

It was my pleasure to review the manuscript. I understand that my comments are coming at a relatively late stage in the review process, which is reflected in the generally high quality of the manuscript at this point. I suspect that many of the initial problems have been worked out by now, and some interesting, additional findings have emerged along the way. In its present state, I would consider this manuscript appropriate for publication with only minor adjustments, as outlined below.

One of the most scientifically interesting findings is the temporal dependence in the differences between the GOSAT and the TROPOMI data (compared through a model), shown in Figure 6. The fact that the largest seasonal features correlate with changes of surface albedo suggests directions for the future improvement of (TROPOMI) retrievals algorithms. I think it would be worth highlighting this finding in the abstract as well.

Like the last reviewer, I am also puzzled by how small the error bars on the posterior flux estimates are. The authors mention that these uncertainties are “optimistic”, as they neglect systematic uncertainties due to the inversion setup. This is not ideal, and should be improved in the future, but does not invalidate the focus of the GOSAT-TROPOMI comparison presented here.

**Minor comments that need to be addressed:**

L326: The WLG site show(s) a relatively low posterior what? (Also, it's not the site itself that is unable to capture variability, as this sentence implies.) How can the simulations simultaneously show “relatively good agreement” and “relatively low posterior (correlation)” at WLG? Relative to what? Is it really “day-to-day” variability that is poorly captured in the case of flask data? I guess the sampling is roughly weekly, although the bizarre x-axis in S9 makes this impossible to say for sure (see comment about the supplement below).

Figure 5 is lacking a colour bar!

L506-507: Based on the error bars, only CSC and NEC show consistent (and not just qualitatively similar) posterior emissions. This is important. Either the uncertainties in the posterior emissions estimates are grossly underestimated (likely true), or the datasets are fundamentally inconsistent with one another (also likely true).

Figure S9: Something very strange is going on with the x-axes. It looks like the observations are spaced equally, and the ticks for the months are adjusted to fit. This happens in such a way that there are always measurements on the first of each month, which is suspicious. This is not a normal way of presenting timeseries data, and should be fixed. Are there really only 23 FTIR measurements from HF for the whole year?

**Technical/stylistic comments/typos:**

L30: Suggested rephrasing: The methane emissions inferred from GOSAT observations are ... higher than those from TROPOMI observations...

L31: These -> The

L38: I think the “Ganges Plain” is usually referred to as the “Indo-Gangetic Plain”. I would recommend changing it throughout.

L67: a number is missing regarding the TROPOMI footprint.

L79: A bit confused by the use of “variable” here. Is it required?

L94: Is it already clear before this study that the two sets of satellite data differ in their “regional accuracy”? If so, a citation is needed! Otherwise, remove this.

L97, but also in abstract and throughout the study: I guess North(ern) India should be capitalized throughout? As is every region of China... This is certainly how Wikipedia does it:

[https://en.wikipedia.org/wiki/North\\_India](https://en.wikipedia.org/wiki/North_India)

L114: are -> is

L126: annual average XCH<sub>4</sub> on the 0.625°× 0.5° grid for GOSAT and TROPOMI -> XCH<sub>4</sub> measured by GOSAT and TROPOMI, annually averaged on the 0.625°× 0.5° grid.

L127: over THE Mongolian...

L129: of multiple measurements fall in -> when multiple measurements fall within

L149: Were the in situ data temporally filtered? Generally nocturnal measurements are not well represented by models.

L163: located distant -> far

L182: more discussed -> discussed further

L183: early -> earlier (or previous)

L190-191: WetCHARTs is an ensemble product, which specific ensemble member was used? Is there any concern about double-counting rice emissions as wetlands?

L204: either “a biased boundary condition” or “biased boundary conditions”

L264: overestimate -> overestimates

L296: estimate -> estimates

L303: applying A traditional regularization

L431: are in India -> either “that are in India” or “in India”

L431: subscript in XCO<sub>2</sub>

I would really like to see a figure of the 600 spatial clusters included in the supplement.

L540: benefit -> improve

Caption of Figure S10, last line: the bias correction factor greater than -> the bias correction factor is greater than