Final Response, Andreas Ostler, Karlsruhe Institute of Technology, Garmisch-Partenkirchen, Germany, 6 October 2014

It is a pleasure to thank both the referee Christian Frankenberg and the Anonymous Referee for very sound and helpful comments which lead to significant improvements and interesting extensions of the paper. We thereafter present our point to point reply.

Review #1: General Comments

"The manuscript by Ostler et al describes comparisons between NIR and MIR uplooking methane retrievals and attributes systematic differences between them to dynamic processes such as strat-trop exchanges as well as subsidence. It is in general well written and suitable for AMT. It can be published once some major comments (below) are incorporated as a proper discussion on what has been done to mitigate the impact of stratospheric methane has not been discussed thoroughly:"

We added a section on **"Mitigation strategies for the stratospheric impact"**, which is now Section 5. As a consequence, the original Section 5 was renamed as Section 6. The new Section 5 discusses all points addressed by the general comments below and is mentioned in the introduction.

The 1st paragraph of the new Sect. 5 reads: "The stratospheric part of CH_4 is defined by the position of the tropopause and the CH_4 vmr gradient in the UTLS. Whereas the CH_4 vmr gradient in the UTLS is the result of large-scale vertical transport (Brewer Dobson circulation), the position of the tropopause depends on synoptic (e.g., polar vortex, STE) as well as seasonal variations (except for the tropics). Due to these facts the stratospheric part of CH_4 can produce a lot of variability and uncertainty (smoothing effects) in the total column CH_4, and, consequently, in XCH_4. This variability can be misleading within the analysis of trends with regard to tropospheric emissions. Therefore, methods have been developed to overcome this problem of stratospheric variability by separating the tropospheric part of CH_4 from the total column."

"The problems with stratospheric variability have been dealt with in previous publications and those should be referenced. This paper (Washenfelder et al., 2003) was the first to retrieve tropospheric CH4 amounts using NIR spectra and the HF method. The paper and the method should be cited and discussed, as it is highly relevant in this study.

It is striking that the HF method, which might help greatly here to detect dynamic events, is not even discussed.

The Sepúlveda paper (Sepúlveda et al., 2012) is cited but not really discussed in terms of how it attempts to isolate the troposphere, which would reduce smoothing errors."

We address and discuss both methods as follows (2nd paragraph of new Sect. 5):

"First of all, Washenfelder et al. (2003) showed that stratospheric tracer-tracer relationships can be used to approximate the stratospheric CH_4 in order to subtract it from the total column CH_4 to get a tropospheric CH_4 with a theoretical precision of ~0.5%. Their study was based on simultaneous NIR retrievals of CH_4 and hydrogen fluoride (HF) that is strongly anti-correlated with stratospheric CH_4. This method has been refined by Saad et al. (2014) by explicitly accounting for averaging kernels. As a result the mean precision of tropospheric CH_4 was improved to ~0.1%. At the same time, Wang et al. (2014) showed that nitrous oxide (N_2O) can also be an appropriate proxy for stratospheric CH_4 with less H_2O dependence compared to HF. In contrast to NIR, the proxy retrieval is not applicable for MIR measurements, as shown by Sepulveda et al (2012). However, the latter extracted tropospheric CH_4 directly from optimized profile retrievals in the MIR.

Nevertheless, all methods mentioned have their limitations, e.g. Sepulveda et al. (2014) found that their tropospheric CH_4 product can be affected significantly by variations in the UTLS. They emphasize that tropospheric CH_4 with a precision of ~0.5% can be derived only by means of an a posteriori correction of the MIR retrievals. The results of the proxy methods in the NIR achieve high theoretical precision indeed, but their accuracy directly depends on the accuracy of the total column CH_4 which is linked to the quality of a priori profiles. E.g., especially in polar vortex conditions NIR proxy retrievals are limited due to smoothing effects (Saad et al., 2014). In the end the smoothing effect described in Sect 4.1.1 directly affects the accuracy of the tropospheric CH_4 which points to the importance of realistic a priori profiles for scaling retrievals. One further problem is introduced by the complex structure of STE events because they do not only affect the stratosphere, but also the troposphere, as shown in Sect. 2.2.1.

Due to these reasons and the fact, that our MIR retrieval strategy is optimized with regard to total columns, we decided to focus on total XCH_4. However, a comparison of tropospheric NIR and MIR retrievals is certainly interesting. It is obvious that such an intercomparison should be performed with retrieval methods dedicated to tropospheric columns, and improved a priori profiles that are able to reproduce the polar vortex subsidence in a realistic way. The NIR proxy retrievals (Saad et al., 2014, Wang et al., 2014) will benefit from these a priori profiles in the same way as the total column retrievals. In addition to that, these profiles could be used by tropospheric MIR profile retrievals (Sepulveda et al., 2014), thereby acting as a common prior within an intercomparison study. ”

“For these reasons (dependence of XCH4 on tropopause height), a proper validation of satellite data (and models) should, in general, take the vertical sensitivities into account, esp. for methane which can be highly depleted in the stratosphere. This should be discussed, esp. as also the satellites have their own AK and in some cases it is advantageous to validate them against a ground-based network with similar sensitivities (for this, NIR TCCON data may be better than NDACC if SWIR sounders are concerned).”

We address this aspect with the following words (new Sect. 5 paragraph 5):

“Attention should also be paid on the dependence of XCH_4 on stratospheric variability with regard to validation of satellite and model data. This means, that the prior and the averaging kernels should be taken into account, particularly as satellites have their own vertical sensitivities. In the case of similar vertical sensitivities we can assume that the smoothing effects from satellite and ground-based retrievals are of nearly equal magnitude. Hence, remaining differences can be attributed to other error sources like systematic errors. Thus, NIR TCCON retrievals may be more valuable than MIR NDACC retrievals if SWIR (short wave infra-red) sounders like GOSAT or SCIAMACHY are concerned. Furthermore, it is important to be very careful with the interpretation of results when evaluating model data from arctic regions.”

“In general though, models can be applied to bound the potential smoothing errors.”

We showed this effect by using the ACTM profiles.

"The analysis of STE and subsidence events is interesting and worthwhile. It would be good though if the authors can make recommendations as how to identify them in a more operational (instead of case study) sense. What kind of stratospheric tracers (e.g. HF) could help?"

We comment on this as follows (new Sect. 5 paragraph 6):

"Instead of separating the stratospheric part from the total column, we tried to detect and exclude situations with high dynamical variability by analyzing meteorological parameters (Sect. 4.1.2 and Sect. 4.2.2). In future, these meteorological criteria for exclusion could be installed as additional information in the prior data (TCCON). Another possibility for the exclusion of affected data has been presented by Angelbratt et al. (2011). They used a multiple regression model with anomalies of HF, carbon monoxide (CO), and tropopause height to reduce the variability in total column CH_4. This concept of anomalies can also be transferred to the retrieval process. Certain threshold values of stratospheric tracers (HF, N_2O) can be implemented as filter criteria for XCH_4. In Appendix D we show that HF is suitable for detecting polar vortex subsidence. In contrast, the situation with STE is much more complex and HF cannot be recommended as an additional filter criterion."

Our recommendations with regard to HF are based on an analysis of NIR XHF retrievals. The results are shown in **APPENDIX D: Mitigation strategy for stratospheric impact using XHF**

"Also: It is seen as a problem that the AK between NIR and MIR are different. A real step forward would be to combine the two regions in a concurrent retrieval setup, which should greatly enhance the degrees of freedom for profile retrievals and might alleviate many of the problems discussed here and could also help atmospheric model if more atmospheric layers can actually be differentiated.

I would encourage the team to look at these aspects in the future but realize that it is probably beyond the scope of the current study.

The potential could be discussed though, esp. because the impact of stratospheric events on differences between NIR and MIR column retrievals is not very surprising and this discussion would add some more novelty."

We comment on this as follows (new Sect. 5 paragraph 7):

"Substantial progress could be achieved when combining the NIR and MIR measurements in a concurrent retrieval setup. In this approach the difference in averaging kernels is not considered as a problem but as an opportunity to differentiate more atmospheric layers. The synergetic potential of such an approach seems to be more promising in the case of polar vortex subsidence because the dynamics of subsidence are not as fine-structured compared to STE events. However, the main objective of combining both retrievals will be to determine the shape of the stratospheric CH_4 profile. Hence, reducing the uncertainty of the stratospheric CH_4 will be beneficial for models and retrievals. We want to construct such a combined method and also want to present the method in a subsequent study."

Review #1: Specific Comments

Page 6746, Lines 9-11: Models have to be improved to detect and quantify local emissions? I would guess that the more diffuse regional through continental scale emissions inversions are more susceptible to transport bias.

We alleviated our statement and changed the text: “The ability to locate CH_4 emissions (anthropogenic and natural) on regional scales will be essential for future climate policy with regard to emission trading schemes. For this purpose, it is necessary to reduce the transport uncertainties of inversions. Furthermore, it is mandatory to increase the network CH_4 observations and to improve the accuracy of CH_4 measurements.”

Page 6750, Line 15: “because of a special smoothing effect as explained in the following” somehow sounds as if something very unusual would be explained. The effect of AKs on retrievals is well known though, so I would just state: “can still arise because of different vertical sensitivities for both retrievals. The smooth. . .”

We changed the wording according to: “Although the use of Eq. (1) in Sussmann et al. (2013) eliminates the impact of differing a priori profiles differences (MIR-NIR) can still arise because of different vertical sensitivities for both retrievals.”

Page 6752, line 16: I think the impact on inversions is somewhat overstated. If the CH4 data from NDACC and TCCON are to be assimilated into the atmospheric inversion models, these would use their respective averaging kernels. While the stratospheric variability might not be perfectly reproduced by the CH4 model, it would at least reduce the bias as the AK corrections would be in general applied. Models could even act as a transfer standard between MIR and NIR column retrievals. You should mention that models can actually take the smoothing error due to the AK directly into account.

We removed the overstatement related to the impact on inversions in Sect. 3.2. In order to consider the possibility of models as a transfer standard, we changed the text in the original Sect. 5 (Summary and conclusions, Page 6761, line 24 ff.):

“We conclude that atmospheric variability (subsidence, STE, and stratospheric variability in general) is a key factor in constraining the accuracy of MIR and NIR seasonal cycles. Different vertical sensitivities for both retrievals give rise to different smoothing effects. We showed that this impact can be mitigated by means of two basically different approaches. Either situations with high atmospheric variability are detected and excluded from further analysis or one has to focus on retrievals of tropospheric XCH_4 (Sect. 5). Nevertheless, NIR and MIR XCH_4 retrievals can be used in inversions without limitation of data. E.g., inversions can explicitly account for averaging kernels and a priori profiles, thereby reducing the bias between MIR and NIR. At the same time, both measurements will be consistent with the model. In addition to that, inverse models are able to take the smoothing error directly into account.

One step forward would be to use models as a transfer standard between MIR and NIR column retrievals. Assuming that only the true CH_4 profile is able to harmonize NDACC and TCCON retrievals of XCH_4, it is possible to construct a concurrent retrieval setup. The synergetic potential of such a combined method is based on the different vertical sensitivities of the retrievals. A study on this subject has already been initiated and is subject of a subsequent publication.

However, NIR retrievals at polar sites may be improved by accounting for stratospheric subsidence in the standard retrieval a priori. Improving the quality of MIR retrievals affected by STE seems to be more complicated due to the diversity of STE processes. We conjecture that more realistic a priori profiles from high-resolution models reflecting small-scale processes could help to reduce MIR and NIR smoothing effects. An alternative method to overcome this problem would be to further improve the FTIR retrievals with the target to achieve a more uniform sensitivity at all altitudes. i.e., if the MIR averaging kernel was more evenly weighted with altitude then the MIR dependence on STE should be reduced. Also, using a formal optimal estimation (OE) inverse technique in GFIT could foreseeably help to improve the sensitivity of NIR retrievals to subsidence.

When using NDACC and/or TCCON XCH_4 data, it is critical to be aware of the effects of dynamical events on the accuracy of the relevant data set. Depending on the requirements on data accuracy, NDACC and/or TCCON XCH_4 data can be used with or without the exclusion of dynamical events. Methods to detect these events have been presented in this study. Given a proper data use based on the findings in this paper, a joint NDACC and TCCON data set will result in wider spatial and longer temporal coverage of XCH_4 data for the validation of top-down estimates, satellite validation, and trend studies."

Page 6752, Line 19ff: Here, you should mention previous paper describing how to reduce the impact of the highly variable stratosphere. Can HF help in this case? It is well known that column CH4 is highly sensitive to tropopauseheight.

At this point we will refer to the new Sect. 5 dealing with the problem of stratospheric variability. In Sect. 5 we mention the previous paper related to this topic and show the potential of HF for mitigating the stratospheric impact. Therefore, we added the following sentence at the end of Page 6752: "In addition to that, in Sect. 5 we discuss the potential of different methods to mitigate the impact caused by stratospheric variability."

Page 6755, line 8: Would you consider these improvements significant? They seem very small.

We alleviated our too optimistic statement. See also answer to review #2.

New Section 5 Mitigation strategies for the stratospheric impact

The stratospheric part of CH_4 is defined by the position of the tropopause and the CH_4 vmr gradient in the UTLS. Whereas the CH_4 vmr gradient in the UTLS is the result of large-scale vertical transport (Brewer Dobson circulation), the position of the tropopause depends on synoptic (e.g., polar vortex, STE) as well as seasonal variations (except for the tropics). Due to these facts the stratospheric part of CH_4 can produce a lot of variability and uncertainty (smoothing effects) in the total column CH_4, and, consequently, in XCH_4. This variability can be misleading within the analysis of trends with regard to tropospheric emissions. Therefore, methods have been developed to overcome this problem of stratospheric variability by separating the tropospheric part of CH_4 from the total column

First of all, Washenfelder et al. (2003) showed that stratospheric tracer-tracer relationships can be used to approximate the stratospheric CH_4 in order to subtract it from the total column CH_4 to get a tropospheric CH_4 with a theoretical precision of ~0.5%. Their study was based on simultaneous NIR retrievals of CH_4 and hydrogen fluoride (HF) that is strongly anti-correlated with stratospheric CH_4. This method has been refined by Saad et al. (2014) by explicitly accounting for averaging kernels. As a result the mean precision of tropospheric CH_4 was improved to ~0.1%. At the same time, Wang et al. (2014) showed that nitrous oxide (N_2O) can also be an appropriate proxy for stratospheric CH_4 with less H_2O dependence compared to HF. In contrast to NIR, the proxy retrieval is not applicable for MIR measurements, as shown by Sepulveda et al (2012). However, the latter extracted tropospheric CH_4 directly from optimized profile retrievals in the MIR.

Nevertheless, all methods mentioned have their limitations, e.g. Sepulveda et al. (2014) found that their tropospheric CH_4 product can be affected significantly by variations in the UTLS. They emphasize that tropospheric CH_4 with a precision of ~0.5% can be derived only by means of an a posteriori correction of the MIR retrievals. The results of the proxy methods in the NIR achieve high theoretical precision indeed, but their accuracy directly depends on the accuracy of the total column CH_4 which is linked to the quality of a priori profiles. E.g., especially in polar vortex conditions NIR proxy retrievals are limited due to smoothing effects (Saad et al., 2014). In the end the smoothing effect described in Sect 4.1.1 directly affects the accuracy of the tropospheric CH_4 which points to the importance of realistic a priori profiles for scaling retrievals. One further problem is introduced by the complex structure of STE events because they do not only affect the stratosphere, but also the troposphere, as shown in Sect. 2.2.1.

Due to these reasons and the fact, that our MIR retrieval strategy is optimized with regard to total columns, we decided to focus on total XCH_4. However, a comparison of tropospheric NIR and MIR retrievals is certainly interesting. It is obvious that such an intercomparison should be performed with retrieval methods dedicated to tropospheric columns, and improved a priori profiles that are able to reproduce the polar vortex subsidence in a realistic way. The NIR proxy retrievals (Saad et al., 2014, Wang et al., 2014) will benefit from these a priori profiles in the same way as the total column retrievals. In addition to that, these profiles could be used by tropospheric MIR profile retrievals (Sepulveda et al., 2014), thereby acting as a common prior within an intercomparison study.

Attention should also be paid on the dependence of XCH_4 on stratospheric variability with regard to validation of satellite and model data. This means, that the prior and the averaging kernels should be taken into account, particularly as satellites have their own vertical sensitivities. In the case of similar vertical sensitivities we can assume that the smoothing effects from satellite and ground-based retrievals are of nearly equal magnitude. Hence, remaining differences can be attributed to other error sources like systematic errors. Thus, NIR TCCON retrievals may be more valuable than MIR NDACC retrievals if SWIR (short wave infra-red) sounders like GOSAT or SCIAMACHY are concerned. Furthermore, it is important to be very careful with the interpretation of results when evaluating model data from arctic regions.

Instead of separating the stratospheric part from the total column, we tried to detect and exclude situations with high dynamical variability by analyzing meteorological parameters (Sect. 4.1.2 and Sect. 4.2.2). In future, these meteorological criteria for exclusion could be installed as additional information in the prior data (TCCON). Another possibility for the exclusion of affected data has been presented by Angelbratt et al. (2011). They used a multiple regression model with anomalies of HF, carbon monoxide (CO), and tropopause height to reduce the variability in total column CH_4. This concept of anomalies can also be transferred to the retrieval process. Certain threshold values of stratospheric tracers (HF, N_2O) can be implemented as filter criteria for XCH_4. In Appendix D we show that HF is suitable for detecting polar vortex subsidence. In contrast, the situation with STE is much more complex and HF cannot be recommended as an additional filter criterion."

Substantial progress could be achieved when combining the NIR and MIR measurements in a concurrent retrieval setup. In this approach the difference in averaging kernels is not considered as a problem but as an opportunity to differentiate more atmospheric layers. The synergetic potential of such an approach seems to be more promising in the case of polar vortex subsidence because the dynamics of subsidence are not as fine-structured compared to STE events. However, the main objective of combining both retrievals will be to determine the shape of the stratospheric CH_4 profile. Hence, reducing the uncertainty of the stratospheric CH_4 will be beneficial for models and retrievals. We want to construct such a combined method and also want to present the method in a subsequent study.

APPENDIX D: Mitigation strategy for stratospheric impact using XHF

In Sect. 4.1.2 and Sect 4.2.2 we applied meteorological criteria for detecting situations with polar vortex subsidence, and STE, respectively. These methods are complex and not really practicable in the operational retrieval process. In contrast to that, it seems to be logical to use simultaneous measurements of stratospheric tracers (HF, N_2O) for detecting polar vortex subsidence and STE. The chemically inert trace gas HF appears to be an excellent candidate for such an intention since it is strongly anti-correlated with stratospheric CH_4.

In the case of polar vortex subsidence it can be expected that the stratospheric CH_4 depletion involves an enhancement in the HF total column. Figure D1 shows the NIR XHF daily mean time series of Ny-Ålesund. Measurement days with polar vortex subsidence were identified in Sect. 4.1.2 and are highlighted in Fig. D1. Indeed, it is clearly recognizable that the XHF values are significantly higher (~ factor 2) at the predominant part of polar vortex situations. Therefore, we conclude that it would be possible to exclude polar vortex situations at Ny-Ålesund with an XHF threshold of ~ 100 ppt.

In contrast to that, the situation is more difficult with regard to STE due to its complex nature. Thus, it is possible that both enhancements and depletions of CH_4 and consequently HF can occur. Figure D2 shows the NIR XHF daily mean time series of Garmisch. Days with and without STE were identified according to Sect. 4.2.2 and can be distinguished in Fig. D2. It is obvious that XHF is in same range of values in situations with STE as in situations without STE. Hence, XHF is not suitable for the detection of STE events.

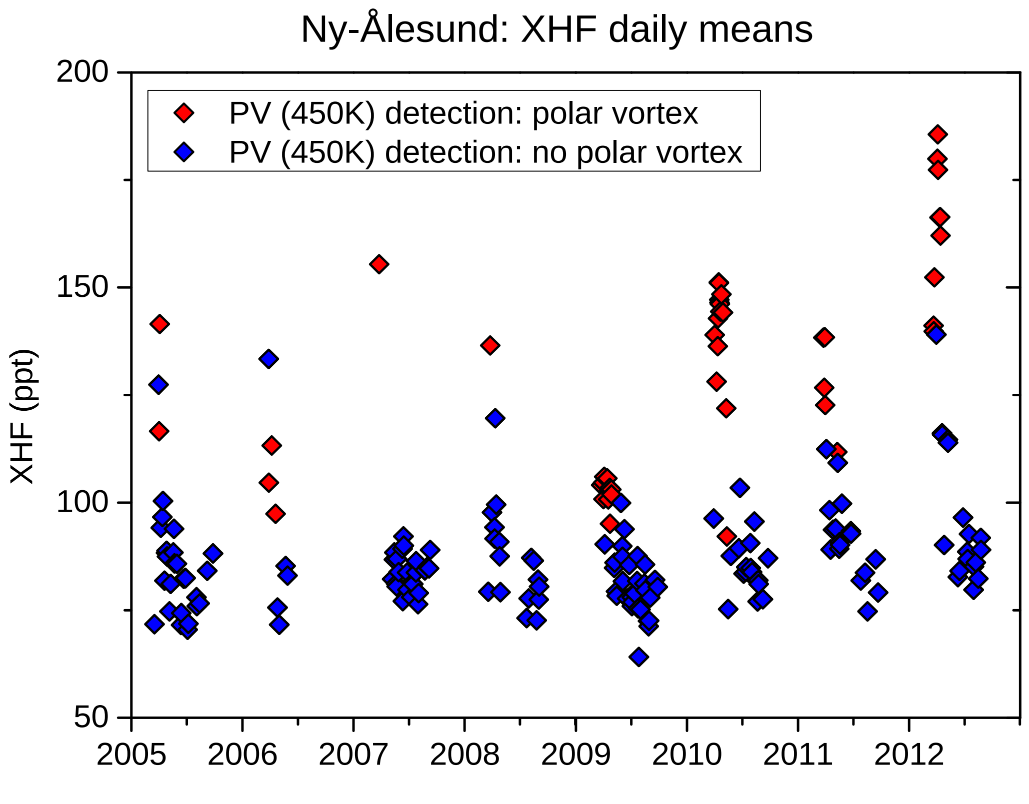

Fig. D1. XHF daily mean time series for Ny-Ålesund obtained from NIR measurements (TCCON). Polar vortex days were detected by using the PV criterion by Nash et al. (1996).

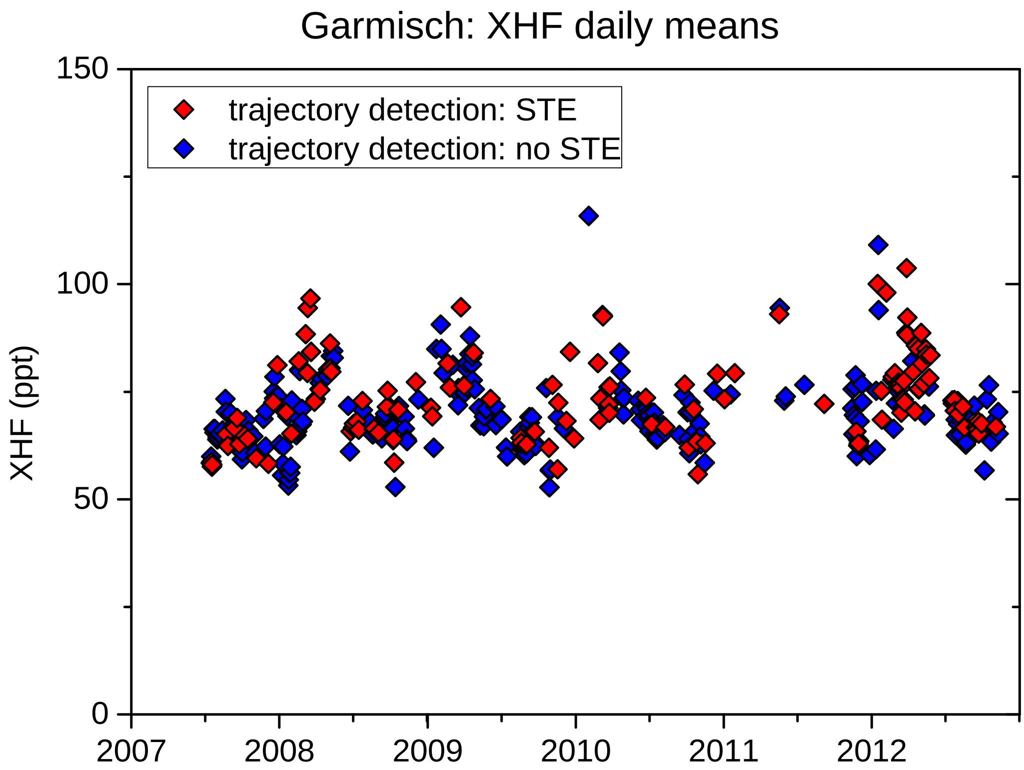

Fig. D2. XHF daily mean time series for Garmisch obtained from NIR measurements (TCCON). The detection of measurement days affected by STE is described in Sect. 4.2.2.

New References:

Angelbratt, J., Mellqvist, J., Blumenstock, T., Borsdorff, T., Brohede, S., Duchatelet, P., Forster, F., Hase, F., Mahieu, E., Murtagh, D., Petersen, A.K., Schneider, M., Sussmann, R., and Urban, J.: A new method to detect long term trends of methane (CH_4) and nitrous oxide (N_2O) total columns measured within the NDACC ground-based high resolution solar FTIR network, Atmos. Chem. Phys., 11, 6167–6183, doi:10.5194/acp-11-6167-2011, 2011.

Saad, K. M., Wunch, D., Toon, G. C., Bernath, P., Boone, C., Connor, B., Deutscher, N. M., Griffith, D. W. T., Kivi, R., Notholt, J., Roehl, C., Schneider, M., Sherlock, V., and Wennberg, P. O.: Derivation of tropospheric methane from TCCON CH4 and HF total column observations, Atmos. Meas. Tech., 7, 2907-2918, doi:10.5194/amt-7-2907-2014, 2014.

Sepúlveda, E., Schneider, M., Hase, F., García, O. E., Gomez-Pelaez, A., Dohe, S., Blumenstock, T., and Guerra, J. C.: Long-term validation of total and tropospheric column-averaged CH4 mole fractions obtained by mid-infrared ground-based FTIR spectrometry, Atmos. Meas. Tech. Discuss., 5, 1381-1430, doi:10.5194/amtd-5-1381-2012, 2012.

Sepúlveda, E., Schneider, M., Hase, F., García, O. E., Gomez-Pelaez, A., Dohe, S., Blumenstock, T., and Guerra, J. C.: Long-term validation of tropospheric column-averaged CH4 mole fractions obtained by mid-infrared ground-based FTIR spectrometry, Atmos. Meas. Tech., 5, 1425-1441, doi:10.5194/amt-5-1425-2012, 2012.

Wang, Z., Deutscher, N. M., Warneke, T., Notholt, J., Dils, B., Griffith, D. W. T., Schmidt, M., Ramonet, M., and Gerbig, C.: Retrieval of tropospheric column-averaged CH4 mole fraction by solar absorption FTIR-spectrometry using N2O as a proxy, Atmos. Meas. Tech. Discuss., 7, 1457-1493, doi:10.5194/amtd-7-1457-2014, 2014.

Washenfelder, R. A., P. O. Wennberg, and G. C. Toon (2003), Tropospheric methane retrieved from ground-based near-IR solar absorption spectra, Geophys. Res. Lett., 30, 2226, doi:10.1029/2003GL017969, 23.

Review #2: Specific Comments

“Page 6749, paras 1, 2, 3 – Since differences between the NIR and MIR XCH4 may be partly due to the spectral line lists, state which version of HITRAN or other line list is used for each set of retrievals. Which version of WACCM? 40-year run? Say how the dry pressure column is obtained for the MIR retrievals. Provide error estimates (precision and accuracy) for NIR and MIR XCH4.”

The description of the retrieval strategies was modified and updated in the original text accordingly (page 6749, line 3 ff.):

“One fixed a priori volume mixing ratio (*vmr*) profile is used per site, derived from the Whole Atmosphere Chemistry Climate Model (WACCM version 5, 40-year run, Garcia et al., 2007). The MIR XCH_4 is calculated by dividing the retrieved total column by the corresponding dry pressure column. To obtain the daily dry pressure column we used the NCEP (National Center for Environmental Prediction) pressure-temperature-humidity profile from 12 UT for calculating the air column and water vapor column.

For the NIR retrievals GFIT uses an iterative method of scaling the a priori profile to provide the best fit to the measured spectrum in the near-infrared spectral region (5938 cm^{-1} - 6076 cm^{-1}). The retrieved total column is divided by the dry pressure column derived from the simultaneously measured oxygen column (Wunch et al., 2011a) and subsequently scaled by the calibration factor 0.976. This calibration is used to account for spectroscopic uncertainties and was determined from various campaigns using coincident airborne in situ measurements calibrated to the WMO scale (Wunch et al., 2010; Geibel et al., 2012). The 2-σ uncertainty of the calibration factor is ~0.2% and can be regarded as the accuracy of TCCON XCH_4. In contrast to that, MIR retrievals are used without calibration, but are optimized to reduce the seasonal bias due to H_2O dependence to < 0.14%. The precision of MIR and NIR retrievals estimated on 1-σ diurnal variation is < 0.3%.

For the MIR retrievals we used HITRAN[1] 2000 including the 2001 update release (Rothman et al., 2003). For the NIR retrievals GFIT uses line lists which are based on HITRAN 2008 (Rothman et al., 2009) including an update by Frankenberg et al. (2008). Further details of the retrieval strategies can be found in Sussmann et al. (2013).

Note that the MIR measurements of Karlsruhe and Izaña were analyzed with the retrieval code PROFFIT instead of SFIT. Differences in these retrieval codes are not expected to have an impact on the MIR retrievals as shown by Hase et al. (2004).”

[1] **HI**gh-resolution **TRAN**smission molecular absorption database

"Page 6751, line 1 – State briefly why monthly means were used. How would the results compare with daily or individual comparisons?"

In analogy to Sussmann et al. (2013) we used monthly means to evaluate the agreement between MIR and NIR retrievals. Because the MIR and NIR measurements are performed in alternating mode and with different scan times, they cannot be compared individually. However, they could be averaged for a defined time slot. The statistical analysis could also been performed using daily means, but the results will not change much. E.g., the linear MIR/NIR slope of the Garmisch data with original a priori profiles is 1.0002 (±0.0012) for monthly means, and 1.0000 (±0.0005) for daily means. Although the scatter is higher for daily NIR-MIR differences it will also be reduced by using ACTM as a common prior. Furthermore, we think that it is more clearly to compare monthly means when looking at seasonal differences.

Therefore, we added a clarifying statement at the end of Sect. 3.1: "The criterion of daily coincidence ensures, that the results of the monthly mean intercomparison will also reflect the agreement between NIR and MIR retrievals on daily and shorter timescales."

"Page 6751, line 11 – State what the terms in brackets represent here, e.g., which decimal places they represent: 1.0002(12) for Garmisch, 1.0010(13) for Wollongong and 0.9996(13) for Karlsruhe"

We changed the text accordingly: "The linear MIR/NIR slopes and their corresponding 2-σ uncertainties are obtained from linear fits forced through zero. Consecutively, the slope uncertainty is illustrated behind the slope value in brackets corresponding to the third and fourth decimal place of the slope value. The linear MIR/NIR slopes are not significantly different from 1 for Garmisch …"

"Page 6755, para 2 and Table 3 – The impact on XCH4 of excluding days influenced by the polar vortex seems very small – comment on this."

Yes, it is true, that this impact is very small. Therefore, we conclude that there are further deviations of the common prior from the true CH_4 profile causing different smoothing effects. The latter could be driven by dynamical variability, like STE events which is the connection to Sect. 4.2. Moreover, systematic deviations of the a priori profiles in the stratospheric CH_4 could also act as a source for smoothing effects.

We have alleviated our too optimistic statement in paragraph 2 and added the argument about smoothing effects due to systematic stratospheric deviations in paragraph 3:

"The parameters from these fits are listed in Table 3. The linear MIR/NIR slope of the data set that is corrected to ACTM as common prior is improved slightly from 0.9940(19) to 0.9950(20) and the stdv is further reduced from 11.5 ppb to 11.0 ppb.

Despite these slightly positive effects of the exclusion of polar vortex situations on the overall intercomparison, there are still significant residual XCH_4 differences (MIR-NIR) for Ny-Ålesund, which vary temporally (see Fig. C2). ... Therefore, the emphasis of Sect. 4.2 lies on the impact of dynamical variability caused by stratosphere-troposphere exchange (STE) processes. Moreover, systematic deviations of the a priori profiles in the stratospheric CH_4 could also act as a source for smoothing effects."

"Page 6758, last para – Explain why none of the trajectories in Figure 8 pass through the blue box."

The trajectories in Fig. 8 do not touch or pass through the detection area because, most likely, there will not be a STE event at Garmisch for the next four days. This clarification has been added.

"Page 6759, para 2 and Table 4 – Is the agreement significantly improved for the original a priori case? The slope for the Garmisch comparison using original a priori gets worse (1.0015 vs. 1.0001) without STE."

Yes, the agreement is improved significantly for the original a priori case with regard to the stdv of NIR-MIR differences. The slope is not improved, but is still in the range of TCCON accuracy (0.2%). To prevent confusion the text was modified to:

"When excluding the affected XCH_4 data the agreement is improved significantly for the original a priori case with regard to the stdv of difference time series NIR-MIR: using the original and common ACTM priors (see Fig. C3) it is reduced from 8.2 ppb to 6.5 ppb and from 6.2 ppb to 4.7 ppb, respectively. The MIR/NIR slope from the intercomparison with original prior (Fig. 9.a) is not improved, but is still in the range of TCCON accuracy. In the case of common prior (Fig. 9.b), the MIR/NIR slope is slightly improved by the exclusion of affected data. All fit parameters for the data sets of the STE analysis at Garmisch are given in Table 4."

"Page 6776, Table 2 caption – Make clear the significance of the (nn) 2sigma terms – what decimal places do these represent? State what the stdv is for."

This has been clarified and the table caption now reads: "Slope of linear scatter plot fits between multi-annual data sets of NIR and MIR XCH_4 retrievals using varied a priori profiles. Slope uncertainties are derived from the fit and are at 2 σ. The slope uncertainty is illustrated in brackets corresponding to the third and fourth decimal place of the slope value, i. e. (12) implies a slope uncertainty of ±0.0012. Standard deviations of NIR-MIR differences are provided as stdv. Data are monthly means constructed from same-day measurement coincidences."

"Page 6780, Figure 2 – This figure is very small and the captions are unreadable. "

Figure 2 is composed of 10 subfigures which are very small. The size of the subfigures will be increased to make it more illustrative and to make the captions readable.

Review #2: Technical Corrections:

All agreed and done.

End of response.

New References:

Frankenberg, C., Warneke, T., Butz, A., Aben, I., Hase, F., Spietz, P., and Brown, L. R.: Pressure broadening in the $2\nu_3$ band of methane and its implication on atmospheric retrievals, Atmos. Chem. Phys., 8, 5061–5075, doi:10.5194/acp-8-5061-2008, 2008.

Rothman, L.S., Barbe, A., Benner, D.C., Brown, L.R., Camy-Peyret, C., Carleer, M.R., Chance, K., Clerbaux, C., Dana, V., Devi, V.M., Fayt, A., Flaud, J.M., Gamache, R.R., Goldman, A., Jacquemart, D., Jucks, K.W., Lafferty, W.J., Mandin, J.Y., Massie, S.T., Nemtchinov, V., Newnham, D.A., Perrin, A., Rinsland, C.P., Schroeder, J., Smith, K.M., Smith, M.A.H., Tang, K., Toth, R.A., Vander Auwera, J., Varanasi, P., and Yoshino, K.: The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001, J. Quant. Spectrosc. Radiat. Transfer, 82, 5-44, 2003.

Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Fally, S., Flaud, J.M., Gamache, R. R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W. J., Mandin, J., Massie, S. T., Mikhailenko, S. N., Miller, C. E., Moazzen-Ahmadi, N., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V. I., Perrin, A., Predoi-Cross, A., Rinsland, C. P., Rotger, M., Šimečková, M., Smith, M. A. H., Sung, K., Tashkun, S. A., Tennyson, J., Toth, R. A., Vandaele, A. C., and Vander Auwera, J.: The HITRAN 2008 molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transfer, 110, 533–572, 2009.