Comment on acp-2022-508
Anonymous Referee #1

The paper by Liang et al. compares regional inverse estimates of methane (CH\(_4\)) surface fluxes in East-Asia for the year 2019. The driver data are GOSAT and TROPOMI satellite observations of the column-average mole fractions XCH\(_4\). Liang et al. describe the methodology based on the GEOS-CHEM transport model and a regularized inversion. They compare inversions of GOSAT and TROPOMI XCH\(_4\) and find good agreement for some, substantial discrepancies for other regions. Comparisons to independent data sets serve as guidance to explain the regional differences.

Scope: I am a bit puzzled of what the overall goal of the study is. Is it a budget report of East Asian methane emissions or is it an evaluation of GOSAT and TROPOMI biases? The former would require more complete error analyses, more than a year of data, and more extensive discussions of previous work. For the latter, I would argue that the manuscript lacks completeness in terms of discussing error sources (see comment below). I recommend making the overall goal of the study clearer and revising the paper in the view of that goal.

Proxy-CH\(_4\): Generally, the main (and, I believe, conceptually limiting) error source of the proxy method (GOSAT) must be discussed more thoroughly. It is the errors of the CO\(_2\) fields that are used to construct XCH\(_4\) from the raw CH\(_4\)/CO\(_2\) ratio. Any (e.g. regionally correlated) errors in the prescribed CO\(_2\) fields (typically taken from models) will map into respective errors in XCH\(_4\). In fact, others [Schepers et al., JGR, 2012, https://doi.org/10.1029/2012JD017549] have compared proxy and full physics methods in the early days of the GOSAT mission. They found that, in a case study for India, erroneous CarbonTracker CO\(_2\) fields caused biases in proxy XCH\(_4\) data [Fig. 9 and 10 and related discussion in Schepers et al.]. The paper must examine and discuss this source of error to balance the discussion of scattering induced errors of the full-physics method (TROPOMI). To the best of my knowledge, the current version of the UoL proxy algorithm uses a model ensemble for CO\(_2\)-rescaling. One could try to estimate the error by looking at the spread of these (and potentially other) models in the investigated regions.
**Setup of the inverse problem:** I wonder about the setup of the inverse problem. If I get it right, the parameter vector contains 600 spatial elements which represent spatially distributed annual surface fluxes. I find this a mismatch of spatial and temporal scales. While the inversion is free to optimize a lot of spatial detail, any sub-annual temporal variability of fluxes is imposed. Given further, that the measurement vector contains daily XCH4 data, I would argue that the temporal resolution of the inversion is at odds. The authors should discuss this aspect and provide sensitivity studies showing that their choice does not induce biases (e.g. by imposed seasonality).

Further, the authors have chosen to represent the prior covariance in relative terms (50%) with respect to the prior. This choice imposes that the spatial structure of posterior fluxes will be very similar to the one of the prior fluxes (simply because changing a small flux by 50% (or likewise) remains a small flux). This is clearly visible when comparing Fig. 2 and 4 (even though the log-scale in Fig. 2 needs some defiant eyeballing). The authors should clearly state the consequences of this assumption.

**Inverse method:** Equation 1 is the cost function of the inverse method. It is the classic regularization setup with a prior mismatch and a least squares measurement term where one term is scaled by a regularization parameter which the authors determine according to Figure S3. If I understand correctly, the condition on the selected regularization parameter is that the scaled least-squares term and the prior term impose equal cost. Why would one set such a condition when aiming at evaluating the information content of different data sets? In my understanding, this particular condition implies that whatever your measurement data are (be it dense or sparse, accurate or not), you force the inversion to deliver roughly the same degrees of freedom (for a given prior constraint). Figure 7 appears to confirm this conclusion: while GOSAT and TROPOMI have vastly different data density, the information content of the inversion is roughly the same. In consequence, the presented findings on degrees of freedom would not in any way represent the “natural” information content of the data but they are driven by design of the inverse method. Generally, I would think that an L-curve method should work better for getting a regularization parameter that actually represents the information content of the data [see the cover (or chapter 4.6) of the book by Per Christian Hansen cited in the manuscript].

**Discussion:** The posterior error bars of the satellite inversions (e.g. line 226f) are very small. I assume that they only represent the propagated measurement errors according to equation (4) (and line 185f) and that model transport errors, representativeness errors, more systematic measurement errors are neglected. When comparing the satellite-derived emissions to other studies (line 230ff), the reported error bars should be representative of the full error budget.