

Review on 'Concurrent assimilation of methane fluxes and concentrations with a 4D LETKF for Germany in 2021 based on ICON-ART'

General Comments

The authors present a concurrent assimilation of methane flux and concentration fields using a 4D LETKF, extending an approach previously explored in aerosol modeling and observation system simulation experiments to a real top-down CH₄ emission estimation framework. This is a meaningful novelty, and the paper is generally well-written.

The most significant limitation is the absence of independent validation: all available ICOS stations are assimilated, and the RMSE skill score (Sections 4.1.1/5.1) is computed on the same stations and days used for assimilation, evaluated only at hours adjacent to the assimilation window. This makes it difficult to assess whether improvements in RMSE reflect genuine predictive skill or in-sample fit. Withholding 2–3 geographically distinct stations for independent validation would substantially strengthen the evaluation (see also Specific comment 9), and I recommend the authors do that, at least for a limited domain or time frame if computational costs are too high to re-run the entire simulation. This is especially relevant in this study as the presented method breaks the relationship between emissions and concentrations that is traditionally at the core of inverse modelling.

Furthermore, the manuscript would benefit from forward references at several occasions in Sections 2 & 3 (see specific comments), and from a short paragraph at the opening of Section 3 explicitly introducing the augmented state vector and the mapping to simulations (i)–(iv), which are currently the last thing introduced but the first thing the reader needs to understand the method.

The method description would benefit from additional detail in the interest of reproducibility, particularly regarding preprocessing of emission and boundary concentration data (see specific comments) or a reference to a description of said processing.

Specific Comments

1. The abstract would benefit from quantitative language: model resolution used and emission adjustments are described only as 'high resolution' and 'significantly higher'.
2. The introduction could more clearly highlight the novelty of the approach: The authors refer to it in l37–39 but then in the next paragraph write about the LETKF vs the previous synthesis inversion of Bruch et al, as though this constitutes the primary innovation.
3. L 72: Please clarify whether CH₄ fields are carried across days or also reinitialized.
4. Section 2.2: what are the original spatial resolutions of the datasets? A brief description of the necessary regridding would be beneficial, too. How are discontinuities handled at country borders?
5. L 79–80: Please comment on the assumption regarding constant monthly fluxes.
6. Table 1: are the natural emissions from wetlands specific to 2021 or climatological averages?
7. 2.3: The CH₄ ICBC are taken from the CAMS global inversion-optimized product, which is itself derived from an atmospheric inversion using surface station observations. Please clarify whether any stations used in that CAMS inversion overlap with the ICOS stations used for assimilation/validation here. While this doesn't affect the RMSE skill score comparisons (Fig. 5, since the boundary is identical across simulations (i)–(iv)), it is potentially relevant to the extent to which the absolute posterior emission estimates can be considered fully independent of prior top-down information.

8. 2.3: The CAMS global dataset has a resolution of approx. 80km. Are any interpolation steps performed to the much finer 6.5km grid? Does this produce any artefacts at the domain boundaries?
9. Section 2.5 specifies the workflow, but it remains unclear whether BACY functions as a scheduler, a postprocessing tool (ICON output → model equivalent), or the LETKF implementation itself. Please specify.
10. Figure 1 / Section 2.5 (page 3): "LETKF," "model equivalents," and "ensemble generation" appear in the figure and in BACY's description before any of these are defined. Section 3.1 is five pages away. At minimum, a sentence like "The LETKF and the construction of model equivalents are detailed in Sections 3.1 and 3.1.3 respectively" would let the reader know to hold these concepts until then.
11. Please justify the linear observation operator assumption in Eq. 3 by noting CH₄ is treated as a passive tracer or an equivalent physical justification. Please also discuss whether the use of 20 independent meteorological ensemble members underlying the 100-member CH₄ ensemble introduces effective nonlinearity/non-Gaussianity into the ensemble-based linearization, since each meteorological realization implies a distinct transport operator.
12. L 130-134: The authors should explicitly define what $W^{a(i)}$ denotes in Eq. (8) — it is not clear from the text whether the superscript (i) refers to the i-th column of the matrix W^a , or to some other member-wise quantity. A sentence such as 'Here $W^{a(i)}$ denotes the i-th column of W^a , representing the analysis perturbation for ensemble member i in ensemble space' would remove the ambiguity.
13. 3.1.3: Is there any reason why observations at 13 and 14 UTC are not assimilated?
14. Sections 3.2 and 3.3 (perturbed field and category split): These describe approaches (iii) and (iv) in detail, but the labeling (i)–(iv) is only formally defined in Section 3.4, two subsections later. A reader encountering 3.2 doesn't yet know what simulation this corresponds to. Please define (i)–(iv) in one sentence at the opening of Section 3, or add a forward reference: "This approach corresponds to simulation (iii), defined in Section 3.4."
15. 3.5: Please describe how spatially correlated grid-cell uncertainties are aggregated to produce national-scale posterior uncertainties in the perturbed field approach.
16. L 243-244: isn't the spatial distinctness of mass increment and emission increment to be expected?: an assimilated observation might lead to updated emissions upwind and/or updated concentrations downwind.
17. 269: If the goal is to reduce boundary data biases, why not optimize boundary and/or initial conditions directly? There is no perturbation of initial CH₄ concentrations. Ensemble spread in concentrations therefore starts at zero and builds only from emission ensemble differences through transport. This could systematically limit early-cycle performance. This is effectively taken care of with the 5 day spin-up but is this period sufficient to transport through the full domain?
18. L 297-299: This appears to duplicate content from 4.1.1. Please specify whether Figure 10 (right) shows the combined effect of emission and concentration adjustments or just the latter?
19. Figure 11: Please add an assessment of how much mass in the domain was changed due to concentration and emission adjustments for (iii) and (iv), respectively. Currently the Figure shows the delta in mass between the two simulations.

20. The argument at L 311–312 is not entirely clear to me. How does one conclude that emissions within Germany are responsible for the reductions in AMI?
21. Fig 12: What does the dashed horizontal line represent?
22. L 399: I propose to change this to '...are close to the diagonal'
23. L 431: The two methods indeed agree well. The authors present this as reassuring for the approaches. However, it can also be interpreted as the system being highly dominated by observations and the choice of prior uncertainty (and to a lesser degree inversion setup) being irrelevant. Please comment on this.
24. L 467–468: above and below what?
25. 5.5: Please comment on the possible order-of-magnitude effect of including an OH sink.
26. 5.5 The month with the lowest (natural) emissions (May) is also the month with the lowest RSME. Please comment on the possibility of transport/modelling errors with a seasonal cycle.
27. 5.6 Please give an indication on how ICON's wet-dry inconsistencies affect modelled CH₄ and whether the assimilation of concentration could be able to prevent this from influencing posterior emissions.

Technical Correction

1. 3.1.1: add a reference to the values used/table 2
2. Line 243: figure number missing
3. Fig 7: the full-stop is missing at the end of the figure caption.
4. Fig 7: suggestion to add the prior uncertainty in the caption.
5. Fig 9: The abbreviations should be explained somewhere, e.g. in the form of a table in the appendix.
6. L 376: B25 and Bruch et al. (2026) refer to the same thing
7. L 400: should be '(KL)' instead of '(NL)' I presume