and black font, respectively.
Please note that not all the details of our response can appear in the main text of the revised manuscript. For the main text we will focus on the presentation and validation of the new method. The details (the following Figs 1-6 and the related text) can then be shown in one or two additional appendices.

The authors propose to combine independent retrievals of CH4 from TROPOMI and IASI, which are in differing orbits, using well-known optimal estimation techniques. They compare the combined and individual L2 retrievals to a number of independent measurements while accounting for the individual sensitivities of the retrievals. Those comparisons are interesting as they show the strengths and weaknesses of IASI and TROPOMI.

Thank you!

They describe in detail the method to combine a column retrieval and a profile retrieval using optimal estimation techniques. This extends previous approaches, e.g., Luo et al, 2013, with a useful twist that could be applicable to other instrument combinations.

Thank you!

Referee comment #1
The value of this combined product, however, is not entirely clear. The theoretical analysis presented in the appendix neglects the fact that IASI is in a morning orbit and TROPOMI is in an afternoon orbit. The coincidence criterion that they propose is reasonable for evaluating L2 products against independent data but it is not adequate for actually combining data.
For this study, they need to directly account for the dislocation error, which could be modeled as a covariance. This could be done in a simulation context using, for example, CAMS methane (https://atmosphere.copernicus.eu/charts/cams/methane-forecasts).

IASI is on an orbit with descending node equator crossing at 9:30 mean local solar time. TROPOMI is on an orbit with ascending node equator crossing at 13:30 mean local solar time. This causes the following typical time difference for observing the same location: at northern high latitudes 0-2 hours, at the equator 3-5 hours, and in southern hemispheric mid-latitudes typically 4-6 hours. Furthermore, there are horizontal dislocations. We combine IASI and TROPOMI observations for which the ground pixel horizontal location can differ by up to 50 km. However, we estimate this dislocation error to be generally smaller than the combined measurement noise error. Actually, for the comparison of the total columns of TCCON and TROPOMI (or the total column of the combined product) the dislocation is more important than for the combination of IASI and TROPOMI. The reason is that the dislocation has a relatively strong effect on CH4 in the lowermost troposphere. This region is important for total column data, but it is not very important for the combination of IASI and TROPOMI, mainly because IASI is rather insensitive to this region. In the following we provide more details:

As recommended by the referee we estimate the corresponding dislocation error using the CAMS methane forecast product at highest available resolution (~ 9 km, Barré et al., 2021). We investigate how the methane profiles vary with time and horizontal distance. The result of this investigation is depicted in Figs. 1-3. Figure 1 shows the RMS of the difference between the reference methane profile and the profiles dislocated with respect to the reference. The dashed black lines indicate our collocation threshold values used for the combination (TROPOMI and IASI are only combined as long as horizontal distance of their ground pixels is within 50 km and the time difference is within 6 hours). Naturally, the respective RMS values are increasing with increasing horizontal distance and time difference. The values are largest in a small layer close to the surface and in the stratosphere, but relatively small in the free troposphere. For a horizontal dislocation of 50 km the RMS value is about 2% very close to the surface, between 0.3 and 0.5% for the rest of the troposphere and then it increases again to about 2% above 25 km altitude. For a time difference of 6 hours the RMS value is about 2.5% in a very small layer above ground, 0.6-0.8% in the free troposphere below 10 km, and it reaches about 1.5% at 15 km and 3.5% at 30 km altitude.

Figure 2 depicts the vertical correlations of the differences between the reference methane profile and the profiles dislocated with respect to the reference. We observe that the vertical correlation lengths (distance where correlation decays to 0.5) are rather short: 100-200 m close to the surface, 1000 m in the middle/upper troposphere, and in the stratosphere they increase to about 6000 m.

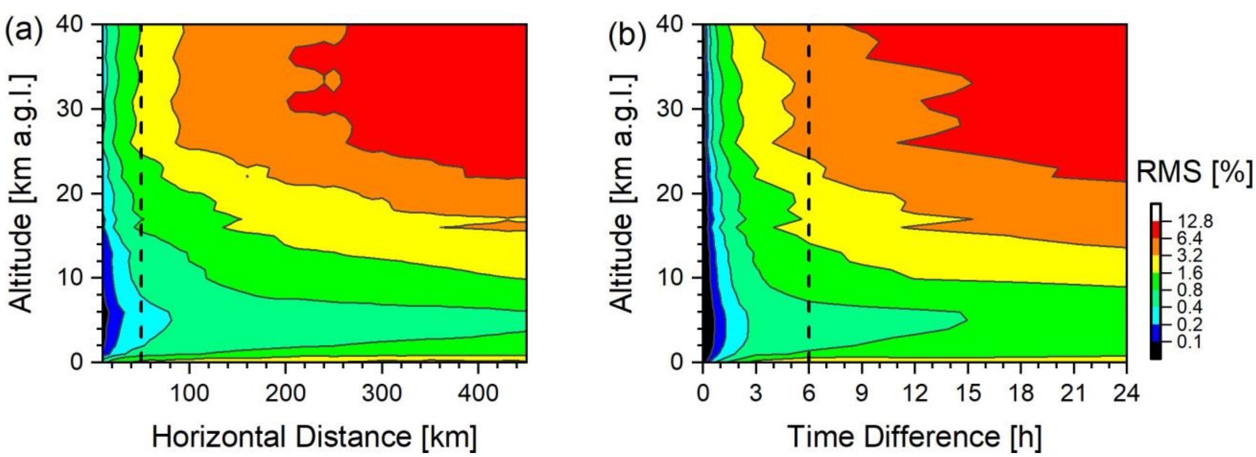

Figure 1: Root-Mean-Squares (RMS) of the difference between the reference methane profile (location 49.1°N and 8.4°E, corresponding to the location of Karlsruhe) and profiles dislocated with respect to the reference: (a) horizontal dislocation; (b) temporal dislocation. The dashed black lines indicate the collocation threshold values used for valid combinations of IASI and TROPOMI.

Figure 3 depicts the same as Fig. 1, but for vertically integrated total and partial columns. Naturally, the RMS values increase for increasing horizontal distance and time difference. For our horizontal collocation threshold values of 50 km the RMS is about 0.2% for the total column data. For our time difference collocation threshold of 6 hours it is about 0.3% for the total column data. For the tropospheric and upper tropospheric / lower stratospheric partial columns the respective relative RMS values are slightly larger.

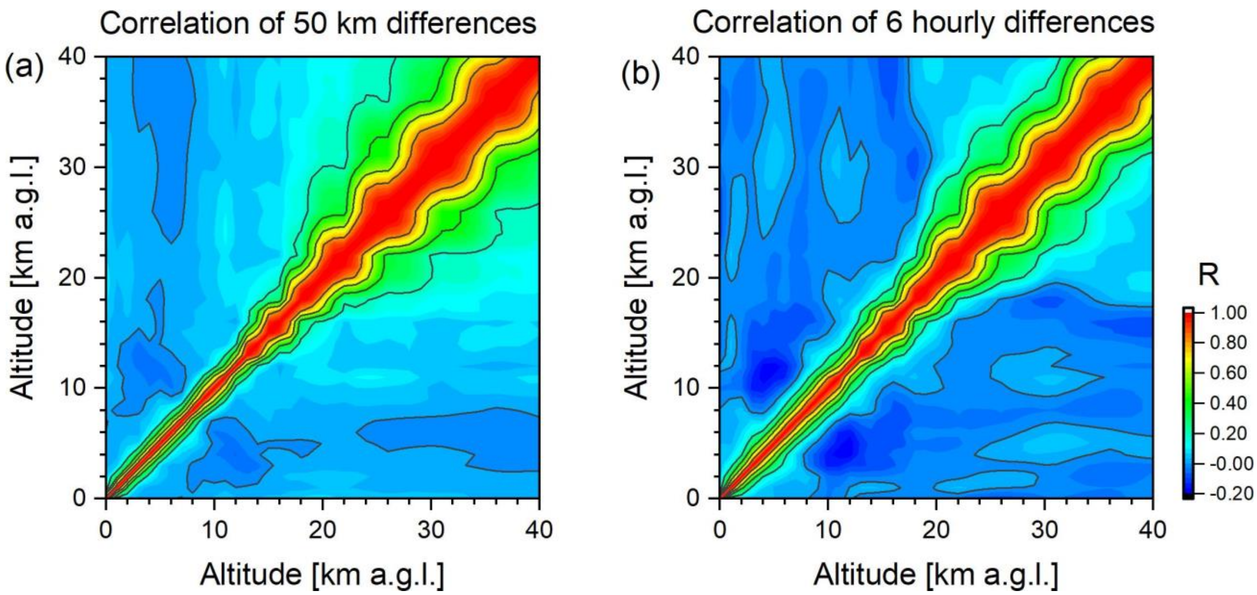

Figure 2: Vertical correlation matrices for the difference between the reference methane profile (location 49.1°N and 8.4°E, corresponding to the location of Karlsruhe) and profiles dislocated with respect to the reference: (a) horizontal dislocation of 50 km; (b) temporal dislocation of 6 hours.

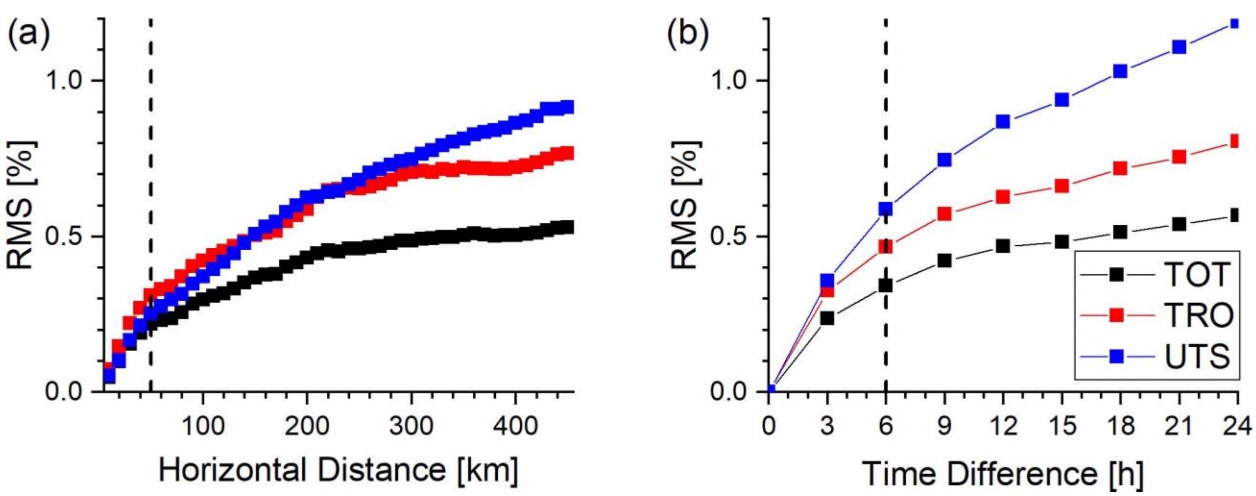

Figure 3: Same as Fig. 1 but for column integrated data. TOT: total column; TRO: tropospheric partial column; UTS: upper tropospheric / lower stratospheric partial column.

For calculating the error in the combined profile due to the horizontal and spatial dislocation between IASI and TROPOMI we set in Eq. (1) of the manuscript $\hat{x}_{\mathrm{I}}$ to $\hat{x}_{\mathrm{I}} + \delta x$, where δx is the dislocation uncertainty of CH4 as shown in Figs. 1 and 2. This means a new term in Eq. (1) that gives the respective error in the combined profile: $(\mathbf{I} - \mathbf{L}^{-1} m a_{\mathbf{T}}^{*T})\mathbf{A}_{\mathbf{I}}^{\mathbf{l}} \boldsymbol{\delta} x^{\boldsymbol{l}}$ (the error covariance is then $\mathbf{A}_{\mathbf{dl}}^{\mathbf{l}}\mathbf{S}_{\delta\mathbf{x}}^{\mathbf{l}}\mathbf{A}_{\mathbf{dl}}^{\mathbf{l}\,T}$, where $\mathbf{S}_{\delta\mathbf{x}}^{\mathbf{l}}$ is the covariance matrix for the CH4 dislocation uncertainty and $\mathbf{A}_{\mathbf{dl}}^{\mathbf{l}} = (\mathbf{I} - \mathbf{L}^{-1} m a_{\mathbf{T}}^{*T})\mathbf{A}_{\mathbf{l}}^{\mathbf{l}}$ is the dislocation averaging kernel.

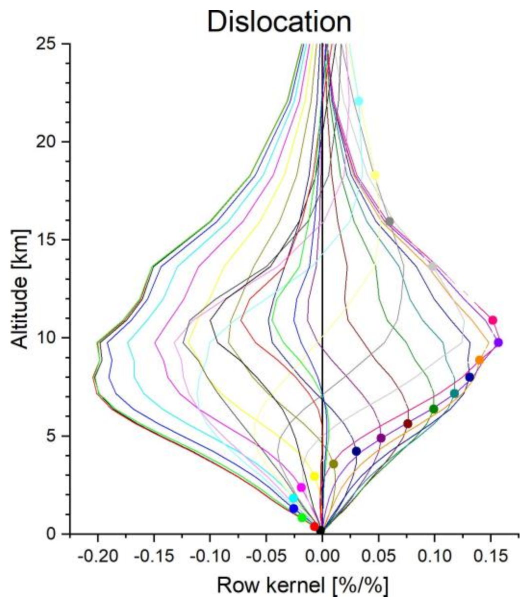

Figure 4: Example of dislocation kernel A_{dl}^{l} for the same late summer observation as used in the context of Figs. 1-3 of the manuscript.

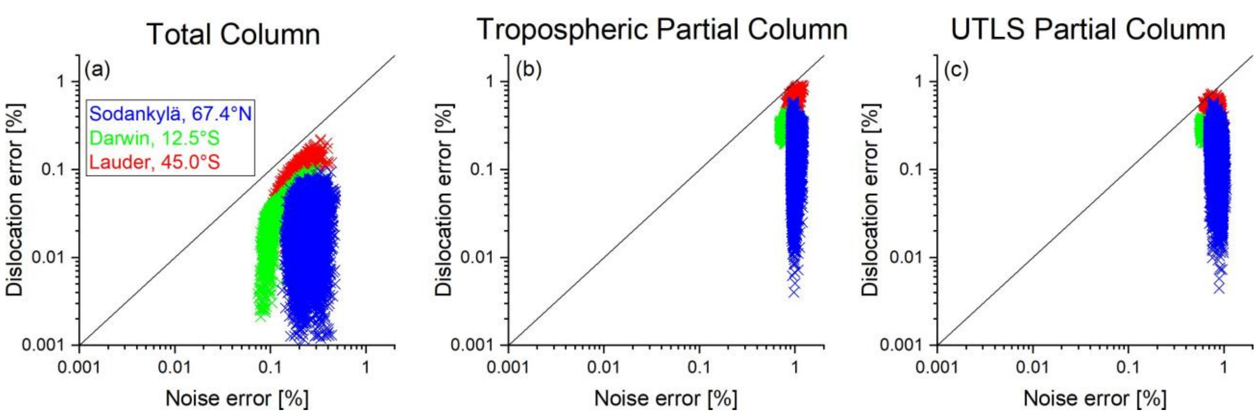

Figure 5: Comparison of the combined products' dislocation error (due to the CH4 dislocation uncertainty) and the noise error (an example of the typical temporal dependencies of the noise error is shown in Fig. 5 of the manuscript). The comparison is depicted for a northern high latitude location (Sodankylä, blue crosses), a tropical location (Darwin, green crosses), and a southern middle latitudinal location (Lauder, red crosses). (a) Total column product; (b) Tropospheric partial column product; (c) Upper tropospheric/lower stratospheric partial column product.

Figure 4 shows an example of this dislocation averaging kernel. For the altitudes where the dislocation uncertainty of CH4 are largest (close to ground and above 20 km) the dislocation kernel has rather low values (i.e. there the combination procedure has only limited sensitivity to the dislocation uncertainty). Figure 5 depicts a comparison of the combined products' noise error (see also Fig. 5 of the manuscript) and the error in the combined data product due to the dislocation uncertainty for three different sites. For the northern high latitude site (where horizontal and temporal dislocation are of similar importance) as well as for a tropical and southern hemispheric middle latitude location (where the temporal dislocation is dominating), the dislocation uncertainty is generally much smaller than the noise error (please note that the axis in Fig. 5 are on a logarithmic scale).

Referee comment #2

This is important in part because the paper neglects a key question: who would want to use this data? Most scientists aren't interested in CH4 concentrations, they are interested in the fluxes that produce them. Given their characterization, IASI and TROPOMI can be used in this context. Models can readily account for the differences in time of day and how winds may shift the origins of morning and afternoon air parcels. The proposed method would improve DOFS but potentially at the expense of the anomalies that a model would exploit infer fluxes. They need to address this issue.

The referee is right, that for many applications CH4 concentrations can be considered as an intermediate product. The emission rates and surface fluxes are often scientifically most intersting. From a purely theoretical viewpoint data assimilation is the method of choice for optimally exploiting emission/flux information of the satellite observations. These assimilation systems are continuously developing and improving, but currently no operational product is available that assimilates IASI and TROPOMI together and estimates methane fluxes at a high resolution (e.g. ≈9 km, which is similar as the resolution of the satellite observation). To our knowledge there are currently two principally different CAMS products that have satellite methane data assimilated:
There is the CH4 inversion production chain of CAMS. It is described in Seegers and Houweling (2020). The scheme works on a 2° x 3° (latitude x longitude) resolution and estimates the CH4 fluxes by assimilating different kind of observations (NOAA surface CH4 and GOSAT satellite XCH4 observations). The assimilation scheme works on a horizontal resolution that is much coarser than the resolution offered by the state-of-the-art satellite IASI and TROPOMI sensors, and provides no information about the fluxes on a horizontal scale smaller than 300 km. Furthermore, in the current setup neither TROPOMI nor IASI data are assimilated by the CAMS CH4 inversion system.
There is the CAMS CH4 assimilation product (Massart et al., 2014), for which currently methane data from the satellite sensors GOSAT and IASI are assimilated. The model used for the assimilation has a resolution of 25 km, which is significantly larger than the ground pixel sizes of IASI and TROPOMI. Furthermore, there is currently no assimilation product that uses IASI and TROPOMI CH4 data simultaneously. Within this CAMS CH4 assimilation procedure only the concentrations are corrected not the emissions/surface fluxes.

Not all research questions can be addressed by CAMS data alone. In particular, the estimation of fluxes on very fine scales and/or estimations that exploit the current state-of-the-art satellite data products from TROPOMI and/or IASI need alternatives/extension to the currently available data generated by the CAMS assimilation schemes. Different approaches are currently in discussion. Barré et al. (2021) proposes to use the CAMS high resolution forecast (≈9 km) for calculating departures of this forecast from actual TROPOMI observations. The departure is an indicator for inconsistencies between CAMS and TROPOMI. Barré et al. (2021) found a bias in the departure of about 20 ppb and a typical departure signal of about 15 ppb (the scatter in the departure values shows a 1 sigma standard deviation of 15 ppb). If the departures were caused due to assuming wrong emission rates/fluxes in CAMS, they could give some hints for constraining uncertainties in the currently assumed emission rates/fluxes. However, the departure might also be due to errors in CAMS that are not related to the errors in the used emission rates (Barré et al., 2021) or the TROPOMI data could have uncertainties. Furthermore, the high resolution forecast is initialized by the CAMS CH4 analyses that contains CAMS emission/flux signals only at the 25 km resolution. This means that the emission rate/flux signals observable in the departure data mainly represent the 25 km scale (not 9 km or the fine scale offered by TROPOMI). Tu et al. (2021) present another method that allows for estimating emission rates by using satellite data together with rather simple plume dispersion calculations (based on horizontal wind data). The method works without comprehensive CH4 model data (like CAMS) and showed to give reasonable results.

We think that application of our IASI+TROPOMI combined product for emission rate/flux studies is beyond the scope of our paper (this paper presents the combination method and discusses their quality). However, we would like to discuss two example of applications. In the Tu et al. (2021) study the combined IASI+TROPOMI data is actually already applied. It enables to prove the robustness of their method, by demonstrating that the estimated emission rates are not due to artefacts caused by upper troposphere and stratospheric methane variabilities (which cause strong signals in the applied TROPOMI XCH4 data). For further advancing with the Barrè et al. (2021) method, the identification of outliers in the TROPOMI XCH4 data would be helpful. Our combination of IASI and TROPOMI can actually be used to identify such outliers. We define TROPOMI outliers as the observations that have a strong difference to the collocated TCCON data. We found that this happens almost exclusively when the shape of the combined CH4 profile is statistically rather improbable. An unrealistic profile shape means that the respective TROPOMI and IASI data are inconsistent, i.e. that there is a significant difference between the IASI and TROPOMI data that cannot be explained by their different vertical sensitivities. So we can use our combination method for identifying and subsequently filtering out the TROPOMI outliers. Actually, we use this new filter method already for filtering out poor TROPOMI data in this study. This partly explains the improved agreement with the references if compared to the figures shown in the manuscript (compare the following Figs. 7 and 8 of this referee response with Figs. 7 and 11 from the manuscript).

Referee comment #3

The paper asserts that a linear optimal estimation combination of L2 products is equivalent to a non-linear combination of L1B products. They never show this. Rather they depend on the mild non-linearity assumption in Rodgers, 2000. However, they never show that this assumption is valid for their problem. They could demonstrate it by showing in a simulation environment where they combine the L1B data in a multi-spectral retrieval and compare it to the equivalent L2 combination.

We will replace equivalence by similarity. Full equivalence (under the condition of a moderately non-linear problem) is only given by applying Eq. A11 of the revised theoretical part (Appendix A of https://doi.org/10.5194/amt-2021-31-AC1), which requires the full TROPOMI a posteriori covariances. However, in the TROPOMI L2 data we only have total column averaging kernels and total column errors available and we cannot reconstruct the full a posteriori covariances and perform a combination according to the aforementioned Eq. (A11).

One method to show the equivalence/similarity of our L2 product combination approach with a multi-spectral retrieval is the study proposed by the referee. However, the strong similarity can also be documented without the rather comprehensive development of a multi-spectral retrieval.

We show this strong similarity in the theoretical part of the paper (see revised Appendix A2.2 of https://doi.org/10.5194/amt-2021-31-AC1). In addition, we show the validity of the linearity assumption empirically by our detailed comparisons to reference data. TROPOMI data consist only of total column methane data. They very well agree with the TCCON total column reference. The IASI profile data have a very good sensitivity in the upper troposphere/lower stratosphere (UTLS), but cannot well capture the TCCON total column variabilities. We document a very good agreement of the IASI UTLS data with the AirCore references. Then we combine the IASI and TROPOMI data (using the linearity assumptions) and the new data (the combined data) show both, a good agreement with the TCCON total columns as well as a good agreement with the UTLS AirCore data. This is actually the proof that non-linearities have no significant impact on the method. Furthermore, in the Appendix B of Schneider et al. (2021) we demonstrate that the assumption of moderately non-linearity is well justified for the MUSICA IASI CH4 data retrieval.

Referee comment #4

The authors are impressively unaware of the literature on combining satellite data for composition. They don't discuss the landmark Landgraf and Hasekamp (2007) or Worden et al, 2007 papers. In addition to Cuesta, there are a number of papers by Fu et al, 2013, 2018 that solve this problem with L1B data. Luo et al, 2013 demonstrate a similar approach for combining TES and MLS L2 data for CO. The authors are encouraged to familiarize themselves with the literature and cite appropriately.

In the manuscript we cite five different papers from the field of satellite trace gas data combination: Ceccherini et al (2009), Constantino et al. (2017), Cortesi et al. (2016), Cuesta et al. (2013), and Worden et al. (2015). Apparently, we missed other important works from this field, for which we would like to apologize. We will dedicate some extra time to improve this reference list. However, our focus will be on literature about L2 data combination. An important work we missed but we will prominently discuss in the revised version is Warner et al. (2014). They use, similar to us, a Kalman filter based approach for the data fusion.

Referee comment #5

As noted by the authors, the value of this approach will be better realized with Sentinel 5, rather than IASI and TROPOMI. I would suggest orienting the paper more towards a proof-of-concept for S5 or similar configurations, e.g, the A-Train. Whether this approach is as good as a combined L1B retrieval with coincident measurements or separately assimilating L2 products remains to be seen. But, the current strategy with IASI and TROPOMI is not clear.

The referee manifests two main concerns about the quality of the combined product. The first concern is that the dislocation error has a dominating impact on the quality of the combined data. The second concern is that non-linearities make a combination on the basis of L2 data not feasible. With our detailed investigations about the dislocation error we address the first concern (see our reply to referee comment #1). We found that the dislocation error is actually of secondary importance. Instead the measurement noise error is dominating in the combined product. Concerning the second concern, we argue that the non-linearities cannot be of major importance, otherwise there would be a strong degradation of the total column and UTLS data quality in the combined data product. However, this is not observed, i.e. the combination works correctly with the linearity assumptions (see our reply to referee comment #3). In addition, in the Appendix of Schneider et al. (2021) we demonstrate the good degree of linearity of the MUSICA IASI CH4 data retrieval.

Furthermore, the referee doubts (at least to our understanding) on the usefulness of generating such combined profile data outside of the assimilation frameworks of big research centers. We also think that a comprehensive assimilation framework can well exploit the synergetic information of different satellite sensors. However, such frameworks are very complex and not very flexible/fast in implementing new measurements. Actually, there is currently no CAMS product, that has the state-of-the-art TROPOMI and IASI CH4 data simultaneously assimilated. Furthermore, the models used for the assimilation have generally a significantly coarser horizontal resolution than the state-of-the-art observational data. In this context, data fusion methods (like the one proposed in our paper) make sense, because they provide something that is not readily available from large data assimilation frameworks.
In this paper we demonstrate the method and document their quality. It is the task of future research to apply the combined data to specific research questions. A first example of an application is given in Tu et al. (2021). Another interesting application is the usage of the method for identifying inconsistencies between IASI and TROPOMI that cannot be explained by the different vertical sensitivities of the two sensors. In some preliminary studies we found that this is very promising for filtering out TROPOMI data of poor quality, which in turn can improve investigations on local uncertainties in the assumed methane sources by the method discussed in Barré et al. (2021).

Further minor referee comments

Line 41: There are quite a few more examples in the literature than those mentioned. In particular they neglect Landgraf and Hasekamp (2007), Worden et al, (2007), Fu et al, (2013), (2018), etc. Authors need to do more diligence with their citatitions

We would like to apologize for that and for the revised paper we will work on an improved reference list. Please see also our reply to referee comment #4.

Line 46: This approach was discussed in detail in Lou et al, 2013 with TES and MLS data.

… yes and also in other studies. Please see also our reply to referee comment #4.

Line 69: The tropopause will be an issue for profile retrievals as well. This argument needs to be more quantitative about how uncertainties in tropopause height will affect the xCH4 calculation.

We are not sure if we correctly understand the comment. We are not talking about problems of the retrievals or for calculating XCH4, instead we are talking about what drives XCH4 variability. It is the variability due to surface emissions/fluxes (mainly observable close to ground), the variability due to tropospheric background signals (observable in the free troposphere), and the variability caused by shifts in the tropopause altitude (observable in the UTLS).

Line 100: I don’t think it’s appropriate to cite a paper in prep. It at least needs to be in review.

In the meanwhile the paper is published.

Line 101: Order

Thanks!

Line 144: These two instruments are in fundamentally different orbits and therefore different local solar times. The difference in LST is already about 4 hours. To the extent to which the variability in either is driven by dynamics, then this difference could be substantial. In particular, the PBL heights could be quite different. The criteria described here appear to be arbitrary. In additional analysis motivating these choices needs to be included (these figure could be added in an appendix).

Please see our reply the referee comment #1.

The assumption that the vertical distribution between morning and afternoon is relatively unchanged is an important assumption. This could be further tested by looking at CAMS CH4 output and showing in an OSSE framework that the sampling assumption here hold.
Please see our reply the referee comment #1.

Caption of Fig. 4: In Fig. 3, it looks like the a priori contribution for TROPOMI is negative. Can you explain why?
The column averaging kernel can have values larger than 1.0. This can cause negative values of $(\mathbf{I}-\mathbf{A})x_a$. In any case we will change this figure and the discussion. Instead of documenting the a priori contribution, we will show an estimation of the uncertainty due to imperfect sensitivity (called "the smoothing error" by Rodgers 2000), because it is the sum of this smoothing error and the noise error that is optimized/minimized by an optimal estimation retrieval. Please also see our reply to the first comment of referee #3.

Line 250: These criteria appear to be driven primarily by pragmatic considerations rather minimization of error from two different locations. The authors need to demonstrate what the theoretical considerations for this to work.
For these criteria we made two considerations. The uncertainty due to collocation mismatch should be ideally smaller than the uncertainty of the noise error of the products that are compared. In addition, the criteria is compromised by requiring a sufficiently high number of remaining collocated data pairs.
For the revised paper we extended the evaluated time series by an additional year (the year 2020). This allowed us to strengthen the spatial and temporal criteria for collocation with TCCON. We require a collocation within 100 km and within 3 hours. According to Fig. 3 of referee comment #1 a spatial collocation mismatch of 100 km and a temporal collocation mismatch of 3h cause a similar uncertainty for the total column comparison (in both cases this uncertainty is 0.25%). We think that this is reasonable, because this uncertainty is in the range of the noise error of the TROPOMI and the combined total column product.
Close to the surface a 250 m altitude difference is about a 20 hPa pressure difference, i.e. a dry air column difference of about 2% (20 hPa / 1000 hPa). Figure 2 of the referee comment #1 shows that close the ground there is a small layer where CH4 varies largely independently from the atmosphere above. Judging from Figure 1 of the referee comment #1 the typical variability in this small layer is in the range of 3%. This means that a mismatch of the ground altitude of 250 m or 20 hPa causes an uncertainty in the XCH4 comparison of less than 0.1% (2% * 3% = 0.06%), which is within the typical noise error of the TCCON data product.

Line 263: Under what basis is this assumption made? Why should this be considered to reasonably capture the error?
Model uncertainties are typically estimated by ensemble runs, i.e. the model is run several times, but with different model setups or initializations and then the root mean square (RMS) value between the different model results could be used. These data are not available for the TM5 model (at least it is not available to us). So we created something like ensembles by looking on the differences when assuming that the model is out of phase.
Alternatively, we could use the RMS of the difference between the TM5 model and the high resolution CAMS forecast. Figure 6 shows the results of these calculations for the surroundings of Karlsruhe. The so-estimated uncertainty of the TM5 a priori model is significantly larger than the uncertainty obtained from our out of phase considerations. We find that the large RMS between the TM5 a priori data and the high resolution forecast of CAMS comes to a large extend from the years 2019 and 2020. After 2018 the TM5 model shows an increase of about 1% per year, but the high resolution forecast shows no significant increase.

We will update this error estimation for the revised manuscript using the RMS in the difference between the TM5 model data and the CAMS high resolution forecast as the uncertainty in the a priori data. Figure 7 shows the error we estimate for the comparison of TCCON and satellite total column products according to the a priori model uncertainty. Because the TROPOMI and the TCCON kernels have both a similar good column sensitivity throughout the troposphere, this uncertainty for the total column comparison is generally within 1‰ (see black squares in Fig. 7). The same is true for the validation of the total column of the combined product (see blue crosses in Fig. 7). For the validation of the total column of the MUSICA IASI product this error is larger, because the total column sensitivity of IASI is significantly different from the respective sensitivity of the TCCON product and the other satellite products (see Fig. 3b of the manuscript). For the comparison of the IASI and TCCON total column data we estimate that the error due to the different sensitivities (of IASI and TCCON) can be above 1% (see red dots in Fig. 7). This error is largest to the end of the time series, because then the TM5 a priori model error is largest (increasing difference between the TM5 model and the CAMS high resolution forecast after 2018).

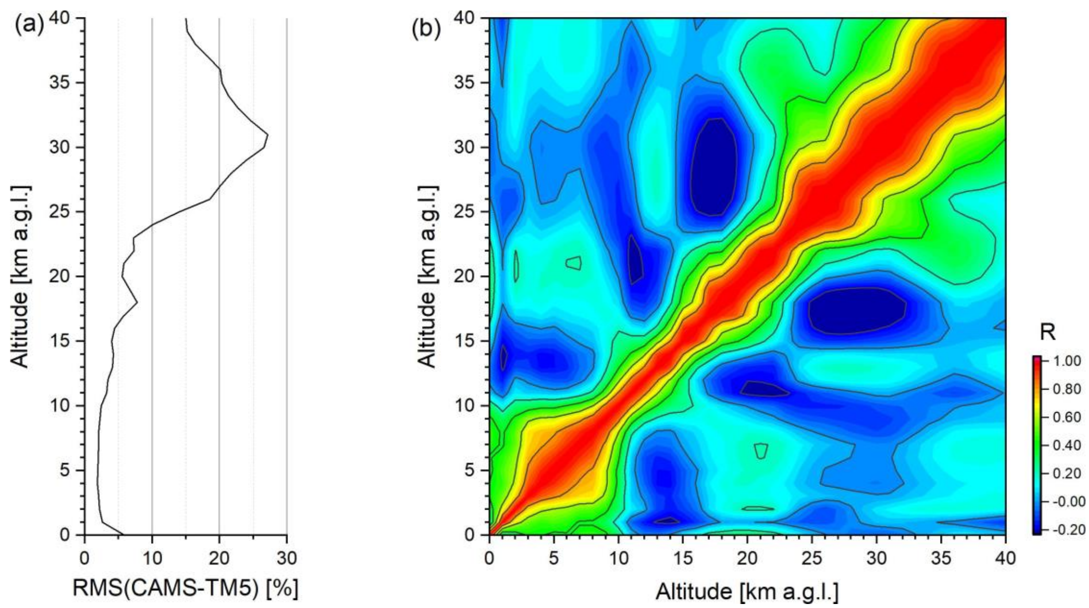

Figure 6: Comparison of the TROPOMI a priori model (TM5) and collocated CAMS high resolution forecasts. (a) RMS of the relative differences; (b) Matrix showing the correlations of CAMS-TM5 differences at different altitudes.

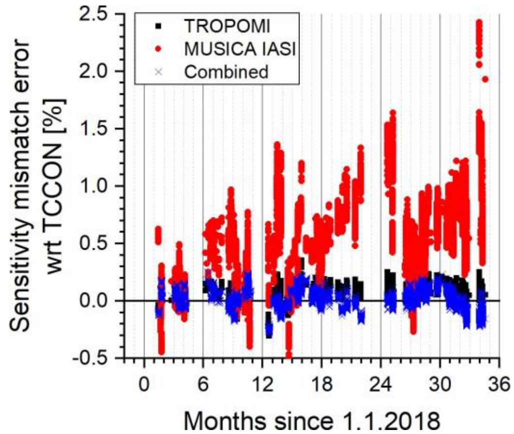

Figure 7: Error in the comparison of TCCON and satellite products due to the a priori model error and the different column sensitivities of the TCCON product and the satellite products.

Line 284: Please include on the other axis the absolute differences ppb.

Please note that there is no uniform relation between relative difference and absolute difference, because the absolute values (total columns and partial columns) vary. So plotting the absolute differences would mean an addition row of plots. Honestly, we do not think that this adds new information and prefer leaving the figure as is: relative values when showing the differences and absolute values when showing the correlations.

Line 285: IASI has a positive bias relative to TCCON but TROPOMI has a negative bias. The combined product is substantially closer to IASI than TROPOMI. Please explain.

We updated the comparison study by including also the year 2020. Furthermore, we refined the filtering of the IASI data (see also our reply to referee comment #3) and take into account differences in the ground pressure of IASI and S5P (we individually account for IASI and S5P ground pressures according to Appendix B of Sha et al., 2021).

Figure 8 shows the updated comparison. Figures 8a-c depict the differences with respect to time. We observe that the bias (and scatter) in the TROPOMI and the combined products are similar.

Line 287: Add a figure comparing the agreement between the a priori and TCCON. Or at least, calculate it and provide the summary statistics for comparison with the retrievals.

We agree with the referee that a correlation of the differences with respect to the a priori is also very interesting, because they show the complementary information of the satellite data (complementary to the a priori knowledge). We will provide this in the revised manuscript (see (g)-(i) of the Fig. 8).

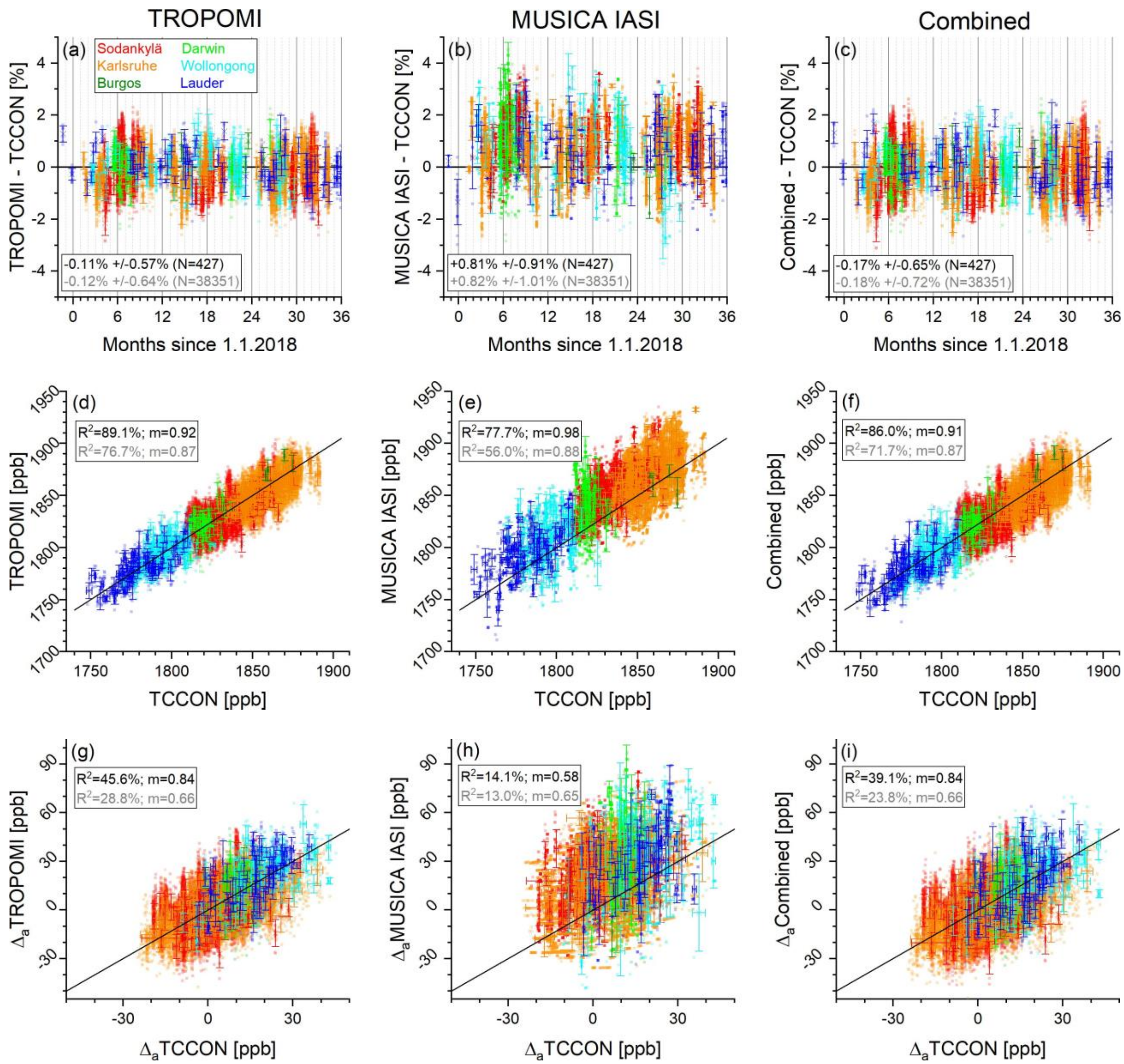

Figure 8: Revised version of Fig. 7 of the manuscript. We will add the additional panels (g)-(i), that show correlations after removing the a priori information (Δ_a means difference with respect to the a priori).

Caption of Fig. 7: Please add the offset in addition to the slope.
Ok.

Line 294: agreement
Thanks!

Line 294: How do you justify that R^2=0.5 is good agreement? The combined product is not appreciably better than the individual products. Is the reason for the low correlation the time-space difference error? Explain.

Please note that the R^2 value reported in the paper for the daily mean correlations is about 0.6 not 0.5.
For the revised paper we will use an improved data filtering and thus get better agreement (scatter within 0.6%). Furthermore, we will consider an additional year of data (the year 2020), which leads to larger variability. Both measures lead to higher R^2 values (for the data amount and filtering method applied for the revised paper the R^2 value is about 0.75, see also Fig. 8d-f). Much more cannot be expected given the uncertainty of the data and the limited natural variability of XCH4. By using data for a longer time period and by the new filtering criteria we achieve a level of agreement with TCCON that is better than the agreement documented by previous studies (e.g. Table 4 of Sha et al., 2021).

Line 343: An important difference is that in this case the bias is not positive relative to AirCore but negative relative to TCCON. Please explain.
One reason is that TCCON data are from six globally distributed sites and that for the TCCON comparison the collocation criteria is rather strict. This is different for the AirCore comparison. AirCore data are only from two sites. Only one of these sites is also used in the TCCON study. In addition, the AirCore collocation criteria is more relaxed, which is a compromise that ensures a reasonable number of comparable data pairs.
Furthermore, there is a relatively high amount of AirCore data measured during spring in Sodankylä, when the ground is probably covered by snow. For these conditions NIR albedo is high and the TROPOMI data have a significant positive bias (Lorente et al., 2021). In the revised paper we developed a more sophisticated data filtering of TROPOMI data. Among others we will filter out the TROPOMI data that are affected by high NIR albedo (which indicates to snow on the surface).

Line 346: Not really clear. Please elaborate as to why the AirCore is substantially worse. One particular point is that TCCON is a remote sensing measurement whereas AirCore is not. The impact of vertical differences will be more pronounced.
The scatter between TROPOMI and AirCore is very similar as the scatter between TROPOMI and TCCON. The R^2 values for the TCCON comparison is larger because in the TCCON data there is much more variability than in the AirCore data. This is because TCCON data are from six different globally distributed sites and AirCore data only from two sites that are not globally representative (middle and high northern latitudes).

Line 374: That would satisfy a “do no harm” case, but how does this show that the combined product is better than the individual product(s)?
This shows that the combination method works robustly and that the linearity assumptions are valid (see also referee comment #3).

Line 400: You can and should assess that by correlating GAW with the IASI prior.
Yes, we think adding additional plots showing the correlations for data after removing the a priori information is helpful. For the revised paper we will add panels that show the correlation between (in situ – a priori) and (satellite – a priori). Figure 9 shows the figure that will replace Fig. 11 of the manuscript. There the panels (g)-(i) show the correlations after removing the a priori information. Then the synergetic effect achieved by the combination method becomes even clearer.

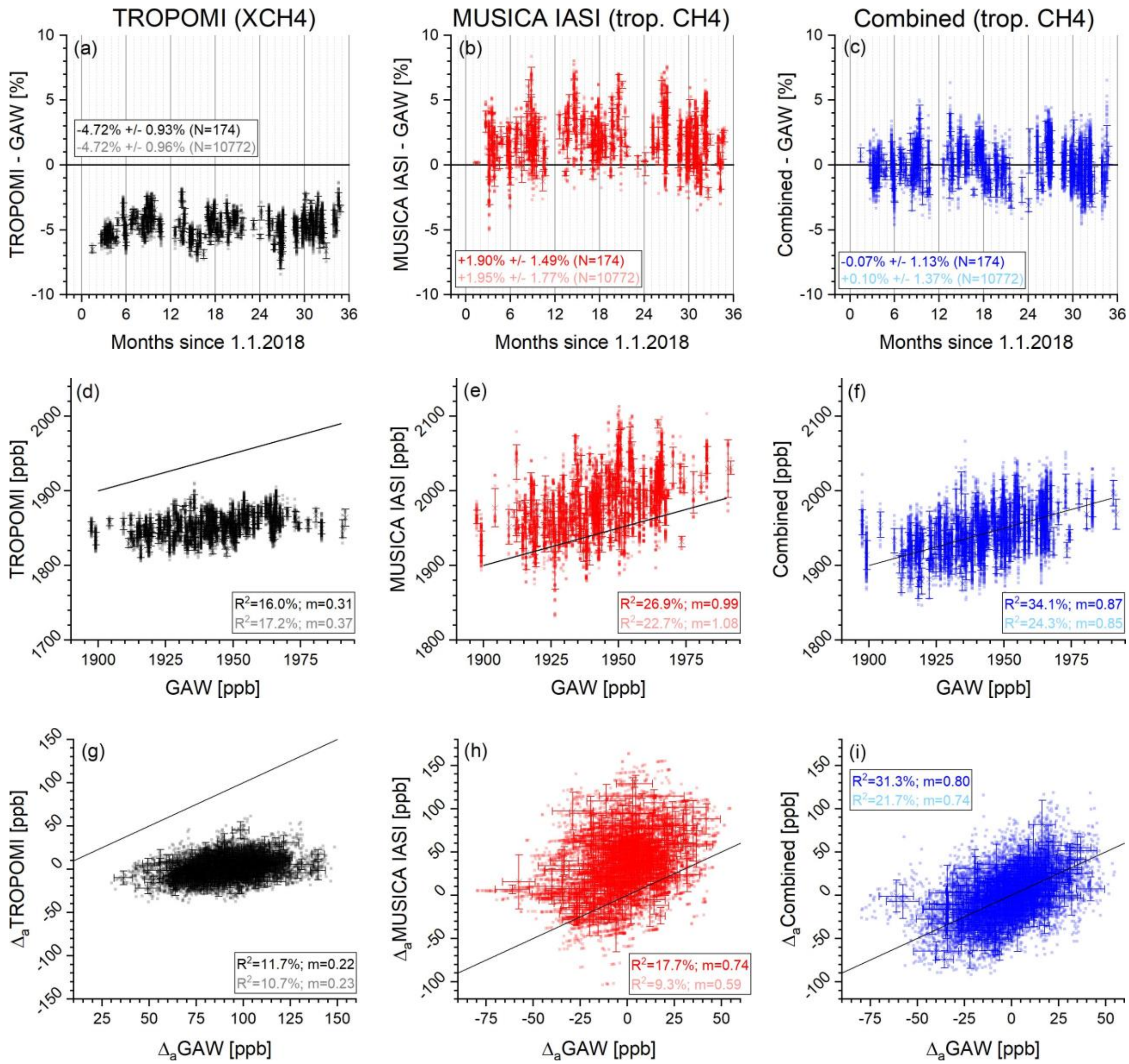

Figure 9: Revised version of Fig. 11 of the manuscript. We will add the additional panels (g)-(i), that show correlations after removing the a priori information (Δ_a means difference with respect to the a priori).

Line 410: That's not been demonstrated. In particular, the assumption is that the problem is linear, which the authors have not shown. Rather, we know that the problem is non-linear, which is not resolved by the derivation in the Appendix.

We think that with the revised theoretical part of the paper (see https://doi.org/10.5194/amt-2021-31-AC1), the empirical studies (see referee comment #3), and the Appendix B of Schneider et al. (2021) we show that non-linearity is of minor importance and that the method can be classified as being very similar to a dedicated combined optimal estimation retrieval.

Line 417: It's not real clear what is being gained here. The combined product suffers from errors due to dislocation that are not quantified. Why do I need a combined product if I already have each product individually, which have the sensitivities to the UTLS and total column?

The fact that in the combined product both the total column and the UTLS data are of good quality proves the robustness of the method (e.g. non-linearities are of minor importance, see referee comment #3). Furthermore, it provides an indirect validation of the tropospheric column product: if total column and UTLS column of one data product is good, then the tropospheric column must also be of good quality. Please note that this is something different than having good total column data from one product and good UTLS data from another product. We cannot simple rest the UTLS product from the total column product, without considering the differences and the correlations in the two sensitivities. The combination method ensures that the characteristics of the sensitivities are fully considered.

Line 418: It would be good to capitulate those improvements here as they are still not clear.
See reply above.

Line 468: For this derivation, the definition of moderate non-linearity is quite important as well as its limitations. The assertions made earlier about equivalence with a radiance-based retrieval hinge on it. Please elaborate and discuss relationship with radiance-based retrievals.
In Appendix B of Schneider et al (2021) we prove that the assumption of moderate non-linearity is justified for the MUSICA IASI CH4 retrieval. Furthermore, as stated in the referee comment #3, we think that our comparison studies provide an empirical prove of the validity of the linearity assumption.

Line 500: While this derivation is well-known, the important assumption is that the two measurement vectors are measuring the same atmosphere. However, for IASI and TROPOMI, that is not the case at all. So, y1=F(x1)+n1 and y2=F(x2)+n2. A more interesting and theoretically necessary question for this paper is the error introduced by x1 not-equal x2. At what point is that error negligible? What are the atmospheric conditions necessary for that to be the case? That could be readily studied by sampling a model such as CAMS and building linear retrievals of both for IASI and TROPOMI in their respective orbit.
We think that our detailed study shown in referee comment #1 fully address this concern of the referee.

Line 618: column
Thanks!

<u>Reference:</u>

Barré, J., Aben, I., Agustí-Panareda, A., Balsamo, G., Bousserez, N., Dueben, P., Engelen, R., Inness, A., Lorente, A., McNorton, J., Peuch, V.-H., Radnoti, G., and Ribas, R.: Systematic detection of local CH_4 anomalies by combining satellite measurements with high-resolution forecasts, Atmos. Chem. Phys., 21, 5117–5136, https://doi.org/10.5194/acp-21-5117-2021, 2021.

Diekmann, C. J., Dubravica, D., Ertl, B., Estruch, C., León-Luis, S. F., Marrero, C., Morgui, J.-A., Ramos, R., Scharun, C., Schneider, C., Sepúlveda, E., Toledano, C., and Torres, C.: Quantification of CH_4 emissions from waste disposal sites near the city of Madrid using ground- and space-based observations of COCCON, TROPOMI and IASI, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2021-437, in review, 2021.

Massart et al., 2014: Massart, S., Agusti-Panareda, A., Aben, I., Butz, A., Chevallier, F., Crevoisier, C., Engelen, R., Frankenberg, C., and Hasekamp, O.: Assimilation of atmospheric methane products into the MACC-II system: from SCIAMACHY to TANSO and IASI, Atmos. Chem. Phys., 14, 6139–6158, https://doi.org/10.5194/acp-14-6139-2014, 2014.

Seegers and Houweling, 2020: https://atmosphere.copernicus.eu/sites/default/files/2020-01/CAMS73_2018SC1_D73.5.2.2-2019_202001_production_chain_v1.pdf

Rodgers, C.: Inverse Methods for Atmospheric Sounding: Theory and Praxis, World Scientific Publishing Co., Singapore, 2000

Schneider, M., Ertl, B., Diekmann, C. J., Khosrawi, F., Weber, A., Hase, F., Höpfner, M., García, O. E., Sepúlveda, E., and Kinnison, D.: Design and description of the MUSICA IASI full retrieval product, Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-2021-75, in review, 2021.

Sha, M. K., Langerock, B., Blavier, J.-F. L., Blumenstock, T., Borsdorff, T., Buschmann, M., Dehn, A., De Mazière, M., Deutscher, N. M., Feist, D. G., García, O. E., Griffith, D. W. T., Grutter, M., Hannigan, J. W., Hase, F., Heikkinen, P., Hermans, C., Iraci, L. T., Jeseck, P., Jones, N., Kivi, R., Kumps, N., Landgraf, J., Lorente, A., Mahieu, E., Makarova, M. V., Mellqvist, J., Metzger, J.-M., Morino, I., Nagahama, T., Notholt, J., Ohyama, H., Ortega, I., Palm, M., Petri, C., Pollard, D. F., Rettinger, M., Robinson, J., Roche, S., Roehl, C. M., Röhling, A. N., Rousogenous, C., Schneider, M., Shiomi, K., Smale, D., Stremme, W., Strong, K., Sussmann, R., Té, Y., Uchino, O., Velazco, V. A., Vigouroux, C., Vrekoussis, M., Wang, P., Warneke, T., Wizenberg, T., Wunch, D., Yamanouchi, S., Yang, Y., and Zhou, M.: Validation of methane and carbon monoxide from Sentinel-5 Precursor using TCCON and NDACC-IRWG stations, Atmos. Meas. Tech., 14, 6249–6304, https://doi.org/10.5194/amt-14-6249-2021, 2021.

Tu, Q., Hase, F., Schneider, M., García, O., Blumenstock, T., Borsdorff, T., Frey, M., Khosrawi, F., Lorente, A., Alberti, C., Bustos, J. J., Butz, A., Carreño, V., Cuevas, E., Curcoll, R., Diekmann, C. J., Dubravica, D., Ertl, B., Estruch, C., León-Luis, S. F., Marrero, C., Morgui, J.-A., Ramos, R., Scharun, C., Schneider, C., Sepúlveda, E., Toledano, C., and Torres, C.: Quantification of CH_4 emissions from waste disposal sites near the city of Madrid using ground- and space-based observations of COCCON, TROPOMI and IASI, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2021-437, in review, 2021.

Warner, J. X., Yang, R., Wei, Z., Carminati, F., Tangborn, A., Sun, Z., Lahoz, W., Attié, J.-L., El Amraoui, L., and Duncan, B.: Global carbon monoxide products from combined AIRS, TES and MLS measurements on A-train satellites, Atmos. Chem. Phys., 14, 103–114, https://doi.org/10.5194/acp-14-103-2014, 2014.