https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

German methane fluxes in 2021 estimated with an ensemble-enhanced scaling inversion based on the ICON–ART model

Valentin Bruch[1], Thomas Rösch[1], Diego Jiménez de la Cuesta Otero[1], Beatrice Ellerhoff[1], Buhalqem Mamtimin[1], Niklas Becker[1], Anne-Marlene Blechschmidt[1], Jochen Förstner[1], and Andrea K. Kaiser-Weiss[1]

[1]Deutscher Wetterdienst, Frankfurter Str. 135, 63067 Offenbach

Correspondence: Valentin Bruch (valentin.bruch@dwd.de) and Andrea K. Kaiser-Weiss (andrea.kaiser-weiss@dwd.de)

Abstract. A reliable quantification of greenhouse gas emissions is important for climate change mitigation strategies. Inverse methods based on observations and atmospheric transport simulations can support emission quantification down to the national scale, yet, they are often limited by the observing systems, transport model uncertainties, and inversion methodologies. Here, we present a system for observation-based, regional methane flux estimation, which has the potential for long-term operational support of national emission reporting. We apply this to Central Europe in 2021 with focus on Germany, where we distinguish emissions from different anthropogenic sectors. The atmospheric transport is calculated with the numerical weather prediction model ICON–ART at 6.5 km resolution, sampling the meteorological uncertainty with a 12-member transport ensemble. We use a priori fluxes from national reporting to facilitate the validation of reported fluxes. Posterior fluxes are estimated with a modified synthesis inversion method, relying on observations from the Integrated Carbon Observation System (ICOS). Compared to the a priori, we find a significant increase in methane emissions in Germany and in the Benelux. We estimate German methane emissions $(32 \pm 19)\,\%$ higher than the anthropogenic emissions in the national inventory, and attribute this difference mainly to the agricultural sector, although separation from Land Use, Land Use Change and Forestry (LULUCF) as well as natural fluxes requires further research. The combination of an ensemble-enhanced numerical weather prediction model for atmospheric transport and good observation coverage paves the way to sector-specific, observation-based national emission estimates.

1 Introduction

Reducing greenhouse gas (GHG) emissions is crucial for mitigating current anthropogenic global warming. UNFCCC (United Nations Framework Convention on Climate Change) compliant national inventories and/or process models quantify anthropogenic GHG emissions for the purpose of monitoring the effectiveness of mitigation as planned, e.g., in the Paris Agreement. In addition to so-called "bottom-up" methods, atmospheric GHG concentration observations are used in "top-down" flux estimations. The latter are complementary, as they are sensitive to the total fluxes (i.e., anthropogenic and natural) and provide options for independent validation of a priori fluxes provided by inventories (IPCC et al., 2019). The usefulness of top-down es-

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

timates has been demonstrated, e.g., for the United Kingdom (Manning et al., 2011), Switzerland (Henne et al., 2016), Europe (Petrescu et al., 2023) and globally (Deng et al., 2022; Petrescu et al., 2024). Although research foundations for top-down methods have been developed in recent decades, see Janssens-Maenhout et al. (2020) and references therein, applications remain limited due to sparse observation coverage and representativeness, and most critically, due to transport model uncertainties (Engelen et al., 2002; Gerbig et al., 2008). The latter is a well-known issue not solved yet (Munassar et al., 2023). Inversions using satellite observations (e.g. Estrada et al., 2024) benefit from larger spatial observation coverage, but the uncertainties of the observations are larger compared to in situ data and the influence on the inversion results was found smaller where in situ coverage is good (Thompson et al., 2025). The benefits of increased model resolution (Agustí-Panareda et al., 2019; Bergamaschi et al., 2022) can be reaped with regional high resolution modeling, and ensembles can cover parts of the meteorological uncertainty (Steiner et al., 2024a). At short time scales, the regional model uncertainties will constitute the main uncertainty, while at longer time scales, the boundary conditions become critical for tracer transport (Chen et al., 2019).

Regional top-down estimates of long-lived GHG can be based on different types of transport models. Lagrangian models calculate trajectories from selected locations by moving with air parcels transported by the wind. They have been widely used for inversions of trace gases like halocarbons, nitrous oxide and methane (CH_4) in European regions, see e.g., Stohl et al. (2009); Ganesan et al. (2015); Henne et al. (2016). In contrast, Eulerian models – such as ICON–ART – continuously transport trace gas concentrations through three-dimensional grid boxes. Although they are computationally more expensive for cases where a relatively small number of trajectories would suffice, they become superior when the amount of data grows and, as Engelen et al. (2002) pointed out, open the road for data assimilation methods as used in numerical weather prediction. Regardless whether Lagrangian or Eulerian or even combined approaches (Rigby et al., 2011) are applied, the top-down estimation requires solving an inverse problem (Enting, 2002). Eulerian transport model based inversions may employ emission ensembles, as in Steiner et al. (2024b) with a localized Kalman filter, and other data assimilation methods. Alternatively, the method of synthesis inversion scales a set of a priori emission categories (Kaminski et al., 2001). Note that Meirink et al. (2008b) compare 4D-Var and synthesis inversion methods.

In this work, we present a modular system for regional top-down estimates of CH_4 fluxes designed to validate national inventories, including the discrimination of economic sectors such as agriculture and industry. We apply this method focusing on German inventories (provided by Umweltbundesamt and Thünen Institute) for the year 2021 using in situ observations collected by ICOS (ICOS RI, 2024). Atmospheric transport is simulated using the numerical weather prediction model ICON (Zängl et al., 2015) extended with the module for Aerosol and Reactive Trace gases (ART) (Rieger et al., 2015; Schröter et al., 2018) with a spatial resolution of $6.5\,\mathrm{km}$. The model is combined with a synthesis inversion approach (Kaminski et al., 2001) which is developed further to make use of the ensemble-estimated transport uncertainty. For minimizing transport errors, we rely on the operational numerical weather prediction at Germany's Meteorological Service (DWD) for meteorological initial conditions, lateral boundaries and transport ensemble calculations. Further, we use the Copernicus Atmospheric Monitoring Service (CAMS) for boundary conditions of methane, and compensate possible biases on the boundaries by deriving a correction field. Benefiting from the numerical weather prediction model and spatially highly resolved a priori fluxes from the

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

inventory agencies, we explore the basis for future operational top-down validation of national emission reporting, with special emphasis on further use in Germany.

In Sect. 2, we explain the transport model and the inversion methodology, while leaving some technical details for the appendices. Section 3 contains an overview of the utilized initial and boundary data as well as the a priori fluxes. Further, we describe the pre-processing of the observations used. Section 4 is dedicated to the details of the method concerning its application for Germany. Section 5 contains the results for our example year 2021, together with the results of the validation and sensitivity tests. The potential of the method is demonstrated by model performance tests using pseudo-observations with known true emissions. In Sect. 6 we discuss limitations and capabilities of the method and compare to other studies, followed by a conclusion in Sect. 7.

2 Method

In the employed offline scaling inversion, we categorize the a priori fluxes and scale each flux category to optimize the agreement between model prediction and observations. We start by defining flux categories which subdivide the fluxes by region and sector. With the Eulerian transport model, the concentrations from each flux category are calculated separately at all grid cells and time points. At the location and time of the observations, the model writes out the predicted concentrations from the flux category contributions and their sum is compared to the observed concentration. Our inversion makes use of the linear relation between surface fluxes and concentrations in the atmosphere. The categorized fluxes are scaled to minimize the mismatch between model prediction and observed concentrations. Thus, the inversion result consists of one scaling factor for each flux category. The a priori fluxes multiplied by the scaling factors yield the a posteriori fluxes.

The described method relies on high quality model predictions as well as accurate concentration observations. Furthermore, we need to estimate the model uncertainty and error correlations to assess whether deviations between model and observations contain information on the fluxes. To match these requirements, we have carefully chosen the configuration of the transport model (Sect. 2.1). Also, selected observational data are employed to remedy model boundary effects and therefore improve the overall model predictions (Sect. 2.2). In this section, we further introduce the Bayesian inversion framework (Sect. 2.3), the model uncertainties (Sect. 2.4), and our strategy for dealing with strong plumes (Sect. 2.5).

2.1 Transport model

The atmospheric transport is simulated using the numerical weather prediction model ICON (Zängl et al., 2015) with the ART module (Rieger et al., 2015; Schröter et al., 2018). The model is run in limited area mode for a domain covering large parts of the European continent (latitudes 34° N to 70° N, longitudes 21° W to 59° E, see Fig. 3) with a horizontal resolution of $6.5\,\mathrm{km}$ (ICON grid R3B8) and 74 vertical levels up to a maximal height of $22.77\,\mathrm{km}$. The surface CH_4 fluxes are provided to the transport model using the online emission module (Jähn et al., 2020; Steiner et al., 2024b).

For long living tracers like methane, the correct treatment of the lateral boundary concentrations is of importance. Therefore, we extended the model by implementing lateral boundary nudging for ART tracers in order to obtain smooth fields and avoid

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

strong gradients. Further, so-called meteogram output has been implemented for ART tracers, providing model output in the vicinity of observation locations with high temporal resolution.

For improved uncertainty estimates, we run a meteorological ensemble of 12 members. Each ensemble member uses slightly different but equally likely parametrizations and meteorological initial and boundary conditions. The construction of this ensemble follows the operational weather prediction at DWD (Schraff et al., 2016). In the following, we therefore distinguish a so-called deterministic model run providing the best estimate of the modeled CH_4 concentration, and the ensemble runs providing 12 different CH_4 concentrations to estimate the uncertainty.

2.2 The need for a far-field correction

For cases where the model predicts almost no influence from our categorized emissions (i.e., clean air cases), deviations between model and observations point to the need for correcting the CH_4 advected across the lateral boundaries – here referred to as "far field".[1] For our regional inversion problem, it is essential to separate the CH_4 emitted within the domain from the far field, in order to avoid model biases which would confound the aspired flux scaling. To minimize potential biases arising from imperfect boundary conditions, we construct a correction field which is added to the modeled far-field concentration in the whole domain after the transport simulation. We require this correction field to be smooth on large length and time scales, chosen in our case as $320\,\mathrm{km}$ (horizontal), $1\,\mathrm{km}$ (vertical) and $16\,\mathrm{h}$ (time). The far-field correction range is usually limited to $\pm 10\,\mathrm{ppb}$ (see Fig. B1 for details). Appendix B comprises the mathematical derivation, parameters, and a statistical overview of the far-field correction.

2.3 Inversion

We use a Bayesian inversion to optimize the agreement of model and observations by scaling the flux categories. This is formulated in the optimization problem

$$s^{\text{post}} = \arg\min_{s} \left\{ \tfrac{1}{2}(y - Hs - x^{\text{ff}})^\top R^{-1}(y - Hs - x^{\text{ff}}) + \tfrac{1}{2}(s - s^{\text{prior}})^\top B^{-1}(s - s^{\text{prior}}) \right\} \tag{1}$$

for the posterior scaling factors s^{post}. Here, the first term penalized the deviation from the concentration observations, and the second term penalizes the deviation from the prior fluxes. In the first term, the vector y of observed concentrations is compared to the model prediction, which consists of the transported fluxes Hs and the modeled far field x^{ff}. The transported fluxes Hs depend linearly on the vector s of scaling factors for the flux categories, which is optimized. The difference between modeled and observed values is weighted by the error covariance matrix R describing the combined uncertainty of the transport model and the observations. With the second term we constrain s by defining a priori scaling factors s^{prior} ($s_i^{\text{prior}} = 1$ for all i) with an error covariance matrix B characterizing the a priori uncertainty.

In Eq. (1), the fluxes are parametrized by the vector s of scaling factors. The flux categories thereby define the low-dimensional space in which the inversion can optimize the fluxes. The model observation operator H connects the space

[1] Technically, in our application to German CH_4 fluxes, the far field also includes the initial CH_4 concentration. But this is hardly relevant due to our generous spin-up period of 17 days.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

of scaling factors (vectors s^{prior}, s^{post}) to the observation space (vectors y, x^{ff}). Computing H requires the transport model which distinguishes the flux categories. The setup is designed for optimizing a low-dimensional vector s^{post} of scaling factors ($\sim 10^2$ degrees of freedom) using a large number of observations ($\sim 10^4$), but an extension to more degrees of freedom and/or more observations is possible.

2.4 Computation of model uncertainties

The inversion in Eq. (1) requires an estimate of the model–observation mismatch uncertainties described by R. In our case, the measurement error is negligible in comparison to the model error, and the non-negligible representativity error is part of R. We simply refer to this as "model error" below. We estimate this uncertainty and the cross-correlations of model equivalents at the observation locations using an ensemble of $M = 12$ different transport model realizations. The potential of using a small transport ensemble for estimating model uncertainties was demonstrated by Steiner et al. (2024a). The main contribution R' of the uncertainty matrix is

$$R'_{ij} = C_{ij} \frac{1}{M-1} \sum_{m=1}^{M} (x_i^m - \bar{x}_i)(x_j^m - \bar{x}_j) + \delta_{ij}\sigma^2_{\text{const}}, \tag{2}$$

where x_i^m is the prediction of ensemble member m for observation y_i, $\bar{x}_i = \frac{1}{M}\sum_m x_i^m$ is the ensemble mean, and $\sigma_{\text{const}} = 10\,\text{ppb}$ is a constant uncertainty added to each observation accounting for any representativity error. Indices i, j label observation data points that are typically distinguished by location, time, and sampling height. By C_{ij} we denote a localization in space and time such that $C_{ii} = 1$ and $C_{ij} = 0$ for any observations i and j we expect to be uncorrelated because of their temporal or spatial separation. We use the notation $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

2.4.1 Uncertainty inflation

The transport ensemble may not necessarily include the full uncertainty of the transport model, and the localization C_{ij} further reduces the simulated uncertainty by suppressing correlations. This motivates an inflation of the uncertainty to avoid overconfidence in the model prediction. We inflate the uncertainty by a factor $f_i > 1$ depending on the observation site of observation i, leading to the matrix $f_i f_j R'_{ij}$. In the application below, we choose $f_i = 2$ except for some stations with known difficulties, for which $f_i = 3$ (see Table A1).

To avoid potential biases through site-specific small-scale features not captured in the model, we aim to base our inversion on many observations. To this end, we limit the influence of individual data points on the inversion result by inflating the uncertainty further in the case of a very large disagreement between model and observation. This is achieved by replacing f_i with $f'_i \geq f_i$ to increase the uncertainty of individual observations until the deviation $\mu = y - H s^{\text{prior}} - x^{\text{ff}}$ between model and observations is at most three standard deviations of the resulting error covariance matrix $R_{ij} = f'_i f'_j R'_{ij}$, i.e., $f'_i = \max\{f_i, \frac{|\mu_i|}{3\sqrt{R'_{ii}}})\}$. This is justified because large deviations between model and observations, $|\mu_i| > 3\sqrt{R_{ii}}$, are likely caused by local pollution or modeling problems that are not captured appropriately in our uncertainty estimate. This correction makes sure that inversion results will be based on many observations and no single measurement can have an extreme impact.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

2.4.2 Flux dependence of model uncertainties

In Eq. (2), we estimate the model uncertainty utilizing the model concentration x_i^m, which itself depends on the fluxes. In principle, one can compute x_i^m using either the a priori fluxes or the a posteriori fluxes. In the latter case, the error covariance matrix R depends on the scaling factors s, and the optimization problem Eq. (1) can only be solved numerically (see Appendix D for details). We will refer to these two variants of the inversion as the prior-R and posterior-R inversion.

Computing posterior-R inversion exactly requires that each ensemble member distinguishes the various flux categories similar to the deterministic model run. To enable the posterior-R inversion without significantly increased computational effort in the ensemble run, we approximate the posterior-based concentration x_i^m in the ensemble members as described in Appendix D2. The posterior-R inversion generally yields slightly higher emission estimates because larger scaling factors lead to higher model uncertainties and thereby reduce the first term in Eq. (1). To provide robust inversion results with cautionary uncertainty estimates, we compute the prior-R inversion and posterior-R inversion separately and interpret the difference of the two methods as methodological uncertainty.

2.5 Plume localization problem

Plumes caused by high emissions in a small area require special treatment to avoid a potential bias in the inversion due to the so-called double penalty issue (Vanderbecken et al., 2023). In cases where our model falsely predicts that the plume reaches an observation site, the inversion will reduce the emissions to improve the agreement with the observation. In the opposite case, when the model fails to predict that a plume reaches the observation, the inversion will not change the plume emission amount but will wrongly increase emissions in other areas instead. This can cause systematic underestimation of fluxes from localized plumes. In the posterior-R inversion, this problem is mitigated because the underestimation of emissions is penalized by a lower model uncertainty. To avoid biases in the inversion results, we suggest to treat strong plumes separately, with their own flux categories. This allows us to quantify the problem (see Sect. 5.6) and to limit the plume penalty influence on other flux categories.

3 Input data and processing

We apply the method to estimate CH_4 fluxes in the year 2021 in Germany and in the surrounding European domain, relying on input data for the transport simulation and CH_4 concentration on the lateral boundary (Sect. 3.1), a priori fluxes (Sect. 3.2), and observations (Sect. 3.3).

3.1 Initial and lateral boundary conditions

The meteorological initial and lateral boundary conditions used to drive our transport model are taken from the archive of DWD's operational numerical weather prediction (NWP), which also employs the ICON model. As we do not assimilate meteorological data in our application, we re-initialize the meteorological fields every night at 0 UTC, using the analysis fields

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Table 1. Input data for a priori CH_4 fluxes. The second column lists where these fluxes were considered. Here, "Germany" refers to all model grid cells that lie fully within the German borders.

Data provider	Domain	Fluxes	Original grid	Time profile	Remarks
Umweltbundesamt (UBA)	Germany	GNFR sectors A–L (excluding LULUCF)	native (ICON)	constant	based on reporting to the UNFCCC (UBA, 2023), spatially distributed using the Gridding Emission Tool for ArcGIS (GRETA 1.2.01) (Feigenspan et al., 2024)
Thünen Institute	Germany	organic and mineral soils	$100\,\mathrm{m} \times 100\,\mathrm{m}$	constant	emissions from organic and mineral soils, including wetlands but excluding artificial ponds (approx. $160\,\mathrm{kt}$ CH_4 per year) (Fuß and Akubia, 2024)
CAMS-REG-ANT, v7.0	model domain excl. Germany	GNFR sectors A–L (excluding LULUCF)	$0.05° \times 0.1°$	constant	based on data reported to the UNFCCC for countries in Western and Central Europe (incl. Finland and the Baltic states) (Kuenen et al., 2021, 2022)
CAMS inversion optimized, v22r2	model domain excl. Germany, excl. oceans	wetlands	$1° \times 1°$	monthly averages	variant using surface air-sample data for the inversion (Segers and Houweling, 2020); Fluxes in model grid cells located over the ocean are set to zero.
Rocher-Ros et al. (2023), version 1.1	full model domain	rivers and streams	$0.25° \times 0.25°$	monthly averages	
Weber et al. (2019)	oceans (full model domain)	oceans	$0.25° \times 0.25°$	constant	

from the operational NWP data assimilation. Lateral boundary conditions for the meteorological fields are taken from the NWP short term forecasts with hourly resolution.

For the CH_4 concentrations, we use initial and lateral boundary concentrations from the global CAMS inversion-optimized dataset (Segers and Houweling, 2020), version v22r2, in the variant based on surface air-sample data for the inversion. The CAMS data have a resolution of $1° \times 1°$ and are interpolated onto our model grid. In contrast to the meteorological fields, the CH_4 concentrations are only transported and never re-initialized.

3.2 A priori CH_4 fluxes

For the inversion, we employ a priori CH_4 fluxes that were compiled from six datasets of anthropogenic and natural fluxes, as detailed in Table 1. We ensured mass conservation when interpolating to our model grid. We generally distinguish between

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

anthropogenic emissions excluding LULUCF, and natural fluxes plus LULUCF. Anthropogenic fluxes excluding LULUCF are split further into 12 GNFR sectors (gridded aggregated NFR, nomenclature for reporting, Veldeman et al. (2013)), but we only use this separation in selected regions of interest. Natural plus LULUCF fluxes of CH_4 are mostly dominated by wetland emissions, for which we do not distinguish between natural and anthropogenic origin.

The datasets in Table 1 were chosen such that our a priori fluxes are based on the national reporting to the UNFCCC. For Germany, the a priori LULUCF fluxes obtained from Thünen Institute cover the emissions from mineral and organic soils. Notably, this excludes emissions from artificial water bodies – such as ponds – amounting to $160\,\mathrm{kt}$ or $8.5\,\%$ of the total German emissions in the national reporting, though these numbers are associated with large uncertainties (UBA, 2024, Table 399). These emissions are missing in our a priori estimate.

3.3 Observations and applied pre-processing

We compare our model predictions to the high quality ground-based in situ observations collected in the European Obspack (ICOS RI et al., 2024), which includes the ICOS stations among others. These observations are assumed to be representative for a larger area (Storm et al., 2023). Table A1 lists all 53 available stations and Fig. 3 shows 50 stations that were used for the inversion. For tower observations, we only consider sampling levels at least 50 m above ground level to reduce the influence of very local emissions. For towers providing more than three sampling heights, we consider the three highest levels. Due to significant model–observation mismatch, we exclude the IPR, FKL and LMP stations. For LUT, BIR and HUN we only consider some seasons, specified in Table A1.

The model data are interpolated horizontally and vertically to the station sampling locations. The vertical sampling locations in model coordinates is derived from the station sampling height and the modeled station elevation, depending on the station characteristics (column “mountain” in Table A1). For high mountain stations, the modeled station elevation is given by the real station elevation above mean sea level. For stations on smaller mountains, we consider the arithmetic mean between real station elevation and model topography as proposed by Brunner et al. (2012) and Henne et al. (2016), and for all other stations the modeled station elevation is set to the model topography.

To make use of observations which are likely well represented by the model, we filter the observations based on the local time of day, wind speed, and model–data mismatch. Table 2 lists how the root mean square error (RMSE) of the model output changes during these pre-processing steps. We start by smoothening both observations and modeled concentrations in a time window of approximately $\pm 1.5\,\mathrm{h}$ around each observation time as depicted in Fig. 1. This allows for some uncertainty in the timing of modeled tracer transport. The resulting correlation of neighboring time steps is automatically considered in the ensemble-based uncertainty estimate.

In the next steps, we filter the data by time in order to keep only observations expected to be representative for large regions. Observations within the planetary boundary layer are most representative in the afternoon hours whereas measurements at high mountains have less local influence at night time (Bergamaschi et al., 2015). We therefore use the time windows 23 h to 5 h (local mean time) for stations on high mountains and 11 h to 17 h for all other stations.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Table 2. Average root mean square error (RMSE), mean absolute bias of the model prediction minus observation, and number of available data points after each processing step (1–6). Each row adds a processing step to all previous steps. Step 7 (uncertainty weighting) is not a processing step in the inversion since it uses only the diagonal of the uncertainty matrix R, but it underscores the importance of accurate uncertainty estimation. Step 8 refers to the result of our inversion. RMSE and bias are computed separately for each station, sampling height and month. The obtained values are weighted by the number of data points and averaged. By taking the mean of multiple RMSEs for different stations, sampling heights and months, we obtain lower numbers than for the RMSE of the combined dataset, which would average squared values and thereby would give higher weight to large deviations between model and observations.

	Step	Mean RMSE	Mean absolute bias	Data points
1	horizontal and vertical interpolation	27.6 ppb	9.6 ppb	$6.02 \cdot 10^5$
2	time average (3 h)	25.8 ppb	9.6 ppb	$6.02 \cdot 10^5$
3	time window 11 h–17 h / 23 h–5 h	23.5 ppb	9.8 ppb	$1.48 \cdot 10^5$
4	minimal wind speed $2\,\mathrm{m\,s^{-1}}$	22.4 ppb	9.7 ppb	$1.30 \cdot 10^5$
5	extreme mismatch exclusion	21.5 ppb	9.4 ppb	$1.29 \cdot 10^5$
6	far-field correction	19.4 ppb	7.2 ppb	$1.29 \cdot 10^5$
7	weight by inverse uncertainty	16.6 ppb	6.6 ppb	$1.29 \cdot 10^5$
8	inversion (posterior fluxes)	14.1 ppb	3.0 ppb	$1.29 \cdot 10^5$

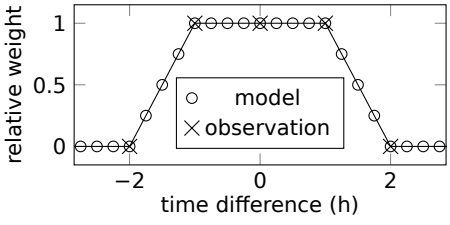


Figure 1. Weighting function for time interpolation of model and observations. For example, an interpolated model point at 16:30 UTC averages over all model output between 15:30 UTC and 17:30 UTC with full weight and another 1 h with linearly decreasing relative weight. The model yields instantaneous values every 15 min, whereas observations are provided as hourly averages, three of which contribute to the observational time average. Reference times are those times for which observations are available.

We furthermore exclude times with no wind to avoid a strong influence of local emissions that are not resolved in the model, motivated by Ganesan et al. (2015). All data points for which the model predicts a wind speed of $< 2\,\mathrm{m\,s^{-1}}$ are excluded, which improves the overall agreement of model and observations as shown in Table 2 (step 4). Figure 2 shows that the RMSE indeed increases significantly at low wind speeds. This increase is partially captured by an increase of the ensemble spread, supporting the idea of an uncertainty estimate depending on wind speed as proposed by Bergamaschi et al. (2022).

In the last filtering step – step 5 in Table 2 – we exclude data points with extreme mismatch between model and observations of more than $200\,\mathrm{ppb}$. Data points where the observations are more than $20\,\mathrm{ppb}$ below the model-predicted far field are also

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

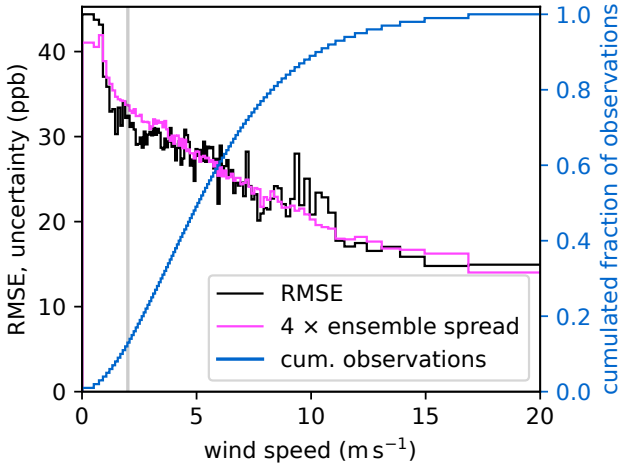


Figure 2. RMSE dependence on wind speed (left axis). All data points from step 3 in Table 2 were ordered by the model-predicted wind speed and split into 100 bins, each containing approximately 1500 data points. The blue line indicates the cumulative fraction of observations (right axis). The figure shows the RMSE difference of model and observation (black line) and the mean ensemble spread multiplied by factor 4 (magenta line) for each of these bins. The ensemble spread is the standard deviation of the model prediction in the 12 ensemble members. It is a main contribution to our uncertainty estimate for the model–data mismatch. Much of the larger RMSE at low wind speed is well captured by the ensemble spread inflated by factor 4. In the inversion, we discard data points with wind speeds below $2\,\mathrm{m\,s^{-1}}$ (gray vertical line).

discarded. Since no strong sinks of CH_4 are expected, the contribution of CH_4 from the lateral boundaries should not exceed the observations. Thus, an observation below the model-predicted far field indicates an error in this far field. Steps 6–8 in Table 2 complete our processing chain by applying the far-field correction (Sect. 2.2), indicating the relevance of the model uncertainty (Sect. 2.4), and finally yielding the inversion results.

4 Application to Germany and neighboring areas for the year 2021

We apply our method to estimate the German CH_4 emissions in 2021. To this end, we consider the domain depicted in Fig. 3 and run the transport model for 2021 with a spin-up period of 17 days, starting on 15 December 2020. In this section, we provide details specific to this application.

4.1 Implementation of CH_4 fluxes

For the inversion, we define the flux categories based on sector and region with the primary aim of providing an accurate estimate for emissions from Germany, resolving federated states where possible, to address the user requirements of potential stakeholders.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

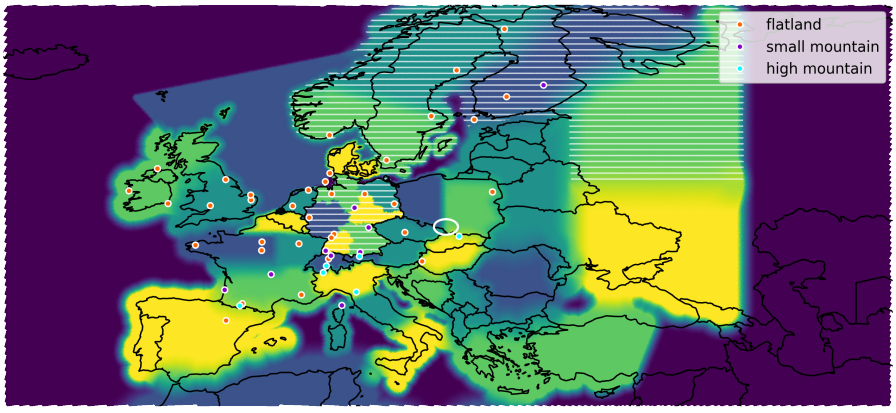


Figure 3. Model domain, colored to distinguish 35 patches defining regional flux categories. Observation sites (dots) are colored by the choice of model equivalent height (see Table A1). Dark blue at the domain boundary indicates regions for which emissions are not categorized and therefore not modified in the inversion. Other colors only distinguish neighboring patches. In white hatched regions, natural fluxes are also categorized and scaled. A white ellipsis marks the Upper Silesian Coal Basin, in which fugitive emissions define their own flux category. In Germany, the map shows the six regions used for the agricultural sector. For other sectors in Germany, we use four regions: south (yellow and light green), west (dark blue), north (light green), and east (dark green and yellow).

4.1.1 Definition of flux categories

For the agricultural sector (GNFR sectors K+L), which contributes roughly two thirds of all German CH_4 emissions, we distinguish six regions within Germany as depicted in Fig. 3. We furthermore try to distinguish the sectors waste (GNFR sector J) and public power (GNFR A) from the sum of all remaining sectors ("other", GNFR B–I). However, we will only present results in which these sectors are combined (see discussion in Sect. 6.3 and Appendix I). For the sectors waste, public power, and other we distinguish four regions, i.e., the federated states south: Baden-Wuerttemberg and Bavaria, west: North Rhine-Westphalia, Hesse, Rhineland-Palatinate and Saarland, north: Lower Saxony, Bremen, Hamburg and Schleswig-Holstein, as well as east: Mecklenburg-Western Pomerania, Brandenburg, Berlin, Saxony, Saxony-Anhalt and Thuringia. Natural plus LULUCF fluxes in Germany are treated as a single flux category.

Outside Germany, we do not distinguish sectoral emissions, with one exception. Agriculture emissions in the Netherlands form their own category, as we found that they strongly influence the CH_4 concentrations in Germany, caused by the proximity and high emission rates in the Netherlands. We define categories by area for anthropogenic emissions excluding LULUCF such that a comparably high resolution is obtained in regions near Germany with high observation coverage. These area-defined categories follow borders as feasible for the inversion. Areas with small expected influence on inversion results for Germany are combined in large categories, such as Spain plus Portugal, Türkiye plus Greece, and large areas east of Poland. All area-defined categories are shown in Fig. 3. We mitigate and analyze the plume problem (Sect. 2.5) in our inversion system

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

by defining a separate flux category for the strongest CH_4 plume in Central Europe. These are fugitive emissions from the Upper Silesian Coal Basin with yearly emissions of $567\,\mathrm{kt}$ in our prior (white ellipsis in Fig. 3).

We treat natural plus LULUCF fluxes separately and categorize them only in Germany, Scandinavia, and the north-eastern part of our domain (hatched regions in Fig. 3). This is motivated by strong CH_4 emissions from wetlands in summer in Scandinavia and northern Russia in our prior (Segers and Houweling, 2020). Uncategorized fluxes – whether natural or anthropogenic – are not scaled in the inversion, but still included in the transport simulation such that no fluxes are discarded. To avoid strong gradients in the concentration fields, the boundaries between different area-defined categories are smoothened as visualized in Fig. 3.

4.1.2 Tracer assignment in the transport model

In the transport simulation, we consider not only the categorized fluxes, but also the CH_4 from lateral boundaries and from uncategorized emissions. Overall, we simulate the transport of 58 tracer fields in the deterministic model run:

(i) **All anthropogenic emissions excluding LULUCF.** This constitutes a single, common tracer.

(ii) **All natural plus LULUCF fluxes.** This constitutes another single, common tracer, which summed with (i) covers all a priori emissions in the domain.

(iii) **Flux categories.** For each flux category an own tracer field is defined. To avoid the accumulation of CH_4 beyond the time scale on which we consider the modeled transport reliable, we set an artificial decay rate of these concentrations. After emission, the concentration in these tracer fields decays with a mean lifetime of five days. In combination with (i) and (ii) this technical feature allows a waning of sectoral attribution over a few days – no CH_4 is lost, though, as full CH_4 transport is modeled without decay in (i) and (ii). This regulates that any attribution of a CH_4 anomaly to a certain region or sector is only attempted if the emission was fresh or a few days ago. Furthermore, this allows us to save computing time by limiting these flux category tracer fields to altitudes below $8\,\mathrm{km}$. In total, 49 categories for anthropogenic emission excluding LULUCF are covered – including 18 sector-resolving categories in Germany, two sector categories in the Netherlands and one plume category – and complemented by five natural plus LULUCF flux categories.

(iv) **Far field.** The far field contains the CH_4 from initial and lateral boundary conditions.

(v) **Auxiliary field for plume detection.** For the purpose of investigating the model uncertainty due to the plume from the Upper Silesian Coal Basin, an auxiliary tracer is added.

In the post-processing of the model data, the categorized concentrations from (iii) are subtracted from the combined tracers (i) or (ii).

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

4.2 Construction of the model error covariance matrix

In Sect. 2.4, we provided a general construction of the model error covariance matrix R based on an ensemble of transport realizations. For the application to Germany, we need a specific choice of the localization matrix C_{ij}, which suppresses spurious correlations arising from the small transport ensemble. Here, we choose C_{ij} to be a Gaussian localization matrix with standard deviations $6\,\mathrm{h}$ (time), $319\,\mathrm{km}$ (horizontal) and $400\,\mathrm{m}$ (vertical). Moreover, we add two further terms to the uncertainty matrix R as follows.

First, we recall that transport errors can cause a bias for a plume (Sect. 2.5). We mitigate this problem by increasing the uncertainty for all observations that may be influenced by the plume from the Upper Silesian Coal Basin. Before the uncertainty inflation, we replace R'_{ij} by $R'_{ij} + 0.25\rho_i^2\delta_{ij}$ where ρ_i denotes the concentration of an auxiliary tracer. This tracer contains only the emissions of the plume, spatially smoothened on the length scale of $0.4°$ (one standard deviation of a Gaussian filter). Next, the model uncertainty is inflated by a factor 2 or 3, depending on the station, see Table A1. After the uncertainty inflation, we account for the uncertainty in the far-field correction. We replace R_{ij} by $R_{ij} + 0.5|c_i c_j|\tilde{C}_{ij}$ where c_i denotes the smooth correction field introduced in Sect. 2.2 and $\tilde{C}_{ij}$ is the Gaussian localization matrix constructed by the length and time scales of the far-field correction (see Appendix B).

To assess whether the estimated uncertainties are reasonable, one can compute the χ^2/N_{dof} value. This value compares the a priori model–observation mismatch to the uncertainty assumed for this mismatch (see Appendix C for details). A value of $\chi^2/N_{\mathrm{dof}} > 1$ indicates that uncertainties are underestimated, whereas values smaller than one indicate the opposite. When comparing the observations to the far-field-corrected model (step 7 in Table 2), we find $\chi^2/N_{\mathrm{dof}} \approx 0.16$ (median of 12 months). In an idealized setup, this indicates that the uncertainties of the model-data mismatch are overestimated by a factor 2.5. This implies that our uncertainty inflation – by a factor 2 for most observations – seems unnecessary in the idealized setup. However, we work with real data that can contain unknown biases in transport and boundary conditions, and simplifying assumptions about the representativity of the low-dimensional state space of the inversion. We contain these potential issues of unknown error components by inflating the uncertainties, for sensitivity tests see Appendix E.

4.3 Inversion time window and temporal aggregation

We apply the inversion separately for each month in 2021. In each monthly time window, we start from fixed a priori scaling factors $s_i^{\mathrm{prior}} = 1$ and use observations within the time window to compute a posteriori scaling factors as explained in Sect. 2.3. However, when aggregating results for the whole year, we treat the uncertainties of the prior or posterior fluxes of different months as correlated because these likely include systematic uncertainties and biases which we cannot fully separate from the statistical uncertainty. We therefore aggregate by adding up absolute emissions and their uncertainties linearly.

4.4 Prior uncertainties

In each inversion time window, we consider uncorrelated a priori scaling factors with a two standard deviation (2σ) uncertainty of $80\,\%$ for most flux categories, corresponding to a $95\,\%$ confidence interval of $\pm 80\,\%$. Throughout this paper, uncertainties

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

will denote two standard deviations or 95 % confidence intervals. Categories resolving emission sectors have a higher prior 2σ uncertainty of $\pm 100\,\%$, and within Germany categories describing the same sector have an a priori uncertainty correlation of 50 % (e.g., uncertainties of agriculture emissions in the German states of Bavaria and Baden-Wuerttemberg are assumed to be correlated). Importantly, the previously defined sectors waste, public power and "other" are combined and denoted "non-agr." in the following. The $\pm 100\,\%$ uncertainty applies to the combination. For the Upper Silesian Coal Basin as well as regions outside of our primary focus in Central Europe and with low observation density, the 2σ uncertainty is set to $\pm 50\,\%$. Figure 4(b) shows these a priori uncertainties on a map. The sensitivity of our results to these choices and many more tuning parameters is tested in Appendix E.

4.5 Posterior uncertainty estimates

Our inversion setup necessarily makes idealized assumptions on uncertainties of a priori and a posteriori fluxes. Based on these assumptions, the inversion yields posterior statistical uncertainties as part of the posterior error covariance matrix. However, these uncertainties neglect possible unknown biases and other systematic errors which we need to anticipate when working with real data. We therefore use an enhanced notion of posterior uncertainty that combines statistical and methodological uncertainties. The main contribution to the posterior uncertainties is the statistical uncertainty, which we provide as two standard deviations (95 % confidence interval). Additionally, we combine the two variants of inversion (prior-R and posterior-R, see Sect. 2.4.2) by taking the arithmetic mean of the two separate inversion results, arriving at the combined scaling factors. Assuming that the difference of the two variants indicates an additional, methodological uncertainty, our combined uncertainty range includes the uncertainty ranges of both individual results.

Furthermore, each observation site may have a bias due to very local pollution or topography that is not modeled properly. We therefore only consider an inversion result reliable if it is robust with respect to the choice of used observations. To represent this in the uncertainty estimate, we repeat the inversions 50 times for both prior-R and posterior-R, excluding each station once. The final posterior uncertainty range is spanned by the lower and upper bounds of the 100 uncertainty ranges, each describing a 95 % confidence interval. This method of compensating for possible methodological weaknesses in a cautionary uncertainty estimate implicates that the posterior uncertainty may not necessarily be smaller than the prior uncertainty. Such cases we interpret as no information gain from the observation-based inversion with respect to the prior.

5 Results

This section presents the inversion results for Germany and the considered European regions, along with examinations of the seasonal cycle, validation of the results, and sensitivity tests. All uncertainty estimates are presented as 95 % confidence intervals as detailed above.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

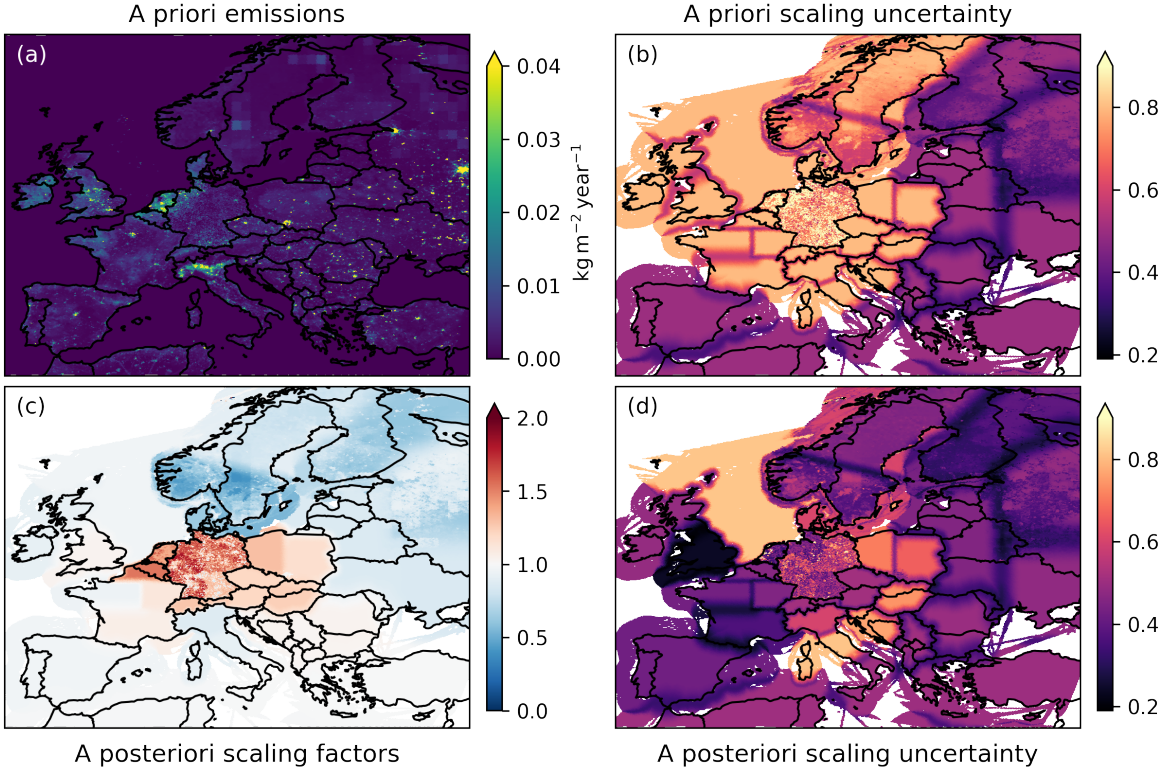


Figure 4. Full-year averages of (a) a priori fluxes, (b) a priori uncertainty on scaling factors, (c) a posteriori scaling factors, and (d) a posteriori uncertainty on scaling factors. Multiplying the a priori emissions (a) with the scaling factors (c) yields the a posteriori emissions. (b) and (d) show half of the 95 % confidence interval of the fluxes relative to the a priori fluxes, i.e., a 2σ uncertainty of $\pm 50\,\%$ on the a priori appears as 0.5 (i.e., 50%) on the color scale. The direct comparison indicates the uncertainty reduction.

5.1 Resulting scaling factors

Figure 4 presents an overview of (a) the a priori CH_4 fluxes, (c) the resulting scaling factors, and the respective uncertainties (b, d), all accumulated over the year 2021. The a posteriori scaling factors (Fig. 4 c) show the correction to the a priori emissions obtained in the inversion. A considerable increase in emissions is found for Germany and the Benelux. Lower emissions compared to the a priori are predicted for Scandinavia (see discussion in Sect. 6.3). The scaling factors should be considered jointly with their uncertainties. The comparison of Fig. 4 (b) and (d) shows a substantial uncertainty reduction for Germany and most of the surrounding countries, for which we chose a high a priori uncertainty. For a more detailed comparison of a priori and a posteriori emissions and uncertainties, we consider selected national emission estimates in Fig. 5.

Reliable inversion results are expected for countries or regions with sufficient observation coverage, strong emission signals, representation in the respective flux categories, and only moderate issues due to complex topography. These criteria are met for Germany, the Netherlands and the United Kingdom plus Ireland as grouped in Fig. 5. For Germany (first line in Fig. 5),

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

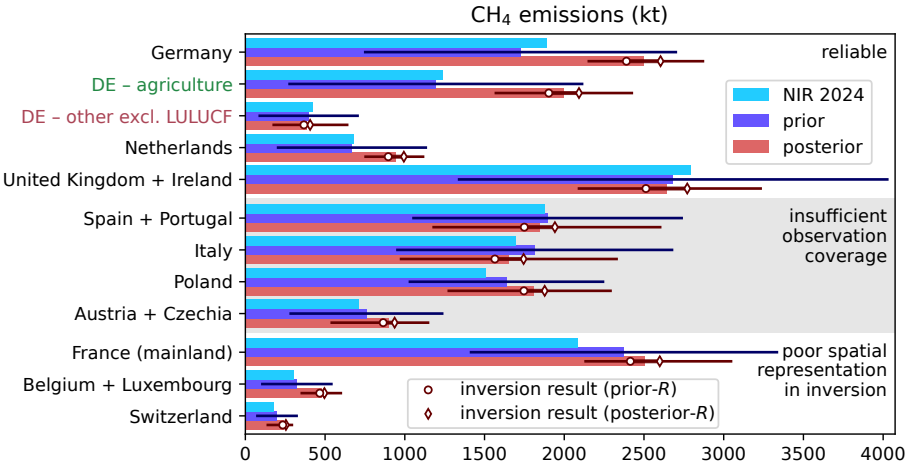


Figure 5. National CH_4 emission estimates comparing reported (NIR), prior, and posterior fluxes for 2021 with horizontal lines indicating $95\,\%$ confidence intervals. Countries are grouped by the expected robustness of their inversion results. Some neighboring countries are combined to obtain more accurate results. For Germany, the inversions can resolve the agricultural sector, though the separation against natural and LULUCF fluxes is difficult. All other anthropogenic sectors are combined in the category "other excl. LULUCF". The inclusion of two inversion methods (prior-R and posterior-R, markers) yields an estimate of the methodological uncertainty. Accumulated fluxes from national inventory reports (NIR) to the UNFCCC submitted 2024 (including LULUCF emissions) are shown for reference (light blue bars, UNFCCC, 2024). For France (Citepa, 2024) and the United Kingdom (Department for Energy Security and Net Zero, 2024), the light blue bars show emission data from the respective inventory agencies excluding overseas territories and crown dependencies. Posterior uncertainties asymmetric with respect to flux estimates such as in Switzerland indicate the strong influence of a single observation site.

the total posterior CH_4 emissions (red bar) are $(32 \pm 19)\,\%$ higher than the anthropogenic emissions (including LULUCF) reported to the UNFCCC in 2024 (light blue bar). The direct comparison to the reporting neglects the unreported natural fluxes, but for Germany these are expected to be small because all relevant soil emissions are included in the LULUCF sector. The inversion significantly increases emission estimates from the agriculture sector while the combined other sectors remain nearly unchanged. Note, however, that the uncertainty in the sector attribution is large (horizontal lines, see further discussion in Sections 5.5.2 and 6.3).

For the Netherlands, we also find significantly higher emissions than in the inventory. Compared to Germany, the attribution to sectors has an even larger uncertainty, associated with fewer observations that could distinguish the sectors. Nevertheless, the total emissions from the Netherlands are comparably well constrained by the observations. For the United Kingdom and Ireland – which we combine to obtain more accurate results – the inversion yields a strong uncertainty reduction while hardly changing the total emissions, indicating a good agreement of observations and national inventory.

In most countries, the observations do not cover the whole country, or the inversion results rely on few observations. In Fig. 5 (gray-shaded part) we provide emission estimates also for countries or regions affected by this issue, though these have a large posterior uncertainty. Another issue arises from the definition of the flux categories, which do not necessarily follow

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

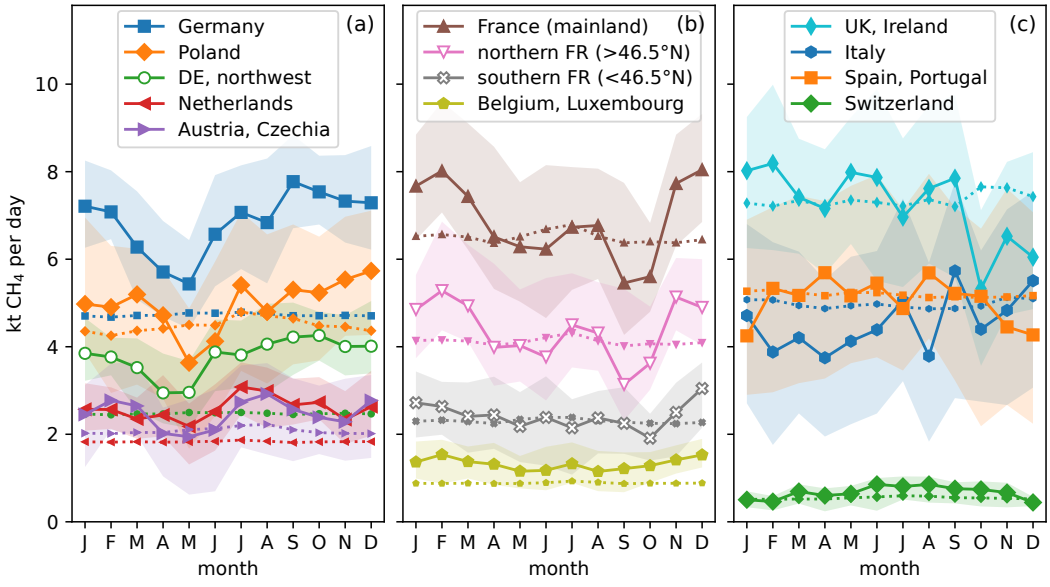


Figure 6. Monthly posterior emission rates for selected countries or regions. Colored areas show the posterior uncertainties, and dotted lines with small markers indicate prior emission rates. In the prior, only the natural and LULUCF fluxes are time-dependent. The panels show (a) countries with minimum in May, (b) countries with a maximum in winter, and (c) other countries and regions. For France and Germany, selected regions are shown additionally (white markers). "DE, northwest" includes Rhineland-Palatinate, Saarland, Hesse, North Rhine-Westphalia, Lower Saxony, Schleswig-Holstein, Bremen and Hamburg.

country borders (see Fig. 3). In France, Belgium, and Switzerland, the inversion scales flux categories overlapping multiple countries[2]. This implies that national emission estimates derived for these countries have an additional uncertainty and artificial correlations with neighboring countries. However, this is of no concern for our application for Germany. The national emission estimates are computed from the gridded posterior fluxes and precisely follow the country borders as shown in Fig. 4. The scaling factors and uncertainties of all flux categories are listed in Fig. A1 for completeness.

5.2 Seasonal cycle

Although the national emission estimates are given for the full year, a closer examination of the seasonal cycle yields additional insights. The posterior fluxes are computed independently for each month and region. Figure 6 shows the monthly emission rates for the countries considered in Fig. 5. While the seasonal cycle is strikingly different depending on the region, we find some recurrent features. For Germany, Poland, the Netherlands, and Austria plus Czechia (panel (a) in Fig. 6), the posterior emission rates have their minimum in May. A local minimum between April and June is also found for northern France and

[2]Technically, the issue also affects Italy because Corsica is combined with parts of Italy in one flux category. But the a priori emissions from Corsica are so low that the effect on the national emission estimate is negligible.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Belgium plus Luxembourg, see panel (b). In most countries, this minimum is followed by a local maximum in July or August, which is most prominent in the Netherlands and Austria plus Czechia (panel (a)).

The differences between the regions become larger in autumn and winter. In September, posterior emission rates reach their maximum in Germany and Italy, and their minimum in (northern) France. France and Belgium plus Luxembourg have their highest emission rates in winter, when Switzerland and Spain plus Portugal have their minimum. For some regions – most notably Italy and the United Kingdom plus Ireland – no clear pattern is found in the seasonal cycle for 2021 (panel (c) in Fig. 6).

The seasonal cycle in the inversion results may be partially influenced by the observation coverage because many stations lack data covering the whole year. To avoid this effect, we repeated the inversion using only stations which provide data for at least 20 days of each month. The seasonal cycle in these results does not change significantly, see Fig. A2. We further note that there is a seasonal cycle in the observations (East et al., 2024), which is captured well by the far field in the model though (see Fig. A3). A possible bias in the lateral boundary conditions determining the far field could also influence the seasonal cycle in the estimated fluxes. Moreover, the different meteorology in summer and winter – especially influencing the planetary boundary layer and vertical mixing (Seidel et al., 2012) – can lead to a seasonal bias in our transport model (Bessagnet et al., 2016; Canepa and Builtjes, 2017). This highlights the need for careful interpretation of the seasonal cycle, as meteorological differences could introduce biases that mask true emission patterns. Another potential contribution to the seasonal cycle could arise from neglecting the OH sink of CH_4 in our limited domain.

5.3 Validation

A straightforward validation of the inversion results is possible using independent validation stations. Having excluded each station once in separate inversion runs, we can use every station as an independent validation site in the respective inversion run. Figure 7 shows histograms of the RMSE statistics evaluating the model–data mismatch before and after the inversion. The validation stations agree on average significantly better with observations when using a posteriori emissions compared to the a priori.

5.4 Sensitivity tests

Our inversion method has various tuning parameters. Above we have described the inversion and its results for one choice of these parameters. We analyzed the sensitivity of these parameters by repeating the inversion 48 times with modified parameters. Table E1 lists these test cases with their ID, parameters, and influence on the inversion results. An overview of the national emission estimates for each test case is provided in Fig. E1. Here, we summarize the main results and refer to Table E1 for details.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

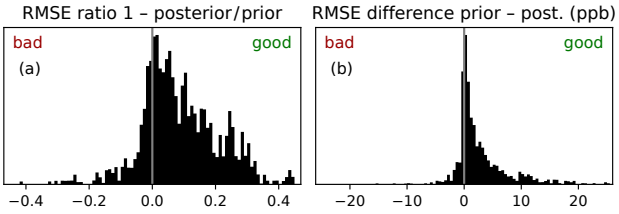


Figure 7. Statistics of the relative (a) and absolute (b) improvement of the model–observation mismatch by the inversion at independent validation stations. Each station and month is considered separately in its own (prior-R and posterior-R) inversion, with the validation station excluded from the inversion to remain independent. The histograms show (a) $1 - r^{\text{post}}/r^{\text{prior}}$ and (b) $r^{\text{prior}} - r^{\text{post}}$ where r^{post} and r^{prior} refer to the RMSE of the model–observation comparison in the case of posterior scaling and prior scaling, respectively. Each time series contributing to the histogram is weighted by the number of its data points. We consider all data points within the daily time window without filtering for wind speed, extreme mismatch, uncertainty weighting or far-field correction (i.e., after step 3 in Table 2) to keep the comparison as close as possible to the original data. Positive values indicate an improvement in the model prediction due to the inversion.

5.4.1 Comparison to observations

The filtering steps listed in Table 2 influence the inversion results significantly. Most prominently, selecting nighttime observations for high mountain stations and afternoon hours for other stations strongly affects the inversion and improves the model representativeness (case 201 in Table E1). This is one of only four sensitivity tests with posterior fluxes deviating from the reference case by $\gtrsim 30\,\%$ of the posterior uncertainty, which we call a strong change in inversion results. Other filtering parameters such as the number of sampling heights per station (case 202) and the minimal wind speed (cases 203–205) affect the inversion results noticeably, although changes are small compared to the uncertainties. Neglecting extreme outliers has only a small effect (cases 206, 207), but limiting the influence of outliers by increasing their uncertainty has a considerable impact (cases 208, 209).

The choice of observation sites is analyzed in cases 601 and 602, which select subsets of stations with good observation coverage over the full year. When using only 27 stations (case 602), the results change strongly compared to the reference case with 50 stations, also because some regions are hardly observed in case 602 (compare also Fig. 6 with Fig. A2). Varying the elevation of high mountain stations has only little impact on the inversion results (case 100). The effect of time-averaging over $3\,\text{h}$ (as chosen in step 2 of Sect. 3.3) is noticeable in the results, but small compared to the uncertainties (case 101).

5.4.2 Uncertainty

The construction of the error covariance matrix R following Sections 2.4 and 4.2 contains numerous tuning parameters. Key parameters are the overall uncertainty inflation factors f_i (cases 302 and 303 in Table E1) and the uncorrelated additive uncertainty σ_{const} of each data point (cases 309, 310). Variations of these parameters change the inversion results considerably. The tuning parameter σ_{const} illustrates the importance of hidden patterns in the considered data. Increasing to $\sigma_{\text{const}} = 20\,\text{ppb}$ effectively reduces the weight of observations with a small ensemble-estimated transport uncertainty. As observations with

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

strong emission signals and high transport uncertainty become more relevant, the emission estimate for Germany is increased by 5 % (case 310 in Fig. E1).

Other important parameters are the correlation scales in the localization for the ensemble-based uncertainty estimate. The overall effect of these scales on the posterior scaling factors is small (cases 304–308), but these parameters also influence the posterior uncertainties. The sensitivity tests indicate that 12 ensemble members are sufficient to estimate the uncertainties and correlations even without a strong localization. In general, we expect that a larger transport ensemble will yield better statistical estimates for uncertainties and their correlations. This reduces the need for a localization which suppresses spurious correlations. The plume localization uncertainty (see Sect. 2.5, cases 300 and 301) arising from the Upper Silesian Coal Basin seems negligible when considering the full domain. However, the additional plume localization uncertainty reduces the negative bias for the plume emissions (see Sect. 5.6).

5.4.3 Far-field correction

The far-field correction explained in Sect. 2.2 strongly influences the results. Removing the correction field leads to striking changes in the inversion results, including unrealistic negative scaling factors for some natural flux categories (case 400). However, changing various tuning parameters of the far-field correction within a reasonable range has much smaller effects. The selection of data points used for the far-field correction (cases 409, 410) and the overall correction strength (cases 401, 402) have modest influence, whereas correlation scales in the correction play a minor role (cases 403–408). The additional uncertainty added to R due to the far-field correction has little influence on the inversion results (cases 412–414). We draw the conclusions that (i) the far-field correction is important for the inversion and (ii) the inversion extracts most information from signals that are hardly affected by the precise form of the far-field correction, especially from strong emission signals ($\gg 20\,\mathrm{ppb}$).

5.4.4 A priori error covariance matrix

Modifying the a priori uncertainty or correlations of the scaling factors (B in Eq. (1)) changes the results quantitatively, but not qualitatively. We notice that a smaller a priori uncertainty (case 500) narrows the ability to discriminate sector emissions, because the sector attribution tends to follow the a priori uncertainties. A coarser spatial resolution in Germany (case 504) and different choices of sectors (cases 503, 506) yield aggregated German sector emissions that agree well with the reference case.

5.4.5 Inversion time windows

In the reference case, we considered each month independently. Increasing the inversion time windows to three months has a considerable influence on the results (case 702). As the inversion time window increases, the overall weight of the observations in the inversion also increases. Thus, posterior uncertainties are reduced and the deviations between posterior and prior are amplified.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

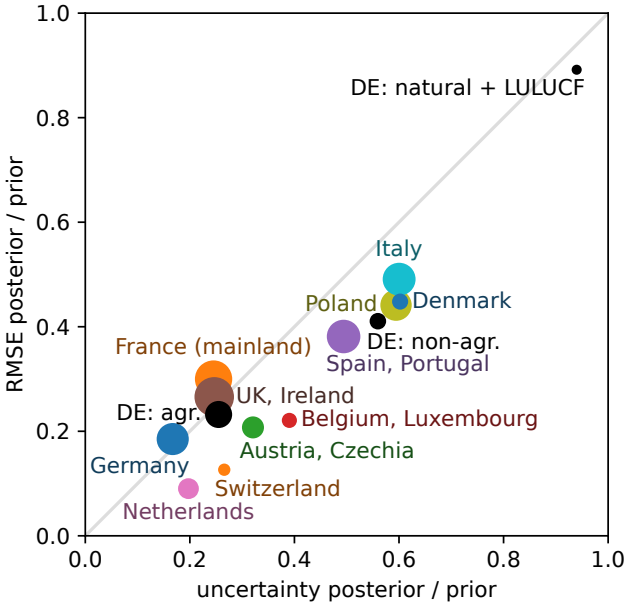


Figure 8. RMSE and mean uncertainty of CH_4 emission estimates in synthetic experiments for selected countries, regions, and sectors. Each of the 100 synthetic experiments generates random true emissions. The vertical axis shows the root mean square (RMS) deviation of the posterior from these true emissions, relative to the RMS deviation of the prior from the truth. Lower values indicate that the inversion improves the emission estimate. The horizontal axis shows the posterior uncertainty relative to the prior uncertainty. The disk size indicates the amount of the prior emissions.

5.5 Potential for detecting emissions

In this section, we complement the uncertainty estimates of our inversion results by separate measures for the sensitivity of the posterior to true emissions. The potential for detecting emissions from different sources can be identified using the posterior error covariance matrix B_{post}. However, the real error reduction is also influenced by the far-field correction and the filtering of observations, which is not fully captured in B_{post}. We therefore use experiments with a "synthetic", i.e., define truth and pseudo-observations to test the full inversion system.

5.5.1 National emission estimates

We first aim to verify that the inversion yields meaningful posterior emission estimates and uncertainties given a perfect transport model. To this end, we generate 100 random vectors of scaling factors following the probability distribution assumed in the a priori uncertainty. These scaling factors define the synthetic truth, and the model prediction for the observations obtained using these scaling factors defines our pseudo-observations. We further add uncorrelated Gaussian noise of standard deviation $2\,\text{ppb}$ to these pseudo-observations. Since the pseudo-observations are inferred from the model data, there is no transport error in these synthetic experiments.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

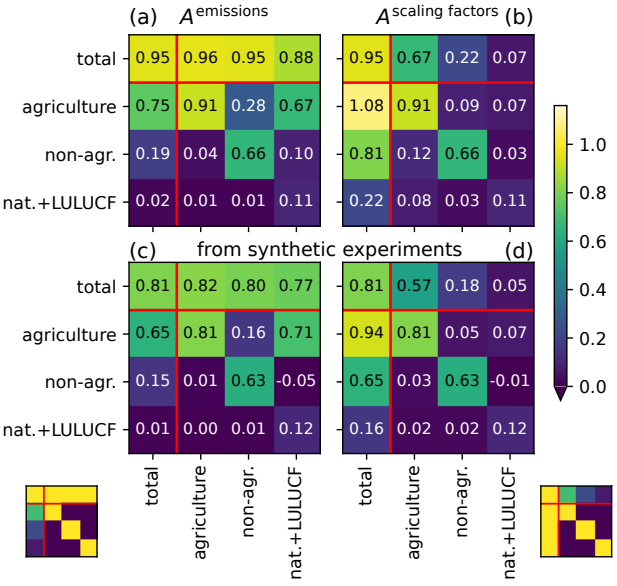


Figure 9. Averaging kernel matrices of German sector emissions (a, c) and scaling factors (b, d). The kernel is estimated using either the posterior covariance matrix (a, b) or 100 synthetic experiments with random truth (c, d). The small matrices on the bottom indicate what we aim for (posterior equals truth). The value 0.96 in the first row ("total"), second column ("agriculture") of panel (a) means that if in reality all German agriculture emissions were $1\,\mathrm{kt}$ higher than in our prior, then we would expect an increase in the posterior total German emissions by $0.96\,\mathrm{kt}$. Similarly, the value 0.67 in the same cell of panel (b) means that increasing real agriculture emissions by $10\,\%$ should increase our posterior total emissions by $6.7\,\%$. All matrices are averaged over the whole year. Red lines separate the individual sectors from their sum ("total"). By "non-agr." we denote anthropogenic emissions excluding agriculture and LULUCF.

The quality of the model prediction for this synthetic truth is shown in Fig. 8 for selected countries and German sectors. By comparing to the synthetic truth, we find the prior and posterior error. Their ratio (vertical axis in Fig. 8) shows a significant improvement by the inversion for all considered regions and German sectors, with the exception of German natural and LULUCF fluxes. The uncertainty reduction of the inversion (horizontal axis) provides a realistic estimate of the real error reduction (vertical axis) for the case of high quality observations, ideal transport modeling and perfect far field. In some cases (Netherlands, Switzerland, Belgium, and Luxembourg), the real error reduction is significantly better than the uncertainty reduction suggests. This is no surprise because in this synthetic setup the transport error as the main source of uncertainty is switched off. Overall, the synthetic experiments confirm the potential for a strong uncertainty reduction in Central Europe.

5.5.2 Distinguishing sectors in Germany

Within Germany, we distinguish agriculture from other emissions. The ability to distinguish sectors can be described by averaging kernel matrices which estimate the dependence of the posterior on the true emissions, $A^{\text{emis}}_{ij} = \partial e^{\text{post}}_i / \partial e^{\text{truth}}_j$ where e_i denotes emissions from sector i. Since the true emissions e^{truth} are generally unknown, the averaging kernels A^{emis} can only

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

be estimated. Figure 9 shows such estimates for A^{emis} (panels a, c) and the averaging kernel for scaling factors, $A_{ij}^{\text{scaling factors}} = \partial s_i^{\text{post}} / \partial s_j^{\text{truth}}$ (panels b, d). Assuming a perfect transport model and perfect far field, the averaging kernel matrix can be estimated by $A^{\text{emis}} \approx I - B_{\text{post. emis}} B_{\text{prior emis}}^{-1}$ (Rodgers, 2000) using the prior and posterior covariance matrices of the emissions from the prior-R inversion (see Appendix H1). I denotes the identity matrix. Figure 9(a) shows this averaging kernel estimate for German sector emissions, extended by a row and column for the total German emissions.

The first row of Fig. 9(a) indicates that the total German posterior emissions follow changes in every sector with high accuracy ($88\,\%$ to $96\,\%$). The diagonal of Fig. 9(a) signifies that changes in the agriculture will be detected very well and also the attribution to the sum of all other anthropogenic sectors excluding LULUCF ("non-agr.") will be mostly correct. However, LULUCF plus natural fluxes will in large parts be falsely attributed to the agriculture (second row, last column). Note that ideally, the first row and the diagonal elements would be close to $100\,\%$ (color-coded in the small matrix bottom left). The averaging kernel $A^{\text{scaling factors}}$ in Fig. 9(b) shows that the influence of LULUCF and natural emissions on the posterior scaling factor for agriculture emissions remains low (second row, last column). But if all emissions are scaled by the same factor (first column), the changes will be mostly attributed to the agriculture sector (as explained in Appendix G).

The averaging kernel matrices in Fig. 9(a) and (b) are estimated based on the prior-R inversion while neglecting the far-field correction. We complement these by a statistical estimate of the averaging kernels using 100 synthetic experiments with random truth (see Appendix H2), shown in Fig. 9(c) and (d). Here, the far-field correction is applied as implemented in our processing chain. While these statistical estimates reproduce all qualitative features in the averaging kernels, the matrix entries estimated using synthetic experiments are generally lower. This is likely due to the far-field correction and indicates that deviations from the prior emissions may be underestimated by our inversion. Importantly, both presented strategies for estimating the averaging kernels assume a perfect transport model. The real sensitivity of the posterior to the true emissions is therefore expected to be lower.

5.6 Simulated transport error

In the following, we estimate methodological uncertainties and biases using synthetic experiments with a simulated transport error. Each of our 12 transport ensemble members yields one prediction for the total CH_4 concentrations from which we can generate pseudo-observations. We use these pseudo-observations for 12 inversion runs in which the true emissions are equal to the prior emissions.

Figure 10 shows how emission estimates computed in this synthetic setup deviate from the synthetic truth with (Fig. 10 a) and without far-field correction (b). The two inversion methods prior-R and posterior-R are shown as horizontal and vertical axis, respectively. As expected, the posterior-R inversion generally yields higher emission estimates. Localized sources that cause a strong plume are underestimated by both methods, though the bias is reduced in the posterior-R inversion as predicted in Sect. 2.5. Considering the average of prior-R and posterior-R inversion for the case with far-field correction, we find posterior biases ranging from $-9\,\%$ for the Upper Silesian Coal Basin to $+0.4\,\%$ for natural plus LULUCF emissions in Germany. For the total German emissions, the expected bias of $-2\,\%$ and the ensemble spread of $-4\,\%$ to $0\,\%$ is well below the uncertainty of the inversion results. Similar results for the case of modified true anthropogenic emissions are shown in Fig. F2.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

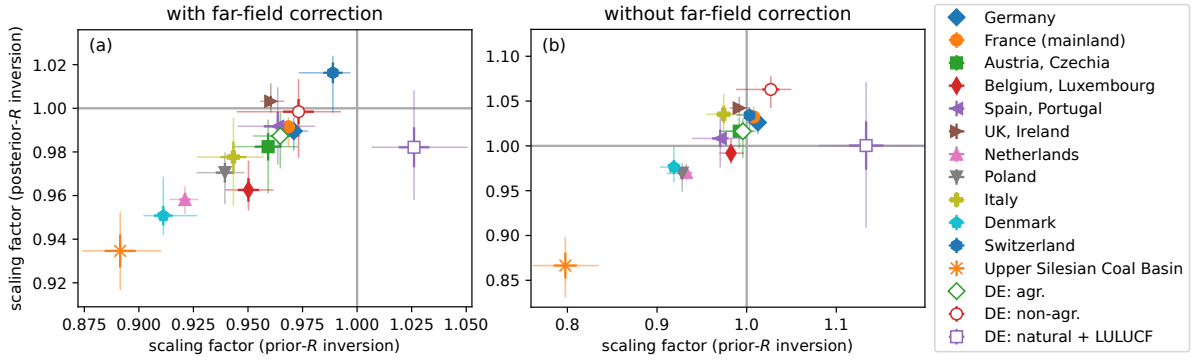


Figure 10. Bias in synthetic experiments with simulated transport uncertainty with (a) and without (b) far-field correction. Note the scale difference between the two panels with scaling factors in (a) much closer to the true value of one. Pseudo-observations were generated from the CH_4 concentration simulated in 12 meteorological ensemble members. Symbols show the mean of inversion results relative to the prior, obtained using either the prior-R (horizontal axis) or the posterior-R inversion method (vertical axis). Thin lines show the spread (minimum to maximum) among the 12 inversions, and thick lines show the 2σ statistical uncertainty of the mean.

In the synthetic setup, the far field is perfect by construction and its correction can be switched off. For this case, Fig. 10(b) shows a significantly larger spread of posterior values compared to the case with far-field correction. This large spread highlights the influence of transport errors on the inversion results. The far-field correction reduces deviations from the prior and introduces a small bias towards lower fluxes. This bias is caused by the filtering of observations with overestimated far field and possibly by the selection of observations for the construction of the correction field.

6 Discussion

Our inversion system combines precise in situ observations, accurate a priori fluxes from national reporting, the ICON–ART transport model at $6.5\,\mathrm{km}$ resolution, and an ensemble-estimated transport uncertainty. We further rely on CAMS boundary conditions and high-resolution meteorological fields from operational numerical weather prediction. This yields in general a good agreement between the model prediction and filtered observations, allowing us robust emission estimates for countries with well-observed emissions, such as Germany. We compare top-down CH_4 emission estimates to the reported German inventory and its agriculture sector with enough accuracy to lay the technical foundations for a future long-term observation-based national inventory verification. This section discusses our main results (Sect. 6.1), including a comparison with other studies (Sect. 6.2). We elaborate the limitations of our approach (Sect. 6.3) and its potential for the development of observation-based national inventory verification to inform climate policy (Sect. 6.4).

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

6.1 Key findings

Firstly, we find that our top-down CH_4 emission estimates are significantly higher than reported for Germany. Secondly, we identify the agriculture sector and possibly LULUCF and natural fluxes as the main source of this discrepancy. Thirdly, we estimated the transport uncertainty using a meteorological ensemble. The same ensemble allowed us to simulate the effect of a transport error on the inversion results, which we estimate to be $2\,\%$ for the total German emissions.

Fourthly, our study points to the relevant tuning parameters. The far-field correction significantly impacts the inversion results, but the specific choices made in its setup have only minor or moderate effects. Switching on the far-field correction lets us err slightly on the side of caution, i.e., our results will be biased towards the prior flux estimates. Without far-field correction, on the other hand, we expect errors from wrongly projecting any boundary bias onto the fluxes. For the station observation filtering (e.g., time window selection, outlier identification), we found a robust setting for our application because of sufficient observation coverage. However, this may become a problem for other applications in less well-observed circumstances. Tuning the model–observation uncertainty parameters proved important, and thus we chose them cautiously to avoid overfitting.

6.2 Comparison to other methods

Our Eulerian approach with sectoral segregation differs from other studies on CH_4 inversions for single countries, e.g., Henne et al. (2016) for Switzerland and Ganesan et al. (2015) for the United Kingdom that use Lagrangian transport models. The latter both qualitatively attribute deviations from the inventory reporting to the agriculture sector by comparing the spatial and/or temporal patterns in the posterior fluxes to sectoral a priori fluxes. A similar strategy for sectoral segregation based on spatial flux patterns is followed by Varon et al. (2022) and analyzed by Cusworth et al. (2021). For deriving sector estimates, some inversions assume a spatial correlation of gridded emissions within each sector (Rödenbeck et al., 2003; Meirink et al., 2008b; Bergamaschi et al., 2010). Based on the same assumption, Steiner et al. (2024b) and Tenkanen et al. (2025) construct ensembles of perturbed a priori fluxes to distinguish natural and anthropogenic fluxes utilizing the CarbonTracker Data Assimilation Shell (van der Laan-Luijkx et al., 2017). Notably, Tenkanen et al. (2025) avoid the lateral boundary problem by simulating transport globally with nested zoom in Europe to estimate Finnish CH_4 emissions on a coarse resolution of $1^\circ \times 1^\circ$. In the present work, we take the next step by validating sectoral emissions reported to UNFCCC and analyzing possible false attributions, making use of a significantly higher model resolution.

Our results are qualitatively in line with the discrepancy of top-down estimates and UNFCCC reporting for Germany and the Benelux found in different regional inversions for the years 2018 and earlier (Petrescu et al., 2023; Bergamaschi et al., 2022, 2018; Steiner et al., 2024b). Furthermore, it appears as a robust feature in our results that emissions from the UK plus Ireland agree well with reported emissions, in line with Bergamaschi et al. (2022) for the year 2018. For the French emissions, our inversion shows a tendency towards slightly higher emissions similar to Steiner et al. (2024b), whereas other inversions suggest significantly higher emissions (Petrescu et al., 2023; Bergamaschi et al., 2022).

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

6.3 Limitations

Though simulating emissions and transport in a large domain, we can only provide reliable emission estimates for selected countries (compare Fig. 5). Regions without notable uncertainty reduction and regions with known modeling difficulties do not benefit from our model setup. In Scandinavia, we find strong wetland emissions with insufficiently modeled fine-scale spatial and temporal variability. Combined with only small signals from non-LULUCF anthropogenic emissions, this leads to a low signal-to-noise ratio, which prevents conclusive results for Scandinavia. Furthermore, the synthesis inversion may be prone to underestimating large localized sources due to transport errors. We quantified this effect (see Sect. 5.6) and reduced the bias by introducing the posterior-R inversion as well as an additional model uncertainty.

Another limitation comes from the challenges for the regional flux inversion caused by biases in the lateral boundary conditions, motivating our far-field correction. The correction effectively reduces the influence of observations with low signal from categorized emissions, while leaving observations with strong signals mostly unchanged. This makes the estimation of well-observed emissions more robust. However, due to the selection of observations for the far-field correction, this correction is more likely to increase the far field rather than decreasing it, leading to a bias towards lower emission estimates, as was proven with synthetic experiments. We expect similar difficulties as long as the bias correction of the lateral boundaries is based on the same observations as the flux estimation. This highlights the difficulty of determining fluxes in a limited area when concentrations at the lateral boundaries cannot be directly inferred from dedicated observations.

In our highly resolved transport simulation, every flux category is numerically expensive. Aiming to validate reported German emissions, we could reduce the state space of the inversion to only 46 scaling factors with monthly time resolution. This substantially limits the spatial and temporal variations that can be represented in the inversion. This approach is justified if the a priori fluxes already provide a realistic spatial distribution of all major CH_4 sources within each flux category. While this may be the case in Germany and neighboring countries, the constant scaling factors for large flux categories in more distant regions may be oversimplified and could lead to less accurate results in these regions. Moreover, adjusting few degrees of freedom may not be sufficient to obtain realistic flux estimates in regions with limited or highly uncertain information on a priori fluxes, such as Scandinavia. The scaling method would further fail to correct zero prior fluxes (Kountouris et al., 2018). However, this is less of a problem for CH_4, as inventories can collect where CH_4-emitting activities are normally located, but emission factors which translate the activities into emissions are generally not well known (Dammers et al., 2024).

When constructing the state space, we unevenly distributed the 46 degrees of freedom on our model domain – using 11 degrees of freedom for Germany and only four for mainland France plus Belgium and Luxembourg. But the choice of flux categories affects the results and can lead to biases depending on the location of the observations (Kaminski et al., 2001). In our application, this effect is small because of the good observation coverage in Germany. Sensitivity tests with 5, 9 and 19 degrees of freedom in Germany (cases 503, 504, 506 in Table E1) indicate robust results for Germany and thereby confirm the potential of the synthesis inversion when focusing on a well-observed region.

We exploit the sectoral discrimination of emission in a well-observed region as a key feature of our inversion method. This relies heavily on an accurate spatial distribution and completeness of the a priori fluxes, which appears to be sufficient for the

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

major emitting sectors in Germany. A general problem in sector attribution is that sectors with large absolute uncertainty – such as agriculture – may be falsely blamed for any change in total emissions when the observations do not clearly distinguish the sectors (see Appendix G). By quantifying this effect in the averaging kernels (see Fig. 9), we confirmed that in Germany agriculture can be distinguished from other anthropogenic emissions excluding LULUCF. Small sectors like natural plus LULUCF fluxes could not be reliably distinguished from large sectors such as agriculture, and we therefore combined smaller sectors like waste and public power into the larger category "non-agr.".

6.4 Implications for future research

We chose the synthesis inversion for the first application of our modular inversion system, but designed this framework to be expandable to other inversion methods. For instance, most of the steps in the inversion can be applied with only minor adjustments when replacing the flux categories by an ensemble of randomly perturbed surface fluxes, similar to Steiner et al. (2024b), or by grid cell clusters as used by Estrada et al. (2024). Such applications with a larger state space are limited by the computational effort of the transport simulation, which is much higher than the computational effort of the inversion itself. Similar to the inversion method, the far-field correction can be replaced by a different strategy for mitigating a boundary bias. For example, one could construct the far field based on an ensemble of boundary concentrations.

Further possibilities of extension concern other observation types, including satellite data. Our Eulerian system allows in principle the handling of large observation datasets without prohibitive computational effort, albeit changes in the construction and handling of R may be required when reaching $\gtrsim 10^5$ observations per time window. This potential is leveraged by many inversion systems that use Eulerian transport simulations (e.g., Varon et al., 2022; Meirink et al., 2008a; Bergamaschi et al., 2013). The increasing availability of satellite data is especially interesting for constraining concentrations and emissions in less observed regions, such as near the boundaries of our domain.

We identified potentials and risks in separating sectors based on highly resolved spatial flux patterns. Extending this by temporal profiles for a priori fluxes offers an untapped potential for improvement. Moreover, our inversion would benefit from a priori emission ensembles reflecting the uncertainty in spatial and temporal distribution of the fluxes. Significant improvements may become possible by distinguishing sectors at the observations using co-tracers such as ethane for fossil CH_4 emissions (Ramsden et al., 2022; Mead et al., 2024) or by distinguishing carbon isotopes (Basu et al., 2022; Thanwerdas et al., 2024; Chandra et al., 2024).

7 Conclusions

We presented a novel system for regional flux inversion designed to validate national CH_4 emission reporting. Applying this method to Central Europe in 2021 with a focus on Germany, we found a significant increase in emissions from Germany and the Benelux. Careful estimation of posterior uncertainties revealed that total German posterior emissions are $(32 \pm 19)\,\%$ higher than the anthropogenic emissions reported to the UNFCCC (submission 2024). This increase is most likely due to emissions from the agriculture sector, possibly with contributions from LULUCF and natural sources. Our results were confirmed by an

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

exhaustive range of sensitivity tests and by validation with independent observation sites. Synthetic experiments with known truth verified the ability to distinguish emission sectors in Germany.

Methodological comparison to other regional inversion systems highlights the advantages of our method for distinguishing emission sectors and fitness for purpose for validation of national emission estimates. The qualitative gap between UNFCCC reporting and our estimates for Germany and the Benelux is consistent with earlier works (Petrescu et al., 2023; Bergamaschi et al., 2022, 2018; Steiner et al., 2024b). We complement these studies by providing an emission estimate for the German agriculture sector that can be directly compared to the national reporting, revealing a significant mismatch.

In this study we only presented the first application of an extensible, novel inversion system. Future developments may include the integration of satellite data, the incorporation of temporal profiles, a more comprehensive treatment of boundary conditions and flux uncertainties using ensemble methods, and an extension of the state space. The close connections to operational numerical weather prediction – especially in the underlying transport simulation – and the modular design establish the potential for long-term operational support of national emissions reporting.

Data availability. A collection of model data, inversion results, and data for reproducing most figures in this work is available at https://doi.org/10.5281/zenodo.15083480 (Bruch et al., 2025).

Appendix A: Extended data tables and figures

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

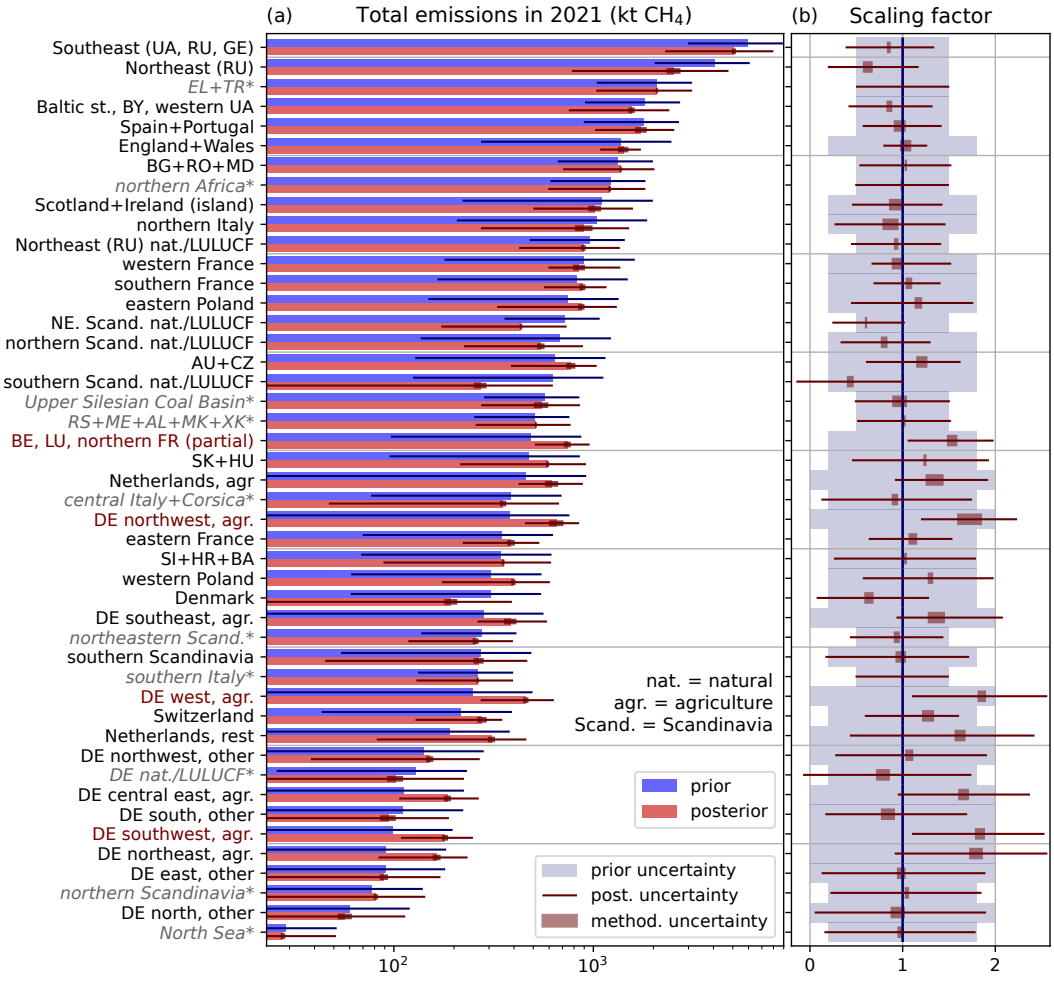


Figure A1. Prior and posterior emissions (a) and scaling factors (b) for all flux categories, ordered by prior emissions. Horizontal lines indicate 95 % confidence intervals. See Fig. 3 for the geographical definition of the flux categories and Fig. 4 for the resulting map of scaling factors. (a) If no sector is explicitly specified, the flux categories contain all anthropogenic fluxes excluding LULUCF. For flux categories marked with an asterisk, the inversion does not reduce the absolute uncertainty. Thus, reliable information is only gained by our inversion for flux categories without asterisk (see Sect. 4.5). Red color of the category names indicates a statistically significant increase of emissions. (b) Scaling factors are the raw results of our inversion, though here they are already combined for the whole year. The posterior scaling factor is defined as the mean of the methodological uncertainty range indicated by brown boxes.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

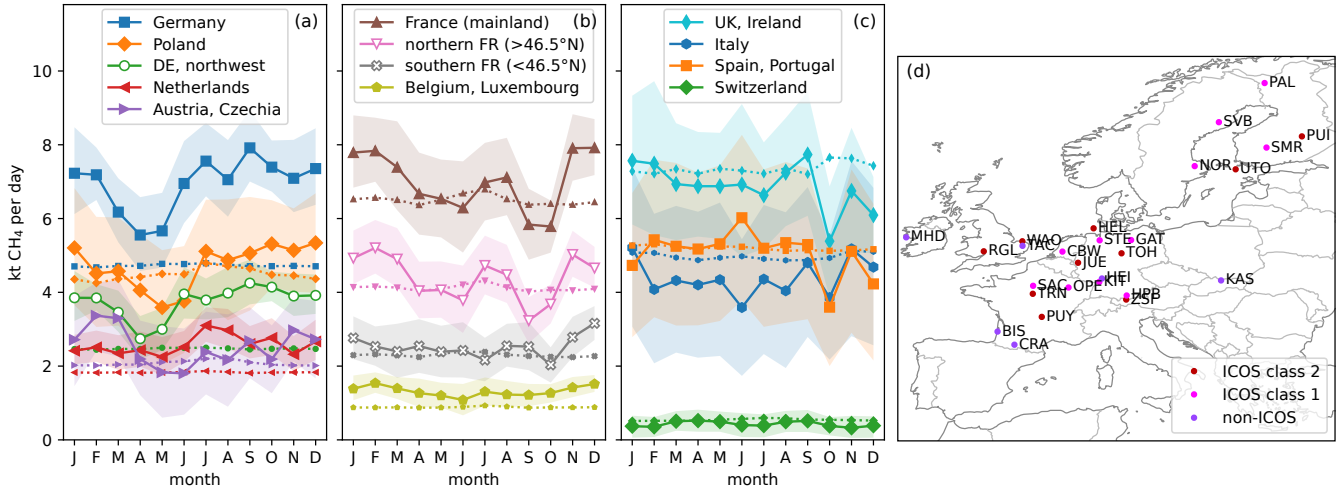


Figure A2. (a–c) Seasonal cycle when using only observations from stations that were active during the whole year. We select those stations and sampling heights, for which we used at least two data points per day on at least 20 days of each month in 2021 in our main inversion. This yields 27 stations shown in (d) with $8.3 \cdot 10^4$ data points for the inversion (step 5 in Table 2), compared to 50 stations with $1.29 \cdot 10^5$ data points in the reference case (compare Fig. 6). Colored areas show the posterior uncertainties (95 % confidence intervals), which were computed without excluding individual stations from the inversion and are therefore smaller than in Fig. 6. Prior emission rates are shown as dotted lines with small markers.

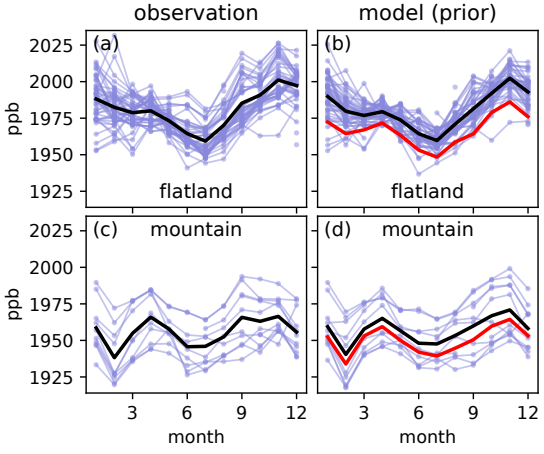


Figure A3. Seasonal cycle in observations at stations with elevation below 500 m above sea level (a, b) and above 1000 m (c, d), supplementary to the discussion in Sect. 5.2. Thin blue lines represent the 10 % quantile of each month, station, and sampling height for (a, c) observations and (b, d) model predictions (prior). The 10 % quantile is chosen to minimize the effect of local pollution. Thick black lines indicate the mean of all selected stations and sampling heights. Thick red lines in (b) and (d) show the 10 % quantile of the modeled far-field concentration. The flatland stations show a pronounced seasonal cycle with minimum in summer for both model and observations. This cycle is dominated by the contribution of the far field. The mountain stations have a weaker seasonal cycle.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Table A1. Observation stations from the European Obspack (ICOS RI et al., 2024). Column 6 ("mountain") characterizes the stations as high mountains, small mountains, and other stations. This serves as a reference for computing the station height in the model and for the daily time window. We indicate the sampling heights used in the inversion (column 7) and mark those sampling heights with an asterisk that were used in Fig. A2 (sensitivity test 602). Column 8 indicates times in which the station was excluded due to modeling problems. Column 9 ("inflation") defines the factor f_i by which the estimated uncertainty is multiplied when constructing the uncertainty matrix R.

Code	Name	Country	ICOS class	Elevation (m)	Mountain	Sampling heights (m)	Limitations	Inflation
BIK	Białystok	PL	–	183	no	90, 180, 300		2
BIR	Birkenes	NO	2	219	no	75	excl. Apr–Aug	3
BIS	Biscarrosse	FR	–	73	small	47*		2
BRM	Beromunster	CH	–	797	no	72, 132, 212		2
BSD	Bilsdale	UK	–	382	no	108, 248		2
CBW	Cabauw	NL	1	0	no	67, 127*, 207*		2
CMN	Monte Cimone	IT	2	2165	high	8		2
CRA	Centre de Recherches Atmosphériques	FR	–	600	no	60*		2
CRP	Carnsore Point	IE	–	9	no	14		2
ERS	Ersa	FR	–	533	small	40		3
FKL	Finokalia	GR	–	250	small	–	excluded	–
GAT	Gartow	DE	1	70	no	132*, 216*, 341*		2
HEI	Heidelberg	DE	–	113	no	30*		3
HEL	Helgoland	DE	2	43	no	110*		2
HPB	Hohenpeissenberg	DE	1	934	small	50, 93*, 131*		2
HTM	Hyltemossa	SE	1	115	no	70, 150		2
HUN	Hegyhátsál	HU	2	248	no	82, 115	incl. Mar–Oct	3
IPR	Ispra	IT	2	210	no	–	excluded	–
JFJ	Jungfraujoch	CH	1	3571.8	high	13.9		2
JUE	Jülich	DE	2	98	no	120*		3
KAS	Kasprowy Wierch	PL	–	1987	high	7*		2
KIT	Karlsruhe	DE	1	110	no	60*, 100*, 200*		2
KRE	Křešín u Pacova	CZ	1	534	no	50, 125, 250		2
LHW	Laegern-Hochwacht	CH	–	840	small	32		3
LIN	Lindenberg	DE	1	73	no	98		2
LMP	Lampedusa	IT	2	45	no	–	excluded	–

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Code	Name	Country	ICOS class	Elevation (m)	Mountain	Sampling heights (m)	Limitations	Inflation
LMU	La Muela	ES	–	571	no	79		2
LUT	Lutjewad	NL	2	1	no	60	excl. Nov–Dec	2
MHD	Mace Head	IE	–	5	no	24*		2
MLH	Malin Head	IE	–	22	no	47		2
NOR	Norunda	SE	1	46	no	58*, 100*		2
OHP	Observatoire de Haute Provence	FR	–	650	no	50, 100		2
OPE	Observatoire pérenne de l'environnement	FR	1	390	no	50*, 120*		2
OXK	Ochsenkopf	DE	1	1022	small	90, 163		2
PAL	Pallas	FI	1	565	no	12*		2
PDM	Pic du Midi	FR	–	2877	high	28		2
PRS	Plateau Rosa	IT	2	3480	high	10		2
PUI	Puijo	FI	2	232	small	84*		2
PUY	Puy de Dôme	FR	2	1465	small	10*		2
RGL	Ridge Hill	UK	2	207	no	90*		2
ROC	Roc'h Trédudon	FR	–	362	no	25, 80, 140		2
SAC	Saclay	FR	1	160	no	60*, 100*		2
SMR	Hyytiälä	FI	1	181	no	67.2*, 125*		2
SSL	Schauinsland	DE	2	1205	small	12, 35		2
STE	Steinkimmen	DE	1	29	no	127*, 187*, 252*		2
SVB	Svartberget	SE	1	269	no	85*, 150*		2
TAC	Tacolneston	UK	–	64	no	54*, 100*, 185*		2
TOH	Torfhaus	DE	2	801	small	76*, 110*, 147*		2
TRN	Trainou	FR	2	131	no	50*, 100*, 180*		2
UTO	Utö - Baltic sea	FI	2	8	no	57*		2
WAO	Weybourne	UK	2	17	no	10*		2
WES	Westerland	DE	2	12	no	14		2
ZSF	Zugspitze	DE	2	2666	high	3*		2

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

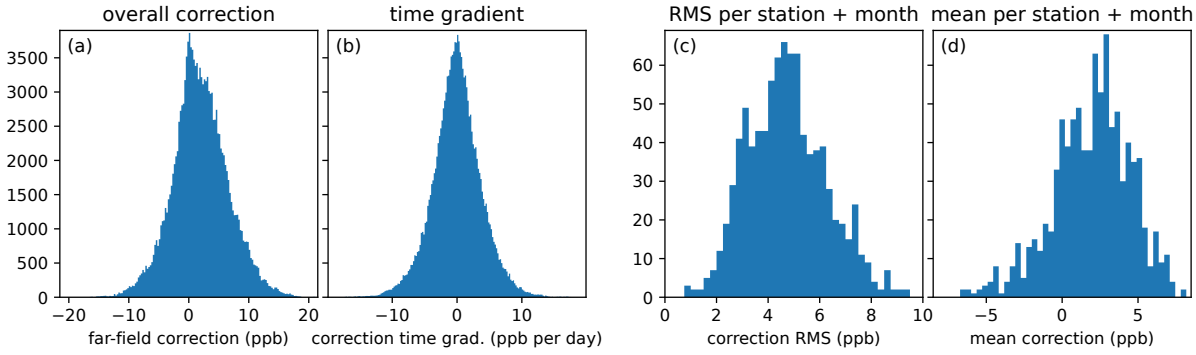


Figure B1. Statistical evaluation of the far-field correction. Considering all data points used in the inversion (steps 5–8 in Table 2), histograms of the far-field correction (a) and its time gradient (b) show that the correction is usually in the range $\pm 10\,\mathrm{ppb}$ and can vary by a few ppb per day. For each station and month, we compute the RMS and the mean (or bias). Histograms combining these values for all stations and months are shown in (c) and (d). The broad distribution of the RMS values indicates significant differences among the stations. The bias shows a tendency towards positive values, which implies that the far-field correction tends to increase the modeled CH_4 concentration. This increase in model concentrations leads to a decrease in emission estimates.

Appendix B: Far-field correction

This appendix provides details for the far-field correction introduced in Sect. 2.2. We correct the computed far field by a smooth field that may only vary on temporal scales $\gtrsim 16\,\mathrm{h}$ and horizontal scales $\gtrsim 320\,\mathrm{km}$. This correction field is determined using all data points where the cumulated signal of all flux categories is at most $20\,\mathrm{ppb}$, the total concentration due to all fluxes in the domain – including natural and uncategorized fluxes – is at most $50\,\mathrm{ppb}$, and natural plus LULUCF fluxes contribute at most $20\,\mathrm{ppb}$. These criteria aim to select only measurements of sufficiently clean air for the far-field correction. Figure B1 shows that this correction lies typically in the range $\pm 10\,\mathrm{ppb}$ and has a bias of few ppb towards higher concentrations.

The far-field correction is realized as a Kalman smoother on the selected data points. Consider the vector of all model predictions x, which is aligned with the observation vector y. By P we denote the projector selecting those data points that shall be used to determine the far-field correction. We aim to find a correction vector c aligned with x and y that minimizes

$$\arg\min_c \left\{ (x+c-y)^\top P^\top \left(P\tilde{R}P^\top\right)^{-1} P(x+c-y) + c^\top P^\top \left(P\tilde{C}P^\top\right)^{-1} Pc \right\}, \tag{B1}$$

where $\tilde{R} = 16I$ is a diagonal matrix and $\tilde{C}$ is an unnormalized (i.e. $\tilde{C}_{ii} = 1$ for all i) Gaussian localization matrix with standard deviations $16\,\mathrm{h}$ (time), $319\,\mathrm{km}$ (horizontal) and $1\,\mathrm{km}$ (vertical). The matrix $\tilde{C}$ ensures that the correction field c is smooth on these scales. For the under-determined Eq. (B1) we use the solution

$$c = \tilde{C}P^\top \left[P(\tilde{C}+\tilde{R})P^\top\right]^{-1} P(y-x). \tag{B2}$$

To prove that Eq. (B2) solves Eq. (B1), we use that Eq. (B1) is a quadratic form and compute its gradient with respect to c:

$$0 \stackrel{!}{=} 2P^\top \left(P\tilde{R}P^\top\right)^{-1} P(x+c-y) + 2P^\top \left(P\tilde{C}P^\top\right)^{-1} Pc. \tag{B3}$$

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

This can be solved by requiring

$$0 \stackrel{!}{=} \left(P\tilde{R}P^\top\right)^{-1} P(x+c-y) + \left(P\tilde{C}P^\top\right)^{-1} Pc \tag{B4}$$

$$= \left[\left(P\tilde{R}P^\top\right)^{-1} + \left(P\tilde{C}P^\top\right)^{-1}\right] Pc + \left(P\tilde{R}P^\top\right)^{-1} P(x-y) \tag{B5}$$

$$\implies Pc = \left[1 + P\tilde{R}P^\top\left(P\tilde{C}P^\top\right)^{-1}\right]^{-1} P(y-x) \tag{B6}$$

$$= P\tilde{C}P^\top\left[P(\tilde{C}+\tilde{R})P^\top\right]^{-1} P(y-x). \tag{B7}$$

Appendix C: Chi-square analysis

In this appendix, we provide the mathematical details for the χ^2/N_{dof} analysis used in Sect. 4.2. The aim of this analysis is to quantify whether the data used in the inversion agree with the assumed uncertainties. The inversion formally relies on the assumption of Gaussian probability distributions of the a priori scaling factors (error covariance matrix B) and the model–observation mismatch (R). For the model–observation mismatch $\mu = y - Hs^{\text{prior}} - x^{\text{ff}}$ this yields the probability

$$P(\mu = y - Hs^{\text{prior}} - x^{\text{ff}})$$

$$= \int_s P(y = Hs + \mu + x^{\text{ff}})\, dP_s \tag{C1}$$

$$\overset{\tau = s - s^{\text{prior}}}{\propto} \int_\tau \exp\left[-\tfrac{1}{2}\tau^\top B^{-1}\tau - \tfrac{1}{2}(H\tau+\mu)^\top R^{-1}(H\tau+\mu)\right] d^{N_{\text{dof}}}\tau \tag{C2}$$

$$\propto \exp\left\{-\tfrac{1}{2}\mu^\top\left[R^{-1} - R^{-1}H(B^{-1}+H^\top R^{-1}H)^{-1}H^\top R^{-1}\right]\mu\right\} \tag{C3}$$

$$=: \exp\left(-\tfrac{1}{2}\mu^\top Q\mu\right). \tag{C4}$$

Such high-dimensional Gaussian probability distributions share various useful properties. Here, we use that when sampling $P(\mu)$ it is very likely to find μ such that $\chi^2 \equiv \mu^\top Q\mu \approx N_{\text{dof}}$ where N_{dof} denotes the number of degrees of freedom, which is the dimension of vector μ. In our case, $N_{\text{dof}} \sim 10^4$ is the number of observation data points used per one-month time window. In the limit of large N_{dof}, one can approximate $P(\chi^2) \sim \mathcal{N}(N_{\text{dof}}, 2N_{\text{dof}})$ (Gaussian distribution with mean N_{dof} and variance $2N_{\text{dof}}$). Thus, in an idealized setup we expect that $\chi^2/N_{\text{dof}} = 1 \pm 0.03$ (95 % confidence interval). Values $\gtrsim 1.05$ imply that uncertainties were underestimated and $\chi^2/N_{\text{dof}} \lesssim 0.95$ indicates that uncertainties were too high. However, in reality we may have biases and other problems such that the assumption of a Gaussian uncertainty in the model–observation mismatch becomes invalid and $\chi^2/N_{\text{dof}} < 1$ does not necessarily imply that uncertainties should be reduced.

Appendix D: Posterior-based model uncertainty estimate

We estimate the model uncertainty using a meteorological ensemble. This leads to a dependence of the model uncertainty on the fluxes. Stronger emissions lead to stronger gradients in the model concentrations and to higher uncertainties in the

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

concentrations due to transport errors. In this appendix, we explain how the uncertainty can be estimated based on the posterior fluxes. As stated in Sect. 2.5, this can reduce the bias for localized emissions.

D1 Optimization problem for self-consistent model uncertainties

In Sect. 2.4, we defined the uncertainty matrix R using the concentrations x_i^m at observation data points i predicted for different ensemble members m. When using the posterior-R inversion method, these concentrations should be consistent with the posterior emissions. In this case, x_i^m and thereby R depend on the scaling factors s of the emissions. The inversion minimizes the cost function Eq. (1) to optimize the agreement with the observations and the prior scaling factors (see Sect. 2.3). For a flux-dependent R matrix, this optimization problem reads

$$s^{\text{post}} = \arg\min_s L(s), \tag{D1}$$

$$L(s) = \tfrac{1}{2}(y - Hs - x^{\text{ff}})^\top R(s)^{-1}(y - Hs - x^{\text{ff}}) + \tfrac{1}{2}(s - s^{\text{prior}})B^{-1}(s - s^{\text{prior}}). \tag{D2}$$

By minimizing $L(s)$ and approximating it in the form $L(s) = \frac{1}{2}(s - s^{\text{post}})B_{\text{post}}^{-1}(s - s^{\text{post}}) + \text{constant}$, we obtain the posterior scaling factors s^{post} and the posterior error covariance matrix B_{post}.

The minimization problem $\arg\min_s L(s)$ can be solved numerically. A quick convergence of a minimization algorithm can be achieved by making use of the analytically computed gradient and Hesse matrix of $L(s)$. We used SciPy's "trust-exact" implementation of a trust-region method (Virtanen et al., 2020; Moré and Sorensen, 1983; Conn et al., 2000). Within each iteration, the incomplete LU decomposition (Li et al., 1999; Li and Shao, 2011) of the sparse matrix $R(s)$ is the most computationally expensive task when the number of observations is large.

D2 Reduced ensemble

When using a priori scaling factors to estimate the model uncertainty, we need only the total concentration $x_i^m(s^{\text{prior}})$ for each ensemble member. Thus, only a single tracer field is required in the ensemble transport simulation. To fully compute $x_i^m(s)$ as function of s, the tracer categories need to be distinguished for each ensemble member, resulting in > 40 tracer fields in the ensemble simulation. To avoid wasting numerical resources, we chose to approximate $x_i^m(s)$ by only a few tracer fields, using additional information from the deterministic model run which distinguishes all tracer fields.

From the deterministic model run we know the operator H mapping scaling factors s to a model prediction $Hs + x^{\text{ff}}$ for the concentrations. For ensemble member m we would ideally know H^m and $x^{\text{ff},m}$ yielding a model prediction $H^m s + x^{\text{ff},m}$. To avoid calculating the full matrix H^m, we group the flux categories into groups $\{g\}$ and denote by P_g the projector of scaling vectors s on the subspace spanned by the flux categories in group g. Using the total concentration from group g, $x_i^{mg} = x_i^m(P_g s^{\text{prior}})$, we estimate the full dependence on the scaling factors:

$$x_i^m(s) \approx \sum_g \frac{(HP_g s)_i}{(HP_g s^{\text{prior}})_i} x_i^{mg} + x_i^{\text{ff},m}. \tag{D3}$$

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Thus, we compute the transport ensemble for a few tracer groups and estimate $x^m(s)$ for arbitrary s by using the ratios of tracer fields within the tracer groups from the deterministic run. Using the approximation in Eq. (D3), we estimate the posterior model uncertainties with only five tracer fields in an ensemble of 12 transport simulations:

1. far field (initial and lateral boundary conditions)
2. total anthropogenic fluxes
3. total natural fluxes
4. total anthropogenic fluxes from Germany with lifetime five days
5. total anthropogenic fluxes from outside Germany with lifetime five days

Appendix E: Sensitivity tests

Table E1 provides an overview of the sensitivity tests. For this table we quantify the impact of a parameter variation on the inversion results by the following, heuristic metric: Consider a fixed region, sector and inversion time window with posterior fluxes F. The normalized deviation from the reference inversion is defined as $\Delta = \frac{2|F-F^{\text{ref.}}|}{F^{\text{ref. upper}}-F^{\text{ref. lower}}}$, where $F^{\text{ref. upper}}$ and $F^{\text{ref. lower}}$ denote the bounds of the posterior uncertainty range. The overall impact is computed as the arithmetic mean of Δ over the (usually monthly) time windows and a selection of regions and sectors. In the regions UK+Ireland, France, Italy, Poland, Austria+Czechia, Netherlands, Belgium+Luxembourg, Switzerland, and Denmark we consider only total fluxes without distinguishing sectors. In Germany we include Δ for the total fluxes in four different regions (north, east, south, west) and additionally for national total fluxes distinguishing the three sectors agriculture, natural plus LULUCF, and other sectors. Effectively, this counts all fluxes in Germany twice and gives them more weight in the impact metric for Table E1.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Table E1. Sensitivity tests for estimating the robustness of the inversion results with respect to tuning parameters. The impact column quantifies the deviation of the inversion results relative to the uncertainties and shall qualitatively indicate the relevance of the modified parameters (see explanation in the text). An impact of $100\,\%$ means that the average deviation from the reference case is as large as the posterior uncertainty. Overall, we see that most test have an impact of $\lesssim 15\,\%$, implying that the effect on the inversion results is small compared to the uncertainty in the reference case. See also Fig. E1 for the posterior emissions in the sensitivity tests.

ID	Test case	Explanation	Impact
0	reference	as explained in Sect. 4	
	Model equivalent calculation		
100	station elevation for mountain stations	treat all mountain stations like small mountains when computing model heights, as proposed by Brunner et al. (2012); Henne et al. (2016); Bergamaschi et al. (2022)	5.3 %
101	no time averaging	average only over 1 h instead of over 3 h	13 %
	filtering observations		
200	fewer hours of day	use time window 12 h–16 h (0 h–4 h for high mountains)	11 %
201	all hours of day	no filtering by time of day, increase uncertainty inflation by factor 1.5	38 %
202	one sampling height per station	use only highest sampling height of each station	16 %
203	no filtering based on wind	include data points with low wind speed	12 %
204	low min. wind speed	minimum wind speed: $1.11\,\mathrm{m\,s^{-1}}$	9.4 %
205	high min. wind speed	minimum wind speed: $3.0\,\mathrm{m\,s^{-1}}$	11 %
206	low max. model-obs. mismatch	discard when absolute deviation exceeds $120\,\mathrm{ppb}$, or model far field minus observation exceeds $12\,\mathrm{ppb}$	3.5 %
207	high max. model-obs. mismatch	discard when absolute deviation exceeds $300\,\mathrm{ppb}$, or model far field minus observation exceeds $30\,\mathrm{ppb}$	1.3 %
208	low max. data point influence	increase uncertainty if $\lvert\mu_i\rvert > \mathbf{2.5}\sqrt{R_{ii}}$ in Sect. 2.4.1	11 %
209	high max. data point influence	increase uncertainty if $\lvert\mu_i\rvert > \mathbf{4}\sqrt{R_{ii}}$ in Sect. 2.4.1	15 %
	uncertainty / error covariance matrix R (model and observations)		
300	no plume uncertainty	no extra uncertainty due to point-like emissions	0.27 %
301	high plume uncertainty	uncertainty due to point-like emissions: use $R'_{ij} + \mathbf{0.5}\rho_i^2\delta_{ij}$ in Sect. 4.2	0.56 %
302	low uncertainty inflation	uncertainty inflation by (1.5, 2.25) instead of (2, 3)	8.6 %
303	high uncertainty inflation	uncertainty inflation by (3, 4.5) instead of (2, 3)	13 %
304	small horizontal error correlation scale	scale $191\,\mathrm{km}$ instead of $319\,\mathrm{km}$	6.0 %
305	large horizontal error correlation scale	scale $510\,\mathrm{km}$ instead of $319\,\mathrm{km}$	8.3 %
306	small vertical error correlation scale	scale $400\,\mathrm{m}$ instead of $1\,\mathrm{km}$	2.3 %
307	short error correlation time scale	scale 4 h instead of 6 h	2.5 %
308	long error correlation time scale	scale 10 h instead of 6 h	2.8 %

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

ID	Test case	Explanation	Impact
309	low uncorrelated uncertainty	$\sigma_{\text{const}} = 5\,\text{ppb}$ instead of $10\,\text{ppb}$ in Eq. (2)	21 %
310	high uncorrelated uncertainty	$\sigma_{\text{const}} = 20\,\text{ppb}$ instead of $10\,\text{ppb}$ in Eq. (2)	22 %
	far-field correction		
400	no far-field correction		97 %
401	weak far-field correction	$\tilde{R} = 100I$ instead of $16I$ in Eq. (B1)	16 %
402	strong far-field correction	$\tilde{R} = 2.78I$ instead of $16I$ in Eq. (B1)	9.2 %
403	small horiz. far-field correction scale	scale $191\,\text{km}$ instead of $319\,\text{km}$	6.8 %
404	large horiz. far-field correction scale	scale $510\,\text{km}$ instead of $319\,\text{km}$	4.5 %
405	short far-field correction time scale	time scale 10 h instead of 16 h	3.7 %
406	long far-field correction time scale	time scale 28 h instead of 16 h	3.8 %
407	long far-field correction time scale	time scale 48 h instead of 16 h	7.1 %
408	low vertical far-field correction scale	scale $400\,\text{m}$	0.92 %
409	strict far-field observation selection	max. signal $10\,\text{ppb}$, max. due to natural fluxes $10\,\text{ppb}$	20 %
410	loose far-field observation selection	max. signal $30\,\text{ppb}$, max. due to natural fluxes $30\,\text{ppb}$, max. uncategorized $80\,\text{ppb}$	14 %
411	unrestricted iterative far-field correction	max. signal $50\,\text{ppb}$, no other selection criteria; $\tilde{C}$ localization scales 10 h, $191\,\text{km}$; iterate far-field correction and inversion 3 times	30 %
412	low correction uncertainty	use $R_{ij} + \mathbf{0.25}\lvert c_i c_j\rvert \tilde{C}_{ij}$ in Sect. 4.2	2.5 %
413	high correction uncertainty	use $R_{ij} + \mathbf{1.0}\lvert c_i c_j\rvert \tilde{C}_{ij}$ in Sect. 4.2	4.2 %
414	uncorrelated correction uncertainty	use $R_{ij} + 2c_i^2\delta_{ij}$ instead of $R_{ij} + 0.5\lvert c_i c_j\rvert \tilde{C}_{ij}$ in Sect. 4.2	3.6 %
	a priori scaling factor error covariance matrix B		
500	low prior uncertainty	1σ prior uncertainty for most areas 0.25, remote and plume 0.2, sector-resolved 0.33	14 %
501	high prior uncertainty in Germany	1σ prior uncertainty such that total sector emissions in Germany have uncertainty 0.6 for each distinguished sector	8.6 %
502	uncorrelated prior, B is diagonal	1σ prior uncert. in sector categories in Germany: 0.75	6.3 %
503	no sector distinction in prior	four regions in Germany with uncorrelated 1σ prior uncertainty of 0.4	7.7 %
504	low spatial resolution in Germany	two initially uncorrelated regions in Germany	15 %
506	distinguish 5 sectors in Germany	see Appendix I	2.1 %
	station selection		
601	stations covering ≥ 10 days each month	use 35 of 50 stations	13 %
602	stations covering ≥ 20 days each month	use 27 of 50 stations as detailed in Fig. A2	33 %
	inversion time windows		
701	2 month inversion window		12 %
702	3 month inversion window		18 %

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

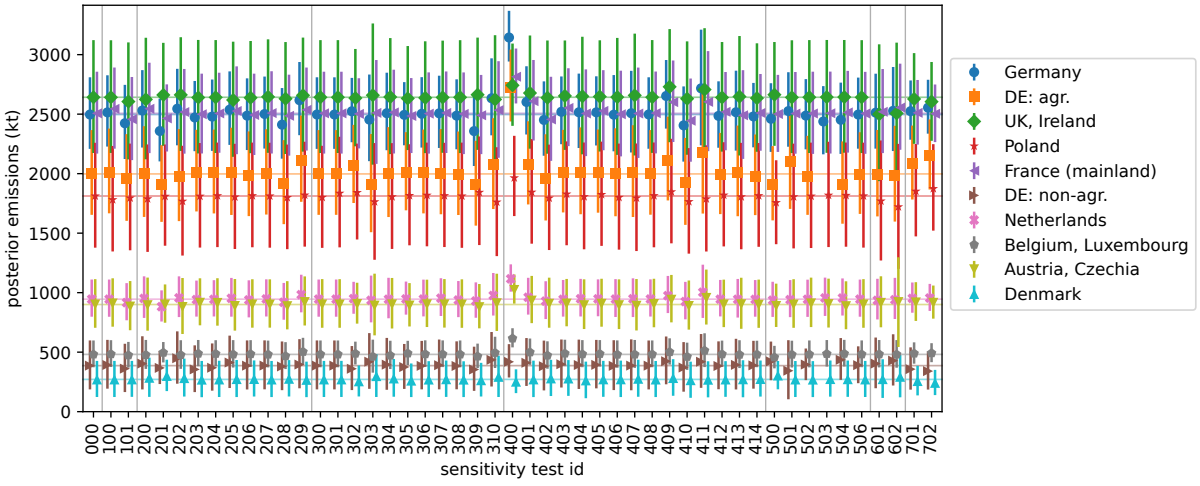


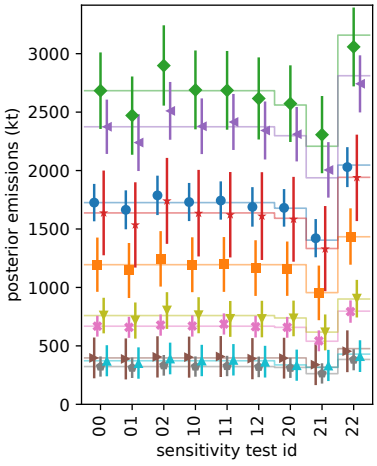
Figure E1. Posterior emissions and uncertainties of selected countries and German sectors for all sensitivity tests. Thin horizontal lines indicate the posterior of the reference case 0. Vertical lines show uncertainties (95 % confidence intervals) that are not extended by excluding stations. The individual tests are listed in Table E1. One can clearly see the strongest deviations for cases 400 (no far-field correction), whereas all other test cases only lead to changes within the uncertainty ranges.

Appendix F: Additional synthetic experiments

In this appendix, we present synthetic experiments evaluating the response of the inversion system to a bias affecting all observations, noise, and all fluxes scaled by a constant factor in the synthetic truth. The posterior emissions for some well-defined scenarios are shown in Fig. F1. In the first scenarios, we shift all pseudo-observations by $-5\,\mathrm{ppb}$ (case 1 in Fig. F1) and $+5\,\mathrm{ppb}$ (case 2). This shift is mostly compensated by the far-field correction with monthly averages of $\pm 2.75\,\mathrm{ppb}$ to $\pm 3.8\,\mathrm{ppb}$, the sign depending on the scenario. Due to this correction, the effect on the German emissions remains within the posterior uncertainty. We also test the effect of correlated and uncorrelated Gaussian noise of standard deviation $5\,\mathrm{ppb}$ added to the observations (cases 10–12), finding that the effect on the posterior emissions is small compared to the posterior uncertainties. The correlated Gaussian noise is a 3d Gaussian random field in flat (longitude, latitude, time) coordinates with a lower cutoff for fluctuations on scales $\lesssim 2.5°$ (horizontal) and $\lesssim 12$ days (time) such that it acts as a slowly varying random bias. The RMS of the noise is normalized to $5\,\mathrm{ppb}$.

Next, we test the effect of an underestimation or overestimation of all emissions. In case 20 of Fig. F1, all natural and LULUCF fluxes are reduced by $40\,\%$ in the truth, and cases 21 and 22 change all anthropogenic emissions excluding LULUCF by $-20\,\%$ and $+20\,\%$, respectively. In all cases, the posterior emissions follow the truth and remain compatible with the synthetic truth within the uncertainties. In Germany, we find that the agriculture follows the scaled emissions stronger than the other sectors (see Sect. 5.5.2 and Appendix G). The same effect is observed when repeating the case of increased anthropogenic emissions with simulated transport uncertainty. Figure F2 shows how results of the two inversion methods deviate from the synthetic truth of $20\,\%$ increased anthropogenic emissions. Agriculture emissions from Germany are almost correct in the

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025



ID	Explanation
00	reference case
01	global bias −5 ppb
02	global bias +5 ppb
10	uncorrelated noise (5 ppb)
11	spatially and temporally correlated noise (5 ppb)
12	correlated + uncorrelated noise
20	natural + LULUCF fluxes reduced by 40 %
21	anthrop. fluxes excl. LULUCF reduced by 20 %
22	anthrop. fluxes excl. LULUCF increased by 20 %

Figure F1. Posterior emissions of selected countries and German sectors for synthetic experiments. Thin horizontal lines indicate the truth. Vertical lines show uncertainties (95 % confidence intervals) that are not extended by excluding stations.

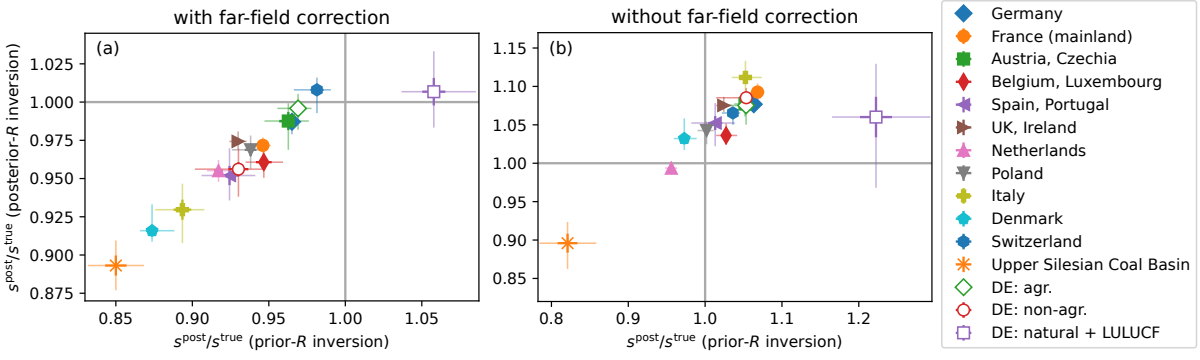


Figure F2. Bias in synthetic experiments with simulated transport uncertainty when all anthropogenic emissions are increased by 20 %. Note the different scales of both axes. The markers show the posterior emissions relative to the truth for selected regions and sectors, averaged over 12 inversion runs with different simulated transport uncertainty. As in Fig. 10, the posterior-R inversion yields higher emissions than the prior-R inversion.

posterior whereas all other shown national emission estimates – with the exception of Switzerland – are underestimated in the posterior. Here, underestimated emissions are expected because the prior is underestimated compared to the synthetic truth.

Appendix G: Relevance of absolute prior uncertainty in sector attribution

When observations can detect a change in total emissions but cannot distinguish between different emission sectors, the sector-resolving inversion will change the sectoral distribution based on the prior uncertainties. To understand this problem qualita-

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

tively, we consider the worst case: We assume that all sectors are uncorrelated in the prior but 100 % spatially correlated such that they cannot be distinguished in the inversion. The a priori probability for an emission vector e of sector emissions e_i is

$$P(e) \propto \exp\left[-\tfrac{1}{2}\sum_i (e_i - e_i^{\text{prior}})^2 \sigma_i^{-2}\right], \tag{G1}$$

where σ_i denotes the a priori standard deviation of e_i. The inversion will yield a result for the total emissions $e_{\text{tot}}^{\text{post}}$ that maximizes the probability $P(e)$ when including information from the observations. But by assumption, these observations do not distinguish between sectors such that Eq. (G1) remains valid. We thus obtain the posterior emissions of the sectors by maximizing Eq. (G1) with the constraint $\sum_i e_i = e_{\text{tot}}^{\text{post}}$. By introducing a Lagrange multiplier, one can show[3] that this yields

$$e_i - e_i^{\text{prior}} = \alpha\sigma_i^2, \qquad \alpha = \frac{e_{\text{tot}}^{\text{post}} - e_{\text{tot}}^{\text{prior}}}{\sum_i \sigma_i^2}. \tag{G2}$$

This shows that sectors with larger absolute a priori uncertainty are disproportionally stronger corrected. Applied to our emission estimates for Germany, this implies that if the observations were unsuitable for distinguishing sectors, the inversion would attribute up to 95 % of the changes in total fluxes to the agriculture sector, which is responsible for 69 % of the total a priori emissions. Fortunately, this worst case scenario is not realistic because the observations do contain information on the different sectors as indicated e.g. by Figs. 8 and 9.

Appendix H: Averaging kernel matrices

As introduced in Sect. 5.5.2, the averaging kernel matrices A^{emis} and $A^{\text{scaling factors}}$ estimate the change in the posterior when changing the truth, $A^{\text{emis}} = \partial e^{\text{post}}/\partial e^{\text{truth}}$ where e denotes the vector of emissions. Here, we summarize how these matrices are estimated using either the prior and posterior error covariance matrices B and B_{post}, or the statistics from inversion runs with synthetic truth.

H1 Analytic estimate using error covariance matrices

We first estimate the sensitivity of the posterior scaling factor to the true emissions under the assumption that the transport model, far field, and the flux pattern within each flux category are perfect. Under these idealized assumptions, the model–observation mismatch is $\mu = y - Hs^{\text{prior}} - x^{\text{ff}} = H(s^{\text{truth}} - s^{\text{prior}})$ where s^{truth} denotes the true scaling factors. Our prior-R inversion will now maximize

$$P(s) \propto \exp\left[-\tfrac{1}{2}(s - s^{\text{truth}})^\top H^\top R^{-1} H (s - s^{\text{truth}}) - \tfrac{1}{2}(s - s^{\text{prior}})^\top B^{-1}(s - s^{\text{prior}})\right] \tag{H1}$$

$$\propto \exp\left[-\tfrac{1}{2}(s - s^{\text{post}})^\top B_{\text{post}}^{-1}(s - s^{\text{post}})\right]. \tag{H2}$$

This yields $s^{\text{post}} = s^{\text{prior}} + A(s^{\text{truth}} - s^{\text{prior}})$ with the averaging kernel $A = I - B_{\text{post}}B^{-1}$ and $B_{\text{post}}^{-1} = H^\top R^{-1} H + B^{-1}$ (Rodgers, 2000). Knowing B and B_{post}, we can compute the averaging kernel to estimate how the posterior scaling factors depend on the true scaling factors.

[3] We define $L(e,\lambda) = -\frac{1}{2}\sum_i (e_i - e_i^{\text{prior}})^2\sigma_i^{-2} + \lambda(e_{\text{tot}}^{\text{post}} - \sum_i e_i)$ and require $\frac{\partial L}{\partial e_i} = 0$, $\frac{\partial L}{\partial \lambda} = 0$.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

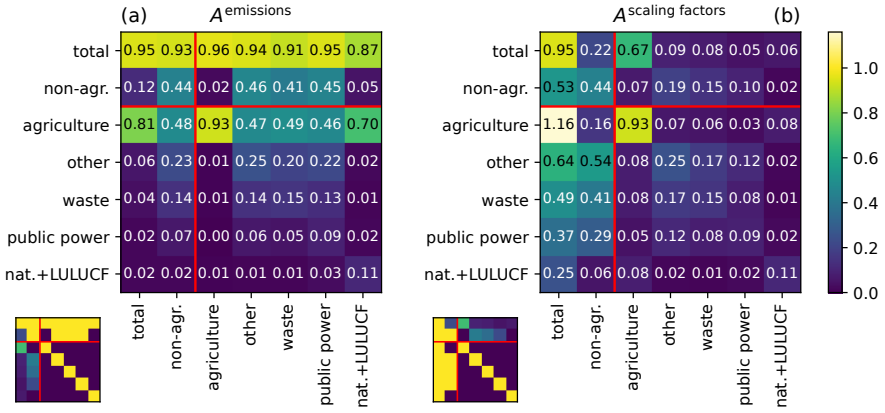


Figure I1. Averaging kernel matrices of German sector emissions when trying to distinguish sectors waste, public power and other, in the same representation as in Fig. 9(a)–(b). Panel (a), third row, shows that increasing true emissions in any sector is expected to cause higher posterior agriculture emissions.

H2 Statistical estimate using synthetic experiments

We aim to estimate the sensitivity of posterior scaling factors $\xi := s^{\text{post}} - s^{\text{prior}}$ to changes in the synthetic truth $\zeta := s^{\text{truth}} - s^{\text{prior}}$. Given a sample of N realizations $\{\xi^n\}_n$ and $\{\zeta^n\}_n$, we aim to find the scaling factor averaging kernel matrix A that solves

$$A = \arg\min_{A'} \sum_{n=1}^{N} \left\| \xi^n - A'\zeta^n \right\|^2 . \tag{H3}$$

For $\|x\|^2 = \sum_i x_i^2$, differentiation by A'_{ij} yields $0 = \sum_{n=1}^{N} \zeta_j^n \left(\xi^n - A\zeta^n\right)_i$ for all i, j and thereby

$$A = \Xi Z^{-1}, \quad \Xi_{ij} = \sum_{n=1}^{N} \xi_i^n \zeta_j^n, \quad Z_{ij} = \sum_{n=1}^{N} \zeta_i^n \zeta_j^n . \tag{H4}$$

Equation (H4) was used to produce panels (c) and (d) of Fig. 9.

Appendix I: Attempt to distinguish five sectors in Germany

Our setup for the transport simulation was designed to separte five sectors in Germany: agriculture, natural plus LULUCF, waste, public power, and the sum of all other sectors ("other"). We test the separation of these sectors in sensitivity test 506 (Table E1 and Fig. E1) and find no notable changes in the posterior emissions compared to our reference setup, in which we combined waste, public power, and other into one larger sector "non-agr.". However, the uncertainties and the averaging kernels change considerably. We assume an a priori 2σ uncertainty of $\pm 100\,\%$ for each sector-resolving flux category. Thus, splitting the total fluxes in more uncorrelated flux categories reduces the a priori uncertainty of the total fluxes.

Figure I1 shows the averaging kernel matrices (introduced in Sect. 5.5.2 and Appendix H) for the inversion separating five sectors. These matrices indicate that waste, public power, and "other" cannot be distinguished: The corresponding columns

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Fig. I1(a) are approximately equal. Thus, trying to distinguish these sectors does not yield any additional information. By comparing the row and column for "non-agr." to Fig. 9, we identify drawbacks of the attempt to distinguish smaller sectors. When trying to distinguish five sectors, the false attribution of emissions to the agriculture sectors is more severe than when distinguishing only three sectors ($48\,\%$ compared to $28\,\%$). Consequently, the expected error reduction in the combined non-agriculture sectors (excluding natural plus LULUCF) is better when considering only three sectors. Qualitatively, this is what we expect from Appendix G for cases where the observations are insufficient to distinguish the considered sectors.

Author contributions. VB and TR conceptualized the inversion method. VB implemented the inversion method and wrote the original draft together with AKW. TR configured the transport model. TR and BE interpolated the a priori flux data which BE collected. DJCO organized data streams of CH_4 concentrations and observations. JF, BM, AMB, DJCO, TR and VB contributed to testing and tuning the transport model. NB contributed to the model–observation comparison. AKW supervised and coordinated the project. All authors reviewed and edited the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. In our simulations we use modified Copernicus Atmosphere Monitoring Service information and ECCAD products for initial and lateral boundary conditions, and for a priori fluxes. We thank Stefan Feigenspan, Christian Mielke, Theo Wernicke, John Akubia and Roland Fuß for helpful discussions and providing a priori emission fields. We thank Roland Potthast, Frank-Thomas Koch, Christoph Gerbig, Dominik Brunner, Michael Steiner, David Ho, Thomas Kaminski, Hannes Imhof and our partners in the ITMS project for very helpful and inspiring discussions. We also wish to thank Peter Bergamaschi, Aurélie Colomb, Martine De Mazière, Lukas Emmenegger, Dagmar Kubistin, Irene Lehner, Kari Lehtinen, Markus Leuenberger, Cathrine Lund Myhre, Michal V. Marek, Simon O'Doherty, Stephen M. Platt, Christian Plaß-Dülmer, Francesco Apadula, Sabrina Arnold, Pierre-Eric Blanc, Dominik Brunner, Huilin Chen, Lukasz Chmura, Łukasz Chmura, Sébastien Conil, Cédric Couret, Paolo Cristofanelli, Grant Forster, Arnoud Frumau, Christoph Gerbig, François Gheusi, Samuel Hammer, Laszlo Haszpra, Juha Hatakka, Michal Heliasz, Stephan Henne, Arjan Hensen, Antje Hoheisel, Tobias Kneuer, Eric Larmanou, Tuomas Laurila, Ari Leskinen, Ingeborg Levin, Matthias Lindauer, Morgan Lopez, Ivan Mammarella, Giovanni Manca, Andrew Manning, Damien Martin, Frank Meinhardt, Meelis Mölder, Jennifer Müller-Williams, Steffen Manfred Noe, Jarosław Nęcki, Mikaell Ottosson-Löfvenius, Carole Philippon, Joseph Pitt, Michel Ramonet, Pedro Rivas-Soriano, Bert Scheeren, Marcus Schumacher, Mahesh Kumar Sha, Gerard Spain, Martin Steinbacher, Lise Lotte Sørensen, Alex Vermeulen, Gabriela Vítková, Irène Xueref-Remy, Alcide di Sarra, Franz Conen, Victor Kazan, Yves-Alain Roulet, Tobias Biermann, Marc Delmotte, Daniela Heltai, Ove Hermansen, Kateřina Komínková, Olivier Laurent, Janne Levula, Chris Lunder, Per Marklund, Josep-Anton Morguí, Jean-Marc Pichon, Martina Schmidt, Damiano Sferlazzo, Paul Smith, Kieran Stanley, Pamela Trisolino and Giulia Zazzeri for providing the atmospheric observations for the stations listed in Table A1. VB, DJCO, NB and AMB acknowledge funding by the German Federal Ministry for Education and Research (BMBF) in the ITMS project (grant 01LK2102B) as well as BE (grant 01LK2104A). Map plots were made with Natural Earth.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

References

Agustí-Panareda, A., Diamantakis, M., Massart, S., Chevallier, F., Muñoz-Sabater, J., Barré, J., Curcoll, R., Engelen, R., Langerock, B., Law, R. M., Loh, Z., Morguí, J. A., Parrington, M., Peuch, V.-H., Ramonet, M., Roehl, C., Vermeulen, A. T., Warneke, T., and Wunch, D.: Modelling CO_2 weather – why horizontal resolution matters, Atmos. Chem. Phys., 19, 7347–7376, https://doi.org/10.5194/acp-19-7347-2019, 2019.

Basu, S., Lan, X., Dlugokencky, E., Michel, S., Schwietzke, S., Miller, J. B., Bruhwiler, L., Oh, Y., Tans, P. P., Apadula, F., Gatti, L. V., Jordan, A., Necki, J., Sasakawa, M., Morimoto, S., Di Iorio, T., Lee, H., Arduini, J., and Manca, G.: Estimating emissions of methane consistent with atmospheric measurements of methane and δ^{13}C of methane, Atmos. Chem. Phys., 22, 15 351–15 377, https://doi.org/10.5194/acp-22-15351-2022, 2022.

Bergamaschi, P., Krol, M., Meirink, J. F., Dentener, F., Segers, A., van Aardenne, J., Monni, S., Vermeulen, A. T., Schmidt, M., Ramonet, M., Yver, C., Meinhardt, F., Nisbet, E. G., Fisher, R. E., O'Doherty, S., and Dlugokencky, E. J.: Inverse modeling of European CH_4 emissions 2001–2006, J. Geophys. Res. Atmos., 115, https://doi.org/10.1029/2010JD014180, 2010.

Bergamaschi, P., Houweling, S., Segers, A., Krol, M., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Wofsy, S. C., Kort, E. A., Sweeney, C., Schuck, T., Brenninkmeijer, C., Chen, H., Beck, V., and Gerbig, C.: Atmospheric CH_4 in the first decade of the 21st century: Inverse modeling analysis using SCIAMACHY satellite retrievals and NOAA surface measurements, J. Geophys. Res. Atmos., 118, 7350–7369, https://doi.org/10.1002/jgrd.50480, 2013.

Bergamaschi, P., Corazza, M., Karstens, U., Athanassiadou, M., Thompson, R. L., Pison, I., Manning, A. J., Bousquet, P., Segers, A., Vermeulen, A. T., Janssens-Maenhout, G., Schmidt, M., Ramonet, M., Meinhardt, F., Aalto, T., Haszpra, L., Moncrieff, J., Popa, M. E., Lowry, D., Steinbacher, M., Jordan, A., O'Doherty, S., Piacentino, S., and Dlugokencky, E.: Top-down estimates of European CH_4 and N_2O emissions based on four different inverse models, Atmos. Chem. Phys., 15, 715–736, https://doi.org/10.5194/acp-15-715-2015, 2015.

Bergamaschi, P., Karstens, U., Manning, A. J., Saunois, M., Tsuruta, A., Berchet, A., Vermeulen, A. T., Arnold, T., Janssens-Maenhout, G., Hammer, S., Levin, I., Schmidt, M., Ramonet, M., Lopez, M., Lavric, J., Aalto, T., Chen, H., Feist, D. G., Gerbig, C., Haszpra, L., Hermansen, O., Manca, G., Moncrieff, J., Meinhardt, F., Necki, J., Galkowski, M., O'Doherty, S., Paramonova, N., Scheeren, H. A., Steinbacher, M., and Dlugokencky, E.: Inverse modelling of European CH_4 emissions during 2006–2012 using different inverse models and reassessed atmospheric observations, Atmos. Chem. Phys., 18, 901–920, https://doi.org/10.5194/acp-18-901-2018, 2018.

Bergamaschi, P., Segers, A., Brunner, D., Haussaire, J.-M., Henne, S., Ramonet, M., Arnold, T., Biermann, T., Chen, H., Conil, S., Delmotte, M., Forster, G., Frumau, A., Kubistin, D., Lan, X., Leuenberger, M., Lindauer, M., Lopez, M., Manca, G., Müller-Williams, J., O'Doherty, S., Scheeren, B., Steinbacher, M., Trisolino, P., Vítková, G., and Yver Kwok, C.: High-resolution inverse modelling of European CH_4 emissions using the novel FLEXPART-COSMO TM5 4DVAR inverse modelling system, Atmos. Chem. Phys., 22, 13 243–13 268, https://doi.org/10.5194/acp-22-13243-2022, 2022.

Bessagnet, B., Pirovano, G., Mircea, M., Cuvelier, C., Aulinger, A., Calori, G., Ciarelli, G., Manders, A., Stern, R., Tsyro, S., García Vivanco, M., Thunis, P., Pay, M.-T., Colette, A., Couvidat, F., Meleux, F., Rouïl, L., Ung, A., Aksoyoglu, S., Baldasano, J. M., Bieser, J., Briganti, G., Cappelletti, A., D'Isidoro, M., Finardi, S., Kranenburg, R., Silibello, C., Carnevale, C., Aas, W., Dupont, J.-C., Fagerli, H., Gonzalez, L., Menut, L., Prévôt, A. S. H., Roberts, P., and White, L.: Presentation of the EURODELTA III intercomparison exercise – evaluation of the chemistry transport models' performance on criteria pollutants and joint analysis with meteorology, Atmos. Chem. Phys., 16, 12 667–12 701, https://doi.org/10.5194/acp-16-12667-2016, 2016.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Bruch, V., Rösch, T., Jiménez de la Cuesta Otero, D., Ellerhoff, B., Mamtimin, B., Becker, N., Blechschmidt, A.-M., Förstner, J., and Kaiser-Weiss, A. K.: German methane fluxes in 2021 estimated with an ensemble-enhanced scaling inversion based on the ICON–ART model, Zenodo [data set], https://doi.org/10.5281/zenodo.15083480, 2025.

Brunner, D., Henne, S., Keller, C. A., Vollmer, M. K., Reimann, S., and Buchmann, B.: Estimating European Halocarbon Emissions Using Lagrangian Backward Transport Modeling and in Situ Measurements at the Jungfraujoch High-Alpine Site, in: Lagrangian Modeling of the Atmosphere, pp. 207–222, John Wiley & Sons, Ltd., Chichester, England, UK, ISBN 978-1-11870457-8, https://doi.org/10.1029/2012GM001258, 2012.

Canepa, E. and Builtjes, P. J. H.: Thoughts on Earth System Modeling: From global to regional scale, Earth-Sci. Rev., 171, 456–462, https://doi.org/10.1016/j.earscirev.2017.06.017, 2017.

Chandra, N., Patra, P. K., Fujita, R., Höglund-Isaksson, L., Umezawa, T., Goto, D., Morimoto, S., Vaughn, B. H., and Röckmann, T.: Methane emissions decreased in fossil fuel exploitation and sustainably increased in microbial source sectors during 1990–2020, Commun. Earth Environ., 5, 1–15, https://doi.org/10.1038/s43247-024-01286-x, 2024.

Chen, H. W., Zhang, F., Lauvaux, T., Davis, K. J., Feng, S., Butler, M. P., and Alley, R. B.: Characterization of Regional-Scale CO_2 Transport Uncertainties in an Ensemble with Flow-Dependent Transport Errors, Geophys. Res. Lett., 46, 4049–4058, https://doi.org/10.1029/2018GL081341, 2019.

Citepa: Format Secten, https://www.citepa.org/fr/secten/, (last accessed: 18 March 2025), 2024.

Conn, A. R., Gould, N. I. M., and Toint, Ph. L.: Trust Region Methods, MOS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), Philadelphia, PA, USA, ISBN 978-0-89871985-7, 2000.

Cusworth, D. H., Bloom, A. A., Ma, S., Miller, C. E., Bowman, K., Yin, Y., Maasakkers, J. D., Zhang, Y., Scarpelli, T. R., Qu, Z., Jacob, D. J., and Worden, J. R.: A Bayesian framework for deriving sector-based methane emissions from top-down fluxes, Commun. Earth Environ., 2, 1–8, https://doi.org/10.1038/s43247-021-00312-6, 2021.

Dammers, E., Tokaya, J., Mielke, C., Hausmann, K., Griffin, D., McLinden, C., Eskes, H., and Timmermans, R.: Can TROPOMI NO_2 satellite data be used to track the drop in and resurgence of NO_x emissions in Germany between 2019–2021 using the multi-source plume method (MSPM)?, Geosci. Model Dev., 17, 4983–5007, https://doi.org/10.5194/gmd-17-4983-2024, 2024.

Deng, Z., Ciais, P., Tzompa-Sosa, Z. A., Saunois, M., Qiu, C., Tan, C., Sun, T., Ke, P., Cui, Y., Tanaka, K., Lin, X., Thompson, R. L., Tian, H., Yao, Y., Huang, Y., Lauerwald, R., Jain, A. K., Xu, X., Bastos, A., Sitch, S., Palmer, P. I., Lauvaux, T., D'Aspremont, A., Giron, C., Benoit, A., Poulter, B., Chang, J., Petrescu, A. M. R., Davis, S. J., Liu, Z., Grassi, G., Albergel, C., Tubiello, F. N., Perugini, L., Peters, W., and Chevallier, F.: Comparing national greenhouse gas budgets reported in UNFCCC inventories against atmospheric inversions, Earth Syst. Sci. Data, 14, 1639–1675, https://doi.org/10.5194/essd-14-1639-2022, 2022.

Department for Energy Security and Net Zero: Final UK greenhouse gas emissions national statistics: 1990 to 2022, https://www.gov.uk/government/statistics/final-uk-greenhouse-gas-emissions-national-statistics-1990-to-2022, (last accessed: 17 January 2025), 2024.

East, J. D., Jacob, D. J., Balasus, N., Bloom, A. A., Bruhwiler, L., Chen, Z., Kaplan, J. O., Mickley, L. J., Mooring, T. A., Penn, E., Poulter, B., Sulprizio, M. P., Worden, J. R., Yantosca, R. M., and Zhang, Z.: Interpreting the Seasonality of Atmospheric Methane, Geophys. Res. Lett., 51, e2024GL108 494, https://doi.org/10.1029/2024GL108494, 2024.

Engelen, R. J., Denning, A. S., and Gurney, K. R.: On error estimation in atmospheric CO_2 inversions, J. Geophys. Res. Atmos., 107, ACL10–1–ACL10–13, https://doi.org/10.1029/2002JD002195, 2002.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge University Press, Cambridge, England, UK, ISBN 978-0-52181210-8, https://doi.org/10.1017/CBO9780511535741, 2002.

Estrada, L. A., Varon, D. J., Sulprizio, M., Nesser, H., Chen, Z., Balasus, N., Hancock, S. E., He, M., East, J. D., Mooring, T. A., Oort Alonso, A., Maasakkers, J. D., Aben, I., Baray, S., Bowman, K. W., Worden, J. R., Cardoso-Saldaña, F. J., Reidy, E., and Jacob, D. J.: Integrated Methane Inversion (IMI) 2.0: an improved research and stakeholder tool for monitoring total methane emissions with high resolution worldwide using TROPOMI satellite observations, EGUsphere [preprint], pp. 1–31, https://doi.org/10.5194/egusphere-2024-2700, 2024.

Feigenspan, S., Wernicke, T., and Mielke, C.: Personal communication, 2024.

Fuß, R. and Akubia, J.: Personal communication, 2024.

Ganesan, A. L., Manning, A. J., Grant, A., Young, D., Oram, D. E., Sturges, W. T., Moncrieff, J. B., and O'Doherty, S.: Quantifying methane and nitrous oxide emissions from the UK and Ireland using a national-scale monitoring network, Atmos. Chem. Phys., 15, 6393–6406, https://doi.org/10.5194/acp-15-6393-2015, 2015.

Gerbig, C., Körner, S., and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, Atmos. Chem. Phys., 8, 591–602, https://doi.org/10.5194/acp-8-591-2008, 2008.

Henne, S., Brunner, D., Oney, B., Leuenberger, M., Eugster, W., Bamberger, I., Meinhardt, F., Steinbacher, M., and Emmenegger, L.: Validation of the Swiss methane emission inventory by atmospheric observations and inverse modelling, Atmos. Chem. Phys., 16, 3683–3710, https://doi.org/10.5194/acp-16-3683-2016, 2016.

ICOS RI: ICOS Handbook 2024, https://doi.org/10.18160/28AV-80QR, 2024.

ICOS RI, Bergamaschi, P., Colomb, A., De Mazière, M., Emmenegger, L., Kubistin, D., Lehner, I., Lehtinen, K., Leuenberger, M., Lund Myhre, C., Marek, M. V., O'Doherty, S., Platt, S. M., Plaß-Dülmer, C., Apadula, F., Arnold, S., Blanc, P.-E., Brunner, D., Chen, H., Chmura, L., Chmura, Ł., Conil, S., Couret, C., Cristofanelli, P., Forster, G., Frumau, A., Gerbig, C., Gheusi, F., Hammer, S., Haszpra, L., Hatakka, J., Heliasz, M., Henne, S., Hensen, A., Hoheisel, A., Kneuer, T., Larmanou, E., Laurila, T., Leskinen, A., Levin, I., Lindauer, M., Lopez, M., Mammarella, I., Manca, G., Manning, A., Martin, D., Meinhardt, F., Mölder, M., Müller-Williams, J., Noe, S. M., Nęcki, J., Ottosson-Löfvenius, M., Philippon, C., Pitt, J., Ramonet, M., Rivas-Soriano, P., Scheeren, B., Schumacher, M., Sha, M. K., Spain, G., Steinbacher, M., Sørensen, L. L., Vermeulen, A., Vítková, G., Xueref-Remy, I., di Sarra, A., Conen, F., Kazan, V., Roulet, Y.-A., Biermann, T., Delmotte, M., Heltai, D., Hermansen, O., Komínková, K., Laurent, O., Levula, J., Lunder, C., Marklund, P., Morguí, J.-A., Pichon, J.-M., Schmidt, M., Sferlazzo, D., Smith, P., Stanley, K., Trisolino, P., Zazzeri, G., ICOS Carbon Portal, ICOS Atmosphere Thematic Centre, ICOS Flask And Calibration Laboratory, and ICOS Central Radiocarbon Laboratory: European Obspack compilation of atmospheric methane data from ICOS and non-ICOS European stations for the period 1984–2024; `obspack_ch4_466_GVeu_v9.2_20240502`, https://doi.org/10.18160/9B66-SQM1, 2024.

IPCC, Calvo Buendia, E., Tanabe, K., Kranjc, A., Baasansuren, J., Fukuda, M., Ngarize, S., Osako, A., Pyrozhenko, Y., Shermanau, P., and Federici, S., eds.: 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, vol. 1, The Intergovernmental Panel on Climate Change (IPCC), https://www.ipcc-nggip.iges.or.jp/public/2019rf/index.html, 2019.

Janssens-Maenhout, G., Pinty, B., Dowell, M., Zunker, H., Andersson, E., Balsamo, G., Bézy, J.-L., Brunhes, T., Bösch, H., Bojkov, B., Brunner, D., Buchwitz, M., Crisp, D., Ciais, P., Counet, P., Dee, D., van der Gon, H. D., Dolman, H., Drinkwater, M. R., Dubovik, O., Engelen, R., Fehr, T., Fernandez, V., Heimann, M., Holmlund, K., Houweling, S., Husband, R., Juvyns, O., Kentarchos, A., Landgraf, J., Lang, R., Löscher, A., Marshall, J., Meijer, Y., Nakajima, M., Palmer, P. I., Peylin, P., Rayner, P., Scholze, M., Sierk, B., Tamminen, J., and Veefkind, P.: Toward an Operational Anthropogenic CO_2 Emissions Monitoring and Verification Support Capacity, Bull. Am. Meteorol. Soc., 101, E1439–E1451, https://doi.org/10.1175/BAMS-D-19-0017.1, 2020.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Jähn, M., Kuhlmann, G., Mu, Q., Haussaire, J.-M., Ochsner, D., Osterried, K., Clément, V., and Brunner, D.: An online emission module for atmospheric chemistry transport models: implementation in COSMO-GHG v5.6a and COSMO-ART v5.1-3.1, Geosci. Model Dev., 13, 2379–2392, https://doi.org/10.5194/gmd-13-2379-2020, 2020.

Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res. Atmos., 106, 4703–4715, https://doi.org/10.1029/2000JD900581, 2001.

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, T. F., and Heimann, M.: Technical Note: Atmospheric CO_2 inversions on the mesoscale using data-driven prior uncertainties: methodology and system evaluation, Atmos. Chem. Phys., 18, 3027–3045, https://doi.org/10.5194/acp-18-3027-2018, 2018.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., and Denier van der Gon, H.: Copernicus Atmosphere Monitoring Service regional emissions version 4.2 (CAMS-REG-v4.2), Copernicus Atmosphere Monitoring Service (CAMS) [publisher], ECCAD [distributor], https://doi.org/10.24380/0vzb-a387, 2021.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., and Denier van der Gon, H.: CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling, Earth Syst. Sci. Data, 14, 491–515, https://doi.org/10.5194/essd-14-491-2022, 2022.

Li, X. S. and Shao, M.: A Supernodal Approach to Incomplete LU Factorization with Partial Pivoting, ACM Trans. Math. Softw., 37, https://doi.org/10.1145/1916461.1916467, 2011.

Li, X. S., Demmel, J. W., Gilbert, J. R., Grigori, L., Shao, M., and Yamazaki, I.: SuperLU Users' Guide, june 2018 edn., 1999.

Manning, A. J., O'Doherty, S., Jones, A. R., Simmonds, P. G., and Derwent, R. G.: Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach, J. Geophys. Res. Atmos., 116, https://doi.org/10.1029/2010JD014763, 2011.

Mead, G. J., Herman, D. I., Giorgetta, F. R., Malarich, N. A., Baumann, E., Washburn, B. R., Newbury, N. R., Coddington, I., and Cossel, K. C.: Apportionment and Inventory Optimization of Agriculture and Energy Sector Methane Emissions Using Multi-Month Trace Gas Measurements in Northern Colorado, Geophys. Res. Lett., 51, e2023GL105 973, https://doi.org/10.1029/2023GL105973, 2024.

Meirink, J. F., Bergamaschi, P., Frankenberg, C., D'Amelio, M. T. S., Dlugokencky, E. J., Gatti, L. V., Houweling, S., Miller, J. B., Röckmann, T., Villani, M. G., and Krol, M. C.: Four-dimensional variational data assimilation for inverse modeling of atmospheric methane emissions: Analysis of SCIAMACHY observations, J. Geophys. Res. Atmos., 113, https://doi.org/10.1029/2007JD009740, 2008a.

Meirink, J. F., Bergamaschi, P., and Krol, M. C.: Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion, Atmos. Chem. Phys., 8, 6341–6353, https://doi.org/10.5194/acp-8-6341-2008, 2008b.

Moré, J. J. and Sorensen, D. C.: Computing a Trust Region Step, SIAM Journal on Scientific and Statistical Computing, 4, 553–572, https://doi.org/10.1137/0904038, 1983.

Munassar, S., Monteil, G., Scholze, M., Karstens, U., Rödenbeck, C., Koch, F.-T., Totsche, K. U., and Gerbig, C.: Why do inverse models disagree? A case study with two European CO_2 inversions, Atmos. Chem. Phys., 23, 2813–2828, https://doi.org/10.5194/acp-23-2813-2023, 2023.

Petrescu, A. M. R., Qiu, C., McGrath, M. J., Peylin, P., Peters, G. P., Ciais, P., Thompson, R. L., Tsuruta, A., Brunner, D., Kuhnert, M., Matthews, B., Palmer, P. I., Tarasova, O., Regnier, P., Lauerwald, R., Bastviken, D., Höglund-Isaksson, L., Winiwarter, W., Etiope, G., Aalto, T., Balsamo, G., Bastrikov, V., Berchet, A., Brockmann, P., Ciotoli, G., Conchedda, G., Crippa, M., Dentener, F., Groot Zwaaftink, C. D., Guizzardi, D., Günther, D., Haussaire, J.-M., Houweling, S., Janssens-Maenhout, G., Kouyate, M., Leip, A., Leppänen, A., Lugato, E., Maisonnier, M., Manning, A. J., Markkanen, T., McNorton, J., Muntean, M., Oreggioni, G. D., Patra, P. K., Perugini, L., Pison, I.,

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Raivonen, M. T., Saunois, M., Segers, A. J., Smith, P., Solazzo, E., Tian, H., Tubiello, F. N., Vesala, T., van der Werf, G. R., Wilson, C., and Zaehle, S.: The consolidated European synthesis of CH_4 and N_2O emissions for the European Union and United Kingdom: 1990–2019, Earth Syst. Sci. Data, 15, 1197–1268, https://doi.org/10.5194/essd-15-1197-2023, 2023.

Petrescu, A. M. R., Peters, G. P., Engelen, R., Houweling, S., Brunner, D., Tsuruta, A., Matthews, B., Patra, P. K., Belikov, D., Thompson, R. L., Höglund-Isaksson, L., Zhang, W., Segers, A. J., Etiope, G., Ciotoli, G., Peylin, P., Chevallier, F., Aalto, T., Andrew, R. M., Bastviken, D., Berchet, A., Broquet, G., Conchedda, G., Dellaert, S. N. C., Denier van der Gon, H., Gütschow, J., Haussaire, J.-M., Lauerwald, R., Markkanen, T., van Peet, J. C. A., Pison, I., Regnier, P., Solum, E., Scholze, M., Tenkanen, M., Tubiello, F. N., van der Werf, G. R., and Worden, J. R.: Comparison of observation- and inventory-based methane emissions for eight large global emitters, Earth Syst. Sci. Data, 16, 4325–4350, https://doi.org/10.5194/essd-16-4325-2024, 2024.

Ramsden, A. E., Ganesan, A. L., Western, L. M., Rigby, M., Manning, A. J., Foulds, A., France, J. L., Barker, P., Levy, P., Say, D., Wisher, A., Arnold, T., Rennick, C., Stanley, K. M., Young, D., and O'Doherty, S.: Quantifying fossil fuel methane emissions using observations of atmospheric ethane and an uncertain emission ratio, Atmos. Chem. Phys., 22, 3911–3929, https://doi.org/10.5194/acp-22-3911-2022, 2022.

Rieger, D., Bangert, M., Bischoff-Gauss, I., Förstner, J., Lundgren, K., Reinert, D., Schröter, J., Vogel, H., Zängl, G., Ruhnke, R., and Vogel, B.: ICON–ART 1.0 – a new online-coupled model system from the global to regional scale, Geosci. Model Dev., 8, 1659–1676, https://doi.org/10.5194/gmd-8-1659-2015, 2015.

Rigby, M., Manning, A. J., and Prinn, R. G.: Inversion of long-lived trace gas emissions using combined Eulerian and Lagrangian chemical transport models, Atmos. Chem. Phys., 11, 9887–9898, https://doi.org/10.5194/acp-11-9887-2011, 2011.

Rocher-Ros, G., Stanley, E. H., Loken, L. C., Casson, N. J., Raymond, P. A., Liu, S., Amatulli, G., and Sponseller, R. A.: Global methane emissions from rivers and streams, Nature, 621, 530–535, https://doi.org/10.1038/s41586-023-06344-6, 2023.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding, vol. 2 of *Series on Atmospheric, Oceanic and Planetary Physics*, World Scientific Publishing Company, Singapore, ISBN 978-981-02-2740-1, https://doi.org/10.1142/3171, 2000.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 1919–1964, https://doi.org/10.5194/acp-3-1919-2003, 2003.

Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Periáñez, A., and Potthast, R.: Kilometre-scale ensemble data assimilation for the COSMO model (KENDA), Q. J. R. Meteorolog. Soc., 142, 1453–1472, https://doi.org/10.1002/qj.2748, 2016.

Schröter, J., Rieger, D., Stassen, C., Vogel, H., Weimer, M., Werchner, S., Förstner, J., Prill, F., Reinert, D., Zängl, G., Giorgetta, M., Ruhnke, R., Vogel, B., and Braesicke, P.: ICON-ART 2.1: a flexible tracer framework and its application for composition studies in numerical weather forecasting and climate simulations, Geosci. Model Dev., 11, 4043–4068, https://doi.org/10.5194/gmd-11-4043-2018, 2018.

Segers, A. and Houweling, S.: CAMS global inversion-optimised greenhouse gas fluxes and concentrations, v22r2, Copernicus Atmosphere Monitoring Service [data set], https://ads.atmosphere.copernicus.eu/datasets/cams-global-greenhouse-gas-inversion, (last accessed: 18 April 2024), 2020.

Seidel, D. J., Zhang, Y., Beljaars, A., Golaz, J.-C., Jacobson, A. R., and Medeiros, B.: Climatology of the planetary boundary layer over the continental United States and Europe, J. Geophys. Res. Atmos., 117, https://doi.org/10.1029/2012JD018143, 2012.

Steiner, M., Cantarello, L., Henne, S., and Brunner, D.: Flow-dependent observation errors for greenhouse gas inversions in an ensemble Kalman smoother, Atmos. Chem. Phys., 24, 12 447–12 463, https://doi.org/10.5194/acp-24-12447-2024, 2024a.

Steiner, M., Peters, W., Luijkx, I., Henne, S., Chen, H., Hammer, S., and Brunner, D.: European CH_4 inversions with ICON-ART coupled to the CarbonTracker Data Assimilation Shell, Atmos. Chem. Phys., 24, 2759–2782, https://doi.org/10.5194/acp-24-2759-2024, 2024b.

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Greally, B. R., Lunder, C., Maione, M., Mühle, J., O'Doherty, S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P. G., Vollmer, M. K., Weiss, R. F., and Yokouchi, Y.: An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, Atmos. Chem. Phys., 9, 1597–1620, https://doi.org/10.5194/acp-9-1597-2009, 2009.

Storm, I., Karstens, U., D'Onofrio, C., Vermeulen, A., and Peters, W.: A view of the European carbon flux landscape through the lens of the ICOS atmospheric observation network, Atmos. Chem. Phys., 23, 4993–5008, https://doi.org/10.5194/acp-23-4993-2023, 2023.

Tenkanen, M. K., Tsuruta, A., Denier van der Gon, H., Höglund-Isaksson, L., Leppänen, A., Markkanen, T., Petrescu, A. M. R., Raivonen, M., Aaltonen, H., and Aalto, T.: Partitioning anthropogenic and natural methane emissions in Finland during 2000–2021 by combining bottom-up and top-down estimates, Atmos. Chem. Phys., 25, 2181–2206, https://doi.org/10.5194/acp-25-2181-2025, 2025.

Thanwerdas, J., Saunois, M., Berchet, A., Pison, I., and Bousquet, P.: Investigation of the renewed methane growth post-2007 with high-resolution 3-D variational inverse modeling and isotopic constraints, Atmos. Chem. Phys., 24, 2129–2167, https://doi.org/10.5194/acp-24-2129-2024, 2024.

Thompson, R. L., Krishnankutty, N., Pisso, I., Schneider, P., Stebel, K., Sasakawa, M., Stohl, A., and Platt, S.: Efficient use of a Lagrangian Particle Dispersion Model for atmospheric inversions using satellite observations of column mixing ratios, EGUsphere [preprint], pp. 1–28, https://doi.org/10.5194/egusphere-2025-147, 2025.

UBA: Submission under the United Nations Framework Convention on Climate Change 2023, German Environment Agency, https://doi.org/10.60810/openumwelt-2570, 2023.

UBA: Submission under the United Nations Framework Convention on Climate Change 2024, German Environment Agency, https://doi.org/10.60810/openumwelt-7441, 2024.

UNFCCC: National Inventory Submissions 2024, https://unfccc.int/ghg-inventories-annex-i-parties/2024, (last accessed: 18 March 2025), 2024.

van der Laan-Luijkx, I. T., van der Velde, I. R., van der Veen, E., Tsuruta, A., Stanislawska, K., Babenhauserheide, A., Zhang, H. F., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C., and Peters, W.: The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: implementation and global carbon balance 2001–2015, Geosci. Model Dev., 10, 2785–2800, https://doi.org/10.5194/gmd-10-2785-2017, 2017.

Vanderbecken, P. J., Dumont Le Brazidec, J., Farchi, A., Bocquet, M., Roustan, Y., Potier, E., and Broquet, G.: Accounting for meteorological biases in simulated plumes using smarter metrics, Atmos. Meas. Tech., 16, 1745–1766, https://doi.org/10.5194/amt-16-1745-2023, 2023.

Varon, D. J., Jacob, D. J., Sulprizio, M., Estrada, L. A., Downs, W. B., Shen, L., Hancock, S. E., Nesser, H., Qu, Z., Penn, E., Chen, Z., Lu, X., Lorente, A., Tewari, A., and Randles, C. A.: Integrated Methane Inversion (IMI 1.0): a user-friendly, cloud-based facility for inferring high-resolution methane emissions from TROPOMI satellite observations, Geosci. Model Dev., 15, 5787–5805, https://doi.org/10.5194/gmd-15-5787-2022, 2022.

Veldeman, N., van der Maas, W., van Aardenne, J., Goodwin, J., Mareckova, K., Adams, M., Ruyssenaars, P., Wankmüller, R., and Pye, S.: 7. Spatial mapping of emissions, in: EMEP/EEA air pollutant emission inventory guidebook 2013, European Environment Agency, https://www.eea.europa.eu/publications/emep-eea-guidebook-2013/part-a-general-guidance-chapters/7-spatial-mapping-of-emissions/view, 2013.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald,

https://doi.org/10.5194/egusphere-2025-1464
Preprint. Discussion started: 15 May 2025

A. M., Ribeiro, A. H., Pedregosa, F., and van Mulbregt, P.: SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Weber, T., Wiseman, N. A., and Kock, A.: Global ocean methane emissions dominated by shallow coastal waters, Nat. Commun., 10, 1–10, https://doi.org/10.1038/s41467-019-12541-7, 2019.

Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, Quart. J. Roy. Meteorol. Soc., 141, 563–579, https://doi.org/10.1002/qj.2378, 2015.