This is a review of the manuscript titled "A multi-year global methane data set obtained by merging observations from TROPOMI and IASI" submitted to ESSD by Shahzadi et al. The authors use the algorithms proposed by Schneider et al. to produce a daily, global, long-term methane dataset from TROPOMI and IASI satellite retrievals. Overall, this work continues their previous studies, fits the scope of the journal, and is timely because methane has recently risen sharply and exerts a strong warming influence on the climate. Therefore, I recommend publication in ESSD after the authors address several issues related to key technical validation details and the presentation of results.

We thank the referee for their interest in our work and the effort spent in reviewing the manuscript. In the following, we reply to each of their comments and write our text in red fonts.

Main comments:

1. Although the authors have assessed the dataset accuracy using TM5 outputs and EDGAR emissions, ground measurements were not used for a more independent validation. I suggest a qualified comparison of the dataset with in situ/ground based results (e.g., TCCON).

We agree that the validation of new data products is very important. However, please note that the validation of the method and the data product has been shown in great detail in Schneider et al. (2022b). Therein, we compare the data to independent references of TCCON, Aircore, and GAW. We will try to make it even clearer in the text that Schneider et al. (2022b) show a detailed validation study of the data product.

Moreover, data inter-comparisons are beyond the scope of a regular ESSD article. The aims and scopes of a regular ESSD article are (according to https://www.earth-system-science-data.net/about/aims_and_scope.html): "Articles in the data section may pertain to the planning, instrumentation, and execution of experiments or collection of data. Any interpretation of data is outside the scope of regular articles. Articles on methods describe nontrivial statistical and other methods employed (e.g., to filter, normalize, or convert raw data to primary published data) as well as nontrivial instrumentation or operational methods. Any comparison to other methods is beyond the scope of regular articles."

2. The manuscript states that IASI is "adjusted to TROPOMI a priori (TM5)", but it is not explicit whether this is (a) a re retrieval with a different a priori, (b) an additive/multiplicative post processing correction, or (c) a linearization/offset mapping. The precise operation matters because it affects bias propagation and degrees of freedom.

We do not perform additional retrievals, which would be computationally extremely expensive. The strength of the method is that it works with already retrieved L2 data products, which makes it computationally cheap, i.e., ideal for very large data sets. Moreover, by working with the L2 products, the method automatically benefits from the retrieval developments made by the dedicated TROPOMI and IASI retrieval experts.

We adjust the IASI a priori to the TROPOMI a priori according to Eq. (B13) of Schneider et al. (2022b), which is also well explained in Rodgers and Connor (2003, https://doi.org/10.1029/2002JD002299, therein Eq. 10), i.e., it is an additive post-processing correction:

$\hat{x}' = \hat{x} + (A - I)(x_a - x_a')$.

Here $\hat{x}$ and $\hat{x}'$ are the original and the adjusted/modified retrieval states, accordingly, A and I are the averaging kernel and the identity matrix, respectively, and x_a and x_a' are the original and the modified a priori states, respectively. We will clarify this in the text.

3. The Kalman update uses instrument posterior covariances as input. It is unclear how those covariances were pretreated (regularized, truncated, or inflated), whether they include forward-model uncertainty, and how numerical inversion (or pseudo-inversion) is handled for large matrices.

For combining the IASI and TROPOMI L2 data, we use the IASI product as the background (and use the IASI retrieval a posteriori covariance as the background covariance) and then update this background with the additional information provided by the TROPOMI observation. The IASI a posteriori background is calculated according to Eq. (A8) of Schneider et al. (2022b):

$S_{\hat{x}} = (I - A)S_a$,

i.e., it depends on the a priori covariences (S_a) and the averaging kernels (A). We calculate $S_a = R^{-1}$ from the constraint matrix used for the IASI retrieval (see Sect. 4.6 of Schneider et al., 2022a):

$R = (\alpha_0 L_0)^T(\alpha_0 L_0) + (\alpha_1 L_1)^T(\alpha_1 L_1) + (\alpha_2 L_2)^T(\alpha_2 L_2)$,

where L_0, L_1, and L_2 are the Tikhonov constraint matrices (Tikhonov 1963, https://zbmath.org/?q=an:0141.11001) and α_0, α_1, and α_2 are diagonal matrices that allow for a vertically dependent strength of the Tikhonov constraints. The inversion $S_a = R^{-1}$ is calculated via the Woodbury identity (Woodbury 1950, https://mathscinet.ams.org/mathscinet/article?mr=38136), i.e., there is an algebraic solution for this inversion.

Please note that our data combination method needs no other matrix inversions: we update the IASI profile product by adding a total column observation (one value only), and for calculating the respective Kalman gain, we only have to invert a scalar (see Eq. (4) of the manuscript). We will make clear in the text that the method needs no numerical matrix inversion (or pseudo-inversion).

4. Displacement errors are crucial for tropospheric products and are only described by reference to prior work. The current manuscript lacks the exact formulas, parameter choices (time/space scales), and tests showing sensitivity to those choices.

We will add the equation for the dislocation kernel (Eq. E3 in Schneider et al., 2022b), then the set of equations will be complete.

Please note that we carefully estimated the dislocation error using CAMS data (Copernicus Atmospheric Monitoring Service data). All this work is documented in detail in Appendix E of Schneider et al. (2022b), including figures showing how the atmospheric CH_4 fields vary in space and time. This detailed explanation of the dislocation error is just one click away, and we think it is sufficient to give a clear reference to this Appendix. In order to avoid any doubt on how the dislocation errors are estimated/calculated, we will revise the manuscript and try to further improve the referencing to Schneider et al. (2022b).

5. The manuscript notes lower DOFS and increased noise in high-altitude areas, but does not quantify the implications for users (e.g., systematic smoothing, negative bias, or overconfidence).

A large noise error means that the data product can be subject to significant uncertainty, and a low DOFS value means that the data product sticks to the a priori information, i.e., data with a low DOFS value provides limited information on top of the a priori data.

For a low surface pressure (high surface elevation), the layer below 450 hPa is rather shallow, and the satellite data contains too little information to be sensitive to it (low DOFS) and to detect it with good precision (high noise error). In order to avoid the shallow layers for high surface elevation, we will revise the data set. For the revision, we will separate the atmosphere 50%/50%, i.e., the lower 50% of the atmosphere will be the tropospheric layer, and the upper 50% of the atmosphere will be the upper tropospheric/stratospheric layer. This will remove the "artificial" dependency of the troXCH_4 DOFS and errors on the surface elevation.

The data set contains detailed information about their uncertainty (noise, dislocation) and their characteristics (averaging kernel, DOFS) for each individual observation. We recommend using the averaging kernel as the observation operators whenever the data are used in flux inversion studies, and also consider uncertainties of the observations. This is the best way to ensure that the information provided by the observation can be correctly interpreted by the inversion system. We will briefly mention this in Section 7 of the manuscript and call this section "Data quality and data reusage recommendation".

Moreover, please note that Table 2 summarizes our filter recommendations for high data quality.

Minor / Technical issues

1. Add a detailed NetCDF variable list (names, units, dimensions), an explanation of averaging kernel format/shape, and a short example (Python/xarray snippet) showing how to read XCH4, troXCH4, the averaging kernel, uncertainty, and the quality bitmask. Provide a recommended QC selection for typical use.

We carefully write all the variables in a NetCDF file with standardized metadata (1.7 CF conform). All information about the variables (names, units, dimensions, etc.) is available in these metadata and can be easily accessed by standard NetCDF viewing tools (e.g., Panoply). The user can download the data and easily get information about the data structure by looking at the metadata. We think that an additional table with this metadata information is not needed.

2. The selection of matching windows (50 km / 6 h, then normalized distance minimization)

should be justified quantitatively. If possible, add sensitivity tests for different spatial/time windows (e.g., 25 km / 3 h, 75 km / 8 h) to show tradeoffs between coverage and displacement error.

The dislocation between TROPOMI and IASI in space and time is described in large detail in Appendix E1 of Schneider et al. (2022b). Figure E1 reveals that allowing for a dislocation in time of only 25 km or up to 100 km instead of 50 km will approximately halve/double the spatial displacement error. Concerning the dislocation in time, a 6-hour displacement causes an error that can already be compared to the error of a 100 km spatial displacement (compare Figs. E1a and b of Schneider et al., 2022b). However, the 6-hour displacement is needed to have sufficient matches in the southern hemisphere. The used matching windows of 50 km and 6 hours are a reasonable compromise: they ensure (1) that the dislocation errors are clearly smaller than the noise error, and (2) that we have sufficient matches at any location.

3. In the Methods section, define each symbol and its dimension immediately below the equation. For Eq. (4), (5), (6), and (9), clarify matrix sizes and transpose conventions, and explicitly state when operations are performed in log space versus linear space (you mention omitting log-transform details for brevity, please explicitly state in Methods whether final arrays are reported on the linear or log scale).

We will work on clearly defining the symbols in order to avoid any doubts. Please note that all reported arrays (variables shown in the Figures, as well as the variables provided in the NetCDF file) are on a linear scale. The logarithmic scale is only used for the IASI data processing. We will make this clearer in the manuscript.

4. For Fig. 10, ensure aerosol/snow masking and other flags are clearly labeled in the panels or legend, not only in the title; please check and apply the same clarity to other figures as needed.

Yes, we will go through all figures and check for clarity.