Contributions of the troposphere and stratosphere to CH\textsubscript{4} model biases

Zhiting Wang\textsuperscript{1}, Thorsten Warneke\textsuperscript{1}, Nicholas M. Deutscher\textsuperscript{1,2}, Justus Notholt\textsuperscript{1}, Ute Karstens\textsuperscript{3}, Marielle Saunois\textsuperscript{4}, Matthias Schneider\textsuperscript{5}, Ralf Sussmann\textsuperscript{6}, Harjinder Sembhi\textsuperscript{7}, David W. T. Griffith\textsuperscript{2}, Dave F. Pollard\textsuperscript{8}, Rigel Kivi\textsuperscript{9}, Christof Petri\textsuperscript{1}, Voltaire A. Velazco\textsuperscript{2}, Michel Ramonet\textsuperscript{4}, and Huilin Chen\textsuperscript{10,11}

\textsuperscript{1}Institute of Environmental Physics, University of Bremen, Bremen, Germany
\textsuperscript{2}Centre for Atmospheric Chemistry, School of Chemistry, University of Wollongong, Wollongong, New South Wales, Australia
\textsuperscript{3}Max Planck Institute for Biogeochemistry, Jena, Germany
\textsuperscript{4}Laboratoire des Sciences du Climat et de l’Environnement, LSCE-IPSL (CEA-CNRSUVSQ), Université Paris-Saclay, 91 191 Gif Sur Yvette, France
\textsuperscript{5}Karlsruhe Institute of Technology, IMK-ASF, Karlsruhe, Germany
\textsuperscript{6}Karlsruhe Institute of Technology, IMK-IFU, Garmisch-Partenkirchen, Germany
\textsuperscript{7}Earth Observation Science, Department of physics and Astronomy, University of Leicester, Leicester, UK
\textsuperscript{8}National Institute of Water and Atmospheric Research (NIWA), Wellington, New Zealand
\textsuperscript{9}Finnish Meteorological Institute Arctic Research Center, FMI-ARC, Sodankylä, Finland
\textsuperscript{10}Center for Isotope Research (CIO), University of Groningen, Groningen, the Netherlands
\textsuperscript{11}Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, CO, USA

Correspondence to: Zhiting Wang (zhiting@iup.physik.uni-bremen.de)

Received: 22 November 2016 – Discussion started: 23 December 2016
Revised: 12 September 2017 – Accepted: 3 October 2017 – Published: 9 November 2017

Abstract. Inverse modelling is a useful tool for retrieving CH\textsubscript{4} fluxes; however, evaluation of the applied chemical transport model is an important step before using the inverted emissions. For inversions using column data one concern is how well the model represents stratospheric and tropospheric CH\textsubscript{4} when assimilating total column measurements. In this study atmospheric CH\textsubscript{4} from three inverse models is compared to FTS (Fourier transform spectrometry), satellite and in situ measurements. Using the FTS measurements the model biases are separated into stratospheric and tropospheric contributions. When averaged over all FTS sites the model bias amplitudes (absolute model to FTS differences) are 7.4±5.1, 6.7±4.8, and 8.1±5.5 ppb in the tropospheric partial column (the column from the surface to the tropopause) for the models TM3, TM5-4DVAR, and LMDz-PYVAR, respectively, and 4.3±9.9, 4.7±9.9, and 6.2±11.2 ppb in the stratospheric partial column (the column from the tropopause to the top of the atmosphere). The model biases in the tropospheric partial column show a latitudinal gradient for all models; however there are no clear latitudinal dependencies for the model biases in the stratospheric partial column visible except with the LMDz-PYVAR model. Comparing modelled and FTS-measured tropospheric column-averaged mole fractions reveals a similar latitudinal gradient in the model biases but comparison with in situ measured mole fractions in the troposphere does not show a latitudinal gradient, which is attributed to the different longitudinal coverage of FTS and in situ measurements. Similarly, a latitudinal pattern exists in model biases in vertical CH\textsubscript{4} gradients in the troposphere, which indicates that vertical transport of tropospheric CH\textsubscript{4} is not represented correctly in the models.
1 Introduction

Atmospheric methane (CH$_4$) is the second most important anthropogenic greenhouse gas. Atmospheric CH$_4$ concentrations began to rise again in 2007 after a decade of near-zero growth (Rigby et al., 2008). Possible explanations for the stability of CH$_4$ concentrations during 1999–2006 include an increase in anthropogenic emissions and coincident decrease in wetland emissions (Bousquet et al., 2006), decreased Northern Hemisphere microbial sources (Kai et al., 2011), and a combination of decreasing-to-stable fossil fuel emissions and stable-to-increasing microbial emissions (Kirschke et al., 2013). Several possible reasons for the renewed growth of CH$_4$ concentrations after 2006 have been proposed, including the increase of wetland emissions during 2007 and 2008 in either the tropics, owing to greater than average precipitation, and/or in the Arctic, owing to high temperatures (Dlugokencky et al., 2009); the anthropogenic contribution at the tropics and midlatitudes in the Northern Hemisphere during the period 2007–2010 (Bergamaschi et al., 2013); an increase of emissions from oil and gas production and use during 2007–2014 (Hausmann et al., 2016); and from agriculture (Schaefer et al., 2016).

Prediction of the evolution of CH$_4$ in the atmosphere requires knowledge of the sources and sinks. Inverse modelling is usually used to retrieve fluxes from observations of atmospheric concentrations. The commonly used measurements include surface measurements from global networks, such as the NOAA/ESRL (Earth System Research Laboratory of the National Oceanic and Atmospheric Administration), and total column data from satellites, such as the SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHArtography) or GOSAT (Greenhouse gases Observing Satellite). However, compared to total column data the surface measurements characterise the boundary layer only and CH$_4$ concentrations in the boundary layer are sensitive to boundary layer height, which is difficult to accurately simulate in a global transport model. The total column measurements are less sensitive to model errors in the vertical distributions of CH$_4$. However, they are also only sensitive to broader-scale signatures. Compared to satellite measurements, surface in situ measurements have poor spatial coverage but are more precise and less subject to biases. Total column measurements of CH$_4$ include a contribution from the stratosphere where the concentrations are influenced by dynamical processes like meridional transport, tropopause variations, and subsidence associated with the polar vortex, and chemistry. If a transport model does not accurately simulate these processes, the retrieved sources and sinks using total column measurements will not be correct (Locatelli et al., 2015a, b). Especially in the polar region, the tropopause height varies strongly and the dynamical processes are complex. Turner et al. (2015) compared GOSAT CH$_4$ with GEOS-Chem simulations, and found large differences at high latitudes. They proposed that the model bias in total column CH$_4$ at high latitudes comes from the stratosphere since the validation with TCCON (Total Carbon Column Observing Network), NOAA surface and aircraft measurements, and HIPPO shows good performances of the model in the troposphere. Ostler et al. (2016) assessed accuracies of models in the stratosphere by replacing modelled stratospheric CH$_4$ with satellite measurements. They found that modelled stratospheric CH$_4$ shows large scatter and the corrected total columns of CH$_4$ show improved or degraded agreements with TCCON measurements depending on the used satellites and models. These results imply that satellite-based stratospheric CH$_4$ is not accurate enough to resolve a possible stratospheric contribution to model biases in total column CH$_4$ as uncovered by TCCON. TCCON-based measurements could fulfill such a role, as presented in Saad et al. (2016) and this study. Using HF as a proxy, Saad et al. (2016) derived tropospheric CH$_4$ products and investigated the impact of stratospheric and tropospheric model biases in GEOS-Chem on inversions. They found an increasing stratospheric mismatch with decreasing tropopause altitudes and a phase lag in modelled tropospheric seasonality. A small bias in the modelled CH$_4$ column could come from counteracting stratospheric and tropospheric model errors. They noted that the tropospheric time lag can produce large errors in posterior wetland emissions at high northern latitudes.

In this study the model biases in the stratosphere and troposphere are assessed with respect to the latitudinal pattern. In order to investigate the accuracy of the models several measurements are used: (i) total, tropospheric, and stratospheric column-averaged CH$_4$ mole fractions measured at the TCCON (Wunch et al., 2011; Wang et al., 2014), which are used to separate stratospheric and tropospheric contributions to model bias in total columns; (ii) total column-averaged CH$_4$ mole fraction measured by GOSAT (Parker et al., 2011) and CH$_4$ profiles measured by TES (Tropospheric Emission Spectrometer) (Worden et al., 2012); (iii) surface CH$_4$ measured within the NOAA network (Dlugokencky et al., 1994); and (iv) in situ CH$_4$ profiles from aircraft campaign HIPPO (HIAPER Pole-to-Pole Observations) (Wofsy et al., 2012). In the following, Sect. 2 presents the measurements, models, and analysis approach, while Sect. 3 presents the results and discussions. Conclusions are drawn in Sect. 4.
Figure 1. Calibration results of FTS-derived tropospheric column-averaged CH$_4$ mole fractions against in situ measurements. The in situ profiles are smoothed using GFIT CH$_4$ averaging kernels in the troposphere as described in Wang et al. (2014). The FTS data are averaged for the in situ measurement periods. The IMECC is an aircraft campaign over Europe (Geibel et al., 2012). The Lamont-AirCore measurements are from Greenhouse Gas Group Aircraft Program (http://www.esrl.noaa.gov/gmd/ccgg/aircraft/). The AirCore data at Sodankylä is from the FTS group there.

Table 1. Overview of TCCON sites used.

<table>
<thead>
<tr>
<th>TCCON site</th>
<th>Latitude (° N)</th>
<th>Longitude (° E)</th>
<th>Altitude (m a.s.l.)</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ny-Ålesund</td>
<td>78.9</td>
<td>11.9</td>
<td>20</td>
<td>Messerschmidt et al. (2010)</td>
</tr>
<tr>
<td>Sodankylä</td>
<td>67.3668</td>
<td>26.6310</td>
<td>188</td>
<td></td>
</tr>
<tr>
<td>Bialystok</td>
<td>53.23</td>
<td>23.025</td>
<td>183</td>
<td>Messerschmidt et al. (2012)</td>
</tr>
<tr>
<td>Bremen</td>
<td>53.10</td>
<td>8.85</td>
<td>27</td>
<td>Messerschmidt et al. (2010)</td>
</tr>
<tr>
<td>Orléans</td>
<td>47.97</td>
<td>2.113</td>
<td>130</td>
<td>Messerschmidt et al. (2010)</td>
</tr>
<tr>
<td>Garmisch</td>
<td>47.476</td>
<td>11.063</td>
<td>740</td>
<td>Sussmann et al. (2013), Sussmann and Rettinger (2014)</td>
</tr>
<tr>
<td>Park Falls</td>
<td>45.945</td>
<td>−90.273</td>
<td>440</td>
<td>Washenfelder et al. (2006)</td>
</tr>
<tr>
<td>Lamont</td>
<td>36.604</td>
<td>−97.486</td>
<td>320</td>
<td>Wunch et al. (2009)</td>
</tr>
<tr>
<td>Izaña</td>
<td>28.3</td>
<td>−16.483</td>
<td>2370</td>
<td>Blumenstock et al. (2014)</td>
</tr>
<tr>
<td>Darwin</td>
<td>−12.424</td>
<td>130.891</td>
<td>30</td>
<td>Deutscher et al. (2010)</td>
</tr>
<tr>
<td>Wollongong</td>
<td>−34.406</td>
<td>150.879</td>
<td>30</td>
<td>Deutscher et al. (2010)</td>
</tr>
<tr>
<td>Lauder</td>
<td>−45.038</td>
<td>169.684</td>
<td>370</td>
<td>Sherlock et al. (2014)</td>
</tr>
</tbody>
</table>

ment within 3.0 ± 2.0 ppb (see Fig. 1). Given the total and tropospheric CH$_4$ columns, stratospheric column-averaged CH$_4$ is derived using knowledge of the tropopause pressure. The TCCON sites used in this study are listed in Table 1, the products are all using the GGG2014 version (Wunch et al., 2015), except for at Ny-Ålesund.

The CO$_2$ proxy retrieval method (Frankenberg et al., 2011) is applied in GOSAT data, which infers dry air columns from the CO$_2$ columns retrieved from the same spectra as used in the CH$_4$ retrieval. This method assumes the CO$_2$ concentrations are known and provided by model simulations (the CarbonTracker model). The GOSAT total column-averaged dry-air CH$_4$ mole fractions used here are version UoL-OCPRv7 and only spectra measured in clear-sky conditions are used (Parker et al., 2011). GOSAT has a ground footprint diameter of about 10.5 km and 4 s exposure duration. The TES instrument measures atmospheric radiances from which atmospheric profiles are inferred using an optimal estimation algorithm subject to a priori constraints. The CH$_4$ retrieval of TES has a DOFS (degree of freedom for signal) of about 0.8–
Table 2. Information on the models and set-up details.

<table>
<thead>
<tr>
<th>Model</th>
<th>Institute</th>
<th>Resolution (lat x lon)</th>
<th>No. of vertical levels</th>
<th>Output time step (h)</th>
<th>Meteorology</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM3</td>
<td>Max Plank Institute for Biogeochemistry</td>
<td>4° x 5°</td>
<td>26</td>
<td>3.0</td>
<td>ERA-Interim</td>
</tr>
<tr>
<td>TM5-4DV AR</td>
<td>European Joint Research Centre</td>
<td>1° x 1° for Europe, 6° x 4° for the rest of the world</td>
<td>25</td>
<td>1.5</td>
<td>ECMWF-IFS</td>
</tr>
<tr>
<td>LMDz-PYVAR</td>
<td>Laboratoire des Sciences du Climatet de l’Environnement</td>
<td>1.875° x 3.75°</td>
<td>39</td>
<td>3.0</td>
<td>Prediction by LMDz with nudging to ECMWF reanalysis</td>
</tr>
</tbody>
</table>

Table 3. FTS and in situ sites used for comparison to FTS tropospheric column-averaged CH₄ and surface/tower CH₄.

<table>
<thead>
<tr>
<th>FTS site</th>
<th>In situ site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Lat (° N)</td>
</tr>
<tr>
<td>Ny-Ålesund</td>
<td>78.923</td>
</tr>
<tr>
<td>Sodankylä</td>
<td>67.367</td>
</tr>
<tr>
<td>Orléans</td>
<td>47.965</td>
</tr>
<tr>
<td>Park Falls</td>
<td>45.945</td>
</tr>
<tr>
<td>Lamont</td>
<td>36.604</td>
</tr>
<tr>
<td>Izana</td>
<td>28.300</td>
</tr>
<tr>
<td>Lauder</td>
<td>-45.038</td>
</tr>
</tbody>
</table>

2.3, which peaks in the tropics and decreases toward high latitudes. The version F07_10 data are applied and measurements with less than 1.4 DOFS are filtered out. Validation of F07_10 data against HIPPO measurements shows a bias of −8–5 ppb with standard deviations of 25–50 ppb below 100 hPa (Herman et al., 2014).

Vertical gradients of tropospheric CH₄ can be qualitatively calculated by using the comparative tropospheric column-averaged CH₄ and surface CH₄. Only long-term timescales are used here, and variations with scales longer than 1.4 years are extracted from the time series of tropospheric and surface CH₄. TCCON and in situ sites are selected to be located close to one another so that both instruments measure similar air masses. The sites and measurements are listed in Table 3.

The CH₄ measurements during HIPPO 1–5 are those made with a quantum cascade laser spectrometer (QCLS). Calibrations derived through comparisons with NOAA Programmable Flask Package measurements are applied.

The models used in this study are TM3, TM5-4DVAR, LMDz-PYVAR; details are given in Table 2. All the three models are optimised against in situ measurements at the surface through inversions of CH₄ surface emissions. The first two models used a common emission a priori for their inversion runs. Detailed information on the inversion methodology is discussed in Bergamaschi et al. (2015). The LMDz-PYVAR uses different a priori and background stations as constraints, the BG–SP (background network – transport parameterisation scheme) set-up described in Locatelli et al. (2015b). The chemical reactions considered in the models are the oxidation by OH in the troposphere and by Cl, OH, and O(1D) in the stratosphere. The fields of the radicals are prescribed monthly with no interannual changes.

Details on the global atmospheric tracer model TM3 can be found in Heimann and Körner (2003) and the inversion method of the Jena CarboScope is described in Rodenbeck (2005). TM5-4DVAR is a four-dimensional data assimilation system for inverse modelling of atmospheric methane emission (Meirink et al., 2008). The system is based on the TM5 atmosphere transport model (Krol et al., 2005). LMDz-PYVAR is a framework that combines the inversion system PYVAR (Chevallier et al., 2005; Pison et al., 2009) with the transport model LMDz (Hourdin et al., 2006).

For evaluation of the models, we interpolate the simulations in time, latitude, longitude, and pressure to match the measurements. For the total and tropospheric column-averaged CH₄ the model profile is integrated taking the a priori and averaging kernel into account according to Rodgers and Connor (2003) using Eqs. (9) and (14) from Wang et al. (2014). In contrast to FTS and GOSAT the transformation of model CH₄ profiles to the counterpart of TES is done in logarithms of a prior and model quantities. The thermal tropopause calculated using the ERA-Interim reanalysis data is used in all calculations, and would not be so accurate for the TM5 and LMDz models, especially for LMDz, which predicts its own meteorology fields through nudging to reanalysis data.
3 Comparison between measurements and models

The CH\textsubscript{4} column meridional distribution is sensitive to the latitudinal distribution of CH\textsubscript{4} sources and sinks, tropopause altitudes, inter-hemisphere transport in the troposphere, and the residual circulation in the stratosphere. Assessing latitudinal variabilities of biases of a model could reveal how well these processes are represented in the model. Another important concern of this study is to determine which of the tropospheric or stratospheric components contributes more to the model biases in the total column. The model to FTS comparison covers the period 2007–2011 when FTS measurements are available and the comparison to GOSAT is for the period 2009–2011.

The latitudinal behaviour of the model bias in total column-averaged CH\textsubscript{4} mole fractions is revealed by comparisons to FTS and GOSAT measurements as presented in Fig. 2, similarly to previous work (Monteil et al., 2013). CH\textsubscript{4} is emitted mainly in the Northern Hemisphere, destroyed mainly in the tropics by OH, and has a slow inter-hemisphere transport with a temporal scale of approximately 1 year. CH\textsubscript{4} is transported into the stratosphere mostly in the tropics and back to the troposphere in the extratropics by the residual circulation. In the troposphere, CH\textsubscript{4} concentrations are higher in the Northern Hemisphere than in the Southern Hemisphere with a gradient throughout the tropics. In the stratosphere, CH\textsubscript{4} has a more or less symmetrical distribution between the two hemispheres. In Fig. 2 the model biases present a clear latitudinal dependence, similar to results revealed by other studies (e.g. Turner et al., 2015 and Alexe et al., 2015). The latitudinal dependence is similar between FTS and GOSAT northward of 50\degree S where FTS measurements are available. The model to measurements difference shows a north–south gradient with positive values at northern high-latitudes, northward of 50\degree S for all the models.

With FTS-derived tropospheric and stratospheric column-averaged CH\textsubscript{4} (Wang et al., 2014), it is possible to examine how the tropospheric and stratospheric partial columns contribute to the model bias in the total column-averaged CH\textsubscript{4}. The partial columns are represented as the tropospheric and stratospheric column-averaged mole fractions scaled by the fraction of the partial air column. Figure 3 shows yearly and seasonal median model biases in the troposphere and stratospheric partial columns. It is clear that model biases in the tropospheric partial column exhibit a north–south gradient with positive values at northern high latitudes during all seasons for all models. The model biases in the stratospheric partial column do not present any clear latitudinal pattern that persists throughout the whole year and shows significant seasonal variabilities for TM3 and TM5-4DVAR. This is consistent with the fact that stratospheric CH\textsubscript{4} distributions cycle between summer and winter hemispheric states.

In the case of LMDz-PYVAR there is a permanent pattern in the stratospheric partial column biases that is more negative in the south. This pattern is consistent with the north–south gradient in the total column biases. Compared to Fig. 2 one
can see that the latitudinal pattern of model biases in total column-averaged \( \text{CH}_4 \) results from both the stratosphere and troposphere for LMDz-PYV AR, but arises from the troposphere for TM3 and TM5. The model biases change signs yearly and seasonally; therefore it is more appropriate to use the amplitudes (absolute model to FTS differences) to evaluate the contributions of the troposphere and stratosphere. The medians of model bias amplitudes over all FTS sites and years are 7.4 ± 5.1 ppb in the tropospheric partial column and 4.3 ± 9.9 ppb in the stratospheric partial column for TM3, 6.7 ± 4.8 and 4.7 ± 9.9 ppb for TM5-4DV AR, and 8.1 ± 5.5 and 6.2 ± 11.2 ppb for LMDz-PYVAR.

Evaluations of the models at the surface using in situ measurements, which are assimilated into the models, show smaller biases than the tropospheric column-averaged \( \text{CH}_4 \). The amplitudes are mostly below 10 ppb in the Northern Hemisphere except for a few outliers and below 5 ppb in the Southern Hemisphere (not shown). The model biases at the surface do not show any significant latitudinal dependence that is present in the model biases of both the tropospheric partial column and column-averaged \( \text{CH}_4 \) (see Fig. A1). It is not clear how the model biases at the surface appear in the regions where no measurements are assimilated. However, it could be true that the overestimation of the tropospheric column-averaged \( \text{CH}_4 \) meridional gradient is due to model biases in the middle and upper troposphere. That would mean that vertical distributions of \( \text{CH}_4 \) in the troposphere are not represented correctly in the models.

Figure 4 presents a comparison of modelled and measured vertical gradients of tropospheric \( \text{CH}_4 \), as qualitatively represented by the difference between the tropospheric column-averaged \( \text{CH}_4 \) and the surface \( \text{CH}_4 \). The vertical gradient is influenced by surface emissions, transport, and OH fields. Generally there are negative vertical gradients in the Northern Hemisphere and positive vertical gradients in the Southern Hemisphere (except for over the southern continents in locations with strong emissions). Here we refer to decreasing \( \text{CH}_4 \) mole fractions with altitude as a negative vertical gradient, while increasing \( \text{CH}_4 \) with altitude is a positive vertical gradient. This occurs because most \( \text{CH}_4 \) is emitted in the Northern Hemisphere and mixed into the southern hemispheric Hadley cell, the southward branch of which prevails in the middle and upper troposphere. In the troposphere, surface emissions cause decreasing \( \text{CH}_4 \) with altitude, while OH oxidation causes a negative vertical gradient. The model biases in the tropospheric vertical gradient are mostly positive at middle and high northern latitudes, and negative at other latitudes. Therefore the overestimated tropospheric \( \text{CH}_4 \) at middle and high northern latitudes could not originate from overestimated emissions, which should result in a more negative vertical gradient in the troposphere.

Figure 5 shows a comparison between model simulations and HIPPO measurements. The results are longitudinally av-
Z. Wang et al.: Contributions of the troposphere and stratosphere to CH$_4$ model biases

Figure 4. Measured (black) and simulated (yellow: TM3, blue: TM5-4DV AR, magenta: LMDz-PYVAR) vertical gradients of CH$_4$ in the troposphere (a) and differences between the measurement and simulations (b) against latitude. The results are averaged for 2007–2011 when FTS measurements are available.

Figure 5. HIPPO-measured CH$_4$ and differences with models in the stratosphere (a) and troposphere (b). The result is an average for five HIPPO missions, averaged for latitudinal bins of 4° and vertical increments of 10 hPa.

averaged for all five HIPPO missions within grids of 4° latitude and pressure increments of 10 hPa. A significant feature is an overestimation of CH$_4$ in the lowermost stratosphere over latitudes higher than 30° S/N, much larger than the biases in the troposphere. It is not clear whether the overestimation arises from the residual transport in the stratosphere, which appears to be too strong, a too high tropopause, an incorrect vertical CH$_4$ gradient across the tropopause or a misrepresentation of stratospheric chemistry. Underestimations dominate in the upper southern troposphere, consistent with the results in Fig. 4 that modelled gradients of tropospheric CH$_4$ are negatively biased as revealed by FTS and surface measurements. There are no significant patterns for the vertical gradient bias in the northern troposphere.

Unlike for the FTS, the model biases in the tropospheric column-averaged CH$_4$ revealed by HIPPO do not show a significant latitudinal trend (Fig. 6, only TM3 are shown there since other models give similar behaviour). This could be because the FTS-measured tropospheric column-averaged CH$_4$ is defined differently to the mean mole fraction between the surface and thermal tropopause. In deriving the FTS tropospheric CH$_4$, the stratospheric CH$_4$ is removed via its linear correlation with N$_2$O. The tropopause in the FTS data therefore has a chemical definition. It is not clear how different from each other the two kinds of tropopause are during this period. A sensitivity test was conducted by shifting the thermal tropopause 200 hPa upward to include the lower stratosphere where CH$_4$ is overestimated by the models. The model biases compared against HIPPO then become closer to those against FTS. However, this difference of 200 hPa between the chemical and thermal tropopause is unrealistically large. In addition, the FTS-measured tropospheric column-averaged CH$_4$ agrees well with in situ measurements in Fig. 1 where the thermal tropopause is applied.

Another possible explanation is that HIPPO sampled the atmosphere mostly in the region 150° E–110° W, over the Pacific Ocean. Apart from Izaña and Ny-Ålesund, the northern FTS sites are located inland. The longitudinal dependence of model biases is investigated with TES-measured CH$_4$ mole fractions at 215, 464, and 680 hPa (the lower panel in Fig. 6). Because the TES profiles have limited vertical resolution, the concentrations at the three levels are not independent. The weighting function of CH$_4$ at 215 hPa peaks around 200 hPa in the tropics and around 300 hPa higher than 50° N/S. The measurements at 464 hPa show the largest sensitivity around 500–600 hPa, and those at 680 hPa have similar vertical sensitivity but fewer weights above 400 hPa. The comparisons are separated into a region representing HIPPO sampling (referred as region I) and the remaining longitudes (referred as region II). Differences between the model biases in the two regions occur northward of 45° N most significantly at the
Figure 6. Comparisons of CH\(_4\) with TM3 and (a) FTS, (b) HIPPO and (c) TES. In the case of HIPPO and FTS the tropospheric column-averaged CH\(_4\) is compared, which is obtained from integration between surface and the tropopause (empty characters) or 200 hPa above the tropopause shifted (solid characters, only in the HIPPO case). For TES CH\(_4\) mole fractions at 215, 464, and 680 hPa are compared with TM3 simulations in a region 110\(^\circ\)W–150\(^\circ\)E (black) and the region beyond it (red) separately. Both TM3 and the measurements are averaged during the HIPPO 1–5 period.

level 215 hPa. Increases in the model biases continue in region II but decrease in region I, which is more or less similar to the differences between model biases revealed by FTS and HIPPO at these latitudes. Consistent with FTS the model–TES difference also shows a north–south gradient northward of 50\(^\circ\) S. However, it is not clear whether the latitudinal pattern comes from the TES retrieval or model errors. Validation of TES tropospheric CH\(_4\) with HIPPO gives near-zeros biases except for latitudes 40–60\(^\circ\) N where the TES biases vary within −10 to −20 ppb (Herman et al., 2014).

4 Conclusions

In this study, three inverse models for CH\(_4\) are evaluated using different observations that cover different scales. The aim is to determine whether most of the model biases are from the stratosphere or troposphere. With FTS stratospheric and tropospheric column-averaged CH\(_4\) derived from the FTS total column measurements, it is shown that model bias amplitudes are 7.4 ± 5.1, 6.7 ± 4.8, and 8.1 ± 5.4 ppb in the tropospheric partial column for TM3, TM5-4DVAR, and LMDz39-PYVAR. The corresponding stratospheric partial column biases are 4.3 ± 9.9, 4.7 ± 9.9, and 6.1 ± 11.2 ppb. The tropospheric partial column model bias exhibits a north–south gradient northward of 50\(^\circ\) S with an overestimation at northern high latitudes for all models. There is no persistent latitudinal pattern with season in the stratospheric partial column model bias for TM3 and TM5-4DVAR.

The evaluation of the models at the surface shows a smaller bias compared to the tropospheric column-averaged CH\(_4\). We assume that the tropospheric model biases are mainly located in the middle and upper troposphere, although comparisons at the surface are only limited to sites where the measurements have been assimilated into the models. A comparison with HIPPO in the troposphere does not show the same latitudinal pattern in model biases as in the comparison with FTS. Two possible reasons are suggested: (i) the difference between the thermal tropopause and that in the FTS tropospheric CH\(_4\) product, and (ii) the latitude patterns of model biases are dependent on longitude. Using an assessment of model biases relative to TES satellite measurements, we propose that the longitudinal dependence of the model performance contributes to the difference between HIPPO and FTS. However, the tropopause altitude could cause differences during short temporal scale processes, e.g. strato-
spheric intrusions where the stratospheric air can sink below the thermal tropopause. Stratospheric air can also detach from the stratosphere completely and enter the troposphere. If the detached air parcels still have stratospheric properties, e.g. CH$_4$ correlates with N$_2$O as in the stratosphere, the FTS-measured tropospheric CH$_4$ excludes these air parcels; however, direct integration from the surface to the thermal tropopause, such as that used for the models and in situ profiles, will include these in the tropospheric CH$_4$. More confusing situations could occur where there is strong mixing across the UTLS (the upper troposphere and lower stratosphere) and both thermal and chemical tropopause are not well defined. Future work will be devoted to clarifying the realistic content in FTS tropospheric column-averaged CH$_4$ and to defining a reasonable approach when comparing it with in situ and model products in these situations.

Data availability. The TCCON data can be obtained from the TCCON Data Archive (http://tccondata.org/). The model outputs are from Marille Saunois (Laboratoire des Sciences du Climat et de l’Environnement, France) for LMDz-PYVAR, Ute Karstens (the Max Plank Institute for Biogeochemistry, Jena, Germany) for TM3, and Peter Bergamaschi (European Commission Joint Research Centre) for TM5-4DVAR. One should contact these authors directly considering the availability the model output. The GOSAT data UoL-OCPRv7, TES data F07_10 and HIPPO data are public available. Surface CH$_4$ measurements from NOAA are publicly available. The in situ CH$_4$ profile measurements by AirCore will become available via the EU project RINGO. Lamont-AirCore measurements have been provided by the Colm Sweeney at the NOAA Carbon Cycle and Greenhouse Gas Group Aircraft Program (http://www.esrl.noaa.gov/gmd/ccgg/aircraft/). The AirCore data at Sodankylä are from the FTS group there.
Appendix A

Figure A1. Latitudinal dependences of yearly averaged model biases in the tropospheric partial column (black) and the tropospheric column-averaged mole fraction (red). The three models are represented by plus (TM3), circle (TM5-4DVAR) and multiplication (LMDz-PYVAR) signs.
**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** This research is funded by EU project InGOS. We acknowledge funding from the European Union’s Horizon 2020 research and innovation programme for the project RINGO (grant agreement no. 730944) as well. Nicholas Deutscher is supported by an ARC-DECRA fellowship, DE140100178. TCCON measurements at Park Falls and Lamont are possible thanks to NASA grants NNX14AF60G, NNX11AG01G, NAG5-12247, and NNG05-GD07G, and the NASA Orbiting Carbon Observatory Program, as well as technical support from the DOE ARM programme (Lamont) and Jeff Ayers (Park Falls). Darwin and Wollongong TCCON support is funded by NASA grants NAGS-12247 and NNG05-GD07G and the Australian Research Council grants DP140101551, DP110103118, DP0879468 and LP0562346, as well as support from the GOSAT project and DOE ARM technical support in Darwin. The EU projects InGOS and ICOS-INWIRE and the Senate of Bremen provide financial support for TCCON measurements at Bremen, Orleans, Bialystok and N-Łęskud, and Orleans is also support by the RAMCES team at LSCE. The Lauter TCCON programme is core-funded by NIWA through New Zealand’s Ministry of Business, Innovation and Employment.

The article processing charges for this open-access publication were covered by the University of Bremen.

Edited by: Ilse Aben
Reviewed by: two anonymous referees

**References**


