Reviewer #1

We thank Reviewer #1 for the thorough assessment and helpful, constructive comments, which have improved the manuscript. Below, our responses appear in italics, and new or modified text in the manuscript is highlighted in green.

General comments

1. The authors argued that COVID lockdown in 2020 could affect the methane emissions, but they show comparisons of the posterior results from prior studies and national UNFCCC reports which were mostly not in the year of 2020. Given that the focus of this manuscript is not investigating COVID lockdown effects, why was the abnormal year of 2020 selected as the target year?

Yes, that is a fair question. There were indeed a number of other reasons for choosing the year 2020, motivated by several methodological and practical considerations:

One main reason for selecting 2020 was our focus on the baseline optimization, following the approaches of Thompson et al. (2025) and Subramanian et al. (2025). Both studies used the 2020 EGG4 initial mole-fraction fields but adopted substantially different uncertainties for the baseline scaling factors. Using the same target year and initial fields allowed us to investigate those different assumptions more directly, without introducing additional differences.

A second consideration was the availability of national emission inventories for comparison. With the exception of Greece, all countries in our domain are non-Annex I Parties, which report their national inventories to the UNFCCC only irregularly, so the choice of target year also determines how many reported estimates are available for comparison. In this respect, 2020 proved to be a rather favourable year: reported emissions were available for six countries, (and for most of the remaining countries, at least reports for neighbouring years exist). This provided a relatively good basis for comparing our inversion results with national inventory estimates across much of the domain, although, of course, differences in reporting years need to be considered when making these comparisons.

Furthermore, 2020 was the most recent year covered by the GFEI inventory available at the time of the study, providing prior information on fossil fuel emissions.

The fact that 2020 is the year of the COVID-19 pandemic made the year also interesting for examining potential changes in methane emissions during this period. However, the year was selected primarily for the methodological and data-availability reasons outlined above, rather than specifically to isolate the effects of lockdown measures. We have clarified this rationale in the revised manuscript.

We added:

We selected 2020 as the target year primarily to facilitate comparison with Thompson et al. (2025) and Subramanian et al. (2025) using the same initial mole-fraction fields. The choice was also supported by the availability of GFEI prior emissions and national inventory estimates for comparison. Since 2020 coincided with the COVID-19 pandemic, our results also provide an opportunity to examine potential changes in methane emissions during this period.

2. As the top-down method is meant to correct our current understanding of methane emissions from process-based bottom-up inventories, we should choose the most up-to-date emission inventories, which is EDGAR 2025 version with GFEI replacement. I suggest the authors switch the reference inversion from EDGAR v8.0 to EDGAR 2025 with GFEI.

Yes, we agree that using the best available prior estimate is desirable (also from a Bayesian perspective) and that the most up-to-date inventory allows a more direct comparison with current bottom-up emission estimates.

Our study explicitly investigates the sensitivity of the inversion results to the choice of prior emissions by considering four different prior configurations, including both EDGAR v8.0 and EDGAR 2025, with and without the GFEI replacement. We compare the (national) results across all four configurations, while our final results are not based on a single reference inversion with a specific prior. Changing the reference prior would therefore not change our final emission results.

However, we would prefer to retain our reference configuration for two illustrative reasons. First, one important outcome of our study is that the downward corrections to EDGAR v8.0 emissions around the Persian Gulf obtained from the inversions are comparable to the updates introduced in the EDGAR_2025 inventory (We think showing smaller adjustments compared to EDGAR_2025 would be less illustrative). Second, we would like to show, that these downward corrections are robust across the different inversion configurations, demonstrating that the satellite observations can identify and correct substantial biases in the prior inventory. We consider the current reference configuration useful for illustrating these findings, while our final results are not affected by the choice of a single prior. However, as we agree with the reviewer that the difference to the most updated inventory is important, we further illustrate and discuss the spatial differences relative to EDGAR 2025 with GFEI emission inventory, and have included the respective posterior emissions and inversion increments in the Supplement.

We added to the main text:

In contrast, when using the EDGAR_2025_GHG + GFEI inventory as prior information (see Fig S8), the increments around the Persian Gulf are substantially smaller, suggesting better agreement between the prior emissions and the observations in this region.

and to the supplements:

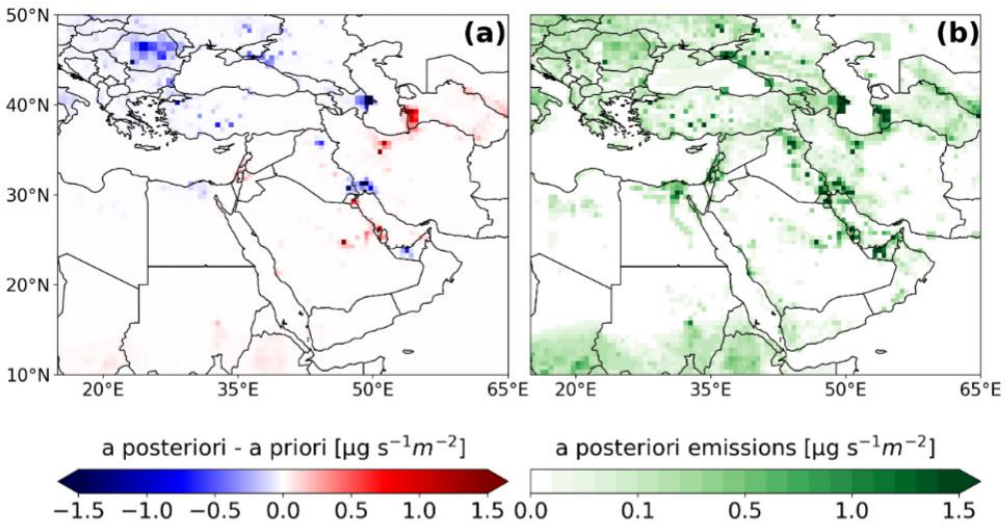


Figure S8. Inversion results, when using the EDGAR_2025_GHG + GFEI as prior inventory . (a) Emission increments from the inversion (a posteriori–a priori emissions), (b) a posteriori emission distribution

Figure S8 shows (a) the inversion increments and (b) the a posteriori emissions obtained using the EDGAR_2025_GHG + GFEI inventory as prior information. Several spatial patterns in the increments are consistent with those in the reference inversion, including negative increments over Romania, Turkey, and regions northwest of the Black Sea, as well as positive increments over Iran and Turkmenistan. However, the adjustments around the Persian Gulf are substantially smaller, suggesting that the prior emissions in this region are more consistent with the observations. Similarly, the increments over Egypt are considerably smaller than in the reference inversion, resulting in similar a posteriori emissions in this region.

3. How is the boundary conditions optimized for the regional inversion over the EMME region?

In our regional inversion, the baseline contribution, which accounts for the initial atmospheric mole fractions and emission contributions from outside the EMME domain, is optimized jointly with the surface emissions as part of the inversion state vector: We use the CAMS EGG4 global greenhouse gas reanalysis to provide the initial three-dimensional mole-fraction fields. FLEXPART is run in backward mode for 20 days to calculate the sensitivity of each observation to the initial mole fractions at the trajectory endpoints. The contribution from the initial mole fractions is calculated by combining these sensitivities with the corresponding EGG4 mole fractions. (so the coupling is in time, rather than in space). This so-called "diffusive backward Lagrangian reconstruction" provides a local approximation to the advection-diffusion equation (e.g. Legras et al., 2003, 2005). Contributions from emissions occurring outside the EMME inversion domain during the 20-day simulation period are then also included in the baseline. To optimize the baseline, we include scaling factors for the initial mole-fraction contributions in the inversion state vector. These factors are optimized jointly with the emission offsets and are defined for four latitudinal bands and three vertical layers, allowing the baseline to be adjusted according to the observations. We have clarified this procedure in the manuscript.

The baseline[6] is optimized using scaling factors, which adjust the contribution from the initial mole fraction fields aggregated over coarse grid cells. For the initial fields we use the CAMS global greenhouse gas reanalysis EGG4 (Copernicus Atmosphere Monitoring Service, 2021; Agustí-Panareda et al., 2023). The scaling factors are defined for four latitudinal bands (90°–30°S, 30°S–0°, 0°–30°N, 30°–90°N) and three vertical layers (0–2 km, 2-10 km, and 10–50 km above ground level).

[6]The baseline is calculated by combining the sensitivities of the observations to the initial mole fractions, derived from the endpoints of the 20-day FLEXPART backward simulations, with the initial mole fraction field (see Sect. 2.3). Contributions from emissions occurring outside the EMME inversion domain but during the simulation period are also included.

4. The authors said the state vector elements for initial conditions (baseline) were for latitudinal bands and vertical layers. Would more state vector elements for initial conditions surrounding the inversion domain better represent the baseline optimization for the inversion results over the inversion domain of the EMME region?

Yes, increasing the number of baseline scaling factors could allow spatial variations in biases in the initial mole-fraction fields to be better represented. However, this would also increase the number of degrees of freedom and could make the baseline optimization less well constrained by the available observations. Our sensitivity tests indicate that the optimized baseline contributions remain already sensitive to the assumed baseline uncertainties, suggesting that the observational constraint on these parameters is limited even with the current coarse grid cells. Introducing additional scaling factors for smaller spatial regions could therefore increase the risk of overfitting and lead to less stable solutions. (The particles are widely dispersed after 20 days backwards simulation so one looses the ability to resolve much finer grid cells). Conversely, using fewer scaling factors would provide a stronger constraint on each parameter but would limit the ability to represent spatially varying biases. A potential alternative would be to introduce spatial correlations between the baseline scaling factors, allowing spatial variations while regularizing the optimization. We consider the current parameterization a reasonable choice, although further investigation of alternative sizes of coarse grid cell and spatial correlation structures would be interesting.

We added:

The choice of the spatial resolution for the baseline scaling factors may also influence the behaviour, although the sensitivity to this choice was not investigated in the present study.

5. There are various spatial resolutions applied for different parameters used in the inversion, including 0.5 degree for the sensitivities H and emission fields, 0.25 to

0.5 degree for the super observations and wind fields, variable resolution for the state vector elements. Why is this, instead of consistent spatial resolutions for emissions, sensitivities, super observations, and wind fields?

The different spatial resolutions serve distinct purposes in the inversion system and can (or should) be selected according to the specific requirements and limitations. Using a common spatial resolution for all components is not advantageous; instead, the resolutions should be chosen to balance computational efficiency, spatial detail, and the available observational information. An inappropriate choice may lead to excessively high computational costs, poorly constrained emission estimates, spatially noisy fields, or a loss of relevant spatial information through large aggregation.

The number of state vector elements strongly influences the computational cost of the inversion. A too fine resolution could also lead to poorly constrained emission estimates and spatially noisy posterior fields, particularly in regions with insufficient observational coverage. We therefore use a variable spatial resolution, aggregating grid cells according to their influence on the observations, retaining finer resolution where the observational information supports a more detailed representation of emission patterns. The resolution of the sensitivities is chosen to balance spatial detail against the computational and storage costs of the LPDM simulations and subsequent inversion. These costs are particularly relevant for satellite-based inversions, which involve large numbers of observations and correspondingly large amounts of data. In addition, for a given number of released particles, a coarser LPDM output grid further from the observations means that it is likely that a greater number of particles are counted in each grid cell, which decreases noise in the source-receptor relationships.

On the observation side, aggregating retrievals into super observations reduces the computational burden and helps to reduce random errors. Using a variable resolution of super-observations enables us to retain relevant spatial information at a higher resolution. In the Lagrangian framework, the particles are released to exactly represent the super-observations and their resolution.

We drive FLEXPART by ERA5 fields at 0.25° in the nested domain, which is close to the native horizontal resolution. Outside the nested domain, 0.5° fields are used, because here finer meteorological detail has less influence on the simulated concentrations.

6. It is said state vector elements are with temporal resolution of 30 days. How is the optimization for each month updated? Would the emissions be reset to the EGG4 initial fields or inherit from the posterior emissions from last month?

The inversion is not performed sequentially month by month. Instead, all state vector elements, i.e. all spatial elements for all 30-day periods are optimized simultaneously in a single inversion. Therefore, there is no reset to initial fields and no posterior estimate is

propagated. Each observation is linked through its FLEXPART sensitivities to emissions over the preceding 20 days, which is handled consistently because all periods are solved for jointly. Temporal correlations between periods are accounted for in the prior error covariance matrix.

We clarified in the manuscript:

The state vector is composed of two key components: offsets to the a priori emissions and initial mole fraction scalars, which are jointly optimized across the whole year in a single inversion.

7. How would methane sink biases affect the inversion result in this study?

Loss from soil uptake is small compared with total methane emissions in the EMME region, and we therefore expect biases in this sink to have a very limited impact on the inferred regional posterior emissions. Atmospheric methane loss is accounted for in FLEXPART through reaction with OH. Biases in the prescribed OH concentrations could generally affect the simulated sensitivities and, consequently, the inferred emissions. However, the backward simulations in our study cover only 20 days. Given the atmospheric lifetime of methane of approximately 10 years, only a small fraction of methane is expected to be lost over this period, suggesting that the impact of OH-related biases is likely to be limited. The effects prior to the 20-day simulation period are not explicitly resolved by these backward simulations and are instead reflected in the prescribed baseline. This provides an advantage over models that explicitly represent atmospheric transport and chemical loss over longer timescales, where OH loss can have a substantial influence on the inferred methane budget (e.g., Zhao et al., 2020)

We added:

Atmospheric methane loss is accounted for in FLEXPART through reaction with OH, although the amount of methane lost during the 20-day backward simulation period is expected to be limited given its atmospheric lifetime.

8. Would dust storms prevalent in the EMME region affect the retrieval success rate and accuracy, and potentially affect the seasonality of methane emissions?

Yes, we believe dust storms could affect the availability of satellite methane retrievals. Aerosol scattering and absorption can alter the optical path of reflected sunlight and can introduce errors in the retrieved column-averaged methane concentrations, while we believe quality filtering likely reduces the number of usable observations during dust events (also documented in e.g.,Schneising et al., 2026). Because dust activity can vary seasonally, these effects could potentially influence the seasonal constraints on inferred methane emissions, however we expect cloud cover to have a larger impact. We have not explicitly quantified the impact of dust storms on retrieval availability or inversion results in this study.

We added:

The observational coverage also varies substantially with season, reflecting variations in observation conditions (e.g. cloud cover, aerosol loading), which may affect the seasonal constraints on the inferred emissions.

And extended to

While the overall decrease is consistent with reduced oil activity during the COVID-19 pandemic, we believe that the apparent seasonal signal should be interpreted with caution, as it may also partly reflect potential methodological biases and variations in observational coverage.

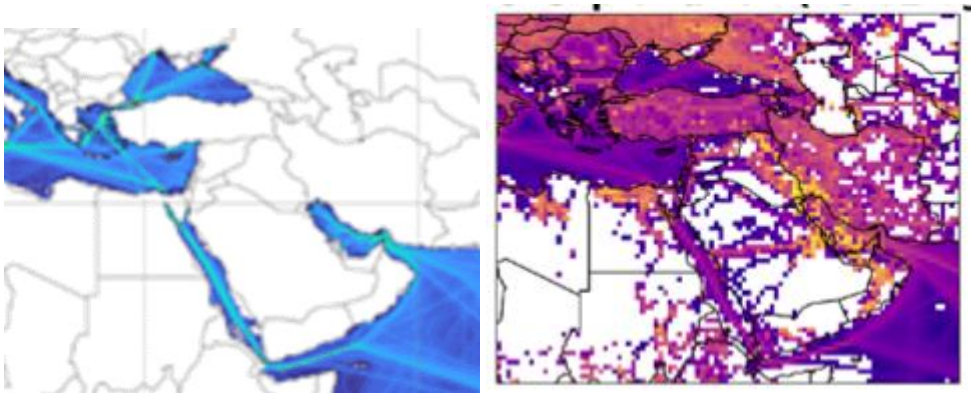
and

While, together, these results suggest that reduced oil activity during 2020 could have led to a measurable decrease in regional CH_4 emissions, the seasonal variations in monthly emissions are more uncertain and should be interpreted with caution, as they may be affected by residual biases in the initial mole fraction fields or the varying observational coverage.

Specific comments

1. Figure 2: The distribution of fuel exploitation from EDGAR v8 exhibits emissions over ocean with suspicious shipping tracks, which were absent from GFEI shown in Figure S4. Are there any shipping emissions included in EDGAR v8? Also, what would be the spatial distribution comparison between the most up-to-date EDGAR 2025 version compared to GFEI inventory?

Yes, there are shipping emissions included in EDGAR v8. Crippa, et al. (2024) states: "Since EDGAR v6.0, international shipping emissions have been distributed using the Ship Traffic Emission Assessment Model (STEAM3) from the Finnish Meteorological Institute (Jalkanen et al., 2012; Johansson et al., 2017) and this approach has remained unchanged in EDGAR v8.0", where also maps of International shipping GHG emissions are shown (Figure 7), that match with the emissions over ocean that are shown in our Figure (see comparison here)



In Table S1 of the supplements, shipping tanker are explicitly included in the Fuel exploitation sector:

Fuel exploitation	Coal mining: brown and hard coal	In-house EDGAR proxy based on EPRTR (http://prtr.ec.europa.eu) and USGS (https://www.usgs.gov/) and Global Energy Observatory (http://globalenergyobservatory.org/)
	gas flaring	In-house EDGAR proxy based on https://www.ngdc.noaa.gov/eog/viirs.html
	Gas pipelines transmission	In-house EDGAR proxy
	oil pipelines	In-house EDGAR proxy
	oil terminals	In-house EDGAR proxy based on World Port Index (PUB 150) (http://msi.nga.mil/MSISiteContent/StaticFiles/NAV_PUBS/WPI/Pub150bk.pdf)
	shipping tankers	In-house EDGAR proxy based on LRIT and Wang et al.

We compared the spatial distribution of the date EDGAR_2025_GHG and GFEI inventory and added:

Figure S5 further shows the comparison of CH_4 emissions from fuel exploitation in the EDGAR_2025_GHG and GFEI inventories for 2020. Both inventories show broadly similar spatial patterns but can differ substantially in individual countries, such as Saudi Arabia, Iraq, or the United Arab Emirates.

and

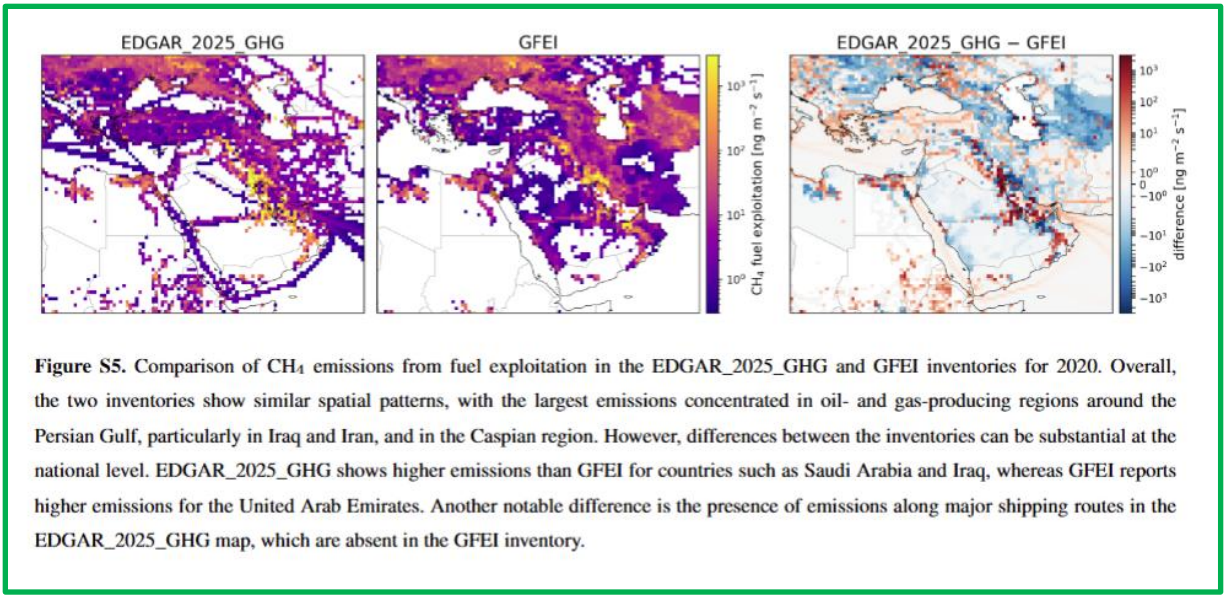


Figure S5. Comparison of CH_4 emissions from fuel exploitation in the EDGAR_2025_GHG and GFEI inventories for 2020. Overall, the two inventories show similar spatial patterns, with the largest emissions concentrated in oil- and gas-producing regions around the Persian Gulf, particularly in Iraq and Iran, and in the Caspian region. However, differences between the inventories can be substantial at the national level. EDGAR_2025_GHG shows higher emissions than GFEI for countries such as Saudi Arabia and Iraq, whereas GFEI reports higher emissions for the United Arab Emirates. Another notable difference is the presence of emissions along major shipping routes in the EDGAR_2025_GHG map, which are absent in the GFEI inventory.

2. Line 158-160: Why is the temporal correlation of 30 to 180 days selected with the absence of temporal correlation days less than 30 days? What is the temporal resolution of prior emission inventories?

The temporal resolution of the prior inventory is monthly (and the inversion state vector is resolved at monthly time steps), while the prior month-to-month variability is relatively small. In FLEXINVERT, the temporal error correlation between two time steps is described by an exponential decay function. For a monthly state vector, correlation lengths substantially shorter than 30 days would therefore result in weak correlations between adjacent months. Such a choice would therefore be almost equivalent to assuming no temporal correlation, leaving each month's emissions largely constrained individually. This would increase noise and could introduce spurious month-to-month fluctuations in the posterior emissions. We therefore considered temporal correlation lengths between 30 and 180 days. The lower limit corresponds to the monthly temporal resolution of the prior information and state vector, while the upper limit imposes strong temporal smoothing, approaching a seasonal constraint.

We added:

We test temporal correlation lengths of 30, 90, and 180 days, similar to Bergamaschi et al. (2022), who tested timescales of 1 and 12 months. While other studies also used smaller temporal correlations (e.g. Tsuruta et al., 2019; Montenegro et al., 2025), correlation lengths substantially shorter than 30 days would be almost equivalent to assuming no temporal correlation for a monthly state vector, leaving each month's emissions largely constrained individually. This would increase noise and could introduce spurious month-to-month fluctuations in the posterior emissions. We therefore use 30 days as the lower limit, consistent with the monthly resolution of the prior information and state vector, while the longer timescales represent progressively stronger temporal smoothing

and

Based on these results, we select a temporal correlation length of 30 days as the reference value, as it preserves the observed month-to-month variability, including the August emission peak, while avoiding the stronger temporal smoothing associated with longer correlation lengths. This choice is also better suited to represent the temporal variability of fossil-fuel emissions and is consistent with the temporal length used by Thompson et al. (2025) and Subramanian et al. (2025).

and

In our inversion, we optimize monthly gridded a priori CH4 fluxes.....

3. Line 220-222: It is said emission offsets (absolute difference?) and baseline scaling factor (ratio?) are optimized in the inversion. Why are different optimization targets (absolute difference/ratio) chosen between emissions and baseline?

The different formulations reflect the different nature of the two state-vector components. The fluxes are optimised as additive offsets from the prior because their magnitudes vary substantially across the domain, including cells with near-zero prior emissions. This formulation allows the inversion to add emissions where the prior is zero or nearly zero and allows the prior error covariances to be defined in flux units. The background concentration is represented by a few scaling factors. Scaling the global-model background preserves its spatial and temporal structure while adjusting its overall magnitude. This is consistent with the aim of correcting potential biases in the background level and expressing its uncertainty in relative terms.

4. For section 2.5: I suggest adding a summary line for the parameter setting of the reference inversion alongside the sensitivity test so that it would be easier for readers to know the comparison reference.

Thank you for this suggestion, we agree! Rather than adding a separate summary line, we have highlighted the reference configuration in bold throughout the list of sensitivity tests to make it easier for readers to identify the comparison reference.

(We added: The reference configuration is indicated in bold in the following list.)

- baseline uncertainties of 0.1, 0.3, 0.5, 0.7, **1.0**, 1.3, 1.5, 2.0, 2.5, 3.5, 5, and 7%.
- four different a priori emission data sets: (i) **EDGAR v8.0**, (ii) EDGAR_2025_GHG, (iii) EDGAR v8.0 + GFEI, (iv) EDGAR_2025_GHG + GFEI (see Sec. 2.2)
- a priori emission errors of 30, 50, **100**, 150, and 200%.
- spatial error correlation lengths of 50, **100**, 175, 250 and 500 km.
- temporal error correlation lengths of **30**, 90 and 180 days.
- a lognormal a priori error distribution in addition to the **normal distribution**.

5. Figure 4: The seasonality between the baseline and the observations is reversed. Would this propagate to the enhanced seasonality with summer maximum in the posterior emissions?

Yes, this is an important question. The "reversed" seasonality is present in the prior baseline, which exhibits a summer minimum (which is not present in the observations). After optimization, the baseline follows a seasonal cycle more consistent with the observations, suggesting that the baseline optimization largely absorbs this "reversed" seasonality of the prior baseline rather than transferring it to the inferred emissions. To investigate this further, we performed sensitivity inversions with different prior baseline uncertainties (Fig. S13). When the baseline uncertainty is small, the baseline is more tightly constrained to its prior values, and a larger fraction of the seasonal mismatch is attributed to emissions, resulting in a strong seasonal cycle and a very larger summer peak. As the baseline uncertainty increases, the peak amplitude decreases and the seasonal cycle becomes flatter. Nevertheless, the August maximum relative to adjacent months persists across all tests, suggesting that it is not solely an artefact of the baseline, although its magnitude depends on the assumed baseline uncertainty. We therefore cannot rule out that residual baseline biases contribute to the inferred emission seasonality. Accordingly, we interpret the seasonal pattern cautiously and discuss this limitation in the Results section.

We extended the discussion to:

The inversion is able to correct this bias, as reflected in the a posteriori baseline, which is shifted toward lower values by approximately 20 ppb during this period, while it follows a seasonal cycle more consistent with the observations.

&

The seasonal signal is particularly sensitive to assumed baseline uncertainties (Fig. S13), making the results less certain, particularly the relatively pronounced decline toward the end of the year when observational coverage also decreases. Consistent with this, sensitivity tests further show that assuming a lognormal error distribution produces a weaker decline in late 2020 compared to a normal distribution (Fig. S27). The magnitude of the August peak is also sensitive to the assumed baseline uncertainty, although the peak persists across all sensitivity tests, suggesting that it is not solely a baseline artefact.

6. Line 409-413: How about a plot for the fuel exploitation emissions versus global oil demand or production?

Yes, thank you for this suggestion. We changed figure S10 accordingly.

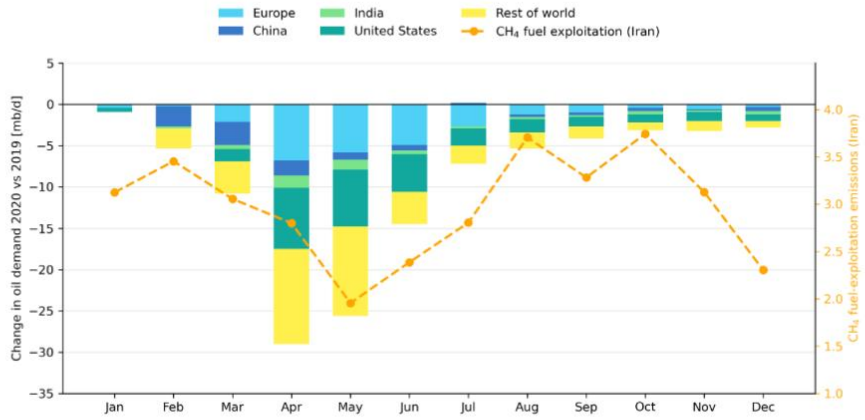


Figure S10. Oil demand in 2020 relative to 2019 (International Energy Agency, 2020), compared with fuel exploitation emissions in Iran derived from our inversion results. To estimate the fuel exploitation component from our total posterior emissions, we apply the sectoral emission shares from the EDGAR_2025_GHG dataset.

7. Line 413-415: the prior emissions in Figure 2 and Figure S3 show minor contributions from agricultural and wetland emissions in the EMME region. It is unlikely that these two sources can explain the summer peak observed.

Yes, we agree and address the comment in the text:

Regional analyses have reported a pronounced seasonal cycle in atmospheric CH_4 over Iran, with a late-summer maximum that may partly reflect enhanced emissions from wetlands and agricultural sources (Mousavi et al., 2024). However, these sources account for a comparatively small fraction of emissions in our prior inventory and are therefore unlikely to fully explain the pronounced summer peak inferred by our inversion. Furthermore, the wetland emission seasonality is complex and may lag behind precipitation and inundation patterns (Liu et al., 2026), complicating any source attribution based solely on seasonal timing.

Technical comments

1. Line 195 and 234: the numbered superscriptions are not referenced anywhere. Do they represent certain citations?

The superscripts refer to footnotes providing additional information and appear at the bottom of the corresponding page.

Literature

Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., Graziosi, F., Muntean, M., Maes, J., Dijkstra, L., Van Damme, M., Clarisse, L., and Coheur, P.: Insights into the spatial distribution of global, national, and subnational greenhouse gas emissions in the Emissions Database for Global Atmospheric Research (EDGAR v8.0), Earth Syst. Sci. Data, 16, 2811–2830, https://doi.org/10.5194/essd-16-2811-2024, 2024.

Legras, B., Joseph, B., and Lefèvre, F.: Vertical diffusivity in the lower stratosphere from Lagrangian back-trajectory reconstructions of ozone profiles, Journal of Geophysical Research: Atmospheres, 108, 4562, https://doi.org/10.1029/2002JD003045, 2003.

Legras, B., Pisso, I., Berthet, G., and Lefèvre, F.: Variability of the Lagrangian turbulent diffusion in the lower stratosphere, Atmospheric Chemistry and Physics, 5, 1605–1622, https://doi.org/10.5194/acp-5-1605-2005, 2005.

Montenegro, N., Berchet, A., Martinez, A., Thanwerdas, J., Bousquet, P., Pison, I., and Saunois, M.: Capability of current observing systems to monitor CH4 emissions from the regional to the global scales, EGUsphere, https://doi.org/10.5194/egusphere-2025-5923, preprint, 2025.

Schneising, O., Bovensmann, H., Buchwitz, M., Buschmann, M., Deutscher, N. M., Griffith, D. W. T., Hachmeister, J., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Ohyama, H., Petri, C., Reuter, M., Robinson, J., Roehl, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Vrekoussis, M., Wang, W., Warneke, T., Weidmann, D., Wunch, D., Zhou, M., and Bösch, H.: TROPOMI/WFMD v2.0: Improved retrievals of XCH_4 and XCO with XGBoost-based quality filtering, Atmos. Meas. Tech., 19, 2407–2435, https://doi.org/10.5194/amt-19-2407-2026, 2026.

Tsuruta, A., Aalto, T., Backman, L., Krol, M. C., Peters, W., Lienert, S., Joos, F., Miller, P. A., Zhang, W., Laurila, T., Hatakka, J., Leskinen, A., Lehtinen, K. E. J., Peltola, O., Vesala, T., Levula, J., Dlugokencky, E., Heimann, M., Kozlova, E., Aurela, M., Lohila, A., Kauhaniemi, M., and Gomez-Pelaez, A. J.: Methane budget estimates in Finland from the CarbonTracker Europe-CH4 data assimilation system, Tellus B: Chemical and Physical Meteorology, 71, 1–20, https://doi.org/10.1080/16000889.2018.1565030, 2019.

Zhao, Y., Saunois, M., Bousquet, P., Lin, X., Berchet, A., Hegglin, M. I., Canadell, J. G., Jackson, R. B., Deushi, M., Jöckel, P., Kinnison, D., Kirner, O., Strode, S., Tilmes, S., Dlugokencky, E. J., and Zheng, B.: On the role of trend and variability in the hydroxyl radical (OH) in the global methane budget, Atmos. Chem. Phys., 20, 13011–13022, https://doi.org/10.5194/acp-20-13011-2020, 2020.

Reviewer #2

We are grateful to Reviewer #2 for the careful evaluation of our work and the valuable suggestions, which have helped us strengthen the manuscript. Below, our responses appear in italics, and new or modified text in the manuscript is highlighted in green.

Lines 65–67: I suggest clarifying the statement that "the computational cost remains independent of the number of observations." As currently written, this seems too general. For adjoint-based inversions, the main computational advantage is that the cost of computing the gradient is largely independent of the number of elements in the control/state vector, rather than of the number of observations. The treatment of observations (e.g. application of the observation operator, calculation of model–observation mismatches, averaging kernels, and data handling) may still scale with the number of observations, which can become relevant for large satellite datasets.

Yes, thank you, we agree! We have revised the statement.

Compared with Lagrangian approaches, these methods can offer computational advantages for inversions involving large numbers of satellite observations, particularly when adjoint techniques are used to efficiently calculate gradients with respect to a large number of control variables.

Line 110: The value of 30 km for the correlation length depends on the characteristics of satellite data retrieval, surface conditions, sources of instrumental error, etc. In this sense, could you briefly explain why this value, taken from a study on NO_2, is applicable to your study?

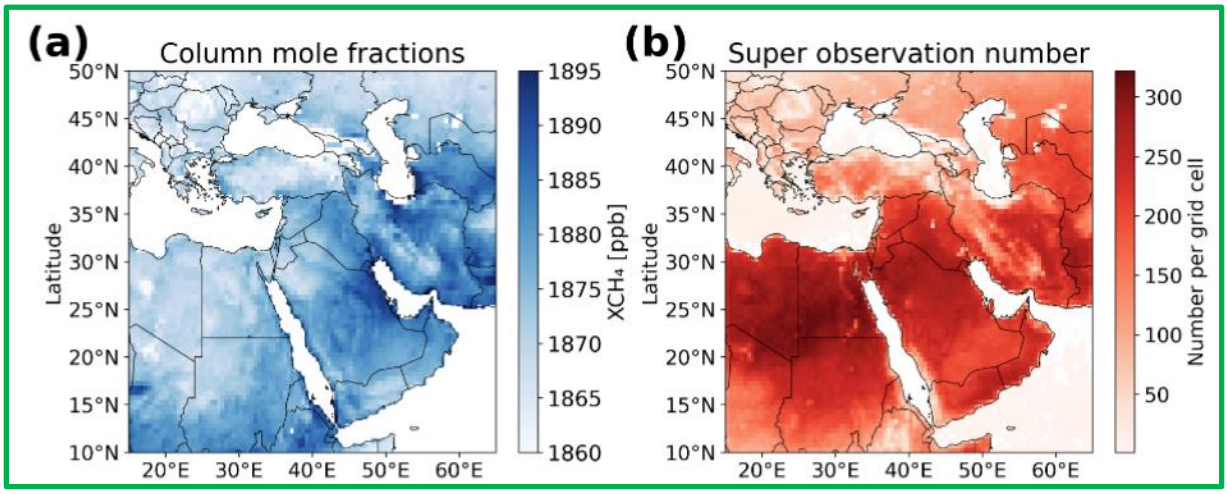
Yes, we agree that the spatial error correlation length depends on the characteristics of the retrieval product. However, to our knowledge, no estimate is available for TROPOMI/WFMD XCH_4, and Rijsdijk et al. (2025) is the only study we identified that estimated spatial error correlations from TROPOMI data. We consider their estimate for NO_2 a reasonable proxy because both products use the same instrument and share aspects of its measurement geometry and calibration. Furthermore, in the EMME region, NO_2 is influenced by oil and gas extraction activities, which are also major sources of CH_4. Thus, some similarities in the spatial distribution and correlation structures of the associated signals might be expected. However, we agree that differences in retrieval algorithms and other error sources might lead to different correlation scales. We have clarified this choice in the revised manuscript.

To account for that, we assume that retrieval errors are correlated within each super-observation cell using a spatial exponential decay function. In the absence of a known spatial error correlation length for TROPOMI/WFMD XCH_4, we use the value of 30 km based on Rijsdijk et al. (2025) for TROPOMI NO_2 as a proxy, as both products use the same

instrument and share aspects of its measurement geometry and calibration. Furthermore, oil and gas activities are important sources of both NO_2 and CH_4 in the EMME region, suggesting some similarities in the spatial patterns and correlation structures of the associated signals.

In Figures 1a and 1b, the addition of political boundaries (country borders) would help the reader to better locate the main features shown in the maps.

Yes – thank you, we agree – done!



Line 158: The choice of spatial and temporal correlations needs to be justified and cited, especially given that each sector has a different spatiotemporal structure.

Yes, agreed. We added and deepened discussions about our choices in the Appendix: Here we wrote:

In the reference inversions, we assume a spatial correlation length of 100 km and a temporal correlation length of 30 days, based on literature values and sensitivity tests (see Sect. C and D), which further explore spatial correlation lengths between 50 and 500 km and temporal correlation lengths between 30 and 180 days

Please see the added text, in response to your comment: Line 596

Line 213: The choice of 20 iterations for the sensitivity experiments and 30 iterations for the final inversions appears to be based mainly on visual inspection of the cost-function evolution. Could the authors provide a quantitative convergence criterion, for example based on the gradient norm.

Yes, we added:

……. where the criteria of a relative cost decrease < 0.005 and relative gradient norm < 0.05 are met.

Lines 250–260: The design of the sensitivity experiments is not entirely clear from the current description. Were these parameters varied one at a time relative to a reference configuration, while all other settings were kept fixed, or were combinations of the different parameter values explored? A full factorial combination would result in a very large number of experiments, so I assume a one-at-a-time sensitivity approach was used. Please clarify the experimental design. If parameters were varied individually, it would also be useful to briefly acknowledge that interactions between uncertain parameters are not explored.

Yes, exactly, we generally used the "one-at-a-time" sensitivity approach. However, we also did additional sensitivity tests in which we varied the baseline uncertainty while using one alternative prior uncertainty and, separately, one alternative correlation length. We clarified that in the text.

Similar to Vojta et al. (2025), we conduct sensitivity tests, where we systematically explore the parameter space by varying one parameter at a time relative to a reference inversion setup, while keeping all other settings fixed. The reference configuration is indicated in bold in the following list:

……..

This approach allows us to assess the individual influence of each parameter, but does not account for potential interactions between uncertain parameters. We, therefore, also conduct additional sensitivity tests in which we varied the baseline uncertainty while using one alternative prior uncertainty and, separately, one alternative correlation length.

Line 275: The sensitivity analysis explores a wide range of uncertainties assigned to the EGG4 baseline, but the inversion still relies on the same baseline product in all experiments. Given the apparent early-2020 bias in EGG4, it would be useful to discuss whether the posterior emissions may also depend on the choice of baseline product itself. If an alternative CH_4 background/initial field is available, a sensitivity experiment using an independent product would provide a complementary test of the robustness of the inferred emissions.

Yes, we agree. We had actually planned to use the CAMS product (https://doi.org/10.24381/ed2851d2) as alternative initial mole fraction field in a sensitivity test. However, this resulted in a substantial low bias in the baseline and the simulated mole fractions, making the product not usable for our purposes (see Fig.1). A colleague is currently working on a bias correction, while we are also planning to generate our own initial mole fraction fields using FLEXPART LCM for comparison. For the present study, however, we needed to retain the EGG4 product.

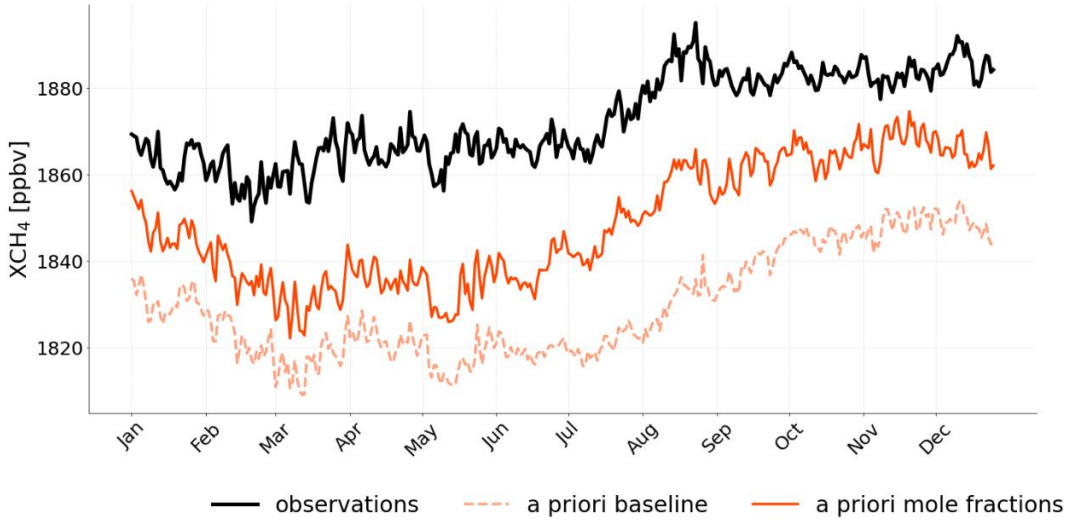


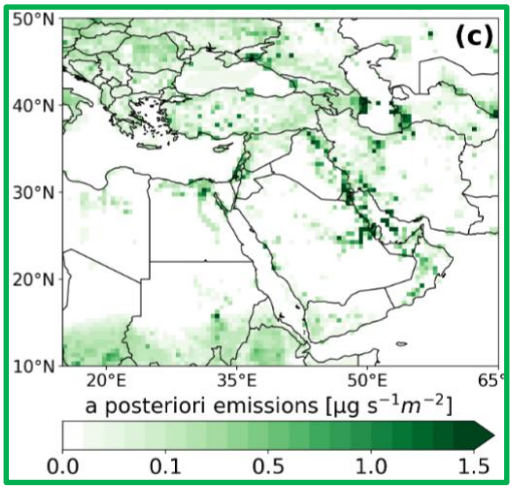
Figure 1: Daily averages of observed column mole fractions (solid black line), a priori modeled column mole fractions (solid red line), and a priori baseline (dashed red line), averaged over the study domain, when using the CAMS product as initial mole fraction field.

279: Replace (b) the relative error reduction by (b) the relative uncertainty reduction

Yes, thank you. Done!

Figure 5: The spatial distribution of posterior emissions is somewhat difficult to interpret because most grid cells appear close to the lower end of the color scale, while only a few emission hotspots are clearly visible. The authors may consider adjusting the color-scale normalization or using units such as mg m^{-2} d^{-1} to improve readability. It would also be helpful to include a figure showing the relative percentage change in the increase.

Yes, we agree. For this region, we found it particularly tricky to make small emissions visible while still conveying the differences in emission magnitudes. After testing several options, we have now adopted a different colour-scale normalization to achieve a better balance.



Line 302-308: Given the large number of sensitivity tests performed, I understand that organizing the corresponding results and discussion is challenging. However, Section 3.3 would be easier to follow if the text more explicitly indicated which figure or supplementary figure supports each of the main findings discussed. In particular, between lines 302 and 308, several conclusions are presented without reference to a specific figure in the Appendix or Supplement. Adding these references would help the reader follow the analysis and identify the evidence supporting each statement.

Yes, that is a great suggestion, thank you. We added the references to the figures supporting the individual findings.

Line 308 and Figure B1: The sensitivity tests using different a priori emission uncertainties are very informative. As shown in Fig. B1, the inversion increments are clearly sensitive to the assumed prior uncertainty, both in magnitude and spatial distribution. Could the authors clarify over which range of a priori uncertainties the posterior emissions can be considered robust? In other words, is there a range of values for which the main spatial patterns and regional emission estimates remain relatively stable? This would help justify the uncertainty chosen for the reference inversion.

Yes, agreed: We added to the discussion:

We test five different uncertainty values ranging from 0.3 to 2.0. This range encompasses sensitivity tests done in previous methane inversions, e.g. 0.5 and 1.0 in Thompson et al. (2025), 0.5-2.0 in Bergamaschi et al. (2022) and Subramanian et al. (2025)

and

When retaining the other settings of the reference inversion, the results are relatively stable for prior emission uncertainties of 1.0-1.5 (Figs. B1 and S19), without spurious negative emission patterns observed in the 2.0 case. We therefore select 1.0 as the reference uncertainty, consistent with Bergamaschi et al. (2010, 2022); Thompson et al. (2025) and Subramanian et al. (2025), which provides a balance between allowing substantial adjustments in response to the observations and avoiding artifacts in the inferred emission patterns.

Figure 6: suggests that the spread among posterior estimates is substantially reduced compared with the spread among the different a priori inventories for several regions. However, the term “remarkably robust” appears to be used qualitatively. Could the authors quantify the robustness of the posterior emissions, for example by reporting the spread, standard deviation, or coefficient of variation across the inversions using different prior inventories? This would also help distinguish regions where the posterior is largely independent of the prior from cases such as Saudi Arabia, where substantial prior dependence remains.
In addition, Saudi Arabia appears to be a notable exception to the overall consistency

across prior inventories. Could the authors discuss why the posterior remains more dependent on the prior in this region.

Yes, thank you for the suggestion. We agree and wrote (Appendix E)

As shown in the main text (Fig. 6), the inversions result in similar posterior emission estimates for Iran, Iraq, Egypt, Turkey, and the total EMME region, even when a priori emissions are very different. This convergence is further supported by Tab. E1 showing the coefficient of variation (CV) across inversions using different prior inventories and the ratio of posterior to prior standard deviations, ($R = \sigma_{post}/\sigma_{prior}$). For all these regions, the CV decreases from the prior to the posterior estimates, from 18–24%, to 3–11% for the individual countries and from 21% to 6% for the EMME region as a whole. Likewise R indicates a reduced spread, with values between 0.16 and 0.31 for these regions.

In contrast, for Saudi Arabia, the inversion shifts emissions to higher values by a similar absolute amount in all cases and the discrepancies in the prior emissions largely persist, resulting in less certain posterior estimates. This behavior might be explained by the definition of the a priori uncertainties as a fraction of the respective emission values, which may provide the inversion algorithm with too little freedom to accommodate large adjustments required for low a priori values. However, Tab. E1 shows a posterior CV of 31% compared to a prior CV of 38%, and R of 1.14, indicating that the spread in posterior emissions is not reduced relative to the prior spread, generally making the inversion results for Saudi Arabia less robust. We assume that this larger variability is related to Saudi Arabia's greater sensitivity to the assumed baseline uncertainty discussed in Sect. A and generally to all other sensitivity tests (Figs. S21, S23, and B1). Relatively large differences in the spatial distribution of prior emissions (see Fig. S5) may further contribute to the persistence of discrepancies between the posterior estimates.

Table E1. Coefficient of variation (CV) of priori and posteriori emission estimates across inversions using different prior inventories, and the ratio of posterior to prior standard deviations, $R = \sigma_{post}/\sigma_{prior}$

Region	Prior CV [%]	Posterior CV [%]	R
EMME	21	6	0.27
Iran	18	3	0.16
Iraq	24	11	0.28
Saudi Arabia	38	31	1.14
Egypt	18	7	0.31
Turkey	23	8	0.24

Line 313-314: Please indicate which figure supports or illustrates this statement

Yes, done

Line 315-316: Same, please indicate which figure supports or illustrates this statement

Yes, done

Line 596: The sensitivity of the inversion to the assumed spatial error-correlation length is clearly demonstrated. However, it is less clear how the tested range (50–500 km), and particularly the value adopted for the reference inversion, were selected. Could the authors provide a physical or statistical justification for these correlation scales? For example, were they informed by the spatial structure of the prior inventory, comparisons among inventories, variogram analyses, or previous inversion studies? Just as an

example, I'm sharing a paper that proposes different values for correlation length depending on the emission sector and also mentions other studies that use different reference values.

Montenegro, N., Berchet, A., Martinez, A., Thanwerdas, J., Bousquet, P., Pison, I., and Saunois, M.: Capability of current observing systems to monitor CH4 emissions from the regional to the global scales, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-5923, 2025.

Yes, agreed - we made the reasons for our choice clear in the manuscript and added:

We test spatial correlation lengths ranging from 50 to 500 km, similar to the ranges explored in previous regional methane inversions, e.g., 50–300 km in Europe (Bergamaschi et al., 2010, 2022) or 100–400 km in North America (Yu et al., 2021). Global inversion studies have generally used longer correlation lengths, often exceeding 500 km (e.g., Cressot et al., 2016; Montenegro et al., 2025). However, such long correlation lengths may be too large for regional applications, as demonstrated by Tsuruta et al. (2019) for Finland. This particularly concerns regions, such as the EMME region, where fossil-fuel emissions that are often concentrated in point sources, account for a substantial share of total emissions. For these emissions, shorter (or no) spatial correlation lengths are likely more appropriate, as they allow for more spatially localized emission adjustments (e.g., Szénási et al., 2021).

&

Our results were relatively stable for spatial correlation lengths of 100–250 km (Figs. C1 and S23), without obvious artifacts in the emission patterns (unlike the 500 km case). Although Thompson et al. (2025) and Subramanian et al. (2026) used correlation lengths of 200 and 250 km, respectively, in similar inversion frameworks, we selected 100 km as the reference value to allow for more spatially localized emission adjustments, particularly given the importance of localized fossil-fuel emissions in the EMME region. This choice is consistent with the 100 km correlation length used by Bergamaschi et al. (2022) for their European methane inversion. Szénási et al. (2021) also suggested correlation lengths in the range of 100–150 km for agricultural and waste emissions, although this recommendation was specific to those sectors.

We also deepened the discussion for temporal correlations, adding:

We test temporal correlation lengths of 30, 90, and 180 days, similar to Bergamaschi et al. (2022), who tested timescales of 1 and 12 months. While other studies also used smaller temporal correlations (e.g. Tsuruta et al., 2019; Montenegro et al., 2025), correlation lengths substantially shorter than 30 days would, be almost equivalent to assuming no temporal correlation for a monthly state vector, leaving each month's emissions largely constrained individually. This would increase noise and could introduce spurious month-to-month fluctuations in the posterior emissions. We therefore use 30 days as the lower

limit, consistent with the monthly resolution of the prior information and state vector, while the longer timescales represent progressively stronger temporal smoothing

and

Based on these results, we select a temporal correlation length of 30 days as the reference value, as it preserves the observed month-to-month variability, including the August emission peak, while avoiding the stronger temporal smoothing associated with longer correlation lengths. This choice is also better suited to represent the temporal variability of fossil-fuel emissions and is consistent with the temporal length used by Thompson et al. (2025) and Subramanian et al. (2025).

Line 326: Could the authors please clarify the total number of inversions underlying the final emission estimate? If each of the eight inversion setups (4 prior inventories × 2 error distributions) includes a Monte Carlo ensemble of 100 members, does this correspond to 800 Monte Carlo inversions in total.

Yes, it is 808 in total. 8 (4 prior inventories × 2 error distributions) inversions for quantifying the central estimate, and 800 inversions for estimating the uncertainties (100 members for each configuration). We added:

This results in a total of 808 inversions underlying the final emission estimates: 8 to determine the central emission estimates and the remaining 800 to quantify the uncertainties.

Figure 7: Given that the inversion specifically targets 2020, I strongly encourage the authors to include comparisons with independent emission estimates for the same year wherever available. This is particularly important because 2020 was an anomalous year, with substantial changes in economic activity and oil and gas production associated with the COVID-19 pandemic. Comparisons with 2019 or earlier estimates are useful for context, but they cannot provide a direct like-for-like evaluation of the 2020 posterior emissions.

Yes, we agree that comparisons with independent emission estimates for 2020 are particularly valuable. The availability of independent estimates for the same year and comparable geographical domain is, however, unfortunately, very limited. We identified, three additional sources for 2020 estimates that we have now included in the manuscript:

1. *Liu et al. (2024) use a TROPOMI-based divergence method to estimate emissions for some subnational regions in the Middle East, including 2020.*
2. *Qu et al. (2022) provide a regional total estimate for 2020 based on GOSAT observations and atmospheric inversions, covering a geographical domain comparable to ours (excluding Greece and Cyprus).*

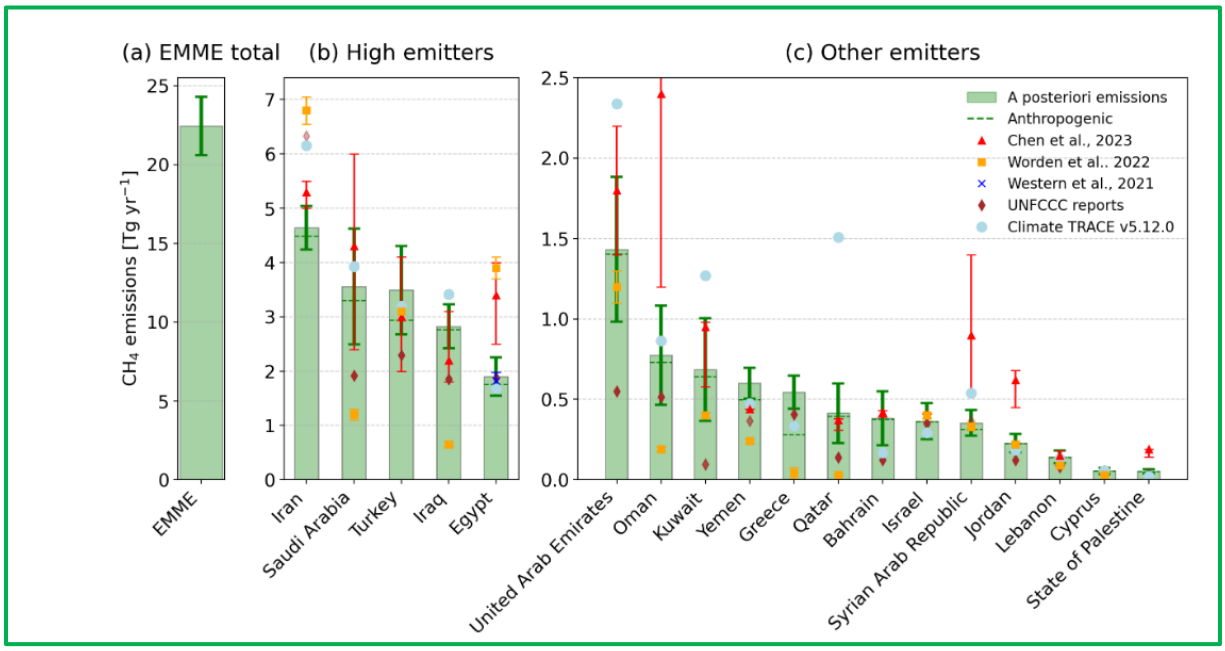
3. *The Climate TRACE inventory provides national, sector-resolved emission estimates for 2020. (however not totally independent from EDGAR v8.0, as we also note in the manuscript)*

We have added comparisons with these estimates to the Results section and expanded the discussion accordingly, while also acknowledging the limitations of the comparisons with estimates from other years:

Similar findings were reported by Liu et al. (2024), who derived lower 2020 methane emissions than EDGAR v7.0 for the coastal regions of Iran and Iraq using a TROPOMI-based divergence method. They found that emissions in this region exhibit substantial temporal variability, with a relatively small contribution from persistent sources, suggesting that bottom-up estimates may overestimate emissions in regions dominated by intermittent releases.

..... while positive adjustments are found in Turkmenistan and Iran, with the increments in Tehran consistent with Liu et al. (2024), who also reported substantially higher 2020 Tehran emissions than EDGAR v7.0.

For a comparable geographical domain, our estimate of 21.8 Tg yr^{-1} (excluding Greece and Cyprus) is also broadly consistent with the 24 Tg yr^{-1} reported by Qu et al. (2022) for 2020 using GOSAT-based inversions.



We compare our national a posteriori emissions to top down estimates from Chen et al. (2023), Worden et al. (2022), and Western et al. (2021), as well as to estimates from the Climate TRACE (v5.12.0; Climate TRACE, 2026) inventory and to national reports submitted to the United Nations Framework Convention on Climate Change .....

Please note that these studies report emission estimates for 2019, 2019, and 2017, respectively, while the inventory years covered by national reports submitted to the

UNFCCC also vary between countries and do not always include 2020 estimates. These sources, therefore, provide a broader context for interpreting our results, but are not directly comparable with our 2020 estimates. This is particularly relevant given the changes in economic activity and fossil-fuel production associated with the COVID-19 pandemic in 2020. We have therefore also included comparisons with the 2020 estimates from Climate TRACE, which provides national, sector-resolved greenhouse gas emission estimates based on a combination of existing inventories (e.g., EDGAR v8.0, CEDS, and FAOSTAT), activity data, and facility-level information, complemented by satellite observations and machine-learning approaches. However, we note that Climate TRACE estimates are not fully independent of EDGAR v8.0, as this inventory is among the data sources used in their compilation.

However, our estimated Egyptian emissions of 1.9 (±0.4) Tg yr^{-1} are almost identical to those reported by Western et al. (2021) for 2017, and are also consistent with the 2020 Climate TRACE estimate and the 2020 value submitted by Egypt in its first Biennial Transparency Report to UNFCCC.

Our estimate is substantially lower than the value reported for 2010 in Iran's Third National Communication to UNFCCC, which may also explain the comparatively high emission estimates in EDGAR v8.0 and Climate TRACE.

For Iraq, we likewise find consistent downward corrections across all tested cases, yielding an a posteriori emission estimate of 2.8 (±0.4) Tg yr^{-1}, which nevertheless remains higher than the 2019 value reported to UNFCCC in the Second National Communication and First Biennial Update Report but is close the 2020 Climate TRACE estimate.

For Turkey, our estimate of 3.5 (±0.8) Tg yr^{-1} is also in good agreement with Chen et al. (2023) and Climate TRACE,.....

For Saudi Arabia, our estimate of 3.6 (±1.1) Tg yr^{-1} is likewise close to that of Chen et al. (2023) and Climate TRACE, ...

Our results are also similar to the 2020 Climate TRACE estimates for many smaller emitters (including Cyprus, Greece, Israel, Jordan, Lebanon, Oman, Palestine, Syria, and Yemen), however substantially smaller for Kuwait, UAE, and Qatar. For these countries, the Climate TRACE estimates closely agree with EDGAR v8.0, suggesting that the differences mainly reflect the downward corrections relative to EDGAR v8.0 (see Fig. S9). In the case of Kuwait, lower emissions are also consistent with the relatively low number of methane plume detections 2019-2020 reported by Lauvaux et al. (2022), potentially reflecting fewer intermittent releases and more stringent maintenance operations.

Lines 414–416: The attribution of the late-summer CH_4 maximum over Iran to enhanced wetland and agricultural emissions appears somewhat stronger than the supporting evidence. Mousavi et al. (2024) identify a pronounced seasonal cycle in XCH_4 over Iran, but the interpretation in terms of wetland and agricultural emissions relies largely on

broader/global literature (Saunois, 2024). Given that wetland CH_4 seasonality can depend strongly on local hydrological and biogeochemical processes, including potentially substantial lags between precipitation, inundation and emissions, I suggest either providing more region-specific evidence for this attribution or softening the statement to indicate that wetlands and agriculture are possible contributors rather than established drivers.

Recent work has also shown that wetland CH_4 emissions can exhibit substantial lags relative to precipitation and inundation, highlighting that the timing of wetland emissions may be highly region dependent.

Yes, thank you , we agree and changed the statement to:

Regional analyses have reported a pronounced seasonal cycle in atmospheric CH_4 over Iran, with a late-summer maximum that may partly reflect enhanced emissions from wetlands and agricultural sources (Mousavi et al., 2024). However, these sources account for a comparatively small fraction of emissions in our prior inventory and are therefore unlikely to fully explain the pronounced summer peak inferred by our inversion. Furthermore, the wetland emission seasonality is complex and may lag behind precipitation and inundation patterns (Liu et al., 2026), complicating any source attribution based solely on seasonal timing.

Line 458: In line with the suggested revisions regarding the comparison with studies from other years, please make the corresponding changes to the conclusions so that they are consistent with the revised discussion.

Yes, we added/changed:

Our total estimate is broadly consistent (<10%) with the GOSAT-based inversion study by Qu et al. (2022), while the spatial pattern of negative adjustments along the coasts of Iran and Iraq, together with the positive adjustments over Tehran, is in line with the TROPOMI-based study by Liu et al. (2024).

Our national results are in good agreement with the TROPOMI-based inversion study of Chen et al. (2023), but show poor agreement with the GOSAT-based study of Worden et al. (2022), both of which use the GEOS-Chem model. This may suggest that the improved observational coverage of TROPOMI can enhance inversion performance and lead to more consistent emission estimates, in addition to the higher inversion resolution used in our study compared to Worden et al. (2022). Nevertheless, these comparisons should be interpreted in the context of the different years considered, as both studies report emissions for 2019, whereas our estimates refer to 2020.

Our national emission estimates are generally in good agreement with the 2020 estimates from the Climate TRACE inventory. For Iran, Kuwait, Qatar, and the UAE, however, our

posterior estimates are lower, reflecting our inversion downward adjustments from similar high prior values

Literature

Bergamaschi, P., Krol,M.,Meirink, J. F., Dentener, F., Segers, A., van Aardenne, J.,Monni, S., Vermeulen, A. T., Schmidt,M., Ramonet,M., Yver, C., Meinhardt, F., Nisbet, E. G., Fisher, R. E., O'Doherty, S., and Dlugokencky, E. J.: Inverse modeling of European CH4 emissions 2001–2006, Journal of Geophysical Research: Atmospheres, 115, D22 309, https://doi.org/10.1029/2010JD014180, 2010.

Bergamaschi, P., Segers, A., Brunner, D., Haussaire, J.-M., Henne, S., Ramonet, M., et al.: High-resolution inverse modelling of European CH4 emissions using the novel FLEXPART-COSMO TM5-4DVAR inverse modelling system, Atmospheric Chemistry and Physics, 22, 13 243–13 268, https://doi.org/10.5194/acp-22-13243-2022, 2022.

Climate TRACE: Climate TRACE Emissions Inventory v5.12.0, https://climatetrace.org/data, accessed: 2026-10-29, 2026.

Lauvaux, T., Giron, C., Mazzolini, M., d'Aspremont, A., Duren, R., Cusworth, D., Shindell, D., and Ciais, P.: Global assessment of oil and gas methane ultra-emitters, Science, 375, 557–561, https://doi.org/10.1126/science.abj4351, 2022.

Liu, M., van der A, R., van Weele, M., Bryan, L., Eskes, H., Veefkind, P., Liu, Y., Lin, X., de Laat, J., and Ding, J.: Current potential of CH4 emission estimates using TROPOMI in the Middle East, Atmospheric Measurement Techniques, 17, 5261–5277, https://doi.org/10.5194/amt-17-5261-2024, 2024.

Qu, Z., Jacob, D. J., Zhang, Y., Shen, L., Varon, D. J., Lu, X., Scarpelli, T., Bloom, A. A., Worden, J., and Parker, R. J.: Attribution of the 2020 surge in atmospheric methane by inverse analysis of GOSAT observations, Environmental Research Letters, 17, 094 003, https://doi.org/10.1088/1748-9326/ac8754, 2022.

Szenasi, B., Berchet, A., Broquet, G., Segers, A., Denier van der Gon, H., Krol, M., Hullegie, J. J. S., Kiesow, A., Gunther, D., Petrescu, A. M. R., Saunois, M., Bousquet, P., and Pison, I.: A Pragmatic Protocol for Characterising Errors in Atmospheric Inversions of Methane Emissions over Europe, Tellus B: Chemical and Physical Meteorology, 73, 1–23, https://doi.org/10.1080/16000889.2021.1914989, 2021.

Yu, X., Millet, D. B., and Henze, D. K.: How well can inverse analyses of high-resolution satellite data resolve heterogeneous methane fluxes? Observing system simulation experiments with the GEOS-Chem adjoint model (v35), Geoscientific Model Development, 14, 7775–7793, https://doi.org/10.5194/gmd-14-7775-2021, 2021.