We thank the reviewers for the thoughtful comments. In the following, we list the reviewer's comments in normal fonts, while we provide our answers in italic fonts. Changes we have made in the manuscript are highlighted in blue.

Review 1:

General comments:

The authors use the FLEXPART transport model and the FLEXINVERT+ inversion framework, and introduce a post-processing constraint that aligns the sum of regional emissions with global total estimates from the AGAGE 12-box model. While this approach is potentially useful, I have several major concerns regarding the methodology and the associated conclusions. In particular, it is unclear whether the introduction of a global-total constraint can effectively stabilize emissions in poorly monitored regions. Such a constraint may help to constrain the aggregate emissions of these regions, but it does not necessarily improve the ability to distinguish emissions among them, especially under limited observational coverage. The authors are encouraged to provide additional justification and supporting analyses (e.g., posterior error covariance, sensitivity tests, or resolution diagnostics) to demonstrate the effectiveness of this approach. Furthermore, the limitations of the inversion framework in resolving regional emissions should be more clearly acknowledged and discussed. Overall, substantial revisions are required before the manuscript can be considered for publication.

Specific Comments

- Abstract: I am not convinced that introducing a global-total constraint can effectively stabilize emissions in poorly monitored regions.While such a constraint, combined with well-constrained regions (e.g., Europe, the USA, and East Asia), may help constrain the aggregate emissions of the remaining poorly monitored regions, it does not necessarily improve the ability to distinguish emissions among these regions (e.g., Africa, South America, South Asia, and Southeast Asia). In other words, the global-total constraint effectively reduces only one degree of freedom (the global sum), but the partitioning of emissions among poorly observed regions likely remains weakly constrained. Have the authors examined the posterior error covariance (or correlation) among these regions? It is likely that substantial negative correlations persist, indicating that the inversion cannot robustly separate their emissions despite the global constraint. Therefore, statements suggesting that the global-total constraint "stabilizes emissions in poorly monitored regions" may be overstated and should be revised or more carefully qualified.

  We generally agree with the reviewer's comment (see also general response). Emissions and especially their distributions in and between regions with sparse observational coverage will remain uncertain and we do not claim that the global constraint can solve that problem (see also the large respective uncertainties).

Nevertheless, by using a global constraint, we have another piece of information that we can use to improve the prior estimates of emissions in those countries, which we think, is worth using! However, we agree that the phrase might be overstated and change it: "We introduce a global-total constraint to align the inversion results with the relatively accurate global total emissions from the AGAGE 12-box model, resulting primarily in emission adjustments in poorly monitored regions.

- Abstract: It is unclear whether the inversion framework can robustly distinguish emissions from South and Southeast Asia, given the limited observational constraints in these regions. The authors are encouraged to provide supporting evidence (e.g., posterior error correlations, resolution metrics, or sensitivity analyses) to substantiate this claim. In addition, uncertainties in the fractional contributions to the global total should be quantified and reported, given their potentially large magnitude.

  We generally agree, and refer to our answer to a related question below as well as our general response for the first part of this comment. For the second part, we added uncertainties to Fig. 5 (pie chart), combined the SSEA country emissions in the figure, and revised the text accordingly. In this process, we identified a small error in the calculation of the fractions and uncertainties, therefore the values in the text differ from those in the original draft.

- Introduction: This statement may give the impression that the present study can close the global gap by robustly constraining regional emissions in South and Southeast Asia, South America, and Africa. However, this may be overstated, as posterior emissions in these regions are likely still strongly influenced by prior assumptions due to limited observational constraints. While the study may provide useful insights, it may not be sufficient to fully resolve emissions in these poorly monitored regions. The authors are encouraged to moderate this statement or clarify its scope.

  We agree and changed to:

  "We introduce a novel method to constrain the total global emission result of the regionally resolved model to the more robust global emission results from the AGAGE 12-box model simulations"

- 2.2: The use of "daily intervals" for emission sensitivities is unclear. Earlier, the observations were averaged to 3-hour intervals, could the authors clarify why a different temporal resolution is used here? Is there any inconsistency between the temporal resolution of the observations and that of the sensitivities, and if so, how is this handled in the inversion?

  The "daily intervals" refers to the output of the emission sensitivities, which are completely independent of the frequency of the observations. In principle, the emission sensitivities could also be stored without any time resolution. However, this would make it impossible to attribute emissions correctly to a specific year. For *instance, with our long backward simulation time an observation made on 10 January provides information on emissions during the first days of the year but also on emissions in the last months of the previous year. To correctly resolve the emission constraint, emission sensitivities must be temporally resolved.*

We added:

"Emission sensitivity fields are stored with daily resolution to be able to correctly attribute emissions to specific years when the backward simulation period spans two years."

- Equation 4: While the formulation shows how the posterior emissions are adjusted after imposing the global-total constraint, it is unclear how the posterior uncertainties for individual grid cells are affected.How does this constraint propagate into the posterior error covariance? In particular, does it reduce uncertainties uniformly or introduce additional correlations among grid cells? The authors are encouraged to provide the corresponding formulation or a more detailed discussion.

  C *represents the assumed uncertainty of the AGAGE 12-box model's global-total emission estimate, and controls the overall magnitude of the adjustment applied to bring the regional sum into agreement with this target. In our approach,* ***C*** *is used only to scale the size of the mean adjustment and does not alter the shape or covariance structure of the underlying distribution. A rigorous propagation of this adjustment into the posterior uncertainty would additionally require treating the regional inversion result and the AGAGE 12-box model global total as independent, as in a standard Bayesian-type update. However, both estimates are constrained by overlapping measurements, so some degree of correlation between their errors is likely. Given this, we consider it more appropriate and conservative to retain the unconstrained uncertainties in our results. In 2.3, we added:*

  "In principle, introducing ***C*** *would also imply a new posterior covariance. However, since this would only amount to a scaling of the existing covariance and the errors of the regional inversion and the AGAGE 12-box model are likely correlated, we chose to retain the unconstrained a posteriori uncertainties also for the constrained emissions."*

- Equation 4:The constraint is applied as a post-processing step using Eq. (4). How sensitive are the final regional emissions to the prescribed uncertainty matrix C? A sensitivity test with, e.g., ±50% changes in C would help readers gauge the influence.

  We agree and have added a sensitivity test for the global and SSEA regions to the supplement, applying factors of 0.25, 0.5, 2, and 4 to the standard uncertainty. We also clarified in Sect. 2.3 that the standard uncertainty corresponds to the uncertainty of the AGAGE 12-box model results reported by Western et al. (2025b). The following text was added to Sect. 2.3:

*"The uncertainty assigned to the total global emission in **C** governs how strongly the a posteriori emissions are pulled from the unconstrained solution towards the AGAGE 12-box model global emissions. We tested the sensitivity of this constraint by scaling the original uncertainty in **C** by factors of 0.25, 0.5, 2, and 4 in figures S5-S6. Smaller uncertainties did not significantly alter the constrained emissions, whereas larger uncertainties progressively relaxed the result towards the unconstrained emissions."*

- The prior emissions are described as "adjusted BU inventory" (EDGAR with modifications for China, India, South Korea). This adjustment brings the prior closer to the global total but also introduces external BU information.I recommend presenting both the original EDGAR and the adjusted prior separately in the main text (e.g., in Fig. 4 and in the regional time series) to allow readers to distinguish the effect of the adjustment from the pure inversion. How does the use of this adjusted prior affect the attribution of the TD BU discrepancy in Section 3.3? The discrepancy is defined against EDGAR, but the inversion prior is already partly corrected. This should be explicitly noted.

  We agree and added in 2.2:

  "The overall gap between a priori emissions and the well known global total TD emissions from Western et al. (2025b) is reduced from 55% to 45% for CF4 and from 53% to 34% for C2F6."

  And added in 3.1:

  "Still, the inversion results are 1.8 ± 0.2 (CF4) and 1.6 ± 0.4 (C2F6) times higher than our custom BU inventory in the most recent period."

  In Fig 4 and the regional time series figures, both the original EDGAR emissions and the adjusted prior emissions were already added. We define the TD BU discrepancy in section 3.3 against EDGAR since we do not want to highlight the increment from the a priori achieved by the inversion but the added information of our study compared to previously published BU data. EDGAR represents the only globally complete BU inventory.We explicitly noted in the text:

  "We chose the original EDGAR inventory for comparison,...."

- The paper notes a pronounced drop in Chinese CF4emissions (approximately 20%) in 2019, identifying this as a potential inversion artifact. While the authors link it to the US–China trade war's impact on aluminum production (around 4%), the 20% reduction far exceeds what can be explained by output fluctuations alone. The authors are advised to further investigate whether this anomaly stems from missing observations at specific sites, anomalies in the meteorological fields, or shifts in the allocation weights of the global constraint equation in that year.

Yes, we agree. We state in the manuscript: "However, the relative magnitude of the 2019 emission drop in all three TD estimates (approximately 20%) is much larger than the decline in aluminum production and cannot be explained by production fluctuations alone; moreover, no BU inventory shows a comparable decrease.", and added:

"Thus, although changes in aluminum production may partly explain the observed CF_4 decrease, its magnitude remains unexplained, and we cannot exclude the possibility that the 2019 reduction is at least partly an inversion artifact."

reflecting the reviewer's comment.

We examined observational coverage and the influence of the global constraint as two possible explanations for the 2019 anomaly. We found no significant gap in observational coverage in China, and the global constraint did not substantially affect the result, given the relatively low posterior uncertainty in China. Assessing the role of meteorological fields, however, would require a much more detailed analysis, as the relationship between meteorological anomalies and inferred emission anomalies is not necessarily straightforward, which is why we consider it to be beyond the scope of our study. (Moreover, Chinese emissions have been investigated in detail by several previous studies.)

- The study uses sectoral emission fractions (aluminum industry: 0.10; electronics industry: 0.40) from Kim et al. (2014) to attribute sources. However, given the evolution of aluminum smelting technologies (e.g., anode effect control) and electronics industry abatement technologies (e.g., end-of-pipe treatment) over the past 15 years, retaining these fixed ratios may introduce biases in sectoral attribution. The authors themselves acknowledge unrealistic allocation results for South Korea and Taiwan. It is recommended that the discussion include an analysis of temporal changes in sectoral emission fractions.

We agree that the sectoral emission fractions are unlikely to remain constant over time, given the evolution and technological improvements in both the aluminum industry and the electronics sector, which likely affect relative sector contributions. We also assume that the unrealistic allocation results obtained for South Korea and Taiwan indeed reflect limitations of our simplified approach. We expanded the discussion in section 3.4 to acknowledge potential biases from constant emission fractions.

"A higher ratio (0.4kgkg−1) indicates predominant emissions from electronics manufacturing, while a lower ratio (0.1kgkg−1) is associated with emissions from aluminum smelting. We present country-level and grid point-level emissions per industry averaged over the late period (2018–2023), assuming that sectoral emission

fractions *are constant in space and time. While constant emission fractions are a commonly used approximation due to limited availability of temporal and regionally resolved sectoral emission data, this assumption introduces additional uncertainty in the attribution. For both industries, there is a wide range of manufacturing methods, while emission abatement technologies have evolved over the past two decades (e.g. Marls et al., 2016), likely affecting relative sector contributions. In addition, other potential PFC sources further contribute to the uncertainty in sectoral attribution.*

For South Korea and Taiwan, where there are no active aluminum smelters, the resulting industry contributions in panels b) and c) are unrealistic and likely reflect the discussed limitations of our simplified approach using constant emission factors."

- The study reports that the share of emissions from South and Southeast Asia (SSEA), particularly India, has surged from 6–10% in the early period to 18–22% in recent years. Given the complete absence of in-situ observational stations in this region, to what extent are these "increases" driven by forced allocation to unconstrained regions via the global constraint equation? The authors are encouraged to explicitly assess the robustness of the growth trends for these regions.

  We agree that it is not fully clear what the increases in the SSEA region are driven by. We present the unconstrained inversion output in fig. 9 as the dashed red line and added in 3.2:

  "Nevertheless, it is clear that uncertainties remain high in the region and we note that our unconstrained inversion results also indicate an increase of emissions from 0.38±0.69kt/yr and 0.23±0.30kt/yr in the early period to 3.01±1.31kt/yr and 0.53±0.36kt/yr in the late period for CF_4 and C_2F_6, respectively."

Technical corrections:

- Section 2.2 mentions the use of a 100-day backward tracking window; whilst this helps to improve far-field constraints, it may also introduce coupling between earlier emission signals and prior emissions. It is recommended that the implications of this choice for interannual variations in emissions and regional allocation be further explained.

  See answer to next comment.

- 2.2: Can the authors further clarify why longer backward simulation periods (e.g., 50 or 100 days) are expected to improve the robustness of the inversion? It is unclear how emissions from such long time scales contribute additional information to constrain regional emissions. Air masses older than ~30 days are likely well mixed and may primarily influence the background (baseline) rather than providing sensitivity to regional sources.

  The choice of a long backward simulation period is motivated by earlier findings (e.g. Vojta et al., 2022), which showed that extending the simulation length can improve *agreement between model and observations. In our framework, the baseline is defined as contributions prior to the backward simulation period, estimated by coupling LPDM trajectories at the end of the simulation to a global concentration field, (which gives a clear separation between contribution from inside and outside the simulation period). Extending the simulation period reduces the fraction of the signal that must be attributed to this estimated baseline and increases the portion that is explicitly resolved by the LPDM. This also reduces the sensitivity to potential biases in the global initial fields, as the influence of localized errors is diluted, as particles are transported over a wider domain.*

However, the information content decreases with increasing backward time, as air masses become more mixed and sensitivities to specific regions weaken. Consequently, contributions from emissions beyond ~30–50 days are usually very small, and the benefit of longer simulations diminishes. The 100-day simulation period was primarily chosen, to assess whether extending beyond 50 days provides additional benefit, which we find to be only marginal. We acknowledge the reviewer's concern, that large simulation periods could in principle have the disadvantage of reducing the distinguishability between long-timescale emission contributions and the baseline. However, baseline contributions are optimized using a small number of parameters (scalars within coarse latitudinal bands), which strongly constrains their degrees of freedom. Even for long simulation periods, these parameters are typically well constrained (e.g. Vojta et al., 2025). As a result, the potential mixing between long-timescale emission contributions and the baseline is expected to be limited.

We further added in the manuscript:

"Additionally, the long simulation period is also expected to reduce the sensitivity of the inversion results to potential biases in the global initial fields."

- The observation uncertainty in R is defined with thresholds (2.0 ppt for CF_4, 0.4 ppt for C_2F_6) and a linear increase above those thresholds. Provide a brief justification (e.g., based on model-observation residuals from a test run) or cite a previous study where this approach was validated.

  This method limits the influence of individual high measurements affected by nearby emissions, for which transport is not well represented by the LPDM simulation at such small scales, resulting in larger model representation errors. Our approach is a simplified version of that used by Bergamaschi et al. (2022), who define the model representation error as a constant value below a certain threshold and, above it, use the observed mole fraction enhancement (observed mole fraction minus baseline) as the error estimate. Because our long backward simulation period and the stability of the two PFCs yield a very smooth baseline, using absolute observed mole fractions instead of the enhancement above baseline does not significantly affect the result, which justifies this simplification.

- The paper provides data citations and some DOIs, but does not explicitly state whether the core inversion code, post-processing scripts and structural descriptions of the gridded emission output files are available. Given EGUsphere's requirements regarding reproducibility, it is recommended that the authors provide this information in the 'Data availability' section.

  All input and output data is available through the links in the "Data availability" section. The FLEXINVERT+ and FLEXPART11 (including LCM) code is available through the links listed in the "Code availability" section.

- South and Southeast Asia (Fig. 9) show a rapid emission increase, but the uncertainties remain large and the error reduction is low. The conclusion that this region contributes 22% (CF4) and 18% (C2F6) to global emissions in 2018–2023 is important, yet the text should more explicitly caution that these numbers are heavily influenced by the global constraint rather than by local measurements.

  We agree that the region's emissions are not influenced by local measurements but we disagree that they are heavily influenced by the global constraint. Looking at the inversion results without the constraint stated in an answer above (and now also in 3.2), the results with and without the constraint differ relatively little (3.39kt/yr to 3.01kt/yr for CF_4 and 0.42kt/yr to 0.53kt/yr for C_2F_6) compared to their uncertainties.

  To ensure that readers are always aware that SSEA emissions are uncertain, we also added in section 3.3:

  "Additionally, we note that the fractional contribution of SSEA to the TD-BU discrepancy is more uncertain than that of other regions, owing to the region's weaker observational constraint."

- The fractional contributions to the TD-BU discrepancy are shown for EDGAR. Given that the prior already incorporated adjustments for some countries, the discrepancy might be partially reduced by construction. I suggest adding a version of Fig. 10 using the original EDGAR inventory as the BU reference, to show the full extent of underestimation.

  The BU reference in Fig. 10 is already the original EDGAR inventory without the adjustments added for it's use as the prior. To emphasize this fact for readers, we added "original EDGAR" instead of "EDGAR" in section 3.3.

- Taiwan is a province of China. Revise Taiwan to Taiwan Province of China. Revise China to China mainland.

  We revised "Taiwan" to "Taiwan, Province of China" and "China" to "China, mainland" at first mention, retaining the short forms thereafter for readability.

Review 2:

This study presents global and regional/national emissions of two perfluorocarbons using measurements from AGAGE, NOAA and the Chinese flask network (previously published data), through 2023.

The manuscript is well-written (I came across almost no typos), which made it very easy to read. The figures and text are polished with appropriate uncertainties, etc.

However, I have two major reservations, which in my opinion, requires the results to be presented differently in order to be considered for publication.

The first is in the level of disaggregation the authors are presenting for the regions with no measurements. The authors present this work in several places as "the first regionally-resolved inversion results"; however, given that there has been no change in the measurements used (over previous studies), the results need to be presented differently. I agree that it may be possible to use a global approach to resolve emissions at higher resolution than the box model resolution of Western et al., (semi-hemispheric) but I am not convinced that it is possible to resolve individual countries (e.g., India, Canada, Russia, Singapore/Malaysia) with those measurements with a global total constraint alone. Much more evidence and testing would be needed. Let alone resolving, as shown in Fig S1, at an even higher spatial resolution within these countries (even with a 100km correlation scale). I'd suggest aggregating the emissions over much larger continental scales. Specific comments on this point:

- Even in the "well-constrained" Europe, the measurements do not constrain the whole EU27 but parts of north-western Europe (see for example Say et al., 2021 only presented Northwest Europe). so care should be taken not to over-extrapolate.
- In the USA, the flask network sensitivity is relatively low compared to the in situ stations, and more information is needed to demonstrate that the whole country can be resolved (looking at the SI Table on measurements, the flask measurement frequency ranges between 2d and 2m, so still quite sparse).

We agree with the reviewer, that the measurements do not constrain the whole EU27 equally well and that the sampling frequency of the flask network in the USA limits the sensitivity to the whole country. The globally sparse measurement system for these species, set, indeed, limitations to our global approach (and is generally a problem in atmospheric inversions). In regions less covered, posterior emissions will be highly governed by the prior information, while also less error reduction will lead to larger overall uncertainties of our inversion results. Consequently, the regions totals have more uncertainty compared to the regions with the best observational coverage (e.g. Northwest Europe). Nevertheless, we think it is useful to provide estimates for USA and EU-27 (even if that means that uncertainties are larger), particularly in the context of a global inversion, and try to make the best use of the existing observation network. We discuss the limitations of our approach and provide additional sensitivity tests to address this in the manuscript.

The second comment is about novelty. As much of this work is already published elsewhere (with the exception of the USA results) I would have expected this manuscript to extensively compare results with those studies. The main difference, for the regions that are well-constrained, is in using a different transport model (which is important). Yet there is little comparison with those studies. I'd suggest focusing this manuscript on the comparisons with other studies for the well-constrained regions (i.e. what is the impact of transport model choice) and doing more on the USA results, since to my knowledge, this is the first study to look at this region.

We follow the reviewer's suggestion by adding the emission timeseries including results from Say et al. (2021) for northwest Europe in the supplement and expanding the section on USA. If there are additional specific studies that the reviewer would like us to include, we would be happy to incorporate them.

I provide some specific comments as well, but as there are major issues to do with the presentation of results, I refrain from presenting detailed specific comments on those aspects at this stage.

- Line 104- why are the instrumental uncertainties not used? They have important time-varying patterns that the model uncertainties sometimes do not capture.

  We agree that time-varying patterns in instrumental uncertainties may provide additional useful information. However, the construction of the observation error covariance matrix (R) is complex and the optimal treatment of different error components remains an active area of research. Separating and combining instrumental and model uncertainties is challenging, and different approaches can lead to substantially different weighting of observations. In this study, we used the described approach, implicitly assuming that model errors dominate over instrumental uncertainties. However, we acknowledge that incorporating time-varying instrumental uncertainties could be a way to further improve uncertainty estimates.

- Line 150 – why is extrapolating C2F6 based on aluminium production only valid when there is a large semiconductor source?

  For India, we adopted a pragmatic approach by using aluminum production as a proxy to extrapolate CF4 and C2F6 emissions, as this sector represents by far the largest source of PFC emissions according to India's latest UNFCCC report (https://unfccc.int/sites/default/files/resource/India%20BUR-4.pdf) and EDGAR's industry specific estimates (https://edgar.jrc.ec.europa.eu/dataset_ghg80#p1). However, we agree that, particularly for C2F6, this approach introduces uncertainty if semiconductor-related emissions are far larger than reported. Certainly, a more detailed sector-specific extrapolation, accounting separately for aluminum and semiconductor emissions, would in principle be preferable. However, according to the current knowledge about India's sector-resolved PFC emissions, the difference would be negligible.

- Line 167 – is a spatial correlation scale of 100km appropriate for point sources? Why would one expect point source emissions uncertainties to correlate as the emissions are episodic.

  We agree with the reviewer that large spatial correlation lengths are not ideal for point sources, particularly given their often localized and episodic nature. However, in practice the choice of correlation length often represents a compromise between providing sufficient constraint on emissions and maintaining the inversion's ability to resolve regional emission patterns. When emission uncertainties are assumed to be fully uncorrelated, the inverse problem becomes poorly constrained, particularly when the observational coverage is limited, which can lead to unrealistic artifacts such as noisy spatial patterns or negative emissions. Introducing spatial correlations acts as a form of regularization, as observational information is effectively spread over multiple grid cells. At the same time, we agree that excessively large correlation lengths smooth the solution and may bias the inversion results. In our setup, we already try to mitigate this issue by using a variable-resolution grid, with grid cell sizes of up to 8° × 8° in data-sparse regions. Nevertheless, additional spatial correlation was still required to ensure a stable and physically plausible solution. We therefore selected a correlation length that is as small as possible while maintaining inversion stability. The chosen value of 100 km is rather within the lower range of values commonly used in the literature.

- Line 175 – what is the slope of the linear increase in model error?

  For CF_4 the slope is 0.1 ppt/ppt, for C_2F_6 the slope is 0.3 ppt/ppt. We added this information in section 2.2 as well.

- Section 3.1 – the discussions about global emissions (e.g., financial crisis) are already presented in other papers and does not need to be discussed again.

  We agree that these aspects have been discussed in previous studies. However, we consider it useful to mention them for context and provide appropriate references, while keeping the discussion concise.

- Fig 4 – it does not make sense to present global UNFCCC numbers when the reporting for UNFCCC does not have global coverage

  We partly agree with the reviewer. However, we think it is still interesting to compare the emissions to the global UNFCCC reported ones, even though many countries are missing, as it shows the global "reporting gap". Also, in principle UNFCCC report totals could also be larger than global totals. The fact that this is not the case shows that UNFCCC report totals are at least not dramatically overestimated. To facilitate a better interpretation, we have added in the figure caption: Notice that UNFCCC reports are missing from many countries and the UNFCCC report totals are therefore expected to be smaller than global emissions.

- Line 328 – is there evidence to back up a shift within the USA to aluminum over semiconductor manufacturing?

  We added a paragraph in section 3.2. further discussing the shift from higher to lower C2F6/CF4 ratios. Evidence from the EPA bottom-up inventory (https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks) shows that emissions from the aluminum industry decreased simultaneously for both species, whereas semiconductor industry emissions diverged, with CF4 increasing and C2F6 decreasing. While we found no U.S.-specific evidence confirming that this divergence is driven by the replacement of C2F6 with NF3, the Semiconductor Industry Association reports this substitution as an ongoing global trend (https://www.semiconductors.org/wp-content/uploads/2023/06/FINAL-Plasma-Etch-and-Deposition-White-Paper.pdf).

General response:

Both reviewers have concerns about the global-total constraint used in our paper. It is absolutely true that the global-total constraint alone does not allow a detailed disaggregation of emissions between "poorly-constrained" regions. Yet, there are large advantages of using this novel approach:

1. The goal of inverse modelling, in general, is to improve a priori emissions. It is commonly accepted that this is the case in "well-constrained" regions. If that's the case, this automatically implies that, even if no improvement in the relative distribution of emissions in "poorly-constrained" regions can be obtained and this relative distribution is largely guided by the a priori patterns, the absolute magnitude of the emissions in these regions is improved, as their collective emissions must be the residual of the global total and the emissions in well-constrained regions.
2. In addition to using the global-total constraint, we also substantially extended the usual duration of Lagrangian backward simulations (often 10-20 days) to 100 days. While the information content decreases with every day of simulation added, this still allows to put weak constraints on the emissions in "poorly-constrained" regions and the agreement between the model and the observation improves, as demonstrated previously (Vojta et al., 2022). For instance, this allows episodes of interhemispheric air mass exchange to be captured. Therefore, some separation of "poorly-constrained" source regions (especially if they are located at different latitudes) is possible when using the global-total constraint (note the analogy to the AGAGE 12-box model), even if the relative emission distribution within these source regions is almost exclusively guided by the a priori emission information.
3. Past inversion studies using Lagrangian models have mostly focused on particular "well-constrained" regions. However, even for these regions a clearly-defined boundary separating them from "poorly-constrained" regions does not exist, and the

> choice of the inversion domain is always somewhat subjective. In fact, when extending such regions too much and having a somewhat unbalanced choice of uncertainties, an inversion could result in emissions that are above the known global total. Even if such a scenario may seem extreme, regional inversions may tend to yield total emissions that are unrealistically high or low in "poorly-constrained" parts of the domain. The use of a global-total constraint effectively prevents this from happening.

Therefore, we argue that, for species for which the global emissions are well known, it is always better to use a global-total constraint as introduced in this paper. The question is: How much better? And to what extent can we trust the emission distribution within and between "poorly-constrained" region? These are certainly questions that are important to explore and we have added some sensitivity studies exploring this, but it is beyond the scope of this first paper to give final answers to these questions, which will require testing using different inversion setups and different models.

The criticism of the reviewers was mainly that we overinterpret the results in the "poorly-constrained" regions. To avoid this, we therefore revised the manuscript to explicitly flag the large uncertainties associated with these regions wherever they are discussed. In addition, to provide a more robust and better-constrained alternative to the full EU emissions, we included supplementary results for Northwest Europe.

Finally, we also want to emphasize that the use of the global-total constraint is only one (and maybe not even the most important) aspect of this paper. We also provide new emission estimates (and for longer periods than past studies) in the well-constrained regions of East Asia and Europe, as well as North America, for which no previous estimates exist. We have the impression that this was not fully acknowledged by the reviewers.