This study investigates the origin of a latitudinal bias that has been reported in a number of inverse modelling studies of CH4 using GOSAT retrievals. Most of those studies point to the stratosphere as the plausible origin of a model bias. This study, however, suggests that the upper troposphere could make a sizeable contribution, pointing to errors in the model representation of vertical transport as a possible cause. The comparisons that are presented make a useful contribution to the discussion. However, in my opinion, and as will be explained below, their interpretation requires further attention. This needs to be solved to make this study suitable for publication in ACP.

GENERAL COMMENTS

As explained on page 6, biases are assessed by taking the absolute difference between model and FTS. The motivation is that biases may change sign seasonally, and therefore may not show up in annual averages when positive and negative contributions cancel out. However, whether this is a good choice or not depends on the kind of bias that is investigated. Here the focus is largely on a latitudinal bias. Suppose that there is no latitudinal bias in the annual mean, but only a latitudinally varying bias in the seasonal amplitude. By taking absolute model to FTS differences across the year you would end up with a latitudinally varying bias. In this case the choice of absolute differences was clearly not appropriate. There may not be a single solution to this problem for the biases that are investigated here, but the meaning of the numbers that are summarized in the abstract and the conclusions for stratospheric and tropospheric contribution to the bias is not clear to me. A relation with a latitudinally vaying bias is suggested, but do these numbers really reflect stratospheric and tropospheric contributions to that bias. This requires more attention, including information on how the absolute differences are calculated (on every data point like an RMS, or on monthly averages, or?).

According to the caption of Figure 3, the tropospheric and stratospheric model biases are scaled with the corresponding contributions of the troposphere and the stratosphere to the total column air mass. However, there is a danger in doing so. Suppose that the model had a latitudinally and seasonally uniform offset in the tropospheric concentration. Then the scaling with the seasonal and latitudinal varying tropopause pressure would introduce a seasonal and latitudinal variation in the bias. In that case, when you look for varying biases within the troposphere in comparison with in situ data you wouldn't find any. This is exactly what seems to be happening here. This problem is attributed to differences in the global representation of the measurements, but could also be caused by differences in the NCEP and N2O derived tropopause heights. Since CH4 shown show a sharp vertical concentration gradient just above the tropopause, the analysis may be quite sensitive to how these heights compare. The uncertainty of this needs to be assessed and discussed.

The comparison with TES is used to investigate longitudinal variations in the bias and the global representativeness of the comparisons with HIPPO which are limited to the Pacific. Apart from the fact that it is not clear that the TES data for the troposphere are accurate enough for this purpose (sizeable offsets are seen in the troposphere, that are not due to the TM3 model), the results do not seem to support the case that is made. If anything, the latitudinal gradient in the offset is stronger in the Pacific longitude band (in red) then at other latitudes. The authors are right that the bias has a longitudinal dependence, but it works on the wrong direction. This needs to be discussed more clearly, and the message of the study should be brought in accordance with this finding.

Looking at Figure 5, the most significant differences between the models and HIPPO seem really at the highest measured altitudes. You might debate whether they are in the troposphere or the stratosphere. I wonder how important this really is. Wouldn't it be better to conclude that the problems show up most strongly at tropopause altitudes. In that case the method of separating the troposphere from the stratosphere may actually not be so appropriate. A plausible cause could be strat-trop exchange. I don't see how the results that are presented here exclude this possibility. Yet, it is not considered as an option.

SPECIFIC COMMENTS

page 4, line 8: Where does the tropopause pressure come from?

page 4, line 13: What model CO_2 fields are used to translate the retrieved ratios into XCH_4?

page 5, line 13: 'The NCEP tropopause ...'. It is less accurate for TM5 also, which doesn't use NCEP either (in TM3 it depends on the meteo that was used). Please reformulate to make this sentence more accurate.

Page 7, line 18: 'underestimations dominate'. There are lower values elsewhere, so it is not clear that they 'dominate' in the SH.

Figure 3: Please add vertical lines between the columns (i.e. models). At the boundary between the models it is not so clear which bar belongs to which model.

Page 6, line 1: It would be fair to add Monteil et al, JGR, 2013 here, since they were among the first to report a latitudinal bias.

TECHNICAL CORRECTIONS

page 2, line 4: 'transport' i.o. 'transports'

page 2, line 19: 'increase' i.o. 'incrase'

page 4, line 11: 'CH_4' i.o. 'CO_2'

page 4, line 11: 'applied to' i.o. 'applied from'

Page 7, line 2: 'except over' i.o. 'except for over'

Figure 4: the dashed zero line is missing in the upper panel

Page 7, line 23: 'show' i.o. 'gives'