We sincerely thank Reviewer #1 for the constructive and thoughtful suggestions. We have provided point-by-point responses to the comments in blue and improved our manuscript accordingly. In the revised manuscript with track changes, the changes are highlighted in red.

The major revisions are summarized below:

1) We have refocused the manuscript on the real-data inversion, shortened the OSSE discussion, and clarified its idealized assumptions and limitations. We have also moderated our claims on inversion performance and sectoral attribution, since the OSSE, which assumes a true sectoral distribution matching the prior and excludes transport and background errors, cannot test sectoral separation or quantify the uncertainty of the real-data estimates.

2) We have added covariance-based sectoral averaging-kernel estimates and expanded the discussion of intersectoral correlations, i.e., the negative posterior correlations between sector totals (Appendix B). We discuss how these correlations limit the separation of individual sectors. We also report normalized innovation chi-squared diagnostics and clarify the empirical basis and limitations of the SARI model-data mismatch settings.

3) We have clarified the assimilation cycle, localization settings, and use of modeled wind speed. We have also streamlined the results, reorganized the figures, added full-domain prior-emission maps and error bars, standardized plotted uncertainty ranges to ±1σ, and corrected citations and dataset attribution statements.

The point-by-point responses are provided below:

General Comments

The authors describe a regional top-down estimation of CH_4 fluxes using in-situ tower observations in East Asia. They employ the established tools CTDAS and WRF-Chem, a combination that has been used before. The manuscript presents a performance analysis of the inversion system using a simplified observation system simulation experiment (OSSE) followed by an application to real data. The results obtained by applying the system to real data fit well in the scope of ACP. Combined with a discussion and performance analysis, these results deserve publication. However, the discussion and interpretation of the results misses some important limitations of the inversion setup.

Problems with idealized OSSE

The results for individual sectors are uncertain and likely biased. This aspect deserves more attention in the discussion of the results. Furthermore, the interpretation of the OSSE results (section 3.1) requires more caution in dealing with the highly idealized pseudo-observations. For example, the OSSE does not demonstrate the ability to distinguish sectors. The authors mention limitation on a very general level (line 850–855) but they do not address the concrete implications of the idealized test case. I suggest to select one of the following two options to deal with this problem (see also point 13 below):

1. Put the focus on the inversion with real data. Shorten the description of the OSSE, but consider the limitation of the setup when discussing the OSSE results. In this case, I recommend that the authors critically check their statements on sector discrimination since the ability of the system to distinguish emission sectors is not demonstrated.

2. Replace or extend the OSSE by new simulations with more complex prescribed true states and analyze

how the inversion system can reconstruct such true states (similar to Steiner et al., 2024). This should lead to more robust conclusions on the capabilities of the method and help with the interpretation of the results.

RE: Thank you for pointing out the concerns regarding our OSSE study. We agree that the current OSSE does not demonstrate reliable sectoral discrimination and that the implications for sector-specific results require clearer discussion. We have therefore adopted your first option, refocused the manuscript on the real-data inversion, and shortened Sect. 3.1 to approximately half of its original length.

The revised Sect. 2.3.1 and Sect. 3.1 explain that uniformly doubling the prior preserves its spatial patterns and relative sectoral contributions; therefore, the experiment does not test the recovery of independently varying sectoral emissions. Using the same transport-model configuration and pseudo-observations without added random noise also leaves the robustness to transport errors and measurement noise untested. We therefore interpret the OSSE as a test of emission adjustments and of the sensitivity of the inversion to its settings. Sect. 3.1.2 clarifies how prior uncertainties, prior emission distributions, and observational sensitivity influence sectoral adjustments. The posterior intersectoral correlations (Table B1) and the estimated averaging kernel matrix (Appendix C, Table C1) indicate limited sectoral separation within this experiment.

For the real-data results, we now emphasize that sector-specific estimates remain strongly influenced by the prior sectoral distribution and may therefore be biased: improved concentration agreement alone does not establish correct sectoral attribution. The current OSSE cannot quantify these biases or independently validate the inferred sectoral contributions. We have revised the discussion accordingly. The relevant revisions are shown below.

(Abstract, Page 2, Lines 44–46):

"*Idealized Observing System Simulation Experiments (OSSEs), using a uniform doubling of prior emissions as the synthetic truth, indicate that posterior adjustments are concentrated mainly in the Yangtze River Delta (YRD), South Korea, and northern Japan.*"

(Sect. 2.3.1, Page 14, Lines 320–326):

"*The synthetic truth is prescribed as twice the prior emissions for every optimized flux category and grid cell ($\boldsymbol{X}_{true} = 2\boldsymbol{X}_{prior}$; Wecht et al., 2012; Yu et al., 2021). It therefore preserves the prior spatial patterns and relative sectoral composition. No random noise is added to the pseudo-observations. The same WRF-Chem configuration and meteorological forcing are used to generate the pseudo-observations and perform the inversion, so this OSSE does not test the effects of transport-model mismatch. It evaluates recovery of the prescribed scaling and sensitivity to inversion settings, but cannot assess sectoral separation under a different true sectoral composition, or the errors of the real-data inversion.*"

(Sect. 3.1, Page 21, Lines 439–443):
"*Using eight OSSE cases, we examine the spatial distribution of posterior emission adjustments (Sect. 3.1.1), compare the responses of individual flux categories within the YRD (Sect. 3.1.2), and evaluate the influence of selected inversion settings (Sect. 3.1.3). Given the idealized setup, the sectoral analysis does not demonstrate that the inversion can fully separate emissions from individual sectors (Sect. 3.1.2). Its results provide methodological context for the subsequent real-data analysis of the YRD (Sect. 3.2).*"

(Sect. 3.1.2, Page 23–24, Lines 489–515):

"The predefined categories show different posterior responses, with YRD-integrated posterior-to-prior ratios of 2.2, 1.6, 1.5, 1.4, 1.1, and 1.0 for natural emissions, rice cultivation, waste, fuel exploitation, livestock, and "Other", respectively (Fig. 7b–i). These differences reflect the combined effects of prior weighting and sector-specific observation-flux sensitivity. The posterior increments balance the fit to the pseudo-observations against the prior penalty. When observations have comparable sensitivities to multiple categories, sectors with larger absolute prior variances can accommodate larger increments at a lower prior cost (Bruch et al., 2025). This mechanism is consistent with the larger adjustment of natural emissions. However, fuel exploitation shows a smaller posterior adjustment than rice cultivation and waste despite its larger relative prior uncertainty, suggesting that prior flux magnitudes and spatial patterns also shape the posterior response.

Because the prior sector perturbations are independent by construction, their true intersectoral correlations are zero. The weak period-mean intersectoral correlations in Table B1 ($|C^b|$<0.10) are consistent with finite-ensemble sampling noise. In the posterior ensemble, stronger negative correlations between rice cultivation and natural emissions (-0.38) and between rice cultivation and waste (-0.33) indicate compensating variations in the estimated sectoral emissions, indicating that these sectors cannot be constrained fully independently. The estimated averaging kernel (AK) matrix provides complementary information on these sectoral responses (Appendix C; Table C1). Natural emissions have the largest diagonal element (0.39), indicating the strongest response within the same sector among the six categories, consistent with their relatively large posterior adjustment. The off-diagonal averaging-kernel elements show substantial coupling between natural emissions and both rice and waste. The natural-to-rice and natural-to-waste sensitivities are 0.27 and 0.19, respectively, while the corresponding reverse sensitivities are 0.19 and 0.21, indicating that these sectoral signals are not fully separable. The livestock and "Other sectors show weaker self-sensitivity, with diagonal elements of 0.068 and 0.093, respectively. The coupling between sectors is asymmetric, as natural emissions exert little influence on the livestock and "Other estimates (0.014 and 0.007), whereas "Other emissions have a considerably stronger influence on the natural-sector estimate (0.25). Overall, these diagnostics indicate limited separation of the six emission sectors by the observing network."

(Sect. 4, Page 40, Lines 726–729):
"*Sensitivity to prior uncertainty and emission aggregation further shows that these improvements depend on the experimental setup. The uniform perturbation experiment assesses regional emission adjustments but does not independently validate sectoral attribution."*

(Sect. 4, Page 41, Lines 755–769):

"Several limitations should be considered when interpreting these results. The OSSE uses uniformly doubled prior emissions and noise-free pseudo-observations generated with the same transport configuration used in the inversion. Therefore, it evaluates emission adjustments under idealized conditions, while independent sectoral recovery and sensitivity to transport and background errors require further testing. Its short duration and unexplored interactions among inversion settings (e.g., prior uncertainty and the MDM) also limit the generalizability of the sensitivity results. Sectoral estimates remain influenced by the prior sectoral emission magnitudes and spatial distributions, and by intersectoral coupling. The extent to which the sectoral responses in these idealized OSSEs carry over to real-data inversions remains uncertain. In the real-data inversion, the chi-squared innovation diagnostics indicate incomplete consistency between the residuals and the R-matrix; neither the

empirical SARI MDM settings nor the network-wide incorrect application of the 10 m s^{-1} wind-speed cutoff has been fully evaluated. Further evaluation of CTDAS-WRF should therefore examine independently perturbed sectoral priors and truths, background optimization, longer analysis periods, and transport error. Such tests, potentially complemented by satellite observations, would provide a stronger basis for interpreting the real-data inversions."

Writing

Overall, the writing is clear but often lengthy. I value the precise description of the model (section 2), but I found the description of the results in section 3 very lengthy. I recommend to shorten the text by avoiding redundancies and removing purely descriptive text in the discussion of figures or tables (see specific comments below).

Specific comments

1. Lines 108–110: *"Steiner et al. (2024) report regional discrepancies between inversions constrained by satellite observations and those constrained by different surface measurement networks.* 'Steiner et al. present an inversion based on surface observations and provide a qualitative comparison to other inversion results from different models and inversion systems, including satellite-based inversions. But here it seems more suitable to cite works that directly compare inversion results based on surface and satellite observations, e.g.: doi:10.5194/acp-26-10423-2026, doi:10.5194/egusphere-2026-2165 or doi:10.5194/acp-14-577-2014 (this list is incomplete).

RE: Thank you for this suggestion. We agree that Steiner et al. (2024) did not directly compare inversions constrained by satellite and surface observations. We have therefore replaced this citation (Cressot et al., 2014; Lu et al., 2022; Sicsik-Paré et al., 2026; Sun et al., 2026; Ward et al., 2026) and revised the sentence accordingly.

(Page 4, Lines 108–113):

"Direct comparisons of methane inversions constrained by satellite and in situ observations show that the two observing systems provide complementary information but can yield different posterior emission estimates across regions because of their contrasting coverage and sensitivity (Cressot et al., 2014; Lu et al., 2022; Sicsik-Paré et al., 2026; Sun et al., 2026; Ward et al., 2026). Such comparisons are therefore important for assessing the robustness of regional inversions and interpreting inconsistencies under different observational constraints. "

2. Lines 121–123: The authors claim that Lagrangian approaches "provide limited constraint on larger regions such as the YRD". Considering the large variety of Lagrangian methods and the application of such methods to regional modeling, this sounds like a strong and general statement. I recommend that the authors either provide a reference that explicitly supports this statement, or they adjust this claim. In Hu et al., 2026 (references as in the paper), I did not find any support for this general statement. Aside from this, I agree with the authors that having a regional Eulerian inversion framework does provides an added value (e.g. due to the simpler extension to satellite observations).

RE: Thank you for pointing this out. We agree that the statement was too general and was not supported by Hu et al. (2026). We have removed the claim that Lagrangian approaches provide limited constraints

on regions such as the YRD. The revised text describes the scope of the previous studies and presents our regional Eulerian framework as a complementary approach for examining emission seasonality, with the potential to incorporate satellite observations.

(Page 4, Lines 120–122):

"Hu et al. (2023, 2026) used tower observations to investigate CH_4 emissions and seasonal variability in the YRD. As a complementary approach, we use a regional Eulerian inversion framework to examine emission seasonality across the YRD."

3. Figure 1: Panels a and b are sufficiently clear, but panels c and d are only explained much later and panel d can lead to confusion. • Panel (c) can either be kept as part of Figure 1 or presented on its own in section 2.2.1. Keeping it in Figure 1 has the disadvantage that the reader will only learn later what is meant by a "tile". • Panel (d) should be presented on its own in section 2.3 since it is only relevant for the definition of one test case. If it is part of a figure that describes the overall inversion system, this can lead to confusion.

RE: Thank you for this suggestion. We have retained Panels (a) and (b) in Figure 1 and moved Panels (c) and (d) to the sections where they are introduced. The tile schematic now accompanies its definition in Sect. 2.2.1, while the land-ocean aggregation is presented with the sensitivity test in Sect. 2.3.1. We have revised the captions and figure references to make clear that the aggregation applies to that test.

(Sect. 2.1.1, Page 6, Lines 165–168):

"Figure 1. Study domain for the regional inversions. (a) WRF domains (parent domain D01 and nested domain D02) and terrain elevation; (b) locations of the stations used in the inversion (Sect. 2.4) and the boundary zones. Elevation data are from Danielson and Gesch (2011)."

(Sect. 2.2.1, Page 11, Lines 263):

"Figure 2: Schematic of the tile-based state vector in the parent domain (D01)."

(Sect. 2.3.1, Page 15, Lines 353–354):

"Figure 4: Region-based aggregation of the state vector over land and ocean used in the sensitivity test."

4. Line 178: When using products from the Copernicus Atmosphere Data Store (ADS), authors are asked to cite also the ADS catalogue entry. See "Citation and attribution" under https://doi.org/10.24381/ed2851d2.

RE: Thank you for pointing this out. We have added citations to the CDS/ADS catalogue entries for all Copernicus products used, including the two ERA5 datasets and the two CAMS datasets, at their first mention in the Methodology. We have also added the required C3S, EDGAR and CAMS attribution statement to the Disclaimer and included the corresponding references in the reference list.

(Page 5, Lines 148–149):

"... ERA5 meteorological fields with 0.25° spatial resolution and hourly temporal resolution (Copernicus Climate Change Service, 2023a, b; Hersbach et al., 2023a, b)."

(Page 7, Lines 178–181):

"Initial and boundary conditions are taken from the CAMS global inversion-optimized greenhouse-gas concentration, assimilated using satellite and ground-based observations, and available from the Copernicus Atmosphere Data Store (Copernicus Atmosphere Monitoring Service, 2020b; Bergamaschi et al., 2013)."

(Page 7, Lines 184–185):

"As summarized in Table 1, we use the EDGAR_2024 (EDGAR, 2024) and CAMS global emission inventories (Copernicus Atmosphere Monitoring Service, 2020a; Granier et al., 2019)."

(Page 7, Lines 188–189):

"EDGAR_2025 (EDGAR, 2025; Crippa et al., 2025) was released later, on 1 October 2025. CAMS-GLOB-ANT v6.2 (CAMS v6.2; Soulie et al., 2024) is the…"

(Page 7, Lines 195–196):

"In the livestock component of the agriculture sector, local peak values in EDGAR_2024 and EDGAR v8.0 (EDGAR, 2023; Crippa et al., 2023)"

(Page 42, Lines 800–807):

"Disclaimer

The results contain modified Copernicus Climate Change Service information 2026 and modified Copernicus Atmosphere Monitoring Service information 2026. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

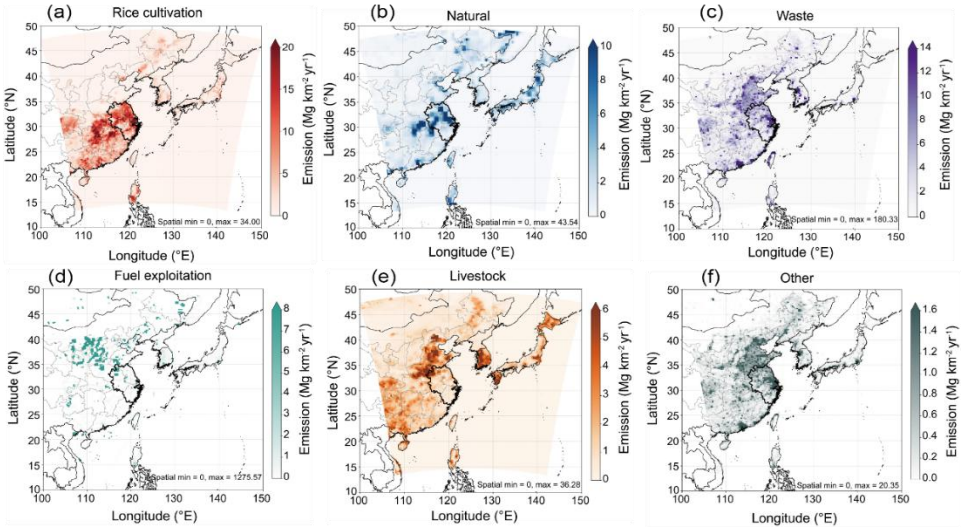
The EDGAR data are © European Union 2023–2025, European Commission, Joint Research Centre (JRC), and are licensed under the Creative Commons Attribution 4.0 International licence. Appropriate credit has been provided, and the modifications made to the ERA5, CAMS, and EDGAR data are described in the Methodology (Sect. 2)."

5. Line 222: Here one could mention that prior fluxes for the YRD region are visualized in Figure S3. In general, I find a visualization of the prior fluxes helpful for interpreting the results. I therefore suggest to include also a figure of the prior fluxes in the whole domain in the supplement.

RE: Thank you for this helpful suggestion. We have added a new figure to the Supplement (Fig. S1) showing the spatial distributions of the annual prior CH_4 emissions for the six source categories across the full model domain. We have also revised the corresponding text.

(Page 8, Lines 225–226):

"*The spatial distributions of the annual prior CH_4 emissions over the full model domain are shown in Fig. S1 and S2."*



"Figure S1: Spatial distributions of annual-averaged prior CH4 emissions from (a) rice cultivation, (b) natural sources, (c) waste, (d) fuel exploitation, (e) livestock, and (f) "Other" sectors over the parent domain (D01) in 2022."

6. Lines 234–244: The description of the data assimilation cycle remained a bit unclear to me, but I think a more detailed explanation could be too long. I therefore suggest to explicitly refer to Peters et al. (2005) for more details on the fixed lag data assimilation cycle.

RE: Thank you for this suggestion. We have reorganized the description of the assimilation cycle to clarify how the two-window configuration advances and how posterior information is retained between cycles. We now explicitly refer to Peters et al. (2005) for details of the general fixed-lag assimilation framework, while retaining the settings specific to this study and removing redundant descriptions

(Page 10–11, Lines 250–262):

"*Following the general fixed-lag assimilation framework described by Peters et al. (2005), each cycle in this study updates the state vector for two consecutive time windows (Fig. 3). In the third cycle, WRF-Chem is run over windows 1 and 2, and the simulated CH_4 mole fractions are compared with the observations assimilated in that cycle to update the state-vector ensemble.*

The next cycle comprises window 2', which retains the posterior flux-scaling-factor ensemble of window 2, and the newly introduced window 3 (Fig. 3). To initialize this cycle, WRF-Chem is rerun over window 1 using its posterior fluxes. The resulting ensemble-mean CH_4 concentration field at the start of window 2' provides the common initial concentration field for all ensemble members.

The newly introduced window (window 3) is initialized with a flux-scaling-factor ensemble centered on $\lambda = 1$. Because the temporal persistence of inventory corrections has not been established in this study, we retain the prescribed emissions for each new period as its prior emissions rather than extrapolating the preceding period's posterior correction. This avoids directly carrying potentially period-specific corrections into the new prior."

7. Lines 291–294: The authors describe a different MDM for different days of the year for station SARI. Without explanation this seems arbitrary. Why was this choice made?

RE: Thank you for catching this. These settings are selected empirically based on preliminary simulations

and sensitivity tests, but the rationale for their selection is indeed insufficiently described in the original manuscript.

The added prior χ^2_{innov} diagnostic (Table S2) shows that, during the three June–July cycles, the SARI values decrease from 1.16–3.70 with 50 ppb to 0.31–1.30 with 100 ppb. It supports reducing SARI's weight during that period, although values below unity in two cycles show that 100 ppb is not established as optimal. The comparison is retrospective and applies to June–July; the MDM schedule and switching dates remain empirical. We have clarified this in Sect. 2.2.2 (Page 13, Lines 299–304).

"SARI observations are assimilated from 9 April to 31 December 2022, with an MDM of 50 ppb during 9 April–27 June and 28 July–5 September, and 100 ppb during 28 June–27 July and 6 September–31 December. These empirical choices reflect larger late-year model-observation mismatches and preliminary sensitivity tests aimed at limiting large negative emission estimates during June–July. Prior normalized innovation chi-squared diagnostics for the June–July tests are provided in Table S2, with the method described in Appendix D."

"Table S2. Prior χ^2_{innov} diagnostics for the 50 and 100 ppb SARI MDM experiments during 28 June–27 July 2022.

10-day cycle (2022)	*SARI MDM (ppb)*	*SARI station prior χ^2_{innov}*	*all station prior χ^2_{innov}*
28 June−7 July	*50*	*1.55*	*2.56*
	100	*0.41*	*2.32*
8 July−17 July	*50*	*3.70*	*2.45*
	100	*1.30*	*1.97*
18 July−27 July	*50*	*1.16*	*2.67*
	100	*0.31*	*2.57*

** The prior χ^2_{innov} diagnostic is described in detail in Appendix D."*

"Appendix D. Diagnostics of prior uncertainty and model-data mismatch

...

The consistency of the innovations with the combined prior prediction uncertainty and prescribed model–data mismatch is assessed using the innovation chi-squared diagnostic of Steiner et al. (2024, Eq. 1). Using the prior observation-space covariance from Eq. (A8), the annual diagnostic is:

The combined prior prediction uncertainty and MDM are evaluated using the normalized innovation chi-squared diagnostic of Steiner et al. (2024). Using the prior observation-space covariance from Eq. (A8), the annual diagnostic is:

$$\chi^2_{innov} = \frac{\sum_c (y_c^o - \bar{y}_c^b)^T (\boldsymbol{P}^b_{yy,c} + \boldsymbol{R}_c)^{-1} (y_c^o - \bar{y}_c^b)}{\sum_c N_{obs,c}}, \quad \text{(D6)}$$

Here, y_c^o contains the assimilated observations, $\bar{y}_c^b$ denotes the corresponding simulated observations for the prior mean state, and $\boldsymbol{P}^b_{yy,c}$ is evaluated before assimilating these observations in cycle c. The MDM covariance $\boldsymbol{R}_c$ follows Appendix A, and $N_{obs,c}$ is the number of assimilated observations in that cycle."

8. Line 305: I do not see how the comparison of a Gaussian and an exponential localization function leads to a factor $\frac{\sqrt{\pi}}{2}$

RE: Thank you for pointing out this ambiguity. The factor $\frac{\sqrt{\pi}}{2}$ arise from comparing the one-dimensional integrals of the two functions (Gaussian localization function and exponential prior error-correlation function:):

$$\rho(d) = exp[-(\frac{d}{L})^2], \text{ and } E_{prior}(r) = exp(-\frac{r}{l}),$$

where *r* is the great-circle distance between the centers of the corresponding spatial regions (tiles; Fig. 2). Their one-dimensional integrals over nonnegative distances are:

$$I_G(r) = \int_0^\infty exp\left[-\left(\frac{d}{L}\right)^2\right] \mathrm{d}d = \frac{\pi}{2}L, \text{ and } I_E(r) = \int_0^\infty exp(-\frac{r}{l})\,\mathrm{d}r = l,$$

We have removed the factor to avoid confusion and clarified the distinct roles of the localization scale and the prior flux-error correlation length. We now describe 900 km as the prescribed localization setting, rather than an optimal multiple of the prior correlation length.

(Page 13, Lines 307–312):

"*To mitigate spurious long-range correlations in the EnKF, we apply covariance localization by multiplying the Kalman gain by a Gaussian localization function (Peters et al., 2005):*

$$\rho(d) = exp[-(\frac{d}{L})^2], \qquad (3)$$

where d is the distance between the state element and the observation, and L = 900 km is the localization length. This parameter controls the attenuation of observation-driven updates, whereas the prior flux-error correlation length of 300 km describes spatial dependence between prior emission errors."

9. Line 318: The prescribed true state is very unlikely in the assumed prior probability density. Is there a motivation for testing such an extreme case?

RE: Thank you for this comment. We chose the uniform twofold perturbation to test where the inversion could correct a large emission bias. Similar prescribed perturbations have been used in methane OSSEs, although with different designs (Wecht et al., 2012; Yu et al., 2021). We agree that this truth is unlikely under the assumed prior distribution. We have therefore clarified that the OSSE is an idealized test of total-emission adjustment, not a demonstration of realistic error recovery or sectoral separation.

(Page 14, Lines 320–326):

"*The synthetic truth is prescribed as twice the prior emissions for every optimized flux category and grid cell ($\boldsymbol{X}_{true} = 2\boldsymbol{X}_{prior}$; Wecht et al., 2012; Yu et al., 2021). It therefore preserves the prior spatial patterns and relative sectoral composition. No random noise is added to the pseudo-observations. The same WRF-Chem configuration and meteorological forcing are used to generate the pseudo-observations and*

perform the inversion, so this OSSE does not test the effects of transport-model mismatch. It evaluates recovery of the prescribed scaling and sensitivity to inversion settings, but cannot assess sectoral separation under a different true sectoral composition, or the errors of the real-data inversion."

10. Line 393: Is this modeled or measured wind speed?

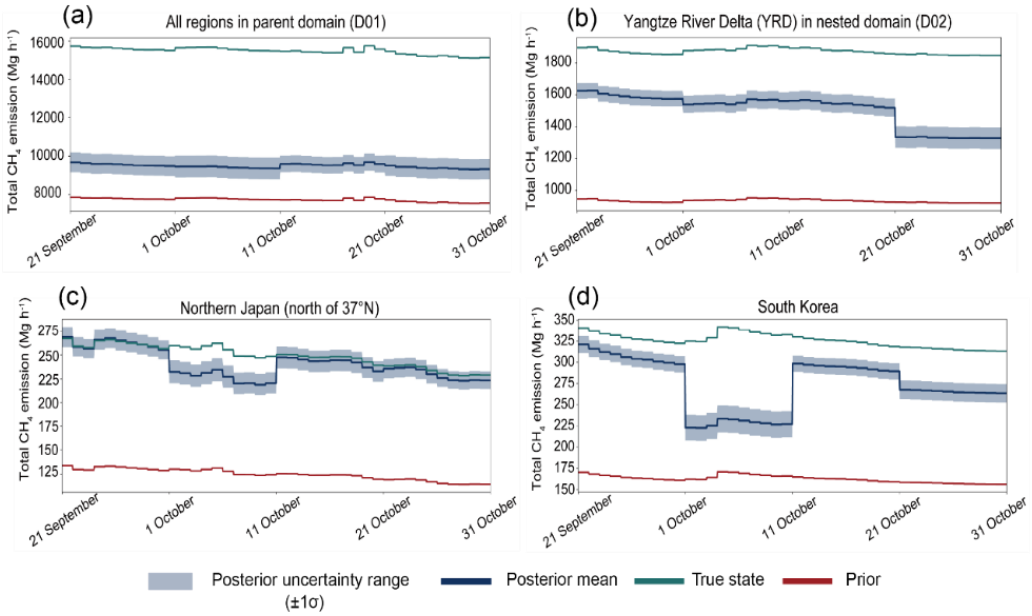
RE: Thank you for pointing this out. We used modeled wind speed. We have clarified this in the revised manuscript.

(Page 16, Lines 385–386):

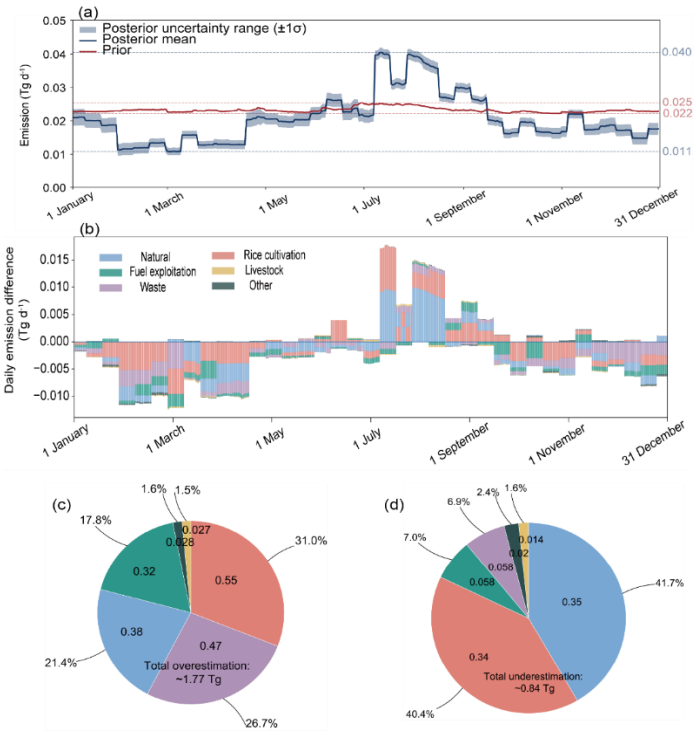
"To reduce local source and sink effects, measurements are also excluded when the modeled wind speed is below 1.5 m s^{-1}."

11. Figure 4 (same for Fig. 10a): The time series show an asymmetric posterior uncertainty range. How is the uncertainty range defined here? As far as I know, the ensemble square-root filter assumes Gaussian uncertainties and should therefore provide symmetric posterior uncertainty estimates. I would usually expect to see one standard deviation (1σ) or the 95% confidence interval (2σ). To avoid ambiguities, the definition of the uncertainty range should be given in the caption.

RE: Thank you for pointing this out. The original shading showed the posterior ensemble minimum–maximum range, which could be asymmetric. In revised Figs. 6 and 11a (formerly Figs. 4 and 10a), it shows the posterior mean ±1σ, calculated from the ensemble at each time step. Both captions now define the range explicitly.



"Figure 6: Temporal variation of hourly CH_4 … The shaded posterior uncertainty range represents ±1 standard deviation (1σ) around the posterior mean."



"Figure 11: Temporal evolution of CH_4 emissions from the prior and posterior estimates … "

12. Lines 457–485 and Figure 5: This paragraph on near-surface concentration fields has little relevance for the rest of the article. Since the article is already quite long, I recommend to make this an appendix or to remove this part.

Thank you for this suggestion. We agree that the discussion of near-surface concentration fields is not essential to the main focus of the manuscript. We have therefore removed the corresponding paragraph and moved Figure 5 (now Fig. S5) to the Supplement.

13. Lines 485–507 and Figure 6: The discussion of the sector separation in the OSSE misses important explanations and conclusions. The OSSE setup is not suitable for validating that sector separation works, but it shows how sector separation fails in a special case:

The true state of the OSSE is far beyond the prior uncertainty range. The inversion system will try to find an emission configuration that fits to the observations while remaining sufficiently close to the prior uncertainty range. It will therefore prefer to adjust sectors with a large absolute prior uncertainty. This explains the strong adjustment in natural fluxes (for which a larger relative uncertainty was assumed). Sectors with small absolute prior uncertainty (livestock, other) will show only small adjustments. The resulting biases affect the sector attribution unless the results are purely data-driven, i.e., unless the observations are sufficient to fully constrain the sector emissions. See Appendix C of doi:10.5194/acp-25-17187-2025 for a more formal explanation.

Thus, the results presented in Figure 6 are to be expected. But the interpretation of the results must take into account that the results are strongly driven by the prior. The results do not show that the system can estimate the natural fluxes. The OSSE only shows that the setup attributes changes in total fluxes to the sector of natural fluxes. To estimate whether the system can distinguish sectors, one should analyze the averaging kernels and conduct further OSSE experiments with different true scaling factors for different sectors. (See the general comments at the beginning of this document.)

In conclusion, the OSSE shows a sector attribution that can be explained by possible biases. It does not demonstrate the correct attribution of sector emissions in a non-trivial setup. Since the mentioned biases remain relevant in the application to real observations, one should be careful when interpreting the posterior sector emissions.

RE: Thank you for this comment. We agree that the original discussion overinterpreted the sectoral adjustments. The current OSSE preserves the prior spatial patterns and sectoral proportions and therefore cannot validate independent sectoral attribution. We have shortened and revised Sect. 3.1.2 and removed statements implying successful sectoral retrieval or strong independent sectoral constraints. Following the reviewer's recommendation, we have also added covariance-based averaging-kernel diagnostics (Appendix C). These diagnostics indicate limited sectoral separation and support the revised interpretation of Figure 5.

We agree that further OSSEs with different sector-specific perturbations are needed to validate sectoral attribution. Such experiments have not been included in this study, and this limitation is now explicitly acknowledged. We have also tempered and shortened the interpretation of sector-specific estimates from the real-data inversion, noting their dependence on the prescribed prior uncertainties (which scale with emission strength) and on intersectoral correlations.

(Sect. 3.1.2, Page 23–24, Lines 489–515):

"The predefined categories show different posterior responses, with YRD-integrated posterior-to-prior ratios of 2.2, 1.6, 1.5, 1.4, 1.1, and 1.0 for natural emissions, rice cultivation, waste, fuel exploitation, livestock, and "Other", respectively (Fig. 7b–i). These differences reflect the combined effects of prior weighting and sector-specific observation-flux sensitivity. The posterior increments balance the fit to the pseudo-observations against the prior penalty. When observations have comparable sensitivities to multiple categories, sectors with larger absolute prior variances can accommodate larger increments at a lower prior cost (Bruch et al., 2025). This mechanism is consistent with the larger adjustment of natural emissions. However, fuel exploitation shows a smaller posterior adjustment than rice cultivation and waste despite its larger relative prior uncertainty, suggesting that prior flux magnitudes and spatial patterns also shape the posterior response.

Because the prior sector perturbations are independent by construction, their true intersectoral correlations are zero. The weak period-mean intersectoral correlations in Table B1 ($|C^b|$<0.10) are consistent with finite-ensemble sampling noise. In the posterior ensemble, stronger negative correlations between rice cultivation and natural emissions (-0.38) and between rice cultivation and waste (-0.33) indicate compensating variations in the estimated sectoral emissions, indicating that these sectors cannot be constrained fully independently. The estimated averaging kernel (AK) matrix provides complementary information on these sectoral responses (Appendix C; Table C1). Natural emissions have the largest diagonal element (0.39), indicating the strongest response within the same sector among the six categories, consistent with their relatively large posterior adjustment. The off-diagonal averaging-kernel elements show substantial coupling between natural emissions and both rice and waste. The natural-to-rice and natural-to-waste sensitivities are 0.27 and 0.19, respectively, while the corresponding reverse sensitivities are 0.19 and 0.21, indicating that these sectoral signals are not fully separable. The livestock and "Other sectors show weaker self-sensitivity, with diagonal elements of 0.068 and 0.093, respectively. The coupling between sectors is asymmetric, as natural emissions exert little influence on the livestock and "Other estimates (0.014 and 0.007), whereas "Other emissions have a considerably stronger

influence on the natural-sector estimate (0.25). Overall, these diagnostics indicate limited separation of the six emission sectors by the observing network."

14. Lines 492–507: This paragraph only repeats what is shown in the figure. Such descriptions should be much shorter (1–2 sentences) and focus only on the most important conclusions that can be drawn from the figure.

RE: Thank you for this suggestion. We have reduced the figure description to one sentence, removing the detailed percentages and sector-by-sector ratios. The remaining discussion focuses on why the sectoral adjustments differ and why they do not demonstrate reliable sector separation.

(Page 23, Lines 489–491):

"The predefined categories show different posterior responses, with YRD-integrated posterior-to-prior ratios of 2.2, 1.6, 1.5, 1.4, 1.1, and 1.0 for natural emissions, rice cultivation, waste, fuel exploitation, livestock, and "Other", respectively (Fig. 7b–i)."

15. Lines 525–565: This lengthy description of Table 4 can be shortened by focusing on the most relevant results. The interpretation could take into account that some results are strongly influenced by the simplified OSSE setup and provide little insights into the performance for real data: A higher prior uncertainty (Case 4) should improve the results because the true state is far from the prior. The region-based aggregation (Case 6) and fewer flux categories (Case 7) implicitly change the prior error structure by changing the state space.

RE: Thank you for this suggestion. We have shortened Sect. 3.1.3 by reducing the description of individual table entries and focusing on the main sensitivities and their interpretation. We now explain that the improvement in Case 4 should be interpreted in the context of the uniformly doubled synthetic truth: larger prior uncertainties permit greater departures from the prior toward this prescribed target. For Cases 6 and 7, we clarify that changing the emission aggregation also changes the state-space dimension and prior error structure, affecting the posterior adjustments. We also clarify that these results are therefore interpreted within the tested OSSE configuration.

(Page 25–26, Lines 523 –538):

"Table 4 summarizes the reduction in total and sectoral flux errors over the parent domain (D01) and the YRD. In the reference case, C_{mae} ranges from 3.0% to 25.0% across the total flux and individual sectors over D01, whereas the total-flux error reduction reaches 55.7% over the YRD.

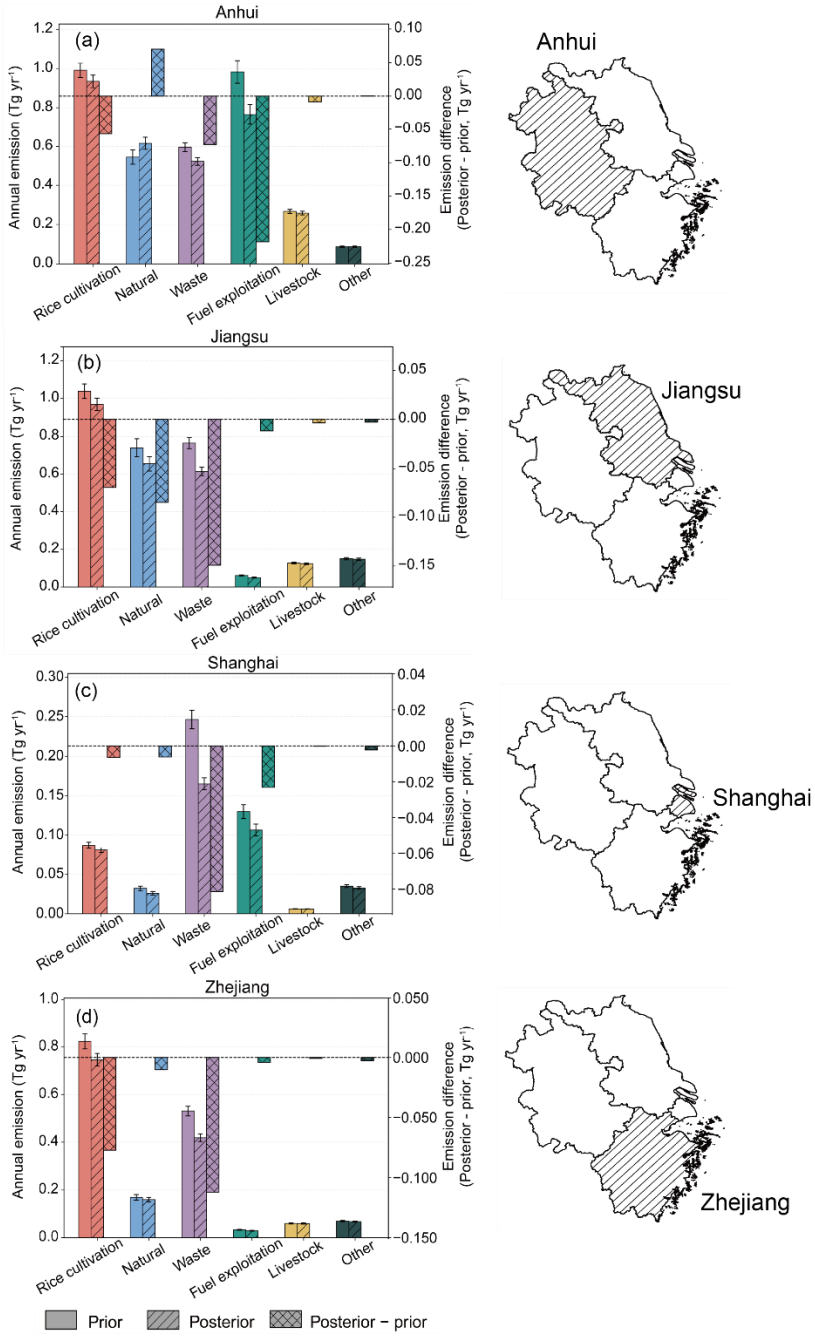
Increasing the ensemble size from 100 to 150 members (Case 2) produces the largest increase in the YRD total-flux C_{mae}, by 6.5 percentage points. Conversely, reducing the localization length from 900 to 600 km (Case 5) decreases the total-flux C_{mae} by 7.1 percentage points over D01 and 3.4 percentage points over the YRD. Shortening the assimilation window (Case 3) slightly decreases total-flux error reduction, whereas adding SARI observations (Case 8) produces a modest increase.

Increasing the prescribed prior uncertainties (Case 4) raises the YRD total-flux C_{mae} to 59.5%. The stronger error reduction is expected because larger prior uncertainties reduce the penalty on departures from the prior, allowing stronger adjustments toward this prescribed target.

Region-based aggregation (Case 6) and the use of fewer flux categories (Case 7) increase the YRD total-flux C_{mae} by 4.7 and 5.7 percentage points, respectively. Both modifications alter the state-vector dimension and the prior error structure, which can influence the magnitude and allocation of posterior emission adjustments. Together with Case 4, these results highlight how prior uncertainty and emission aggregation affect error reduction under the prescribed OSSE conditions. "

16. Figure 8: I recommend to add uncertainties (error bars) to all bars (prior and posterior).

RE: Thank you for this helpful suggestion. We agree and have added error bars to all prior and posterior bars in Fig. 8 (now Fig. 9). The error bars represent ±1 standard deviation (1σ).



"Figure 9: Annual-averaged CH_4 emission … The error bars represent ±1 standard deviation (1σ)."

17. Lines 683–690: See explanation above on biases in the sector attribution.

RE: Thank you for the comment. We have revised this passage to describe the inferred sectoral adjustments more cautiously and acknowledge their dependence on prior assumptions and intersectoral coupling. The added averaging-kernel analysis in Appendix C further informs this interpretation.

(Page 34, Lines 629–642):

"These reductions in ensemble spread are accompanied by changes in intersectoral dependence. Pearson correlations between YRD-integrated sector totals become more negative after assimilation (Appendix B; Table B2), most notably between rice cultivation and natural emissions (-0.29), between rice cultivation and waste (-0.21), and between natural emissions and waste (-0.16). These negative correlations indicate that different combinations of sectoral emissions remain compatible with the observational constraints, so the sectors cannot be fully constrained individually.

The estimated averaging kernel (AK) matrix further characterizes these sectoral responses (Appendix C; Table C2). Its period-mean diagonal elements are largest for waste (0.42), natural emissions (0.32), and rice cultivation (0.30), indicating relatively greater sensitivity of their posterior estimates to changes in the same sectors. The smaller values for livestock (0.080) and "Other (0.097) are consistent with their modest uncertainty reductions. Off-diagonal responses, including those from natural emissions to rice cultivation (0.20) and from rice cultivation to natural emissions (0.16), also indicate coupling between sectors. These diagnostics suggest that reduced emission uncertainty within the YRD coexists with uncertainty in the allocation of emissions among individual sectors."

18. Lines 691–698: The authors describe the posterior error correlation of the sector emissions. This is one indicator that shows the sectoral trade-offs. This paragraph misses a clear conclusion, i.e., that the sectoral trade-offs are relevant and lead to higher uncertainties in emission estimates for individual sectors. For a more detailed analysis of the sector separation, one could analyze the averaging kernels (see, e.g., doi:10.1142/3171 or doi:10.5194/egusphere-2026-2165).

RE: Thank you for this suggestion. We have revised Sect. 3.2.3 to clarify that negative posterior intersectoral correlations indicate compensating adjustments, leaving sectoral attribution uncertain even when ensemble spread decreases. We have also added estimated averaging-kernel diagnostics in Appendix C; their diagonal and off-diagonal responses support a more cautious interpretation of sector-specific estimates. The manuscript revisions are provided in our response to Comment 17; the text added to Appendix C is reproduced below.

"Appendix C. Covariance-based sectoral averaging-kernel estimates in the YRD

A sectoral averaging kernel describes the sensitivity of retrieved sector emissions to changes in true emissions. Diagonal entries represent responses within the same sector, whereas off-diagonal entries represent cross-sector responses. We estimate these responses using the covariance expression of Rodgers (2000) and Bruch et al. (2025b), with aggregation analogous to Sun et al. (2026).

The same calculation is used for the OSSE and the real-data inversion. For cycle c, the original prior at lag 1 of cycle (c -1) is paired with the final posterior at lag 0 of cycle c. They represent the same physical emission window and use identical prior emission weights in Eqs. (B7)–(B8). Synthetic true emissions are not direct inputs to this covariance calculation.

Using the sector totals defined in Appendix B and the ensemble convention in Appendix A, the prior covariance is:

$$[\boldsymbol{P}^{b}_{E,c}]_{pq} = \frac{1}{n-1}\sum_{i=2}^{n}(\boldsymbol{E}^{b}_{c-1,l,p,i} - \boldsymbol{E}^{b}_{c-1,l,p,1}) \times (\boldsymbol{E}^{b}_{c-1,l,q,i} - \boldsymbol{E}^{b}_{c-1,l,q,1}) \quad (l = 1), \qquad (C1)$$

The posterior covariance is

$$[\boldsymbol{P}^{a}_{E,c}]_{pq} = \frac{1}{n-1}\sum_{i=2}^{n}(\boldsymbol{E}^{a}_{c,l,p,i} - \boldsymbol{E}^{a}_{c,l,p,1}) \times (\boldsymbol{E}^{a}_{c,l,q,i} - \boldsymbol{E}^{a}_{c,l,q,1}) \quad (l = 0), \qquad (C2)$$

Here, n = 100, and member 1 represents the mean state. Applying the referenced covariance expression gives:

$$\boldsymbol{A}_c = (\boldsymbol{P}^{b}_{E,c} - \boldsymbol{P}^{a}_{E,c})\,(\boldsymbol{P}^{b}_{E,c})^{-1} = \boldsymbol{I}_6 - \boldsymbol{P}^{a}_{E,c}(\boldsymbol{P}^{b}_{E,c})^{-1}, \qquad (C3)$$

Rows refer to retrieved sectors and columns to sector-total perturbations. Under the linear Bayesian assumptions, the aggregated response interpretation follows the prior covariance structure. For localized ensemble updates, these covariance-based estimates are not assumed to equal exact retrieval derivatives.

Reported means are calculated as:

$$\bar{\boldsymbol{A}} = mean_c(\boldsymbol{A}c), \qquad (C4)"$$

19. Figure 11: For the column observations in SH and HF, the model-observation mismatch shows a bias. Is this bias a physical signal of the emissions? Or is it an artifact of missing averaging kernels or bad model representation of the column? These questions may be difficult to answer, but it they are relevant for the interpretation of the data. I therefore recommend to briefly discuss this aspect.

RE: Thank you for this comment. The original SH comparison omitted AK smoothing, which could bias the comparison. We have revised the procedure and recalculated the columns (new Sect. 2.4.2), using native TCCON retrieval information at SH and locally generated priors, dry-air pressure weights, and transferred HF AKs at SH. The SH comparison remains approximate because its original retrieval information is unavailable.

The revised results show lower posterior MAE and RMSE at both sites, with greater improvement in temporal correlation at HF. We now discuss differences in seasonal coverage and proximity to assimilated stations as possible explanations for the contrasting residual patterns and model adjustments at SH and HF. These interpretations remain tentative because the SH comparison relies on transferred HF kernels and locally generated priors.

(Page 17–18, Line 409–432):

"At SH, retrieval averaging kernels and auxiliary profiles are unavailable. Local CH_4, H_2O, pressure and gravity profiles are generated using GINPUT 1.6.3 (Laughner et al., 2023) and GEOS-FP data (Lucchesi, 2018). The wet priors are converted to dry mole fractions:

$$x_{a,dry,i} = \frac{x_{a,wet,i}}{1 - q_{H_2O,wet,i}}, \qquad (6)$$

$$q_{H_2O,dry,i} = \frac{q_{H_2O,wet,i}}{1 - q_{H_2O,wet,i}}, \quad (7)$$

Following the pressure-integration method of Laughner et al. (2024), normalized dry-air weights are calculated as:

$$h_i = \frac{\Delta p_i / [g_i (M_{air} + q_{H_2O,dry,i} M_{H_2O})]}{\sum_j \Delta p_i / [g_i (M_{air} + q_{H_2O,dry,i} M_{H_2O})]}, \quad (8)$$

Here, j is the summation index over all levels included in the column integration, Δp_i is the pressure integration weight, g_i is gravitational acceleration (m s^{-2}), and M_{air} = 0.0289644 kg mol^{-1} and M_{H_2O} = 0.01801528 kg mol^{-1} are the molar masses of dry air and water, respectively.

For each SH observation, HF averaging kernels from measurements within ±3 days (inclusive) are considered. A kernel is retained only if:

$$\frac{|X_{HF} airmass_{HF} - X_{SH}/cos(SZA_{SH})|}{X_{SH}/cos(SZA_{SH})} \leqslant 0.03$$

Here, X_{HF} and X_{SH} are the observed dry-air column-average CH_4 mole fractions. SH air mass is approximated by $1/cos(SZA_{SH})$. The SH solar zenith angle (SZA_{SH}) is calculated from observation time and location without atmospheric refraction. It should be noted that the ±3-day window and 3% threshold are empirical choices. The window increases the number of available HF kernels. The threshold selects HF observations with slant XCH_4 close to that at SH because averaging kernels vary with solar zenith angle and CH_4 column abundance (Laughner et al., 2024).

Using unscaled local priors and the comparison structure of McGee et al. (2025), the simulated dry-air column was calculated as:

$$X_a = \sum_i h_i \, x_{a,dry,i}, \quad (9)$$

$$XCH_{4,ak} = X_a + \sum_i h_i \, a_i \left(x_{model,dry,i} - x_{a,dry,i}\right), \quad (10)$$

"

(Page 35 Line 659–666):

"The different residual patterns at SH and HF may partly reflect unequal seasonal coverage. Given the dominant waste-source influence at SH (Sect. 3.2), its missing summer observations may omit periods of stronger local enhancements and associated model underestimation. By contrast, HF includes summer measurements that frequently exceed simulated concentrations, giving negative residuals greater weight in its annual distribution. The larger mean prior-to-posterior change at SH may also reflect its co-location with an assimilated surface station, while HF lies farther from the assimilated stations in the northwestern YRD. These explanations remain tentative because the SH comparison relies on transferred HF kernels and locally generated priors rather than its original retrieval averaging kernels and auxiliary profiles."

20. Section 3.3.2: This section contains some repetitive statements that can be removed to shorten the text. To make clear what I mean, the following comments are deliberately more detailed than usual:

• Lines 753–754: *"In the discussion below, we use the ensemble mean value."* This statement seems obvious to me and can be removed.

• Lines 760–761: *"All estimates are reported in* Tg yr−1*."* The units are (and should be) provided together with the numbers, thus this sentence can be removed.

RE: Thank you for these suggestions. We have removed both redundant statements. The units are retained alongside the numerical values throughout the discussion.

• Lines 762–782: These paragraphs only repeat what is shown in Figure 13. I suggest to reduce this to 2–3 sentences that highlight the most important conclusions from Figure 13.

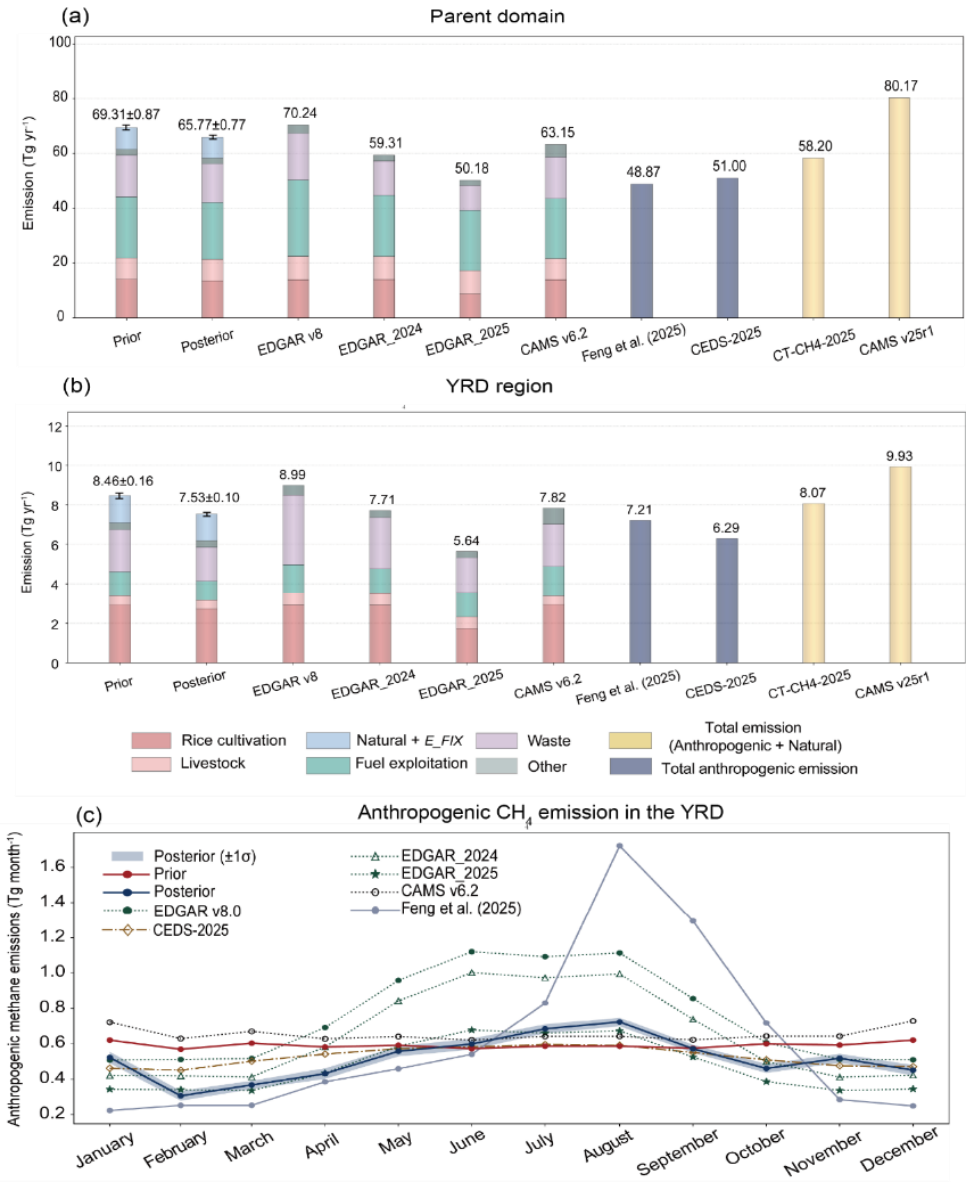
RE: Thank you for this suggestion. We have condensed the two paragraphs, highlighting the regional emission reductions and their main sectoral contributions, the comparison of YRD anthropogenic emissions with existing inventories, and the differences from global inversion estimates. Detailed numerical descriptions that repeat Fig. 13 (now Fig. 14) a and b have been removed.

(Page 36–37, Line 686–692):

"The inversion reduces total CH_4 emissions to 65.77 ± 0.77 Tg yr^{-1} over the parent domain and 7.53 ± 0.10 Tg yr^{-1} in the YRD, representing decreases of 5.1% and 11.0% relative to the prior, respectively, driven mainly by reductions in fuel exploitation and waste emissions (Fig. 14a, b). The posterior anthropogenic estimate in the YRD is 6.18 Tg yr^{-1}, closest to CEDS-2025, differing by only 1.7%. The YRD posterior anthropogenic estimate is also relatively close to EDGAR_2025 but lower than the other anthropogenic estimates by 14.2–31.2%, while posterior total emissions are 6.7% and 24.2% lower than CT-CH4-2025 and CAMS v25r1, respectively."

21. Figure 13: An error bar is provided for the posterior emissions, but not for the prior. I recommend to provide also the prior uncertainty in Figure 13 (a) and (b).

Thank you for this suggestion. We have added error bars representing the prior emission uncertainties to Fig. 13 (now Fig. 14) a and b, allowing direct comparison with the posterior uncertainties. The posterior uncertainty range in panel (c) has also been changed to ±1 standard deviation (±1σ).



"Figure 14: Annual total CH_4 emissions from different inventories … "

Technical and minor issues

22. In the HTML abstract, a ± sign is missing.

Thank you for pointing this out. We have corrected the missing ± sign and revised the formatting to avoid rendering issues in the HTML abstract.

23. line 52: "**an** in situ station"

Thank you for the correction. We have revised the text to "an in situ station."

(Page 2, Lines 51–53):

"the summertime underestimate in the prior. Independent evaluation at an in situ station further supports the posterior improvement… "

24. line 96 and 97: "pixel size" is needed in the first parentheses: "pixel size ~5.5×7 km"

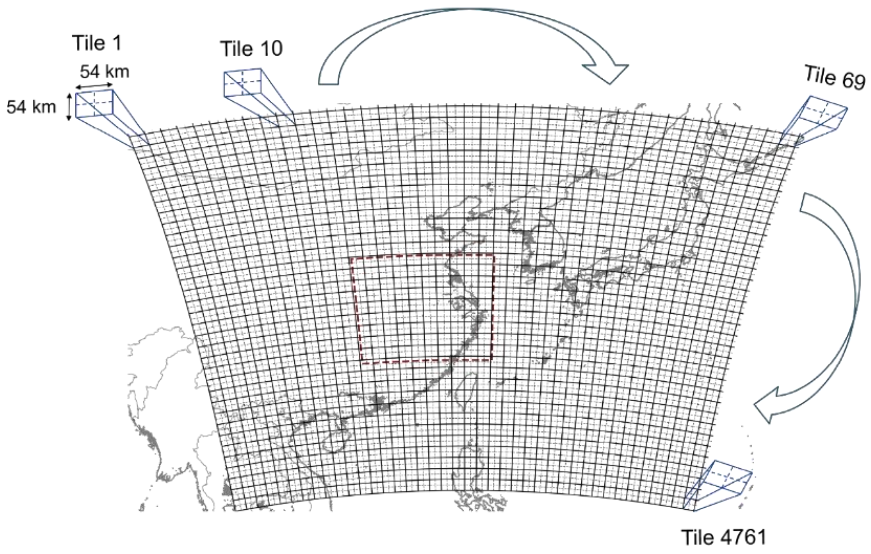
Thank you for the correction. We have removed the passage describing the IMI 2.0 application during revision, so the parenthetical requiring clarification no longer appears in the manuscript.

(Page 3–4, Lines 94–97):

"For example, Estrada et al. (2025) developed the Integrated Methane Inversion version 2.0 (IMI 2.0) and applied it to estimate 2023 CH_4 emissions over the contiguous United States using the Harvard blended TROPOMI+GOSAT product (Balasus et al., 2023), with a native transport-model resolution of approximately 25 × 25 km."

25. Figure 1: The resolution of Figure 1(c) should be improved.

Thank you for this suggestion. We have increased the resolution of the original Fig. 1(c), now presented as Fig. 2, as shown below.



"Figure 2: Schematic of the tile-based state vector in the parent domain (D01).

26. Table 3: In this table it should be mentioned that SH and HF are total column observation sites. For example, this could be indicated in the "Site type" column.

27. Table 3: In the column "Notes" I assume that "2022" means that the full year was considered? Then why not call the column "Time span" and write "full year" instead of "2022"? Moreover, the line break in "7 July" is confusing.

Thank you for these comments. In Table 3, we have identified SH and HF as total-column observation sites in the "Site type" column, renamed "Notes" to "Time span", replaced "2022" with "Full year 2022", and removed the line break in "7 July". The revised table is shown below.

(Page 20):

Table 3: The observation sites used in this study. Site types are defined based on geographic location.

Site ID	Location	Reference	Longitude (° E)	Latitude (° N)	Elevation (m)	Sample altitude (m)	Site type	MDM (ppb)	Purpose	Time span
HAT	Hateruma Island	Tohjima (2025b)	123.81	24.06	10	47	Island	15	Inversion	Full year
YON	Yonagunijima	Takatsuji (2024b)	123.01	24.47	30	50	Island	15		
SARI	Shanghai	Wei et al. (2020)	121.59	31.18	--	~20	Urban	50, 100		9 April–31 December
SORPES	Nanjing	Ding et al. (2016)	118.95	32.12	40	112	Urban	50		1 February–16 April and 7 July–13 November
GSN	Gosan	Yang (2024b)	126.16	33.29	71.4	77.4	Island	25		Full year
AMY	Anmyeon-do	Yang (2024a)	126.33	36.54	42	108	Coastal	40		
ULD	Ulleungdo	Yang (2024c)	130.90	37.48	220.9	230.9	Island	15		
RYO	Ryori	Takatsuji (2024a)	141.82	39.03	260	280	Mountain	15		
COI	Cape Ochiishi	Tohjima (2025a)	145.50	43.16	42.4	93.9	Coastal	15		
DMS	Damingshan		119.00	30.02	1497	1512	Mountain		Validation	
SH	Shanghai		121.59	31.18			Total column observation			
HF	Hefei	Liu et al. (2023)	117.17	31.91						

28. Line 432, grammar: "Because …" should not start a new sentence here but is a continuation of the previous sentence.

Thank you for the correction. We have joined the "because" clause to the preceding sentence and revised the paragraph as follows:

(Page 22, Lines 469–470):

"Taiwan is not included because the posterior adjustments there are weaker and less spatially coherent, despite the proximity of the HAT and YON stations."

29. Line 445: *"Northern Japan shows the best agreement"*: This is redundant, shorten the text here.

Thank you for this suggestion. We have removed the redundant statement and shortened the paragraph as follows:

(Page 22, Line 471–474):

"At the parent-domain scale, the posterior total is 22.5% higher than the prior. Larger relative adjustments occur in the three focus regions, where the posterior totals are 60.8%, 67.9%, and 92.7% above the prior in the YRD, South Korea, and northern Japan, respectively. These adjustments represent different degrees of recovery toward the prescribed 100% emission increase."

30. Lines 572: *"The real-data inversion shows that existing bottom-up inventories are generally biased high […]"*: It is a matter of personal style how scientists present their results. But in general I would suggest to use weaker formulations for results that are subject to substantial uncertainties, e.g. write "indicate" instead of "show". The same applies to lines 589–598.

Thank you for this suggestion. We have revised the wording in the manuscript, using more cautious formulations, such as "indicate" rather than "show", where appropriate to reflect the uncertainties in our results.

(Page 28, Line 545–546):

"The real-data inversion indicates that existing bottom-up inventories are generally biased high and display substantial spatial discrepancies of posterior CH_4 emission (Sect. 3.2.1)."

(Page 38, Line 702–703):

"EDGAR v8 and EDGAR_2024 indicate a similar seasonal tendency but with substantially higher summer emissions,"

Added references:

(Page 55, Line 1044–1049):

Cressot, C., Chevallier, F., Bousquet, P., Crevoisier, C., Dlugokencky, E. J., Fortems-Cheiney, A., Frankenberg, C., Parker, R., Pison, I., Scheepmaker, R. A., Montzka, S. A., Krummel, P. B., Steele, L. P., and Langenfelds, R. L.: On the consistency between global and regional methane emissions inferred from

SCIAMACHY, TANSO-FTS, IASI and surface measurements, Atmos. Chem. Phys., 14, 577–592, https://doi.org/10.5194/acp-14-577-2014, 2014.

(Page 66, Line 1364–1067):
Sicsik-Paré, A., Fortems-Cheiney, A., Pison, I., Broquet, G., Opler, A., Potier, E., Martinez, A., Schneising, O., Buchwitz, M., Maasakkers, J. D., Borsdorff, T., and Berchet, A.: Assessment of the differences in European CH_4 emission estimates from three TROPOMI products, Atmos. Chem. Phys., 26, 10423–10454, https://doi.org/10.5194/acp-26-10423-2026, 2026.

(Page 67, Line 1394–1399):

Sun, S., Zhang, Y., Chen, W., Pan, S., Pu, D., Tong, X., Chen, H., and Houweling, S.: Assessment of methane emissions from European countries using satellite and surface observations, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2026-2165, 2026.

(Page 70, Line 1464–1466):

Ward, R. H., Western, L. M., Tunnicliffe, R. L., Fillola, E., Tsuruta, A., Aalto, T., and Ganesan, A. L.: Uncertainty and retrieval sensitivity in TROPOMI-based methane inversions over the North Slope of Alaska, Atmos. Meas. Tech., 19, 813–837, https://doi.org/10.5194/amt-19-813-2026, 2026.

(Page 54–55, Line 1032–1043):

"Copernicus Atmosphere Monitoring Service (CAMS): CAMS global emission inventories, Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store (ADS) [data set], https://doi.org/10.24381/1d158bec, last access: 23 September 2026, 2020a.

Copernicus Atmosphere Monitoring Service (CAMS): CAMS global inversion-optimised greenhouse gas fluxes and concentrations, Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store (ADS) [data set], https://doi.org/10.24381/ed2851d2, last access: 23 September 2026, 2020b.

Copernicus Climate Change Service (C3S): ERA5 hourly data on pressure levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.bd0915c6, last access: 23 September 2026, 2023a.

Copernicus Climate Change Service (C3S): ERA5 hourly data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.adbb2d47, last access: 23 September 2026, 2023b.

(Page 56, Line 1077–1091):

EDGAR (Emissions Database for Global Atmospheric Research): Community GHG Database (a collaboration between the European Commission, Joint Research Centre (JRC), and the International Energy Agency (IEA), comprising IEA-EDGAR CO2, EDGAR CH4, EDGAR N2O, and EDGAR F-GASES), version 8.0, European Commission, JRC [data set], https://edgar.jrc.ec.europa.eu/dataset_ghg80, last access: 23 September 2026, 2023.

EDGAR (Emissions Database for Global Atmospheric Research): Community GHG Database (a collaboration between the European Commission, Joint Research Centre (JRC), and the International Energy Agency (IEA), comprising IEA-EDGAR CO2, EDGAR CH4, EDGAR N2O, and EDGAR F-GASES), version EDGAR_2024_GHG, European Commission, JRC [data set],

https://edgar.jrc.ec.europa.eu/dataset_ghg2024 (last access: 23 September 2026), 2024

EDGAR (Emissions Database for Global Atmospheric Research): Community GHG Database (a collaboration between the European Commission, Joint Research Centre (JRC), and the International Energy Agency (IEA), comprising IEA-EDGAR CO2, EDGAR CH4, EDGAR N2O, and EDGAR F-GASES), version EDGAR_2025_GHG, European Commission, JRC [data set], https://edgar.jrc.ec.europa.eu/dataset_ghg2025, last access: 23 September 2026, 2025.