Review on Thanwerdas et al., egusphere-2025-5804

General comments

The manuscript investigates how expanding the atmospheric methane monitoring network in Italy could improve top-down emission estimates derived from inverse modelling. Using the ICON-ART Eulerian transport model and the Community Inversion Framework, the authors assess twelve network scenarios, including both existing stations and eight candidate sites, and evaluate their impact using two complementary evaluation metrics. To reduce random dependence on a single "emission truth," five synthetic emission scenarios are generated. In addition, they investigate two scenarios with varying observation uncertainty (model-data mismatch). Chieti and Mount Venda are identified as being the most useful additional stations to the existing ICOS network. The need for such study is increasing and timely as there is an urgent demand for improving national GHG budget monitoring. As the study presents, expanding monitoring network is one way, but the choice of stations should be done carefully with a very good reason to invest. The study presents a case for Italy, but methods are applicable for other countries, contributing to wide community in this field.

The study is well structured and clearly written. The manuscript is interesting, insightful and generally suitable for publication after addressing the comments outlined below.

- Introduction discusses a lot of anthropogenic emissions. However, the rest of the paper does not focus on it as much as you would assume based on the introduction. In addition, it is shown that the natural emissions are better constrained than the anthropogenic emissions, with smaller prior uncertainty in anthropogenic emissions than natural ones in your setup (Section 2.7). If the aim is to compliment the NGHGI, then more emphasis should be placed on improving the anthropogenic emissions. If that is not the main aim, then the introduction should also focus on natural emissions. In addition, as NGHGIs only report the anthropogenic emissions, the partition between anthropogenic and natural emission estimates of the inverse modelling is an important aspect.
- Do we understand correctly that you carried out all the simulations using both uniform and realistic model–data mismatch? What is the reasoning behind using the uniform model–data mismatch as the main results while you would use the realistic mdm in a "real" inversion run? The only argument given is "it allows us to isolate and evaluate the influence of station location on the inversion result". However, this approach neglects the

characteristics of station location and the transport model's ability to model the region near the station, which I would think is an important part of selecting the station. The main conclusion would not change, but I feel it's more natural to present the results from realistic mdm scenarios as the main results.

- Seasonal variation of the resulting metrics is discussed, but was slightly hard to comprehend without knowing the temporal variation of the prior fluxes in different categories and regions. Please add details of it in method section. Additional figures illustrating seasonality of fluxes would be helpful as well.

Specific comments

Introduction

- What is special in Italy, and what is scientific reasoning of you to study Italy specifically? Authors mention about some of anthropogenic emissions and also monitoring sites, but why do you think Italy has been poorly constrained? Anything to worry about natural emissions - earlier studies show strong influence of geological sources in inversion estimates, which makes it difficult to compare with bottom-up inventories.
- Why is the paper is specifically about expansion of the ICOS network and not expanding the measurement network in general? What is the benefit of having an ICOS station? Being an ICOS station means that the measurements have certain data quality criterions and that the data is shared openly and near-real-time etc. However, this is not mentioned or discussed.

Materials and methods

- What is the used temporal resolution of the prior fluxes?
- The largest anthropogenic sources are mentioned. How about the largest natural sources?
- It is mentioned with many candidate sites that they already are or were measuring atmospheric CH_4. Does this mean that they already have the needed infrastructure and getting them the ICOS status would only help the data users to access the data?
- Do you estimate only one set of scaling factors for the total emissions and then use those to estimate the uncertainty/error reductions of both the anthropogenic and natural emissions?
- Having 5 different truth scenarios to assess the effect of randomness on the result. Could you have done this differently? For example, multiplying the prior fluxes, so that you would have known that the true scaling factors are not randomly close to one. Is there any part of Italy that is suspected to change emissions in the future (due to anthropogenic or natural

factors), that needs better monitoring? Maybe a discussion on that could be added (either here or in the discussion/conclusion).

- Do you consider atmospheric sinks at all in the transport model?
- It would be helpful if you could add brief explanation of what prior emissions (especially of the natural ones) contribute to which part of Italy. Maybe a bar figure similar to Fig. 1 with breakdown of emissions could be added in Appendix. As you have poorest constraint in Southern Italy, what emission category needs better refinement (part of this could be in Introduction and Conclusion as well)?
- Could you possibly add maps of scaling or emissions for the 5 different emission scenarios (could be in Appendix or Supplementary)?
- L222-223: Authors discuss that sampling height can have influence on the results and conclusion of this study (Section 4), but here you do not give reasoning of the choice. Please add rational behind the choice. Are they somewhat realistic in terms of modelling and technical implementation of the station/measurements?

Results

- Are the error reduction values in Figure 4 and Section 3.1 calculated from only one emission scenario? If so, which one?
- Please consider adding maps of emission sensitivity to the stations or trajectory analysis. It would be helpful to comprehend your arguments.
- Section 3.5 focuses on refined mdms, and discussion on the differences in emissions and metrics between the scenarios are limited. Although overall argument do not change, I would like to understand better how the results are compared. For example, I've checked the relative differences in MERs between the mdm scenarios, and the case 2 had relatively small differences in MERs for all four regions on total fluxes (e.g. whole Italy = (17-20.7)/20.7 = -0.17), meaning either 1) uniform mdm setup was not a bad guess 2) stations have good power to constrain the fluxes even if mdm are increased. For Case 3 and 4, although CHI has much larger realistic mdm compared to VND, the relative differences in MERs seem to be smaller compared to VND (Table in the supplementary do not have the first digit, so my conclusion may not be valid, but you could check with more accuracy). Does this mean that CHI is extremely important considering poor constraint in southern Italy by existing ICOS stations; although we put large observation uncertainty, the information gain from the station is substantial? On the other hand, the relative MERs showed smallest

differences in central Italy for Case 3, rather than whole or northern Italy. So could this be used to emphasise the importance of that station to constrain emissions in central Italy? I was wondering this because both sites had realistic mdms above average, and certainly much larger than most of the ICOS stations, possibly meaning that the influence of local emissions are large and transport model has difficulties in representing the measured mole fraction variations, but still you consider them to be the most informative ones.

- L353: "POT is an exception, providing stronger coverage thanks to its central position."
  - I do believe, but you did not test without LMP, so Scenario 1 tells combination of POT and LMP. Please add a bit more explanation why you think LMP won't be contributing much to the reduction of MER.
  - POT is mentioned for the first time in the text here. Please add full name as well (see also technical comment).
- L392-395: Broadly speaking, do you think the differences in spatial distribution of anthropogenic and natural fluxes are good enough in Italy to examine them separately even if not using isotope information? I see that at least hotspots are not overlapping very much (Figure 1), but there are regions with possibly equally high emissions from both categories.
- L425-426: Are wetland emissions largest contributor to the Italy's natural emissions or their seasonal cycle? Does any of anthropogenic emissions have sub-annual variations? As you discuss seasonality, it would be helpful to add that information in Section 2.3.1 (see also general comments). I suppose seasonality of natural emissions are different in different parts of Italy. Could this explain part of differences in Fig. 7?

Discussion

- Authors mention about computational cost for extra simulation to examine the validity of this study. Do you think methods such as machine learning or emulators could help on this aspect – for example, you could use the results of this study to train such model and calculate the similar metrics for different years / sampling height / random perturbations of true fluxes? Or, do you think those methods cannot give conclusions to such sensitive study? I understand that this is totally out of the scope of this study, but could you kindly share me your opinion? If you think they could be of help, could you consider adding speculation of that in conclusion?

- As authors mention, sampling height could have strong influence on the conclusion of this study. Do you think model representation would improve if you used higher altitudes? How would you best optimize the sampling height for the planned sites?
- L529-533: Do you think having similar results to Villalobos et al. (2025) also indicates the potential of those sites in constraining natural fluxes rather than fossil fuel related emissions?
- L533-540: Please phrase this carefully considering our point in general comment regarding constraining anthropogenic fluxes.

Conclusions

- Would you have any concrete recommendations how to perform an OSSE to select good candidates for atmospheric measurement stations?
- Could you comment on how to better monitor / constrain anthropogenic fluxes and their possible changes in Italy? Are we urgent in adding stations?
- Adding good high quality ICOS-type stations is best to improve the flux constraint, but it is demanding. Could you comment if satellite data could be a complementary choice? You mentioned that carrying OSSE with Eulerian-based model can be done "without additional computational burden". So do you think in the future, it is worthwhile comparing the performance / information gain with satellite-based inversions?

Technical comments

Whenever you mention station codes in the text for the first time, please add full names. Although the information can be found in the table, it'd be reader friendly.

L10: "multiple "emission truth" scenarios"
The phrase is understandable for those who have read the manuscript, but may not be for those who only reads the abstract. Could it just be "multiple emission scenarios"? I think putting the word "truth" and "scenarios" together makes it a bit confusing.

L20-21: "Currently, countries report their emissions using bottom-up inventories, which are based on 1) socioeconomic and environmental data and 2) source-specific emission factor." This is probably the way it is mostly done, but NGHGIs also use other methods, such as modelling LULUCF related emissions.

L29: “Top-down methods provide independent, consistent, and timely information”.
What is meant with consistent and timely information?

L31-33: “Observational in-situ networks should therefore be carefully designed to minimize uncertainties and improve our ability to track the temporal evolution of GHG fluxes at regional and national scales.”
Minimize uncertainties of what? The above mentioned NGHGIs or the total CH_4 balance?

L53-54: “More stations, introduced in Sect. 2.4, are also monitoring CH_4 in Italy, although they are not part of the ICOS network.”
Is the data available from these sites? Although information is written in Section 2.4, please also add here briefly about the data availability.

L73: “Most of the Eulerian models used in the inversion community have now been coupled to CIF”.
Can you name them?

L144: “represent the soil uptake emissions from peatlands and inundated wetlands”. Probably should be “soil uptake **and** emissions”

L165: “is expected to begin measuring GHGs in 2025–2026.”
Did it already or is it still expected to begin measuring in 2026?

L220: Blue **and yellow** circles

L302: “NUTS (Nomenclature of Territorial Units for Statistics)” -> Nomenclature of Territorial Units for Statistics (NUTS)

L307-308: “Reductions in domain-total flux error can be misleading, as they may result from compensating errors across spatial or temporal domains.”
Can you clarify this sentence?

L310: Figures 2c and 2d -> Figures **3**c and **3**d

L346-347: “While Northern and Central Italy shows similar results under the ideal scenarios, Southern Italy remains less constrained.”
What is meant with “similar results”? For me, the MERs do not seem more similar between Northern and Central Italy than between Southern Italy.

L394-395: “When natural and anthropogenic emissions are co-located and occur simultaneously, the optimization process cannot effectively separate them, resulting in poor agreement between the posterior and true fluxes.”
Poor agreement between the fluxes of different sectors or the total fluxes?

L458-459: “we recalculate the RMSE, which serves as a refined estimate of the model–data mismatch for these stations.”
Does this mean that you recalculate RMSE based on posterior mole fractions, and rerun inversion after refining RMSE, so first inversion is not used for analysis? Please clarify.

L469-474: Do I understand correctly that this section mostly talk about the model–data mismatches? As the sentence “As a result, both MER...” is about results of the simulation and not about the mdms themselves, I would put it in the end, before the sentence “Nevertheless,...”.

L501-505: I think there is also value in studying only the location that would be feasible, i.e., whether it is realistic to get all the infrastructure needed for the station.

L506-513: Is the only advantage of using Eulerian model in this study that it could be used with satellite data in the coming studies?

Tables and Figures

Figure 1: Is this calculated from the prior datasets? Please specify in the caption.

Figure 2 caption: Could you anyway add reference to Table A2 for the names of the ecosystem stations?

Figure 4:

- Please change ‘case’ to ‘scenarios’ for consistency.
- “Stations from Scenario 1 are marked with yellow circles” Except for the Scenario 1 panel, I think they are for Scenario 2. Please revise.

Figure 5: Having both blue and green markers for scenarios and then the “mean over truth scenarios” is a bit confusing. I suggest changing the colours.

Table A1 and related text in section 2.4. For ECO and LMT, the current status or plan are missing. Please consider adding information for all the sites, and possibly in Table A1 as well. For active sites, please also indicate if they are measuring any GHGs at current.

Table A3 caption: may need a bit of explanation about “fake inlet height”.

Supplementary: you probably need to add title of to the document.