Empa
Materials Science and Technology

# Exploring atmospheric $CH_4$ monitoring network expansion in Italy using inverse modelling

Joel Thanwerdas[1], Paolo Cristofanelli[2], Angela Fiore[3], Rianne Droge[4], Sophie Van Mil[4], Yohanna Villalobos[5], Zhendong Wu[6], and Dominik Brunner[1]

[1]Empa, Swiss Federal Laboratories for Materials Science and Technology, Dubendorf, Switzerland

[2]National Research Council of Italy—Institute of Atmospheric Sciences and Climate (CNR-ISAC), Bologna, Italy

[3]Italian Institute for Environmental Protection and Research (ISPRA), Rome, Italy

[4]TNO, Air Quality and Emissions Research, Utrecht, Netherlands

[5]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

[6]ICOS Carbon Portal, Department of Physical Geography and Ecosystem Sciences, Lund University, Lund, Sweden

# Response to referees' comments

We thank the two referees for their invaluable insights, which have greatly enhanced the quality of the paper. Here, we provide a comprehensive response to the comments received. Referee#1's comments are in blue and Referee#2's comments are in red. For each comment, an answer is provided in normal text and the **modifications from the new version of the manuscript are provided in bold and small text**. Note that modifications have been included only when deemed substantial enough. Attached to this response, we also provide a new version of the manuscript and a track-changes document.

Before addressing the referees' comments, we would like to report two writing errors related to inlet heights that were not identified in the initial version of the manuscript. These are writing errors only and do not affect the results or conclusions. However, we consider it important to clarify them.

The inlet height at CHI was set to 50 m, consistent with the value used in Villalobos et al. (2025), which corresponds to the projected configuration at the time of the study, before the station started operating. The inlet height at VND is 100 m, as the 142 m tower installed there easily allows sampling at this level. However, we incorrectly reported these values in reverse in the first version of this manuscript. There was also a text error in Table A1: the inlet height for CGR is 100 m, not 2 m, as this is a lowland site. We deeply apologize for not identifying these errors before the submission.

## General comments

- 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.

We agree with the referee that this aspect was not well presented or sufficiently discussed in the manuscript. This comment is partly addressed in several specific responses below; here, we provide only a consolidated summary of the responses.

First, we acknowledge that the manuscript places too much emphasis on anthropogenic emissions. This focus arises from the fact that a primary objective of inversion frameworks is to better constrain anthropogenic emissions to inform effective mitigation policies. However, constraining natural emissions is equally important because, in regions where anthropogenic and natural emissions are co-located, it is not possible to accurately estimate anthropogenic emissions without first adequately constraining natural sources. Italy is a good example of such a region, as it exhibits both significant anthropogenic and natural emissions. We have clarified this point in the introduction.

Our methodology is designed to assess whether individual emission categories can be effectively constrained. However, the results indicate limited success, particularly for anthropogenic emissions. In contrast, natural emissions appear to be better constrained. As discussed in the manuscript, this outcome is largely due to the use of realistic spatial error characterizations. Anthropogenic emissions are associated with smaller error correlation lengths, as suggested by recent literature, reflecting the higher spatial heterogeneity of their sources. This increases the number of degrees of freedom, making them inherently more difficult to constrain.

In this work, we aim to identify the optimal station that could improve these constraints. While we emphasize this objective, we acknowledge that we initially underemphasized an important complementary finding: that the addition of a single station does not substantially improve constraints. We have now strengthened the discussion on this point, explicitly highlighting that constraints on anthropogenic emissions remain weak.. Furthermore, we stress that simply expanding the observational network is insufficient. Meaningful improvements will likely come from the integration of additional sources of information, such as isotopic measurements, co-emitted species, and satellite data, which should be further evaluated in similar OSSE frameworks. We have also this conclusion in the abstract.

Because this comment concerns several parts of the manuscript, we direct the referee to the corresponding specific comments, where the associated revisions are described in detail.

- 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.

Yes, we performed all our inversions using both uniform and realistic mdm. We agree that the rationale behind this choice was not clear enough in the original text. Initially, we performed inversions with a uniform mdm for two main reasons. First, as stated in the manuscript, a uniform mdm allows us to isolate and quantify the influence of station location on the inversion results, independently of assumptions related to model–data mismatch. In other words, it provides a consistent and fair basis for comparing candidate stations, which would not be possible with a realistic mdm that varies across locations. Second, the interpretation of the results presented in this study is inherently complex. For this reason, we chose to introduce uniform mdm results first, as they offer a clearer and more straightforward framework for comparison, before incorporating the additional complexity associated with a realistic mdm. We fully acknowledge that this simplification neglects site-specific characteristics as well as the transport model's ability to represent regional processes. However, there is no guarantee that the mdm estimated using our methodology will accurately reflect real conditions once a station becomes operational. In practice, experience shows that trying to predict how well a model will reproduce observations from a new station is highly challenging. In this context, presenting a uniform mdm first represents a conservative choice that provides a relatively good baseline for station selection. It is not sufficient on its own, which is why we present results from both approaches. The fact we obtain similar results with a realistic mdm suggests that the conclusions are robust.

We have added these explanations to the text.

**We chose to introduce results with a uniform model-data mismatch first, before incorporating the additional complexity associated with a realistic model-data mismatch. This choice serves several purposes in the initial phase of our study. First, it provides a consistent and fair basis for comparing candidate stations, which would not be possible with a realistic model-data mismatch that varies across locations. Second, there is no guarantee that the model-data mismatch estimated using our methodology will accurately reflect real conditions once a station becomes operational. In practice, experience shows that trying to predict how well a model will reproduce observations from a new station is highly challenging. Finally, it facilitates the interpretation of the results. Using a uniform model-data mismatch is not sufficient on its own, which is why we present results from both approaches.**

- 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.

We have updated the flux descriptions to include more detail regarding their temporal variation. To support this, we have added **Figure S2** to the Supplementary Information illustrating the seasonal cycles of various emission subcategories for each region. We have also added a sentence to introduce this figure in Sect. 2.3.1.

We also wish to clarify that our analysis of seasonal metrics is performed because we generate a unique set of scaling factors for each 10-day window, using a temporal correlation that follows an exponential decay with a three-month decorrelation length. It is already explained here in Sect. 2.7:

**To account for a potential temporal variability of the mismatch between prior and true estimates, we generate a new set of true scaling factors for each 10-day period throughout the year 2018. To maintain seasonal coherence, we impose a temporal correlation using an exponential decay with a temporal correlation length of three months.**

Because these factors evolve continuously, our primary objective is to evaluate how effectively the system can reduce the mismatch between the prior and true seasonality. These clarifications have been integrated into the revised manuscript, at the beginning of Sect. 3.4.

**As explained in Sect. 2.7, different sets of true scaling factors were generated for successive 10-day windows for both natural and anthropogenic emissions to account for a potential temporal variability in the associated uncertainties. This introduces a 'true' seasonality in anthropogenic, natural, and total emissions that differs from the prior. We assess the inversion's ability to recover this true seasonality by analyzing the seasonal behavior of MER and TUR.**

One other important point is that we substantially revised Section 3.4, which focuses on the seasonality of the metrics, because we realized that our previous interpretation of the results was incorrect. The results for MER and TUR should not be influenced by the seasonality of prior emissions, as the seasonality of true emissions differs from that of prior emissions, as explained above. When analyzing a single truth scenario, the inversion system can more effectively attribute and locate underlying sources when (1) the station captures the emission signal and (2) the resulting signal difference between prior simulated values and pseudo-observations, caused by discrepancies between prior and true emissions, exceeds the model–data mismatch. Larger signal differences occur when true and prior emissions differ substantially, rather than when prior emissions alone are larger. For example, if the discrepancy between prior and true emissions peaks during winter and the signal can be adequately captured by the stations, a higher MER will be obtained during that season

However, it is important to emphasize that this mechanism is neutralized when evaluating performance across multiple truth scenarios. Since each scenario exhibits a distinct seasonal profile, averaging performance across scenarios removes the systematic influence of individual mismatches between prior and true emissions. This process instead leaves only the influence of other drivers, such as the seasonality of wind speed, wind direction, and convection. Consequently, we believe that the seasonality of prior emissions does not influence our overall results, and we have completely revised our interpretation accordingly.

We apologize for this error in interpretation and thank the referee, whose comments helped us identify and correct this issue. We refer to the revised manuscript for the updated section.

- It would be good if you could highlight why you chose Italy as a case study and reflect on the limitations of the current state of the network in constraining $CH_4$ emission in this country to underline the need for additional sites. So far, this has only been briefly mentioned in the abstract 'Compared to other countries, Italy remains poorly covered by ICOS atmosphere sites […]', yet I would assume that other countries within

ICOS face equally severe, if not more severe, limitations. Regarding the current state of the ICOS network, I think an important point worth mentioning that limits the estimations of CH4 emissions in Italy is not only the number of the sites, but also the high altitudes at which many of the stations are located, both within and close to Italy.

We agree that this point is not sufficiently explained in the manuscript.

The current ICOS network over Italy includes five stations: PRS, CMN, IPR, LMP, and POT, each with distinct characteristics and limitations. PRS and CMN have been measuring atmospheric $CH_4$ in Northern Italy since 2005 and 2008, respectively. As pointed out by the referee, they are mountain stations and primarily sample background air. As such, they are particularly valuable for constraining background concentrations, but they have limited sensitivity to regional emissions compared to continental lowland stations such as POT. The latter recently began monitoring $CH_4$ (Lapenna et al., 2025) and is expected to provide valuable coverage in Southern Italy in the future. However, it remains difficult to assess how well observations at this location can be reproduced by transport models. IPR has been measuring since 2017, but is located in a valley surrounded by complex terrain, making it challenging to simulate accurately. For this reason, a relatively large model–data mismatch is typically assumed for this site (see Table S1). LMP is a marine remote site located on Lampedusa island, deep in the Mediterranean sea and distant from continental Italy, monitoring since 2008. While it provides useful information on background concentrations, its ability to constrain Italian emissions is inherently limited due to its distance from continental sources. These various limitations help explain why top-down estimates over Italy often show substantial discrepancies (see e.g., Ioannidis et al., 2026) and motivate our focus on Italy to improve the observational network.

We have added these explanations to the manuscript, in the introduction.

**Complementing these estimates with inverse modelling requires the observational network in Italy to provide a good coverage of the country. At present, five ICOS sites are monitoring continuously CH4 in Italy: Plateau Rosa (PRS; 45.9◦ N, 7.7◦ E), Ispra (IPR; 45.8◦ N, 8.6◦ E), Potenza (POT; 40.6◦ N, 15.7◦ E), Monte Cimone (CMN; 44.2◦ N, 10.7◦ E) and Lampedusa (LMP; 35.5◦ N, 12.6◦ E). PRS and CMN have been measuring atmospheric $CH_4$ in Northern Italy since 2005 and 2008, respectively. They are mountain stations and primarily sample background air. As such, they are particularly valuable for constraining background concentrations, but they have limited sensitivity to regional emissions compared to continental lowland stations such as POT. The latter recently began monitoring $CH_4$ (Lapenna et al., 2025) and is expected to provide valuable coverage in Southern Italy in the future. However, it**

**remains difficult to assess how well observations at this location can be reproduced by transport models. IPR has been measuring since 2017, but is located in a valley surrounded by complex terrain, making it challenging to simulate accurately. LMP is a marine remote site located on Lampedusa island, deep in the Mediterranean sea and distant from continental Italy, monitoring since 2008. While it provides useful information on background concentrations, its ability to constrain Italian emissions is inherently limited due to its distance from continental sources. These various limitations help explain why top-down estimates over Italy often show significant discrepancies (see e.g., Ioannidis et al., 2026) and motivate our focus on Italy to improve the observational network.**

- It would be really helpful if you added more references to the figures you are describing, and checked that you are referencing the correct ones (see more in the specific comments). For example, you should label the sub-figures of Figures 5 and 6 (e.g. a, b, c, d) and refer to them in the text; otherwise, it is difficult to identify which sub-figure is being referenced and comprehend the described conclusions. For instance, on page 17, you write, 'However, while averages over the truth scenarios suggest that ECO, LMT, and CGR generally outperform VND in the South, a notable fraction of the realizations (20–40%) indicate better performance for VND.', and I think you are referring to figures 5 and 6, 'Southern Italy'. However, it is often unclear where the reader should look for the evidence presented in the text. The same goes for Figure 7, which is not mentioned at all in the corresponding section.

We agree and apologize for not facilitating the reading further in the first version. We have added labels in most figures and better referenced the figures in the text.

# 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.

We agree that this point is not sufficiently explained in the manuscript.

The current ICOS network over Italy includes five stations: PRS, CMN, IPR, LMP, and POT, each with distinct characteristics and limitations. PRS and CMN have been measuring atmospheric $CH_4$ in Northern Italy since 2005 and 2008, respectively. They are mountain stations and primarily sample background air. As such, they are particularly valuable for constraining background concentrations, but they have limited sensitivity to regional emissions compared to continental lowland stations such as POT. The latter recently began monitoring $CH_4$ (Lapenna et al., 2025) and is expected to provide valuable coverage in Southern Italy in the future. However, it remains difficult to assess how well observations at this location can be reproduced by transport models. IPR has been measuring since 2017, but is located in a valley surrounded by complex terrain, making it challenging to simulate accurately. For this reason, a relatively large model–data mismatch is typically assumed for this site (see Table S1). LMP is a marine remote site located on Lampedusa island, deep in the Mediterranean sea and distant from continental Italy, monitoring since 2008. While it provides useful information on background concentrations, its ability to constrain Italian emissions is inherently limited due to its distance from continental sources. These various limitations help explain why top-down estimates over Italy often show substantial discrepancies (see e.g., Ioannidis et al., 2026) and motivate our focus on Italy to improve the observational network.

We have added these explanations to the manuscript, in the introduction.

**Complementing these estimates with inverse modelling requires the observational network in Italy to provide a good coverage of the country. At present, five ICOS sites are monitoring continuously $CH_4$ in Italy: Plateau Rosa (PRS; 45.9◦ N, 7.7◦ E), Ispra (IPR; 45.8◦ N, 8.6◦ E), Potenza (POT; 40.6◦ N, 15.7◦ E), Monte Cimone (CMN; 44.2◦ N, 10.7◦ E) and Lampedusa (LMP; 35.5◦ N, 12.6◦ E). PRS and CMN have been measuring atmospheric $CH_4$ in Northern Italy since 2005 and 2008, respectively. They are mountain stations and primarily sample background air. As such,**

**they are particularly valuable for constraining background concentrations, but they have limited sensitivity to regional emissions compared to continental lowland stations such as POT. The latter recently began monitoring $CH_4$ (Lapenna et al., 2025) and is expected to provide valuable coverage in Southern Italy in the future. However, it remains difficult to assess how well observations at this location can be reproduced by transport models. IPR has been measuring since 2017, but is located in a valley surrounded by complex terrain, making it challenging to simulate accurately. LMP is a marine remote site located on Lampedusa island, deep in the Mediterranean sea and distant from continental Italy, monitoring since 2008. While it provides useful information on background concentrations, its ability to constrain Italian emissions is inherently limited due to its distance from continental sources. These various limitations help explain why top-down estimates over Italy often show significant discrepancies (see e.g., Ioannidis et al., 2026) and motivate our focus on Italy to improve the observational network. Additional stations, introduced in Sect. 2.4, also monitor $CH_4$ in Italy. However, these sites are not part of the ICOS network. Their measurements are not necessarily continuous, and data access is not public, requiring direct requests to the respective Principal Investigators (PIs).**

In addition, we have modified the sentence in the abstract about the weak constraint over Italy.

**Top-down estimates often show substantial discrepancies over this country, suggesting that the current network provides weak observational constraints on methane fluxes.**

Additionally, too much emphasis has been placed on anthropogenic emissions. Accurately estimating natural emissions in Italy is equally important, as bottom-up estimates suggest they contribute substantially to total emissions. Reliable estimates of anthropogenic emissions cannot be obtained without adequately constraining natural emissions. For a more comprehensive discussion, we refer to other dedicated comments.

- 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.

In the inversion community, ICOS data are widely relied upon for several key reasons:

- Instrumentation is highly harmonized across sites
- Measurements are calibrated against a common reference scale
- Measurements follow strict protocols for calibration, sampling, and data processing
- Data undergo both automated and expert quality control
- Data are open-access and provided in standardized formats

> Data are typically accessible with a delay of approximately one day for near-real-time products (ICOS Research Infra-structure, 2018), whereas quality-checked fast-track data releases become available after a delay of a few months (ICOS RI et al., 2026).

Together, these features ensure direct comparability between observations, low observational uncertainties, long-term consistency, and easy access and integration into inversion systems.

We have added these explanations to the manuscript, in the introduction.

**Today, the Atmosphere network consists of 47 stations (39 labeled and 8 candidate stations) measuring GHG atmospheric concentrations across Europe. In the inversion community, ICOS data are widely relied upon for several key reasons:**

**– Instrumentation is harmonized across sites**

**– Measurements are calibrated against a common reference scale**

**– Measurements follow strict protocols for calibration, sampling, and data processing**

**– Data undergo both automated and expert quality control**

**– Data are open-access and provided in standardized formats**

**– Data are typically accessible with a delay of approximately one day for near-real-time products (ICOS Research Infrastructure, 2018), whereas quality-checked fast-track data releases become available after a delay of a few months (ICOS RI et al., 2026).**

**Together, these features ensure direct comparability between observations, low observational uncertainties, long-term consistency, and easy access and integration into inversion systems, as opposed to data from stations that are not part of the ICOS network.**

- P2, L36: I agree that the ICOS network is very advanced, still I think that the term 'world-class' is too subjective for a scientific publication.

We agree with this comment and have removed the adjective.

## Materials and methods

- What is the used temporal resolution of the prior fluxes?

The prior fluxes used in this study have varying temporal resolutions, **as already documented in Section 2.3.1.** Natural fluxes from LPJ-GUESS are provided at daily resolution, while other natural fluxes from the Global Methane Budget are available at monthly resolution. Biomass burning emissions from GFED are also provided at monthly resolution, whereas anthropogenic emissions from the TNO-AVENGERS dataset are available at hourly resolution. To ensure consistency, all fluxes are resampled to an hourly resolution prior to spatial remapping and aggregation.
We have added a paragraph at the end of Section 2.3.1.

**Natural and anthropogenic datasets are resampled to hourly resolution, conservatively remapped to the ICON-ART spatial grid using the Emiproc package (Lionel et al., 2025) and aggregated to produce prior estimates of natural and anthropogenic emissions over Italy.**

- The largest anthropogenic sources are mentioned. How about the largest natural sources?

**We have included Figure S1 in the Supplementary Information** that illustrates the contribution of each category to total anthropogenic and natural emissions for each region. Geological emissions largely dominate the prior natural emissions across all regions of Italy, followed by emissions from wetlands. We have also added a paragraph in Section 2.3.1.

**According to these estimates, geological emissions largely dominate the prior natural emissions across all regions of Italy (ranging from 70% in Southern Italy to 86% in Central Italy), followed by emissions from wetlands (see Fig. S1).**

- 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?

We agree that this point is not clearly explained in the manuscript. For stations already measuring $CH_4$, the ICOS label would require compliance with ICOS standards for measurement systems, quality-assurance protocols, as well as the adoption of a common data format for the near-real-time transmission of instrumental and diagnostic raw data to the ICOS Atmospheric Thematic Center (Hazan et al., 2016). Moreover, a sustained long-term commitment from national funding agencies is required to support the operational costs of the stations, as well as the national contributions required to ensure the operation of ICOS-labelled stations and their integration within the ICOS Research Infrastructure. This represents a substantial effort compared to simply helping data users access the data.

Regarding the required infrastructure, the main constraint was to remain consistent with the inlet heights assumed in Villalobos et al. (2025) in that study, a low inlet height of 2 m is used for mountain candidate sites, consistent with the configuration of stations that have previously measured $CH_4$ (CUR, MDN). A good sampling at existing ICOS sites is typically performed at around 100 m above ground level (see Table A2). It generally allows to reduce local influences and obtain a good sensitivity to regional fluxes. At present, however, candidate stations that have previously measured $CH_4$ (ECO, LMT, CGR) sample below 100 m. Reaching such sampling heights would therefore require the installation of tall towers. Although this represents a substantial investment, implementation is expected to be more feasible at these locations because they already possess relevant infrastructure and experienced personnel conducting $CH_4$ measurements. This consideration motivated their selection as candidate sites.

We have included both these explanations in Sections 2.4 and 2.6, respectively.

Section 2.4:

**ECO, LMT, CGR, CUR, MDN, and MRG are not part of ICOS but contribute to the regional Global Atmospheric Watch (GAW) programme of the World Meteorological Organization (WMO), which motivated their selection as candidate sites. Further information about these stations is provided in Table A1. Locations are displayed on a spatial map in Fig. 2 (e.g., Scenario 11). For stations already measuring $CH_4$, the ICOS label would require compliance with ICOS standards for measurement systems, quality-assurance protocols, as well as the adoption of a common data format for the near-real-time transmission of instrumental and diagnostic raw data to the ICOS Atmospheric Thematic Center (Hazan et al., 2016). Moreover, a sustained long-term commitment from national funding agencies is required to support the operational costs of the stations, as well as the national contributions required to ensure the operation of ICOS-labelled stations and their integration within the ICOS Research Infrastructure.**

Section 2.6:

**Regarding the required infrastructure, the main constraint was to remain consistent with the inlet heights assumed in Villalobos et al. (2025). In that study, a low inlet height of 2 m is used for mountain candidate sites, whereas a higher inlet height of 100 m is used for lowland candidate sites. Note that CHI is an exception, both here and in Villalobos et al. (2025), as it is a newly established station. At the start of this study, the projected inlet height was 50 m. We retained this value in our analysis because the installation of a substantially higher inlet in the coming years is considered unlikely.**

**The rationale behind the selection of 2 m and 100 m inlet heights is that model performance generally improves with increasing sampling altitude, provided that the surrounding terrain is not excessively complex and does not generate strong local flow disturbances. Higher sampling altitudes reduce the influence of very local processes, which require high spatial resolution to be accurately represented. Elevated sampling also increases the spatial footprint of the measurements, allowing observations to integrate signals from more distant source regions and thereby increasing sensitivity to large-scale flux patterns. However, sampling at excessively high altitudes, such as mountain stations located above the boundary layer, also has disadvantages. Air masses sampled at these altitudes are well mixed, making it more difficult to disentangle contributions from individual source regions. Consequently, such measurements are generally more effective at constraining background concentrations than regional emissions.**

**For lowland sites, a good sampling at existing ICOS sites is typically performed at around 100 m above ground level (see Table A2). This height helps reduce the influence of local processes, and obtain a good sensitivity to regional fluxes, while limiting costs and engineering challenges. At present, however, candidate stations that have previously measured $CH_4$ (ECO, LMT, CGR) sample below 100 m. Reaching such sampling heights would therefore require the installation of tall towers. Although this represents a substantial investment, implementation is expected to be more feasible at these locations because they already possess relevant infrastructure and experienced personnel conducting $CH_4$ measurements. This consideration motivated their selection as candidate sites.**

**For mountain sites, nighttime measurements are generally representative of the free troposphere, even when sampled close to the surface, owing to the elevated location of the stations. As a result, a sampling height of 2 m is sufficient, and this requirement is already met by existing mountain stations performing $CH_4$ measurements (MDN and CUR).**

- 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?

No, we use a different set of scaling factors for each category of emissions (natural and anthropogenic), which allows us to calculate the metrics independently for each category. We clarified this point in Section 2.7.

**Distinct sets of true scaling factors are generated for anthropogenic and natural emission categories, and true total fluxes are computed as the sum of the true anthropogenic and natural fluxes.**

- 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.

There are several possible approaches to perturb prior fluxes. While the use of scaling factors already implies modifying the prior fluxes, we assume that the referee is referring to the application of a spatially uniform scaling factor. Although this could be an option, it would be inconsistent with the error statistics (i.e., variance and spatial correlations) assumed for the prior fluxes and prescribed in the inversion framework.

Importantly, note that there is always a possibility that the prior fluxes are close to the true state. Imposing a uniform scaling factor different from one would systematically exclude this scenario. In contrast, using randomly generated perturbations allows us to construct synthetic "true" fluxes that remain fully consistent with the assumed prior error statistics.

We therefore argue that this approach is better suited to assess the ability of the inversion system to recover realistic fluxes. As explained in the manuscript, the main limitation is the influence of randomness, but this effect is mitigated by considering an ensemble of scenarios, which ensures robust and meaningful results.

We have added a sentence in Section 2.7.

**A major advantage of this approach is that it generates true fluxes consistent with the error statistics of the prior fluxes.**

- 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).

Italy signed the Global Methane Pledge at the 26th Conference of the Parties (COP26) in November 2021, committing to voluntary actions aimed at achieving a collective reduction in global methane emissions of at least 30% from 2020 levels by 2030. Based on Caputo et al. (2022), the largest reductions in $CH_4$ emissions in Italy are expected in the waste sector, driven by evolving waste legislation, improved waste management practices, and increased treatment of waste in mechanical–biological and composting plants, as well as in anaerobic digesters. The second-largest reductions are projected in the agricultural sector, primarily due to more sustainable manure management practices (particularly anaerobic digestion with biogas production) and a decline in cattle and swine populations. In contrast, only limited changes are currently anticipated for natural emissions. While wetland emissions (mainly located in southern regions) may respond to climate change (e.g., Koffi et al., 2020; Zhang et al., 2017), they represent a relatively small fraction of total methane emissions compared to other sources, such as geological emissions, which are not expected to significantly change in the next decades.

We have added this information in the introduction, but do not consider a more detailed discussion necessary, as emissions from the waste and agriculture sectors are relatively dispersed across the country and do not make the addition of stations in specific regions more relevant in our opinion.

**Italy signed the Global Methane Pledge at the 26th Conference of the Parties (COP26) in November 2021, committing to voluntary actions aimed at achieving a collective reduction in global methane emissions of at least 30% from 2020 levels by 2030. Based on Caputo et al. (2022), the largest reductions in $CH_4$ emissions in Italy are expected in the waste sector, driven by evolving waste legislation, improved waste management practices, and increased treatment of waste in mechanical–biological and composting plants, as well as in anaerobic digesters. The second-largest reductions are projected in the agricultural sector, primarily due to more sustainable manure management practices (particularly anaerobic digestion with biogas production) and a decline in cattle and swine populations. In contrast, only limited changes are currently anticipated for natural emissions. While wetland emissions (mainly located in southern regions) may respond to climate change (e.g., Koffi et al., 2020; Zhang et al., 2017), they represent a relatively small fraction of total methane emissions compared to other sources, such as geological emissions, which are not expected to significantly change in the next decades.**

- Do you consider atmospheric sinks at all in the transport model?

We do not consider atmospheric sinks as the air (along with emissions) is flushed out from our European domain in less than 20 days. We therefore neglect the effects of $CH_4$ oxidation in the atmosphere, as typically done in European inversions.

We have added this explanation in Section 2.1.

**We do not consider atmospheric sinks as the air (along with emissions) is flushed out from our European domain in less than 20 days. We therefore neglect the effects of $CH_4$ oxidation in the atmosphere, as typically done in European inversions.**

- 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)?

We have added Fig. S1 to the Supplementary Information, showing the contribution of emission subcategories to total natural and anthropogenic emissions in each region. According to these estimates, geological emissions largely dominate the prior natural emissions across all regions of Italy (ranging from 70% in Southern Italy to 86% in Central Italy), followed by emissions from wetlands. In Southern Italy, the largest contributions to total emissions arise from the waste and agriculture sectors. A similar pattern is observed in other regions, although the influence of waste emissions is particularly pronounced in the South. Consequently, waste and agriculture should be the primary targets for improved constraints on anthropogenic emissions. We believe additional text addressing this comment has already been included in response to other comments.

- Could you possibly add maps of scaling or emissions for the 5 different emission scenarios (could be in Appendix or Supplementary)?

We have added Figure S3 in the Supplementary Information showing the scaling factors (total, anthropogenic and natural) for the five emission scenarios, averaged over the year.

- 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?

The rationale behind the selection of 2 m and 100 m inlet heights is that model performance generally improves with increasing sampling altitude, provided that the surrounding terrain is not excessively complex and does not generate strong local flow disturbances. Higher sampling altitudes reduce the influence of very local processes, which require high spatial resolution to be accurately represented. Elevated sampling also increases the spatial footprint of the measurements, allowing observations to integrate signals from more distant source regions and thereby increasing sensitivity to large-scale flux patterns. However, sampling at excessively high altitudes, such as mountain stations located above the boundary layer, also has disadvantages. Air masses sampled at these altitudes are well mixed, making it more difficult to disentangle contributions from individual source regions. Consequently, such measurements are generally more effective at constraining background concentrations than regional emissions.

For lowland sites, a good sampling at existing ICOS sites is typically performed at around 100 m above ground level (see Table A2). This height helps reduce the influence of local processes, and obtain a good sensitivity to regional fluxes, while limiting costs and engineering challenges. At present, however, candidate stations that have previously measured $CH_4$ (ECO, LMT, CGR) sample below 100 m. Reaching such sampling heights would therefore require the installation of tall towers. Although this represents a substantial investment, implementation is expected to be more feasible at these locations because they already possess relevant infrastructure and experienced personnel conducting CH4 measurements. This consideration motivated their selection as candidate sites.

For mountain sites, nighttime measurements are generally representative of the free troposphere, even when sampled close to the surface, owing to the elevated location of the stations. As a result, a sampling height of 2 m is sufficient, and this requirement is already met by existing mountain stations performing $CH_4$ measurements (MDN and CUR).

This explanation has been added to Section 2.6, also addressing a related comment on infrastructure requirements.

**Regarding the required infrastructure, the main constraint was to remain consistent with the inlet heights assumed in Villalobos et al. (2025). In that study, a low inlet height of 2 m is used for mountain candidate sites, whereas a higher inlet height of 100 m is used for lowland candidate sites. Note that CHI is an exception, both here and in Villalobos et al. (2025), as it is a newly established station. At the start of this study, the projected inlet height was 50 m. We retained this value in our analysis because the installation of a substantially higher inlet in the coming years is considered unlikely.**

**The rationale behind the selection of 2 m and 100 m inlet heights is that model performance generally improves with increasing sampling altitude, provided that the surrounding terrain is not excessively complex and does not generate strong local flow disturbances. Higher sampling altitudes reduce the influence of very local processes, which require high spatial resolution to be accurately represented. Elevated sampling also increases the spatial footprint of the measurements, allowing observations to integrate signals from more distant source regions and thereby increasing sensitivity to large-scale flux patterns. However, sampling at excessively high altitudes, such as mountain stations located above the boundary layer, also has disadvantages. Air masses sampled at these altitudes are well mixed, making it more difficult to disentangle contributions from individual source regions. Consequently, such measurements are generally more effective at constraining background concentrations than regional emissions.**

**For lowland sites, a good sampling at existing ICOS sites is typically performed at around 100 m above ground level (see Table A2). This height helps reduce the influence of local processes, and obtain a good sensitivity to regional fluxes, while limiting costs and engineering challenges. At present, however, candidate stations that have previously**

**measured $CH_4$ (ECO, LMT, CGR) sample below 100 m. Reaching such sampling heights would therefore require the installation of tall towers. Although this represents a substantial investment, implementation is expected to be more feasible at these locations because they already possess relevant infrastructure and experienced personnel conducting $CH_4$ measurements. This consideration motivated their selection as candidate sites.**

**For mountain sites, nighttime measurements are generally representative of the free troposphere, even when sampled close to the surface, owing to the elevated location of the stations. As a result, a sampling height of 2 m is sufficient, and this requirement is already met by existing mountain stations performing $CH_4$ measurements (MDN and CUR).**

- P5, L131-L137: I think the description of the TNO-AVENGERS inventory can be shortened here, since the additional GHGs and available temporal resolutions are not relevant to the study. Otherwise, please state which temporal resolution was used for the CH4 emissions (I assume monthly?) and include the definition of LULUCF for complicity.

We agree and have shortened the description of the TNO-AVENGERS inventory. In response to another comment from Referee #2 regarding the temporal resolution of fluxes, we retained this information and specified that the input fluxes are resampled to hourly resolution before being used in the model. We have also added a definition of LULUCF in the introduction when we first introduce the acronym.

**Natural and anthropogenic datasets are resampled to hourly resolution, conservatively remapped to the ICON-ART spatial grid using the Emiproc package (Lionel et al., 2025) and aggregated to produce prior estimates of natural and anthropogenic emissions over Italy.**

- P5, L139: 'According to this dataset […]' is misleading here because it sounds like you are still referring to the GFED dataset from the sentence before. (Whereas you are referring to the total anthropogenic emissions, I assume?)

We agree and have modified the sentence.

**According to these estimates of anthropogenic $CH_4$ fluxes, about …**

- P5, Section 2.3.1: Would it be possible to include a short explanation or reference on how the prior datasets are adjusted to the icosahedral grid?

Naturally. We have included this information.

**Natural and anthropogenic datasets are resampled to hourly resolution, conservatively remapped to the ICON-ART spatial grid using the Emiproc package (Lionel et al., 2025) and aggregated to produce prior estimates of natural and anthropogenic emissions over Italy.**

- P8, L194-L195: 'Includes only ICOS sites located in countries bordering Italy' is misleading, because those would only include stations in France and Switzerland, which is not in line with Figure 2.

We agree and have modified the sentence.

**Scenario 1 includes only ICOS sites located in countries neighboring Italy.**

- P10, L292: Unfinished sentence 'We briefly present here.'?

We suppose the referee refers to the sentence "We briefly present IT here". Although the "it" already refers to the approach described in the previous sentence, we have modified the sentence to make it clearer.

**We briefly present this approach here.**

- P12, L310: I think you mean Figures 3c and 3d?

Yes, indeed. We apologize and thank the referee for pointing out this error. We have corrected it.

## Results

- Are the error reduction values in Figure 4 and Section 3.1 calculated from only one emission scenario? If so, which one?

We apologize for not providing this information in the submitted manuscript. The MERs in Figure 4 and Section 3.1 are the averages over the 5 emission scenarios. We have corrected this error.

- Please consider adding maps of emission sensitivity to the stations or trajectory analysis. It would be helpful to comprehend your arguments.

Unfortunately, it is not possible to provide emission sensitivities, as footprints are not explicitly computed during the inversion process with an Eulerian model. Doing so would require the full Jacobian of the observation operator, which would necessitate one forward simulation per optimized variable (i.e., a cell of the domain).

A key advantage of the Ensemble Square Root Filter is that it provides an approximation to the Kalman filter using an ensemble representation, thereby avoiding the explicit computation of the Jacobian matrix. This makes the method computationally efficient and well suited to high-dimensional state spaces.

In this context, this can again be seen as an advantage of Lagrangian modelling compared to Eulerian approaches for this type of application, as it enables efficient estimation of sensitivities without requiring an explicit and computationally expensive adjoint or Jacobian calculation.

We mainly used animations of horizontal winds calculated by our models to build our understanding.

We have added some explanation in the discussion on the comparison between Eulerian and Lagrangian modelling (see dedicated comment in the technical comments).

- 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 all of Italy, the MER obtained with refined MDMs is 16.6% in Scenario 2. It leads an absolute difference of 4.1% (a relative difference of −0.20). Compared to the absolute differences between the various cases using uniform mdms (1-2% across cases 3-10), this represents a substantial reduction in performance. This result is consistent with the mean refined mdm being greater than 20 ppb, indicating that the uniform value was not a very good assumption. With these higher mdms values, the stations have reduced ability to constrain the fluxes as the confidence is lower.

To facilitate comparison, we have added **Figures S7 and S8** to the Supplementary Information, which are similar to Figures 4 and C2 but using realistic mdms.

- 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?

The realistic mdms for VND and CHI are 35 and 46 ppb, respectively. While the value for CHI is higher, we would not characterize it as "much larger". Nevertheless, the difference is not negligible and leads to a different level of confidence assigned to CHI compared to VND. Using uniform MDM values, Scenario 3 (VND) and Scneario 4 (CHI) yield MERs of 23.0% and 23.6%, respectively (for total emissions over all of Italy). When realistic mdms are applied, these values decrease to 18.3% and 18.5% (i.e., approximated to 19% in Table S2). The relative reduction is therefore slightly larger for Scenario 4 (CHI), consistent with its higher refined MDM.

To facilitate comparison, we have added **Figures S7 and S8** to the Supplementary Information, which are similar to Figures 4 and C2 but using realistic mdms.

- 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.

For Scenario 3, the relative differences between uniform and realistic mdms are (23.0 − 18.3) / 23.0 = 0.20 for all of Italy, (30.5 − 24.3) / 30.5 = 0.20 for Northern Italy, and (21.1 − 16.7) / 21.1 = 0.21 for Central Italy. This indicates that the reduction in performance is similar across regions for this case. Therefore, we do not consider this result sufficient to emphasize the specific importance of this station for constraining emissions in Central Italy. As far as we understand your point, it would work better with Scenario 4, which shows a stronger relative reduction in Central Italy than in Northern Italy. This primarily reflects the fact that CHI provides fewer constraints in Northern Italy than in Central Italy.

However, we argue that similar conclusions can be drawn from the results obtained using uniform mdms. The relevance of a station for a given region should be assessed based on its contribution relative to the base network (Scenario 2), i.e., the extent to which it improves MER (or TUR) when added. This approach was used in Section 3.1 already. In response to the final point, we refer to these stations as the most informative because they provide better overall performance (MERs and TURs) than the others.

We hope that the addition of Figures S7 and S8, together with the clarifications provided above, addresses the points raised in these comments. We have avoided overcomplicating this section, as the analysis based on uniform mdms is already complex enough for the reader. We have therefore only slightly revised the final sentence to emphasize that the same conclusions hold, namely that Scenarios 3 and 4 yield higher MER and TUR values than the other scenarios for all of Italy.

**Nevertheless, the conclusions drawn in the previous sections remain unchanged: Scenarios 3 (VND) and 4 (CHI) continue to be the optimal choices for all of Italy when comparing MER and TUR values to other scenarios, with CHI retaining a marginal advantage.**

- 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).

We believe the referee is referring to ‘contributing to the increase in MER’ rather than its reduction. We find that POT adds many constraints and contributes to improving MER in Scenario 1 and the other scenarios, as the ER (see Fig. 3c) and UR (see Fig. 3e) show strong performance around the station. This improvement could not be achieved by LMP, which is a remote marine station located far from mainland Italy. While LMP provides valuable information on large-scale background conditions, its ability to constrain Italian emissions is likely limited due to its distance from continental sources.

We have added the given explanation about LMP in the introduction.

**LMP is a marine remote site located on Lampedusa island, deep in the Mediterranean sea and distant from continental Italy, monitoring since 2008. While it provides useful information on background concentrations, its ability to constrain Italian emissions is inherently limited due to its distance from continental sources.**

We have also added the explanation about ER and UR in Sect. 3.1.

**Geographical placement: Most Southern sites are not centrally located, reducing overlap between their footprints and the Italian landmass. Potenza (POT) is an exception, providing stronger coverage thanks to its central position (see ER and UR high values around POT in Fig. 3c and Fig. 3e).**

Although POT (Potenza) is already defined in the Introduction, we have chosen to reintroduce the station at the beginning of the Results section, in response to several comments requesting clarification of abbreviations. We did the same for the other stations.

- 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.

Both spatial and temporal variability can be used to help disentangle natural and anthropogenic emissions. However, we do not believe these factors alone provide sufficient information for a robust separation. We present this breakdown because it is feasible within our framework and serves to illustrate the potential of quantification of the method. In future work, one could perform a similar analysis while incorporating additional constraints, such as isotopic information. The inclusion of such data would likely improve the MERs for both natural and anthropogenic sources.

We want to emphasize that the objective of this study was to find the best candidate for improving the ICOS network in Italy, based on the results for total, anthropogenic and natural emissions. It was not to demonstrate that adding a station to the network would necessarily lead to a major improvement in the estimates of both natural and anthropogenic emissions. Although we observe a low improvement in anthropogenic emissions, achieving close agreement between top-down and bottom-up approaches will require additional information (such as isotopic constraints) to more effectively bridge the gap. As pointed out by this comment and the first general comment, we put too much emphasis on constraining the anthropogenic emissions.

We have completely reshaped Sect 3.3 to emphasize the limited performance in retrieving anthropogenic emissions using these candidate sites, and added a paragraph in the Discussion to highlight that improved separation and estimation require additional constraints, such as isotopic measurements or co-emitted species.

Sect. 3.3:

**Figures 5 and C2 present MER (averaged across all network scenarios) and TUR for total, anthropogenic, and natural emissions.**

**At the national scale, MER and TUR for anthropogenic emissions reach about 10% and 15–20%, respectively, across all candidate stations. These emissions are best constrained in Northern Italy, where they are most intense: adding VND (network scenario 3) increases MER to 15% and TUR to nearly 20%, representing the strongest performance in this region. In contrast, anthropogenic emissions remain poorly constrained in Central and Southern Italy, with both MER and TUR below 10%. Compared to total emissions, the weaker performance for anthropogenic emissions arises from two main factors. First, separating natural and anthropogenic contributions in the observed signal is intrinsically difficult. Without additional constraints (e.g., isotopic information), the inversion relies only on spatial and temporal differences. 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 for each category. Second, anthropogenic emissions are assigned a shorter spatial correlation length in the prior error covariance (Sect. 2.3.1), following recent literature. While more realistic, this choice increases the number of degrees of freedom and thus the complexity of the inverse problem. In other words, the true scaling factors for anthropogenic emissions vary more spatially than those for natural and total emissions. As a result, constraining the higher heterogeneity of anthropogenic scaling factors with the same number of observations is more challenging, leading to poorer performance.**

**Natural emissions are therefore better constrained than anthropogenic emissions. They are also more spatially diffuse and characterized by a larger spatial correlation length, making the metrics less sensitive to favorable wind conditions (i.e., transport from source regions to stations). Their performance closely follows that of total emissions, with MER values typically about 4% higher across most regions, except in Southern Italy where they are nearly identical.**

**Across all regions, the conclusions for anthropogenic and natural emissions remain consistent with those for total emissions. Although anthropogenic emissions are only weakly constrained, CHI and VND consistently outperform the other sites across all categories, with VND performing best in Northern Italy and CHI in the remaining regions. The separation between anthropogenic and natural emissions is discussed further in Sect. 4.**

Discussion:

**Additionally, substantially improving constraints on anthropogenic emissions using $CH_4$ measurements alone appears challenging. In this study, we adopt realistic conditions, including a shorter error correlation length for anthropogenic emissions. Even when adding the optimal station locations identified here, improvements in anthropogenic emission estimates remain limited. While the results show relatively good performance for total emissions, one of the primary objectives of top-down approaches is to complement bottom-up estimates by providing comparable information to better quantify anthropogenic emissions. Our results, consistent with previous literature, indicate that achieving this goal requires additional observational constraints, such as isotopic measurements or co-emitted species. Without such information, the ability to accurately constrain anthropogenic emissions in Italy remains limited.**

- 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?

We have added Fig. S2 in the Supplementary Information, which shows the seasonal cycles of each subcategory. Geological emissions are dominant but wetlands and soil uptake are the only contributors to the seasonal variability in the bottom-up estimates we use, with a stronger influence in Southern Italy, where most of these emissions are located. For anthropogenic emissions, seasonality is dominated by energy-related emissions, which is the only subcategory exhibiting monthly variations.

However, following our response to one of the general comments formulated by the referee, we also completely revisited Section 3.4 because we realized that our previous interpretation of the results was incorrect. The seasonality of prior emissions should not influence the results presented in Section 3.4. We refer to the revised manuscript for the updated section.

- P17, Figure 5: This figure is, unfortunately, quite hard to interpret since there is a lot of overlapping of the different scenarios. Could you possibly reconsider the design?

Following both referee's comments, we have tried to redesign Figure 5 and S4, to improve the readability and the interpretation.

- P18, L406: 'Fig 7 for TUR' do you mean figure C2?

Yes, we apologize for this error. We have corrected it.

- P20, Section 3.4: There are a lot of abbreviations in the section. I think it would be beneficial to use the names of the different seasons, instead of the abbreviations of the months.

We have followed this advice, replaced the abbreviations with the names of the different seasons and added a sentence in Section 3.4 to clarify the correspondence.

**To limit the number of abbreviations in this section, DJF, MAM, JJA, and SON are hereinafter also referred to as winter, spring, summer, and autumn, respectively.**

- P33, Figure C3 and P34, Figure C4: Please copy the caption of Figure 7 and adjust it accordingly, as it is more reader-friendly than having to scroll up and down in the publication.

This has been applied.

### 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?

This is a very interesting yet complex question. If the purpose of an emulator is to evaluate different years, sampling heights, or random perturbations of the true fluxes, then it must be trained across a sufficiently wide range of configurations to remain representative. Constructing such a training dataset would require a massive number of simulations, making the approach computationally expensive. Even with such an extensive training dataset, it remains unclear whether the emulator would properly capture the boundary layer height, tracer mixing processes, and their variability across configurations. Without dedicated validation experiments, we cannot draw firm conclusions. However, we believe that the emulator itself would provide only limited added value compared with the ensemble already generated for the training dataset. Additionally, such an emulator would not generalize to other regions of the world.

Machine learning could also be beneficial for a different purpose: emulating the transport model itself. The simulation of tracer transport is by far the most computationally demanding component of the workflow. Developing an accurate emulator for this step could significantly accelerate the inversion process and greatly reduce the cost of an OSSE study. However, training such an emulator would also require a very large volume of training data and substantial computational resources. Similarly, it also remains unclear whether sub-scale transport processes could be adequately reproduced.

We believe the methodological aspects involved here are complex and cannot be adequately addressed in a brief discussion. For this reason, we prefer not to discuss it in the manuscript.

- 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?

At mountain sites, the free troposphere is already sampled at night, so the model representation is unlikely to improve further. At lowland sites, model representation generally improves when the influence of local phenomena is reduced, as these processes require high spatial resolution to be properly simulated. Sampling at 100 m already reduces the influence of local phenomena. Moreover, because only afternoon observations are assimilated, the boundary layer is generally well mixed, further reducing the benefit of sampling at greater heights. However, this likely depends strongly on the specific site, and further investigation would be needed. The present work helps identify stations with the greatest potential, while future studies could further examine the influence of sampling height at these sites. It would require a dedicated OSSE study conducted at much higher spatial resolution and for different inlet heights at the most promising sites identified here.

We have added some discussion in the paragraph focusing on the influence of sampling heights.

**We assumed an inlet height of 100 m for most lowland candidate stations, implying the construction of tall towers at these sites in addition to existing infrastructure. This sampling height was chosen because ICOS continental stations typically operate at such elevations to reduce local influences, increase model representativeness, and enhanced sensitivity to regional fluxes. Our results should therefore be interpreted in light of this assumption. Sampling at higher levels is unlikely, as it would drastically increase costs while likely providing limited added value, given that local influences are already dampened at 100 m. Moreover, because only afternoon observations are assimilated, the boundary layer is generally well mixed, further reducing the benefit of sampling at greater heights. In contrast, sampling at lower heights, which is more likely in practice, could increase model–data mismatch and reduce the spatial footprint, potentially leading to lower performance. Nevertheless, this effect may remain limited during the afternoon, when vertical mixing reduces differences between observations sampled at 100 m and at lower levels. Note, however, that these considerations are likely site-dependent and would require further investigation through dedicated OSSE studies conducted at much higher spatial resolution and for different inlet heights at the most promising sites identified here.**

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?

Natural $CO_2$ emissions are more spatially widespread than natural $CH_4$ emissions and are not necessarily co-located, as they arise from different underlying processes. Therefore, we do not consider this interpretation to be supported.

- L533-540: Please phrase this carefully considering our point in general comment regarding constraining anthropogenic fluxes.

We have removed the final statement of the paragraph to avoid giving the impression that a single station would be able to properly constrain anthropogenic emissions and narrow the gap between bottom-up and top-down methods.

### Conclusions

- Would you have any concrete recommendations how to perform an OSSE to select good candidates for atmospheric measurement stations?

Our main concrete recommendations are to apply the methodology introduced in this work, to use the metrics derived here, which we find robust and informative, and to employ an ensemble of truth scenarios to reduce the influence of random effects.

We slightly modified the text to clarify those points are recommendations for future work

**The methodology we present can be readily applied to other Eulerian models and adapted to different countries or regions. We therefore strongly recommend adopting a similar approach, using the robust and informative metrics introduced here, and employing an ensemble of truth scenarios to reduce the influence of random effects. While computationally demanding, this type of OSSE study offers valuable guidance for decision-makers and atmospheric scientists when selecting candidate sites and optimizing observational coverage.**

- Could you comment on how to better monitor / constrain anthropogenic fluxes and their possible changes in Italy? Are we urgent in adding stations?

One of the most promising candidate stations (CHI) is already planned for integration into the Italian network and could potentially be added to the ICOS network in the future. Although the additional constraints will likely help to better estimate total emissions, this addition will likely not significantly resorb the substantial gap between bottom-up and top-down estimates of anthropogenic emissions. We have made this conclusion more explicit in the revised version. This raises the question: what alternatives should be considered? Simply adding a new station may not be the most effective solution. One option is to enhance measurements at existing stations, for example by incorporating isotope observations. Another approach would be to measure co-emitted species, such as ethane. Improving bottom-up estimates of natural emissions, and reducing their associated uncertainties, could also help clarify the anthropogenic signal. Finally, the use of satellite data represents another option, which we discuss in the following comment.

**Although we identified optimal locations among a set of candidate sites, a substantial number of additional stations would still be required to substantially reduce the discrepancy between bottom-up and top-down estimates. In this study, we only evaluated the extension of the surface $CH_4$ monitoring network. Further constraints could help improve the partitioning between natural and anthropogenic sources, such as isotopic measurements or co-emitted species (e.g., ethane).**

**These approaches warrant investigation in future OSSEs to assess whether deploying new stations or enhancing existing ones with additional observational capabilities (e.g., isotopes or co-emitted tracers) would be more effective. Importantly, a national capacity for stable isotopic carbon measurements has been established in Italy through the deployment of Cavity Ring-Down Spectroscopy (CRDS) isotopic analysers at CMN, POT, LMT, and LMP. Moreover, ethane measurements were initiated at CMN in 2023 within the framework of ACTRIS (Aerosol, Clouds and Trace Gases Research Infrastructure). Satellite data also holds strong potential for covering regions that are poorly sampled by ground-based stations, providing information at finer spatial scales. Although satellite observations are subject to larger uncertainties and are more challenging to assimilate into transport models, they could complement surface measurements and substantially improve estimates of total, natural, and ultimately anthropogenic emissions.**

- 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?

We strongly believe that satellite data could provide a valuable complementary approach. The spatial coverage offered by satellites is much larger than what can be achieved with ground-based observations. However, satellite data are more challenging to compare with transport models, model–data mismatches are generally larger with these data, and systematic biases exist. It would therefore be valuable to investigate how surface and satellite observations can be combined to improve emission estimates. While some studies have already assessed the information gain provided by satellite data relative to surface observations, especially at global scale, none have focused specifically on Italy, to our knowledge. In addition, the assimilation of satellite data remains relatively immature within the inversion community compared to using only surface measurements, and further work is needed to ensure their proper use. Nevertheless, exploring whether satellite data could provide greater benefits than adding new stations would be a worthwhile direction for future research. More broadly, the most robust outcomes are likely to emerge from exploiting the synergies between these different approaches. We have incorporated elements of this discussion into the revised manuscript.

**Although we identified optimal locations among a set of candidate sites, a substantial number of additional stations would still be required to substantially reduce the discrepancy between bottom-up and top-down estimates. In this study, we only evaluated the extension of the surface $CH_4$ monitoring network. Further constraints could help improve the partitioning between natural and anthropogenic sources, such as isotopic measurements or co-emitted species (e.g., ethane).**

**These approaches warrant investigation in future OSSEs to assess whether deploying new stations or enhancing existing ones with additional observational capabilities (e.g., isotopes or co-emitted tracers) would be more effective. Importantly, a national capacity for stable isotopic carbon measurements has been established in Italy through the deployment of Cavity Ring-Down Spectroscopy (CRDS) isotopic analysers at CMN, POT, LMT, and LMP. Moreover, ethane measurements were initiated at CMN in 2023 within the framework of ACTRIS (Aerosol, Clouds and Trace Gases Research Infrastructure). Satellite data also holds strong potential for covering regions that are poorly sampled by ground-based stations, providing information at finer spatial scales. Although satellite observations are subject to larger uncertainties and are more challenging to assimilate into transport models, they could complement surface measurements and substantially improve estimates of total, natural, and ultimately anthropogenic emissions.**

## 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.

Full names were already introduced in the Introduction and in Sect. 2.4. For clarity, we have also repeated the full names when stations are first mentioned in the Results section.

- 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.

We agree and have modified the text in the abstract. We retained the expression "truth scenarios" in the main text because we believe that using "emission scenarios" without explicitly referring to the truth would reduce understanding.

- 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.

We agree and have modified the text.

**Currently, countries primarily report their emissions through bottom-up inventories, which typically rely on 1) socioeconomic and environmental data, 2) source-specific emission factors, and, for some sectors such as Land Use, Land Use Change and Forestry (LULUCF), 3) process-based models**

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

We agree this is not clear. We have modified the text.

**Top-down methods provide independent estimates that are comparable across regions and over time, are available more quickly than bottom-up estimates, and can be produced at global, regional, national, and sub-national scales.**

- 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?

We referred to the above-mentioned NGHGIs. We have modified the text.

**Observational in-situ networks should therefore be carefully designed to minimize the uncertainties of existing bottom-up inventories.**

- 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.

We have added some information.

**Additional stations, introduced in Sect. 2.4, also monitor $CH_4$ in Italy. However, these sites are not part of the ICOS network, their measurements are not necessarily continuous, and data access is not public, requiring direct requests to the respective Principal Investigators (PIs).**

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

We have added the models.

**Most of the Eulerian models used in the inversion community have now been coupled to CIF (LMDz, CHIMERE, ICON-ART, WRF, STILT, FLEXPART, TM5)**

- L144: "represent the soil uptake emissions from peatlands and inundated wetlands". Probably should be "soil uptake and emissions"

Yes, indeed. Thank you for pointing out this error. We have modified the text.

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

It started measuring in March 2026. We have updated the sentence.

**It is operated by the University of Chieti and located on the Adriatic coast, and hast just begun measuring $CH_4$ in March 2026. At present, only "pilot" measurements are conducted, as calibrations are still not fully implemented and WMO compatibility goals are not achieved.**

- L220: Blue and yellow circles

Yes, indeed. Thank you for pointing out this error. We have modified the text.

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

This has been modified.

- 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?

We have included an example.

**Reductions in domain-total flux error can be misleading, as they may result from compensating errors across spatial or temporal domains. For example, a positive error in Northern Italy (overestimation of emissions) and a negative error in Southern Italy (underestimation of emissions), if of equal magnitude, will result in an accurate estimate of emissions at the national scale.**

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

This has been modified.

- 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.

Yes, we apologize, this statement is wrong indeed. We only compared the results for scenario 12, which is the optimal one. MERs in Northern Italy and Central Italy are 44% and 40%, respectively, while MER in Southern Italy is 29% which shows that this region is less constrained than the other two regions. We modified the statement.

**While Northern and Central Italy shows similar results under Scenario 12, Southern Italy remains less constrained.**

- 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?

We have modified the sentence.

**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 for each category.**

- 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.

Yes, we recalculate RMSE based on posterior mole fractions and then perform a new inversion using this updated estimate of model-data mismatch. We modified the sentence.

**For the real stations, we first perform a forward simulation for the year 2018. Simulated values are then sampled and compared to observations to compute an initial model–data mismatch using the root mean square error (RMSE). This initial estimate is used to perform an inversion for 2018. Following this initial inversion, we recalculate**

**the RMSE using the posterior mole fractions, which provides a refined estimate of the model–data mismatch at these stations, and then perform a new inversion using this updated estimate of model-data mismatch. The results of this second inversion are analyzed.**

- 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,...".

It is a good idea. We have moved the sentence there, but we have also changed "As a result" to " Using these new estimates of model-data mismatches ".

**Using these new estimates of model-data mismatches, both MER (Table S2, Fig. S4, and Fig. S7) and TUR (Table S3 and Fig. S5, and Fig. S8) decrease by 2–6\% compared to the experiments assuming a uniform model–data mismatch.**

- 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.

CHI has only recently begun operations (March 2026), whereas at VND a tower already exists and a dedicated feasibility study for installing a station has been conducted. For the remaining sites, constructing tall towers would be less realistic, although still more feasible than building such infrastructure in remote areas lacking existing facilities. Notably, the two best candidate sites also appear to be the most practical to implement. This point has been clarified in the conclusions.

**Our results show that CHI and VND are the most promising candidate stations for improving emission constraints in Italy, with CHI having a slight advantage when sampling at 50 m. While CHI provides stronger constraints in Central and Southern Italy, VND is particularly effective in Northern Italy. Importantly, these stations are also the most likely to be implemented in the future with the inlet heights considered in this study.**

- 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?

We have entirely rewritten this paragraph because we forgot to include one important advantage of the Eulerian approach. We have also improved the other arguments.

**Although an Eulerian model is employed in this work, Lagrangian models can also be used to conduct network expansion studies. Lagrangian frameworks offer several advantages over Eulerian frameworks, but also have notable limitations. In Lagrangian frameworks, footprints can be readily computed, providing an intuitive and effective means of investigating the spatial and temporal sensitivity of measurements to surrounding fluxes. Once computed, these footprints can be reused, enabling multiple site configurations to be tested with negligible additional computational cost. Moreover, the evaluation metrics introduced in this work can be derived just as easily within a Lagrangian framework. Using these footprints, the full Jacobian of the observation operator can be computed efficiently, making the direct application of Kalman filter equations (i.e., an analytical inversion) a natural choice, in contrast to the ensemble-based approach adopted here. However, this approach does not scale well to very large control or observation vectors, as it requires inversion of the associated error covariance matrices. While it is possible to assimilate large numbers of observations sequentially under the assumption of uncorrelated errors, the computational cost increases rapidly with the number of optimized variables. This effectively limits the dimensionality of the inversion problem and, consequently, the amount of information that can be extracted from the observations. For this reason, in the present case study focusing on surface stations, the primary advantage of an Eulerian framework lies in its ability to optimize emissions at high spatial (e.g., model resolution) and temporal (e.g., daily or weekly) resolution. In addition, Eulerian frameworks are well suited for the assimilation of satellite data, which requires the calculation of column-averaged mole fractions using averaging kernels. In contrast, Lagrangian models require particles to be released throughout the atmospheric column to estimate sensitivities of column-averaged mole fractions to emissions, substantially increasing computational demand and making their application to satellite data inversions more restrictive.**

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

Yes, we have added this information to the caption.

**Spatial distribution of total, natural and anthropogenic fluxes in Italy (upper panels) and contributions of natural and anthropogenic emissions to total emissions in Northern, Central, Southern, and all of Italy (lower panel), based on the prior datasets. Numbers displayed at the center of bars represent the contribution of each category to the total emissions in a specific region.**

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

We have added a reference to Table A3 (Table A2 lists the existing ICOS atmosphere stations).

**For readability, ICOS ecosystem station names are not displayed in Scenario 12, but are listed with their locations in Table A3.**

- Figure 4: Please change 'case' to 'scenarios' for consistency.

We apologize for this mistake. It has been modified.

- "Stations from Scenario 1 are marked with yellow circles" Except for the Scenario 1 panel, I think they are for Scenario 2. Please revise.

We apologize for this mistake. It has been corrected.

- 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.

Following your advice and addressing a similar comment for Referee #1, we reshaped the figure to improve readability.

- 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.

We have added the current status in Section 2.4 and Table A1, for all the sites.

- Table A3 caption: may need a bit of explanation about "fake inlet height".

We have modified this.

*** Tested inlet height denotes the inlet height used to generate the synthetic observations assimilated in this study. It represents a preliminary estimate for a potential ICOS station that could operate at this location.**

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

As far as we know, this is automatically added when the paper is published and should not be added beforehand.

- P1, L2: Missing comma: 'Here, we focus on Italy, [...]'

We apologize for this mistake. It has been corrected.

- E.G. P1, L4 and P2, L38: Careful with the grammar: 'ICOS Atmosphere site/network' is usually written with a capital A, to indicate that the station is part of ICOS Atmosphere. Otherwise, 'atmospheric site/measurement/network' should be used. Please check throughout the manuscript including the captions of the figures and tables.

We thank the referee for pointing out this error. This has been corrected throughout the manuscript.

- P14, L336: I wouldn't say 'skill' is the correct word here, maybe 'In contrast, the capacity of the network is slightly weaker in Central Italy'?

This has been applied.

- P15, Figure 4: Better: '[...] while the national value is displayed in bold in the lower-left corner of each panel.'

This has been modified.

- One comma too many: '[...] no monitoring stations because of a lack of information.'

We thank the referee for pointing out this error. We have corrected it.

# References


Caputo, A., Di Cristofaro, E., Gonella, B., and Taurino, E.: Il metano nell'inventario nazionale delle emissioni di gas serra: l'Italia e il Global Methane Pledge, ISPRA Rapporti, 374/2022, Roma, Italia, ISBN 978-88-448-1129-7, 2022. Available at: https://www.isprambiente.gov.it/files2022/pubblicazioni/rapporti/r374-2022-1.pdf.

Fratticioli, P., Trisolino, P., Maione, M., Calzolari, F., Calidonna, D., Biron, S., Amendola, M., Steinbacher, P., and Cristofanelli, P.: Continuous atmospheric in-situ measurements of the CH4/CO ratio at the Mt. Cimone station (Italy, 2165 m a.s.l.) and their possible use for estimating regional CH4 emissions, Environmental Research, 232, 116 343, https://doi.org/10.1016/j.envres.2023.116343, 2023.

Hazan, L., Tarniewicz, J., Ramonet, M., Laurent, O., and Abbaris, A.: Automatic processing of atmospheric CO2 and CH4 mole fractions at the ICOS Atmosphere Thematic Centre, Atmospheric Measurement Techniques, 9, 4719–4736, https://doi.org/10.5194/amt-9-4719-2016, 2016.

Lapenna, E., Buono, A., Mauceri, A., Zaccardo, I., Cardellicchio, F., D'Amico, F., Laurita, T., Amodio, D., Colangelo, C., Di Fiore, G., Gorga, A., Ripepi, E., De Benedictis, F., Pirelli, S., Capozzo, L., Lapenna, V., Pappalardo, G., Tripetta, S., and Mona, L.: ICOS Potenza (Italy) Atmospheric Station: A New Spot for the Observation of Greenhouse Gases in the Mediterranean Basin, Atmosphere, 16, 57, https://doi.org/10.3390/atmos16010057, 2025.

ICOS RI: ICOS Near Real-Time (Level 1) Atmospheric Greenhouse Gas Mole Fractions of CO2, CO and CH4, growing time series starting from latest Level 2 release (version 1.0), https://doi.org/10.18160/ATM_NRT_CO2_CH4, last access: 18/05/2026, 2018.

ICOS RI, Adame, J., Apadula, F., Biermann, T., Blessing, C., Charrondière, C., Colomb, A., Conil, S., Couret, C., Cristofanelli, P., De Mazière, M., Delmotte, M., Di Iorio, T., Emmenegger, L., Forster, G., Frumau, A., Harris, E., Haszpra, L., Hatakka, J., Heliasz, M., Hensen, A., Hermansen, O., Hoheisel, A., Kneuer, T., Komínková, K., Kubistin, D., Larmanou, E., Laurent, O., Lehner, I., Lehtinen, K., Leskinen, A., Lindauer, M., Lopez, M., Lund Myhre, C., Lunder, C., Mammarella, I., Manca, G., Marek, M. V., Marklund, P., Meinhardt, F., Miettinen, P., Molnár, M., Montaguti, S., Müller-Williams, J., O'Doherty, S., Piacentino, S., Pichon, J.-M., Pitt, J., Platt, S. M., Plaß-Dülmer, C., Ramonet, M., Rivas-Soriano, P., Roulet, Y.-A., Scheeren, B., Schmidt, M., Sferlazzo, D., Sha, M. K., Stanley, K., Steinbacher, M., Sørensen, L. L., Vítková, G., Yela, M., Ylisirniö, A., Yver-Kwok, C., Zazzeri, G., Zwerschke, E., di Sarra, A., ICOS ATC, ICOS-CAL-840 CRL, and ICOS-CAL-FCL: ICOS Atmosphere Release 2026-2 of Level 1 Fast Track Greenhouse Gas Mole Fractions of CO2, CH4, N2O, CO, meteorology data and flask samples analysed for CO2, CH4, N2O, CO, H2, SF6, 14CO2, O2/N2, δ13C-CO2 and δ18O-CO2, https://doi.org/10.18160/ABTF-SD2Q, 2026.

Koffi, E. N., Bergamaschi, P., Alkama, R., and Cescatti, A.: An observation-constrained assessment of the climate sensitivity and future trajectories of wetland methane emissions, Sci. Adv., 6, eaay4444, https://doi.org/10.1126/sciadv.aay4444, 2020.

Zhang, Z., Zimmermann, N. E., Stenke, A., Li, X., Hodson, E. L., Zhu, G., Huang, C., and Poulter, B.: Emerging role of wetland methane emissions in driving 21st century climate change, P. Natl. Acad. Sci. USA, 114, 9647–9652, https://doi.org/10.1073/pnas.1618765114, 2017.

Zazzeri, G., Apadula, F., Henne, S., et al.: Methane record at Plateau Rosa confirms its role as background station with episodic sensitivity to European emissions, Communications Earth Environment, 7, 260, https://doi.org/10.1038/s43247-026-03294-5, 2026.

Empa

Materials Science and Technology