European CH_4 inversions with ICON-ART coupled to CarbonTracker Data Assimilation Shell

Michael Steiner[1], Wouter Peters[2,3], Ingrid Luijkx[2], Stephan Henne[1], Huilin Chen[3], Samuel Hammer[4], and Dominik Brunner[1]

[1]Empa, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland
[2]Environmental Sciences Group, Dept of Meteorology and Air Quality, Wageningen University and Research, Wageningen, the Netherlands
[3]University of Groningen, Centre for Isotope Research, Groningen, the Netherlands
[4]Institut für Umweltphysik, Heidelberg University, Heidelberg, Germany

Correspondence: Michael Steiner (michael.steiner@empa.ch) and Dominik Brunner (dominik.brunner@empa.ch)

Abstract.

We present the first application of the atmospheric chemistry and transport model ICON-ART in inverse modelling of greenhouse gas fluxes with an Ensemble Kalman smoother. For this purpose, we extended ICON-ART to efficiently handle gridded emissions, generate an ensemble of perturbed emissions during runtime, and use nudging on selected variables to keep the simulations close to analyzed meteorology. We show that the system can optimize total and anthropogenic European CH_4 fluxes on a national-scale in an idealized setup using pseudo-observations from a realistic network of measurement stations. However, we were unable to constrain the sum of the natural emission sources of comparatively low magnitude. Also regions with low emissions and regions with low observational coverage could not be optimized individually for lack of observational constraints. Furthermore, we investigated the sensitivities towards different inversion parameters and design choices with 15 sensitivity runs using the same idealized setup, demonstrating the robustness of the approach when regarding some minimal requirements of the setup (e.g., number of ensemble members). Subsequently, we applied the system to real in-situ observations from 28 European stations for three years, 2008, 2013 and 2018. We used a priori anthropogenic fluxes from the EDGARv6 inventory and a priori natural fluxes from peatlands and mineral soils, inland waters, the ocean, biofuels and biomass burning and geology. Our results for the year 2018 indicate that anthropogenic emissions may be underestimated in EDGARv6 by ca. 25% in the Benelux countries and, to a smaller degree, in northwestern France and southern England. In the rest of the domain, anthropogenic fluxes are corrected downwards by the inversion suggesting an overestimation in the a priori. For most countries, this means that the a posteriori country-total anthropogenic emissions are closer to the values reported to the United Nations Framework Convention on Climate Change (UNFCCC) than the a priori emissions from EDGARv6. Aggregating the a posteriori emissions across the EU27 + UK results in a total of $17.4 \, \mathrm{Tg \, yr^{-1}}$, while the a priori emissions were $19.9 \, \mathrm{Tg \, yr^{-1}}$. Our a posteriori is close to the total reported to UNFCCC of $17.8 \, \mathrm{Tg \, yr^{-1}}$. Natural emissions are reduced from their a priori magnitude almost everywhere, especially over Italy and Romania/Moldova, where a priori geological emissions are high, and over the United Kingdom and Scandinavia where emissions from peatlands and wetlands were possibly unusually low during the hot and dry summer 2018. Our a posteriori anthropogenic emissions for the EU27 + UK fall within the range estimated

by global top-down studies, but are lower than most other regional inversions. However, many of these studies have used
observations from different measurement stations or satellite observations. The spatial pattern of the emission increments in
our results, especially the increase in the Benelux countries, also agrees well with other regional inversions.

1 Introduction

To achieve the long-term goal of the Paris Agreement to limit global temperature increases to well below 2°C, global green-
house gas (GHG) emissions will have to be reduced drastically in the coming decades. The implementation of the Paris Agree-
ment requires all parties to commit to mitigation measures (described in National Determined Contributions) and to regularly
report their anthropogenic GHG emissions, in the form of National Inventory Reports (NIR), to the United Nations Framework
Convention on Climate Change (UNFCCC). NIRs are developed from socioeconomic statistics, activity data and emission
factors following the guidelines of the Intergovernmental Panel for Climate Change (Eggleston et al., 2006). Following these
guidelines, most parties also report uncertainties in emissions, which they calculate based on uncertainties in activity data,
emission factors, or uncertainties in the underlying data.

Complementary to these bottom-up emission inventories is the "top-down approach", where atmospheric inversions are used
for emission estimation from observations (Bergamaschi et al., 2018; Nisbet and Weiss, 2010). Due to the Paris Agreement, the
interest in high-resolution inversions with country-scale emission estimates has grown recently. However, top-down emission
estimation is still subject to large and poorly quantified uncertainties due to insufficient coverage of measurements, errors
in simulated atmospheric transport, representation errors, measurement biases, and other factors. To advance the field, it is
therefore paramount to reduce these errors as much as possible and to build modeling systems that properly account for the
remaining uncertainties.

Inversions for CH_4 have already been made in previous studies. European CH_4 emissions have been estimated and compared
to bottom-up values in both regional (Bergamaschi et al., 2018, 2022; Petrescu et al., 2023) and global (Deng et al., 2022;
Petrescu et al., 2023) inversions, using both surface in-situ measurements and satellite observations. Bergamaschi et al. (2018)
compared a total of 7 inversion models with a regional setup for the period 2006 to 2012, all models using harmonized
observations. The mean of the estimates of the models for the anthropogenic emissions for EU27 + UK was higher than the
reported values for all years, but still within the calculated uncertainty range. Another regional inversion study for Europe was
recently presented by Bergamaschi et al. (2022) using a new, nested high-resolution inversion system Flexpart-COSMO TM5
4DVAR. They also compared the results with the Flexpart extended Kalman filter (FLExKF) (Brunner et al., 2012) and with
TM5-4DVAR (Meirink et al., 2008) inversions. All three inversion models resulted in higher emissions for 2018 for Germany,
France and Benelux than the sum of UNFCCC reported and natural (estimated by GCP) emissions.

European CH_4 emissions were also estimated from global inversions operating at lower resolution and often assimilating
a smaller set of observations available over Europe compared to the regional systems. Deng et al. (2022) and Petrescu et al.
(2023) compared a number of global inversions presented previously by Saunois et al. (2020) with the reported values. While
Deng et al. (2022) included all Annex 1 countries (with periodic emission reports) and non-Annex 1 countries (with only

sporadic reports) worldwide, Petrescu et al. (2023) focused on EU27 + UK and also compared regional inversions with the reported values. The results showed that anthropogenic CH_4 emissions estimated in regional inversions were generally higher than reported emissions while global inversions were mostly lower. This general tendency was found irrespective of whether only ground in-situ measurements or satellite observations were assimilated. These results show that there is still little consistency between different inversion results and further work is needed to identify the causes of the discrepancies. Therefore, a new model intercomparison experiment was established by the Atmospheric Tracer Transport Model Intercomparison Project (TransCom), which requires all participating groups to follow a common data protocol ensuring maximum consistency in terms of the usage of observation data, boundary conditions and a priori fluxes. The results presented here for the real data application are based on simulations following this protocol.

In inverse modelling, measured atmospheric dry air mole fractions are linked to emissions using an atmospheric transport model (ATM). The most likely set of emissions is determined by minimizing a Bayesian cost function with an inversion algorithm given a prior constraint (usually a bottom-up inventory or flux model) and uncertainties. Different inversion techniques exist (see e.g. Chap. 11 of Brasseur and Jacob, 2017), such as synthesis (Gurney et al., 2002; Baker et al., 2006; Butler et al., 2010), geostatistical (Michalak et al., 2004; Gourdji et al., 2012), Kalman smoother (Bruhwiler et al., 2005), Ensemble Kalman Filter (EnKF), smoother (EnKS) or square root filter (EnSRF) (Peters et al., 2005; Tsuruta et al., 2017), and 4-D variational inversion (4D-var, Chevallier et al., 2005; Baker et al., 2010; Bergamaschi et al., 2022). They have been developed to address different trace gases, observations types and spatial and temporal scales. 4D-var and EnKF methods are computationally expensive but have become standard methods today to address large inversion problems.

A limiting factor for analytical synthesis and geostatistical inversions is the dimension of the inversion problem (both the control and observation space), which needs to be sufficiently small to store the related covariance error matrices in computer memory and calculate their algebraic inverse. The 4-D variational approach, on the other hand, where the cost function is minimized by the calculation of its gradient and through an iterative descent, requires an adjoint model, which is often not available for a given ATM. Ensemble Kalman filter, or smoother data assimilation (Evensen, 1994, 2003; Burgers et al., 1998) has the advantages that it can deal with large inversion problems, that no adjoint ATM is required (Kalnay, 2010) and that they return an approximate error covariance matrix. The disadvantage is that the covariance error matrix and the Kalman Gain are only approximated based on a finite ensemble.

Peters et al. (2005) developed such an ensemble Kalman smoother, which was further implemented in the CarbonTracker Data Assimilation Shell (CTDAS) (van der Laan-Luijkx et al., 2017). It was designed to optimize biospheric and oceanic CO_2 fluxes from different biomes and ocean regions on a weekly time-scale by assimilating the CO_2 observations from a global network of stations. CTDAS has been applied in subsequent studies to investigate, for example, the carbon budget over North America (Peters et al., 2007), Europe (Peters et al., 2010; Smith et al., 2020), South America (van der Laan-Luijkx et al., 2015) and globally (van der Laan-Luijkx et al., 2017). More recently, CTDAS has been coupled with a Lagrangian particle dispersion model for studying regional carbon budgets (He et al., 2018) and has also been applied to other species like methane (Bruhwiler et al., 2014; Tsuruta et al., 2019).

A critical requirement for accurate emission estimates by inverse modelling is the quality of the ATM. An attractive new atmospheric model is the global ICOsahedral Nonhydrostatic (ICON) atmospheric modeling framework (Wan et al., 2013; Zängl et al., 2015; Pham et al., 2021), which can be extended with the "Aerosols and Reactive Trace gases"-model (ART), developed at the Karlsruhe Institute of Technology (KIT) (Rieger et al., 2015; Weimer et al., 2017; Schröter et al., 2018) to simulate aerosols and trace gases. ICON-ART can be run from global scale down to cloud resolving scale and has attractive transport properties such as mass conversation, positivity of tracers and the use of recent developments in subgrid-scale transport.

Here we present the first application of ICON-ART in inverse modelling of GHG emissions with CTDAS. For this purpose, we extended ICON-ART with modules for efficient handling of emissions and online (i.e, during runtime) generation of the ensemble of perturbed fluxes, and with a nudging scheme to keep the simulations close to analyzed meteorology. Using an idealized setup with synthetically generated observations, we analyze how well the new, computationally efficient model is suited to constrain European anthropogenic CH_4 emissions for individual countries with observations from a European observation network. The sensitivity of the system to different parameters is analyzed in a set of sensitivity experiments. The system is then applied to real observations from a harmonized set of CH_4 dry air mole fraction observations from 28 European stations to assess the performance of ICON-ART in terms of atmospheric transport and to demonstrate the capability of the new system to constrain European emissions using this network. A detailed description of our model setup and the methodology is given in Sect. 2. In Sect. 3 we present the results of both applications, with pseudo-observations as well as with real observations. Sect. 4 provides conclusions.

2   Model description and methodology

2.1   ICON-ART model and simulation setup

2.1.1   Weather and climate model ICON

ICON is a highly versatile non-hydrostatic atmospheric model for global and regional weather and climate simulations developed jointly by the German Weather Service (DWD) and the Max Planck Institute of Meteorology (Wan et al., 2013; Zängl et al., 2015; Pham et al., 2021). It has been used at DWD for operational weather prediction since 2016 and has been coupled with other models including an ocean and a land surface model for climate simulations (Giorgetta et al., 2018). ICON is based on an icosahedral triangular grid, where 20 equilateral triangles of an icosahedron are iteratively split into smaller triangles up to the desired resolution. With such a grid, the problem of singularity at the poles is avoided. To zoom into a specific region, refined grids can be nested into the parent grid, with one additional edge bisection. The model equations are fully compressible and the vertical discretisation is in generalised smooth-level vertical coordinates (SLEVE) (Leuenberger et al., 2010). Tracers in ICON are transported with perfect mass conservation by solving the continuity equation of mass for each tracer consecutively in the vertical with a finite volume method and in the horizontal direction with a simplified flux-form semi-Lagrangian method (Miura, 2007; Lauritzen et al., 2011; Rieger et al., 2015).

2.1.2 ART extension for trace gases and aerosols

The ART-model was developed as an extension for ICON at the Karlsruhe Institute of Technology (KIT) (Rieger et al., 2015; Weimer et al., 2017; Schröter et al., 2018) with the aim of simulating aerosols as well as passive and chemically reactive trace gases. The ART module is coupled online with ICON and allows a flexible definition of tracers and processes to be included (Schröter et al., 2018). Since only CH_4 was simulated in the present study, all tracers are passive tracers, i.e. they are only transported without radiative feedback on the meteorology and without degradation by the hydroxyl radical (OH). Depletion by reaction with OH is assumed to be negligible given the short residence time of the air masses within the domain of no more than a few days compared to the CH_4 lifetimes of about 10 years.

To simplify and accelerate the treatment of emissions during the simulations, we implemented the Online Emissions Module (OEM) into ICON-ART, which was originally developed for the regional weather and climate model COSMO (Jähn et al., 2020). Unlike the standard offline approach, where numerous input files have to be provided at discrete model time steps, OEM requires only a small number of files at the beginning of a simulation. These files contain annual mean sector-specific 2D emission fields as well as the temporal and vertical profiles for individual emission categories and countries. During the simulation, these profiles are applied online to update the hourly emissions for each species. OEM has recently become an official component of ART (since ART version 2.6.3).

To project the inventory data to the ICON grid, we extended the stand-alone Python package emiproc (also described in Jähn et al., 2020). Emiproc projects emission data of various inventories to the model grid in a mass-conserving manner by calculating the overlap of the source and target grid at every grid cell. It also generates the temporal and vertical scaling profiles.

2.1.3 General setup of ICON-ART forward simulations

The ICON-ART simulations were performed in limited-area-mode (LAM) on a grid covering Europe (see Figs. in Sect. 2 and 3). The horizontal grid was R3B06 (see the established grid notation in Zängl et al., 2015), which corresponds to a mean grid spacing Δx between neighboring triangles of about 26 km and yields a total of 21,344 grid cells. Vertically, 60 levels were used between the surface and about 23 km altitude. The time step was 120 s. We used ICON in the Numerical Weather Prediction (NWP) configuration with a single-moment microphysics scheme including graupel and the tile approach for soil switched on, considering subgrid-scale land-cover variability with 6 tiles (3 land plus 3 water types).

The meteorological fields were initialised at the beginning of every simulation with the reanalysis data of ERA5 (Hersbach et al., 2020). During the simulation, the meteorological fields were weakly nudged in the entire domain towards the 3-hourly reanalysis data to keep the simulated meteorology close to the analyzed meteorology. This required a modification of the ICON code, since ICON in the LAM configuration only allows nudging towards meteorological boundary conditions (density, virtual potential temperature, Exner pressure, specific humidity and wind) near the borders of the domain. Additionally, the LAM-grid was created such that the boundary zone extends over the entire domain. As in the standard scheme, the nudging strength α_{nudge} was set to decrease with the distance of the cell row r from the lateral boundaries (r_0), but the nudging was not restricted to the 8

cell rows closest to the boundaries. Instead it decreased exponentially with an e-folding width of 2 cell rows towards a minimum

155 dimensionless nudging strength of 0.001 applied in the main part of the domain: $\alpha_{nudge} = 0.069\ exp(-\frac{r-r_0}{2}) + 0.001$.

The modeled CH_4 dry air mole fractions were constructed with three different types of tracers representing the CH_4 background, emissions, and ensemble members. The tracer representing CH_4 from emissions (CH_4^E) within the model domain was initialized with a value of zero (or rather with an offset value which is later subtracted in the post-processing) at the simulation start and was updated by OEM with the emissions from the inventories (described in Sect. 2.2.4) at every model time step. The

160 ensemble members used for the optimization scheme in the EnKS are represented as tracers in the model, similar to CH_4^E, but with perturbed emissions (see Sect. 2.2). The background CH_4 dry air mole fractions (CH_4^{BG}) were represented as a separate tracer, which was initialized and updated at the lateral boundaries with data from the CAMS v19r1 inversion product (available via https://ads.atmosphere.copernicus.eu/, last access: 4 December 2023).

For every 10 d simulation window of CTDAS (see Sect. 2.2.1), one ICON-ART simulation was performed. The model runs

165 were initialized 24 h before each 10 d window to allow for model spin-up. At the end of the 24-hour meteorological spin-up period, all CH_4 tracer mole fractions were overwritten by the initial CH_4 conditions produced by the CTDAS system (see Sect. 2.2.1).

We carried out the simulations on the supercomputer "Piz Daint" of the Swiss National Supercomputing Center (CSCS, https://www.cscs.ch/, last access: 4 December 2023), namely on the XC40 compute nodes, each with two Intel Xeon E5-

170 2695 v4 2.10GHz processors (2x18 cores, 64/128 GB RAM). A simulation with the described setup and spread across 16 nodes required around 10 node hours per 11-day simulation. For the inversion of a whole year, the total costs (including the comparatively low costs for CTDAS) amounted to around 1300 node hours or 36 x 1300 = 46'800 core hours.

2.1.4 Adaptations of ICON-ART to CTDAS

To couple ICON-ART with CTDAS in a robust and efficient way, we made a few adaptations to ICON-ART and the simulation

175 setup. CTDAS requires a large ensemble of CH_4 tracers to be simulated, each ensemble member corresponding to one specific perturbation of the state vector (e.g., fluxes and boundary conditions). This is usually achieved by generating an ensemble of input fields, typically one set of perturbed hourly emission maps and boundary conditions per member. For a large ensemble with a few hundred members and hourly emissions, this may results in a very large number or size of input files and correspondingly expensive I/O during the simulation. To overcome this problem, we extended the OEM module with the option

180 to generate an ensemble of perturbed fluxes and corresponding tracers online during the simulation. With this extension, the only input required at the start of the simulation is the ensemble of perturbed scaling factors (λ) provided by CTDAS, which greatly simplified and accelerated the simulations. Each scaling factor scales the flux of one emission category (one tracer can experience the emissions of multiple categories) in one region. The regions can be any combination of grid cells (including individual cells) and are defined by a region mask provided as input for OEM (see Sect. 2.2.2). The setups used in this study,

185 e.g. the emission categories and regions, are described in Sect. 2.2.

The generation of the flux ensemble permits negative fluxes, which could result in negative mole fractions of the CH_4 tracers. Since none of the available transport schemes in ICON allows for negative tracer mole fractions, we bypassed this
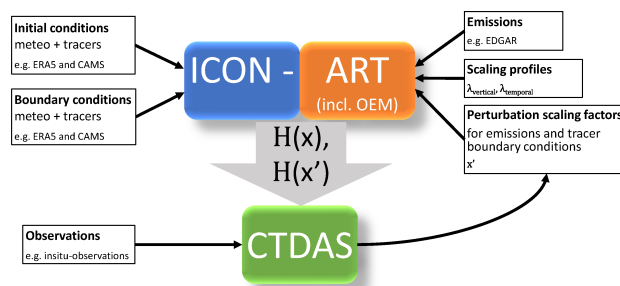
Figure 1. Schematic of the coupling of CTDAS with ICON-ART.

problem by adding a constant offset mass mole fraction of $1.2e^{-6}$ to all CH_4^E tracers and subtracting the same offset afterwards from the output. Since the transport characteristics within the numerical model ICON-ART exhibit non-linearity, the tracer concentrations are not identical when transported with or without an offset. However, our tests showed that these differences are negligible and we consequently expect no significant impact on the outcome of the simulations.

Our implementation of ICON-ART in CTDAS not only allows for the perturbation of fluxes, the ensemble can at the same time also hold perturbed background mole fractions. This allows for the optimization of the background mole fractions, which is described in 2.2.6. Technically, this is achieved by perturbing the CH_4^{BG} tracer mole fractions in different regions of the lateral boundary, with scaling factors provided by CTDAS in a separate file. Similar to the generation of the flux ensemble, a lateral boundary region can be any combination of grid cells (in this case of the lateral boundary cells) defined by a region mask.

Figure 1 gives a schematic overview of how CTDAS is coupled to ICON-ART. ICON-ART is driven by initial and boundary data from ERA5 for meteorology and CAMS for tracer concentrations and adds CH_4 emissions via OEM. Using a priori perturbation scaling factors from CTDAS, an ensemble of CH_4 tracers is generated. The simulated concentrations of the individual ensemble members (H(x')) and the ensemble means (H(x)) are sampled at the stations and read into CTDAS. CTDAS optimizes the emissions based on these values and the observations and uses the optimized scaling factors to generate the scaling factors for the next cycle.

2.2 CTDAS inversion setup

In this study, we used CTDAS to estimate anthropogenic CH_4 emissions either in an idealized setup using synthetically generated atmospheric observations, or in a real data application using quasi-continuous in-situ observations and very few discrete flask samples of CH_4 dry air mole fractions. ICON-ART acts as an observation operator, i.e. it connects the surface fluxes to atmospheric CH_4 dry air mole fractions. In the following, we describe the setups for the two applications. The description of the idealized setup refers to the reference inversion in which we only optimize emissions but otherwise stay as
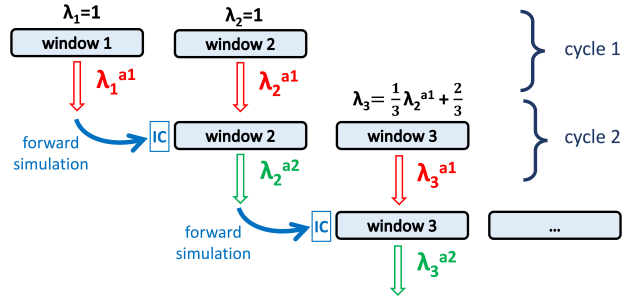
Figure 2. Schematic of the ensemble Kalman smoother. The schematic shows the first two assimilation cycles. All remaining cycles are identical to the second one. Each assimilation window has a length of 10 days. One cycle is composed of two windows.

close to the real setup as possible. In addition to the reference inversion, we performed sensitivity experiments testing different aspects of the inversion setup as described in Sect. 2.3. These include experiments where we optimized background mole fractions in addition to emissions (see Sect. 2.2.6). Background optimization is also applied in the real data application.

2.2.1 Optimization scheme

CTDAS applies an Ensemble Kalman smoother with a fixed-lag assimilation window (Peters et al., 2005). A schematic of the configuration used here is shown in Fig. 2. Since our model domain is limited to Europe, the fluxes can affect observed CH_4 mole fractions only over a couple of days. In our simulations, almost all mass of the emission tracer has left the domain after 20 days. Accordingly, CTDAS has been set up to optimize 10-day mean fluxes with a fixed lag of 2, resulting in a total assimilation window length of 20 days. Observations in a given 10-day window can thus constrain the fluxes of the previous and the present 10-day window.

The s scaling factors are optimized using an Ensemble Kalman smoother as described in Peters et al. (2005), which is based on the Ensemble Square Root Filter presented by Whitaker and Hamill (2002). In the filter, the error covariance matrix $\mathbf{P}$ (both a priori and a posteriori) of size $[s \times s]$ is represented by information in a smaller dimension N, which corresponds to the number of ensemble members, which is set to 192 in our applications. The ensemble is generated from randomly perturbed state vectors with the magnitude and correlation structure of the perturbations, as determined by the a priori error covariance matrix.

Since we start with a priori scaling factors with a value of one, the initial perturbation of the scaling factors for the first two windows of the simulation has a mean of $\lambda_{(t=1,2)} = 1$. After assimilation of the observations during these two windows, the first window is simulated again with the optimized scaling factors λ_1^{a1} (superscript $a1$ indicates analyzed/optimized once) to provide updated initial CH_4 mole fractions for the second cycle. From there, the second cycle uses the optimized scaling factors λ_2^{a1} from the first cycle and continues with the third window, which inherits the scaling factors from the previous window with the following state propagation model: $\lambda_{t+1} = 1/3 + 2/3\,\lambda_t^{a1}$. A priori information is thus inserted with a weight of 1/3 while a posteriori information from the previous assimilation step is propagated with a weight of 2/3. This weighting

is based on the study of van der Laan-Luijkx et al. (2017), where the average of the optimized scaling factors from the two previous windows were taken into account with a weight of 2/3. Their choice of 2/3 was a design choice based on the concept that the timescale to revert a system to the prior state should approximately be 8 weeks in case that no data is available to constrain it. Full propagation of a posteriori information is analyzed in a sensitivity inversion. We apply the state propagation model only to the mean scaling factors and not to the individual ensemble members. Each window starts with a new a priori covariance again, hence no information on the uncertainty reduction in the previous windows is taken into account. During the second cycle, observations of the third window are assimilated. The second cycle is completed by simulating the second window again with the now twice optimized scaling factors λ_2^{a2}, to provide updated initial mole fractions for the third cycle. The third and all following cycles follow the same principle as the second cycle (Peters et al., 2005, Fig. 1). The main output is thus the sequence of twice-optimized scaling factors (and their uncertainties) for each 10-day window of the simulation period. Error covariances are discussed separately in Section 2.2.5

2.2.2 State vector

In our setup, CTDAS separately optimizes anthropogenic and natural fluxes. For both categories, the emissions are optimized for 21,344 individual regions, each region corresponding to one grid cell in the domain. The formalism of CTDAS is described in detail in Peters et al. (2005). As we use a lag of 2, the state vector $\mathbf{x}$ has 85,376 flux elements in our implementation (2 windows x 2 categories x 21,344 grid cells). Only flux elements are optimized in the reference setup of the idealized study. Simultaneous optimization of fluxes and background mole fraction is analyzed in additional sensitivity inversions and then applied in the real data application where 8 state vector parameters are used to optimize the background mole fractions (see Fig. 4).

2.2.3 Ensemble size

Previous applications of CTDAS used 200, 1500 (Peters et al., 2005), 500 (Tsuruta et al., 2017, 2019) and 150 ensemble members (van der Laan-Luijkx et al., 2017). In Peters et al. (2005) 100-200 model ensemble members were sufficient to represent a state vector of 14,400 parameters, however with substantial spatial correlations to reduce the degrees of freedom. In our setup we use an ensemble size of 192, wherein every ensemble member is represented by one tracer. The sensitivity to a smaller and larger size of 50 and 300 members was also analyzed. While the performance was significantly reduced with only 50 members, there was only a marginal improvement with 300 members. Therefore, we chose 192 members for our inversions. The degrees of freedom (calculated with Eq. 21 in Peters et al., 2005) in our a-priori covariance error matrix is 425.

2.2.4 A priori fluxes

In the application of our system, we followed the "protocol for the intercomparison of national CH_4 emissions estimated by inverse modelling system for Western Europe", an intercomparison effort of the TransCom modelling community. The protocol prescribes the prior fluxes to be used, which were pre-processed and provided on a 0.25° x 0.25°grid. The anthropogenic fluxes
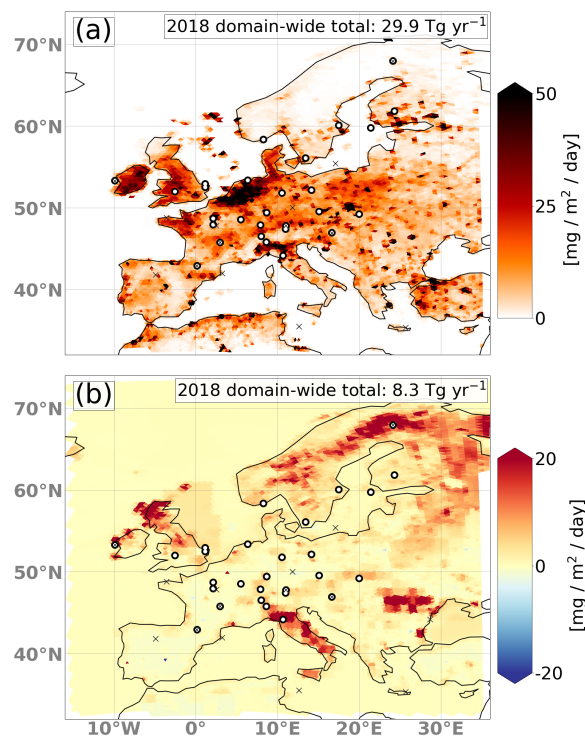
Figure 3. Spatial distribution of CH_4 emissions from anthropogenic (a) and natural sources (b) used in the real data application for the period of 11-21 June 2018. Overlaid are the measurement locations used in the inversion. The numbers for the domain-wide total fluxes are given for the entire year 2018. The circles show the in-situ measurement locations while the small crosses show the locations of flask sampling.

(agriculture, waste and fossil fuels) are taken from the EDGAR v6.0 (Crippa et al., 2021) inventory. The following natural
265   fluxes are taken into account: peatlands and mineral soils from JSBACH-HIMMELI (Raivonen et al., 2017; Reick et al., 2013) (version 2), inland water (provided by Université Libre de Bruxelles to the GCP-CH4 data set; Saunois et al., 2020), termites (Saunois et al., 2020), ocean (Weber et al., 2019), biofuels and biomass burning (GFED-4.1s; van der Werf et al., 2017) as well as geological emissions (Etiope et al., 2019) (scaled to a global total of 15 Tg). The pre-processed fluxes with monthly resolution are temporally interpolated to the 10 d windows of our inversion system, such that we have a separate emission
270   file for every 10 d simulation. Since we optimize anthropogenic and natural emissions separately (see Sect. 2.2.2), the various natural fluxes were merged to one category. An example of the anthropogenic and natural emissions is shown in Fig. 3 for the period of 11-21 June 2018.

All CH_4 emissions are considered to be constant in time over 10-days and are emitted between 0 and 20 m above the surface.

2.2.5 A priori error covariance matrix

We apply an uncertainty of 100% for each flux in each grid cell. This takes into account that EDGAR applies its own methodologies uniformly to the whole domain and, in contrast to the European TNO inventory, does not scale the total national emissions to the officially reported values. The scaling factors corresponding to the same category but belonging to different regions are correlated depending on the great-circle distance between the centers of the grid cells assuming an exponential decay of the correlation with a length scale L=200 km (see Gaspari and Cohn, 1999; Peters et al., 2005), regardless of country borders. Between different categories we assume no correlation. Also, temporal a-priori correlations between consecutive cycles is also not applied, however, inheriting a posteriori information from the previous assimilation step with a weighting of 2/3 (see Sect. 2.2.1) has a comparable impact as applying temporal correlations.

2.2.6 Background optimization

In the reference inversion of the pseudo-observation application, only emissions but no background mole fractions are optimized. However, if background mole fractions provided at the lateral boundaries from a global model are biased, an inversion system without background-optimization will try to compensate this bias by increasing or decreasing the emissions, which ultimately leads to biased emission estimates.

To address this problem, we implemented the option to optimize background mole fractions alongside the emissions. For this purpose, 8 additional state vector parameters were introduced to optimize the background mole fractions from 8 different inflow regions where CH_4 from the global CAMS model enters our model domain (see Fig. 4). We test the capability of this approach to correct for different magnitudes and types of biases in three dedicated sensitivity experiments. In these experiments we introduce different artificial biases with magnitudes of 1%, which we think to be a realistic bias for re-analysis products, and set the uncertainty correspondingly. Background-optimization is also applied in the real-data application, where we apply an uncertainty of 0.05%, which is roughly 1 ppb CH_4. This uncertainty is smaller than what we assumed in the sensitivity simulations. We choose a smaller uncertainty in the real data application as the CAMSv19r1 product used here, which is the result of a global inversion system, proofed to have almost no biases. We apply a different state propagation model (see Sect. 2.2.1) for the 8 background parameters than for the emission parameters. We propagate the a posteriori with a weight of 100% to the next window assuming that biases are changing only slowly with time and are therefore similar in subsequent assimilation windows. This allows the system to adjust the background concentration by roughly 1 ppb every 10 days.

2.2.7 Localization

To avoid erroneous state vector updates due to spurious covariances between observations and far distant grid cells, the Kalman Gain can be modified using a method called localization. As shown by Houtekamer and Mitchell (1998), localization reduces spurious correlation and reduces the number of required ensemble members. Spatial localization with damping factors decaying exponentially with distance from, or normally distributed around, each observation is possible. An alternative is to only update state vector parameters whose correlation passes a two-tailed t-test, but this proved to be computationally too expensive for our
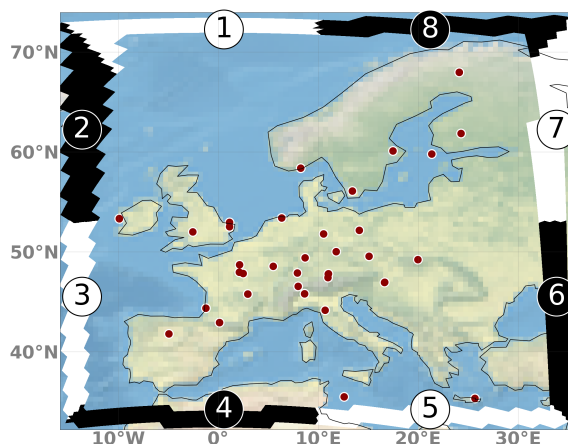
Figure 4. Map showing the 8 inflow-regions (labelled 1-8) used for the background optimization.

large problem size. To reduce correlations with far distant grid cells, in all our inversions we apply a damping to the Kalman Gain with factors normally distributed around the observation location with a 1σ standard deviation of 600 km. The optimal length scale is dependent on the setup and the number of the ensemble members. Our choice of 600 km was tested in sensitivity inversions. The effect is illustrated in Fig. 5 for one observation and one emission category. The values of the Kalman Gain matrix can be interpreted as the sensitivity of the observation to upstream emissions from that category.

2.2.8 Observations

In our application, a harmonized set of measurements including quasi-continuous in-situ observations and a few discrete flask samples, is used. The in-situ measurements are from 28 stations while the flask samples are taken at 10 different locations. The flask samples account for ca. 0.5% of the total number of observations. Another 4 in-situ stations are used for validation. Most stations are from the atmosphere network of the Integrated Carbon Observation System (ICOS) (Heiskanen et al., 2022). All 32 in-situ stations and 10 flask sample locations are listed in Table 1, where the assimilated sites are indicated in the column "T" with "x" and the validation sites with "V". It is important to note that not all time series are complete. We distinguish mountain sites from sites in flat terrain. For the sites in flat terrain, observations in the afternoon from 11 to 16 UTC were assimilated, as usually done in atmospheric inverse modelling to avoid the difficulties in simulating the effects of shallow nocturnal boundary layers. In contrast, only night-time values between 23 and 06 UTC were assimilated for mountain sites. At this time such stations are more representative of free tropospheric conditions and are not influenced by pollution from daytime up-slope valley winds, which would require a very high model resolution in order to be reasonably represented.

Pseudo-observations were synthetically generated for the first two months of the year 2018 (the synthetic inversions are performed in the period 02.01.2018 - 03.03.2018) by sampling the output of a forward ICON-ART model simulation at a
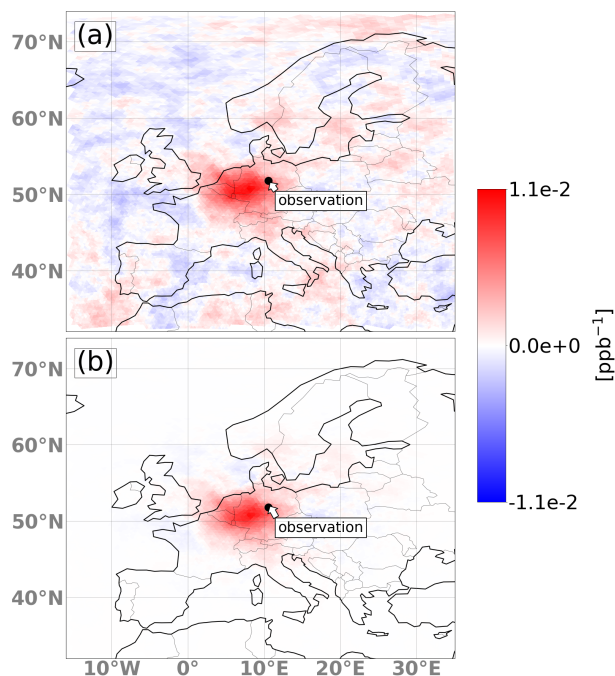
Figure 5. The values of the Kalman Gain Matrix for anthropogenic emissions for one single observation at the Torfhaus station in Germany (29 January 2018 12:00 - 16:00) before (a) and after (b) applying localization.

325   prescribed set of locations. To make the synthetic experiment comparable to an inversion with real observations, these locations correspond to actual CH_4 measurements stations that are used for the real data inversion for the year 2018. Furthermore, the pseudo-observations were generated only for the times when actual observations were available. The 4 stations that are used for validation purposes in the real application were also not assimilated in the reference inversion but the effect of assimilating their observations was tested in a separate inversion. We added gaussian noise to the pseudo-observations with σ=2 ppb.

330   The simulation used to produce the synthetic observations followed the reference setup described in Sect. 2.1.3 with one exception. Emissions from the two categories were scaled with a pre-defined field of scaling factors in order to create a "true" emission field that is systematically different from the a priori. The performance of the inversion can then be assessed in terms of its ability to reconstruct the true emission field. The field of scaling factors was randomly drawn from a normal distribution but spatially correlated assuming an exponential decay with a length scale of 200 km. In one sensitivity inversion, a different

335   field of scaling factors was used to scale the so-called "true" emissions in which 11 European regions were scaled uniformly with different scaling factors.

Table 1. List of observation locations used in this study. The column "M" indicates mountain sites while the column "F" indicates if it is a flask sampling location. The column "T" indicate if this station is assimilated ("x") or used as a validation site ("V").

ID	station name	M	F	latitude	longitude	elevation [m]	inlet height [m]	network	T
BIR	Birkenes (NO)	No	No	58.39	8.25	215	3	EBAS	x
CIB	Centro de Investigación (ES)	No	Yes	41.81	-4.93	845	5	NOAA	x
CMN	Mt Cimone (IT)	Yes	No	44.17	10.68	2165	7	WDCGG	x
GAT	Gartow (DE)	No	No	53.07	11.44	69	341	ICOS	V
HEI	Heidelberg (DE)	No	No	49.42	8.68	113	30	INGOS	x
HPB	Hohenpeissenberg (DE)	No	No	47.8	11.01	934	131	ICOS	x
HPB	Hohenpeissenberg (DE)	No	Yes	47.8	11.02	936	5	NOAA	x
HTM	Hyltemossa (SE)	No	No	56.1	13.42	115	150	ICOS	x
HUN	Hegyhatsal (HUN)	No	No	46.96	16.65	248	96	INGOS + HMS	x
HUN	Hegyhatsal (HUN)	No	Yes	46.95	16.63	248	96	NOAA	x
IPR	Ispra (IT)	No	No	45.81	8.64	210	16	ICOS	x
JFJ	Jungfraujoch (CH)	Yes	No	46.55	7.99	3570	10	WDCGG	x
KAS	Kasprowy Wierch (SVK)	Yes	No	49.23	19.98	1987	2	AGH	x
KIT	Karlsruhe (DE)	No	No	49.09	8.42	110	200	ICOS	V
KRE	Kresin u Pacova (CZE)	No	No	49.57	15.08	534	250	ICOS	x
LIN	Lindenberg (DE)	No	No	52.17	14.12	73	98	ICOS	x
LMP	Lampedusa (IT)	No	Yes	35.51	12.61	45	5	NOAA	x
LUT	Lutjewad (NL)	No	No	53.40	6.35	1	60	ICOS	x
MHD	Mace Head (IRL)	No	No	53.33	-9.90	5	10	AGAGE	x
MHD	Mace Head (IRL)	No	Yes	53.33	-9.90	5	21	NOAA	x
NOR	Norunda (SE)	No	No	60.09	17.48	46	100	ICOS	x
OPE	Observatoire pérenne (FR)	No	No	48.56	5.50	390	120	LSCE	x
ORL	Orleans (FR)	No	Yes	47.83	2.50	170	1467-1634	LSCE	x
OXK	Ochsenkopf (DE)	Yes	Yes	50.03	11.81	1022	163	NOAA	x

2.2.9 Model-data mismatch

We use a model-data mismatch of 10 ppb + 30% of the mean (over the entire inversion period) emitted signal (CH_4^E) in the forward model simulations using the prior emissions for all observations. We assume uncorrelated errors (diagonal observation-error covariance matrix $\mathbf{R}$). The reasoning for scaling the model-data mismatch with the mean instead of the instantaneous emission signal is that the latter approach would generally results in higher uncertainties for measurements with high mole fractions and, hence, those measurements that contain the largest regional signal would receive the lowest weight in the inver-

ID	station name	M	F	latitude	longitude	elevation [m]	inlet height [m]	network	T
PAL	Pallas (FIN)	No	No	67.97	24.12	560	7	WDCGG + ICOS	x
PAL	Pallas (FIN)	No	Yes	67.97	24.12	565	5	NOAA	x
PDM	Pic du Midi (FR)	Yes	No	42.94	0.14	2877	10	LSCE	x
PDM	Pic du Midi (FR)	Yes	Yes	42.94	0.14	2877	0	LSCE	x
PUY	Puy de Dome (FR)	Yes	No	45.77	2.97	1465	10	ICOS	x
PUY	Puy de Dome (FR)	Yes	Yes	45.77	2.97	1465	10	LSCE	x
RGL	Ridge Hill (UK)	No	No	52.00	-2.54	204	90	DECC	x
SAC	Saclay (FR)	No	No	48.72	2.14	160	100	WDCGG	x
SMR	Hyytiala (FIN)	No	No	61.85	24.29	181	125	ICOS	x
SNB	Sonnblick (AU)	Yes	No	47.05	12.96	3106	5	WDCGG	V
SSL	Schauinsland (DE)	No	No	47.90	7.92	1205	6	WDCGG	x
SVB	Svartberget (SE)	No	No	64.26	19.78	235	150	ICOS	V
TAC	Tacolneston (UK)	No	No	52.52	1.14	56	185	DECC	x
TOH	Torfhaus (DE)	No	No	51.81	10.54	801	147	ICOS	x
TRN	Trainou (FR)	No	No	47.96	2.11	131	180	LSCE	x
UTO	Uto (FIN)	No	No	59.78	21.37	8	57	ICOS	x
WAO	Weybourne (UK)	No	No	52.95	1.12	17	10	UEA	x
ZSF	Zugspitze (DE)	Yes	No	47.42	10.98	2667	3	WDCGG	x

sion. More importantly, mismatches between real and simulated pollution events could then lead to overestimated uncertainties where an event occurred in the model but not in the observations and underestimated uncertainties in the opposite situation.

345 However, we still want to account for the overall higher uncertainty at sites which are more strongly exposed to local emissions. The use of a model-data mismatch based on instantaneous tracer signals is tested in one sensitivity experiment.

2.3 Sensitivity experiments

Numerous choices such as the length of the assimilation window, the size of the ensemble or the a priori error correlation length scale have to be made before starting an inversion. The settings for the reference inversion were motivated by previous

350 studies but are to some extent still arbitrary. To evaluate the impact of these choices, additional sensitivity experiments were conducted by systematically varying individual settings. An overview of these experiments is presented in Table 2. We did not test a longer assimilation window of 20 instead of 10 days, which would constrain fluxes over 40 instead of 20 days, as tests showed that after 20 days of simulation usually almost no regional tracer mass is left in the model domain. This means that after 20 days no more flux signal can be "seen" by the observations. Like the reference inversion, all sensitivity inversions are

355 run for the period 02.01.2018 - 03.03.2018. The short period of 2 months is justified as the results of the performance in all synthetic inversions equilibrated after 2 assimilation windows and didn't change significantly afterwards.

– The setup of the reference inversion described in Sect. 2.2 is labeled as case 1.

– In three further inversions we test the sensitivity towards a lower and larger number of ensemble members (cases 2 and 3) as well as towards a smaller model-data mismatch uncertainty (case 4), which only corresponds to the 2 ppb

360 measurement noise added to the pseudo-observations and which represents the ideal scenario in which no transport error has to be accounted for.

– The influence of the Kalman Gain localization is investigated by not applying any localization (case 5) or increasing the σ from 600 to 1200 km (case 6).

– In case 7, a different field of true scaling factors is applied to produce the pseudo-observations, in which 11 large regions

365 are uniformly scaled (labelled as "by region" in Table 2) instead of the randomly drawn but spatially correlated true scaling factors of the reference inversion (labelled as "spatially correlated" in Table 2).

– A different state propagation model for the mean state is tested in case 8, which propagates the once optimized state vector of the previous window by 100% instead of 66%.

– Three cases (9-11) are used to test the optimization of the background mole fractions as described in Sect. 2.2.6. We

370 set the relative standard deviation for the 8 dimensionless state vector parameters to 0.01. In case 9, the 8 background tracer mole fractions are uniformly scaled by 0.99. In case 10, instead, 8 different scaling factors (randomly drawn with a standard deviation of 1%) are applied to the 8 inflow regions. Case 11 is similar to case 10 but the scaling factors for the 8 regions additionally vary with time. This is probably the most realistic experiment capturing the possibility that large-scale biases in a global model vary both geographically and with time. For this case, a second set of "true"

375 background scaling factors is selected. The true scaling factors linearly move with every 10 d window from the first set to this second set.

– Case 12 tests the sensitivity to the assimilation of 4 more stations (which are the validation sites in our real data application), 3 in central Europe and 1 in Scandinavia. These stations (Gartow (DE), Karlsruhe (DE), Svartberget (SE) and Sonnblick (AT)) are indicated in Table 1 with "V" in the column "T".

380 – In case 13 we use a model-data mismatch which is computed from instantaneous $\mathrm{CH_4^{emis}}$ values instead of the mean values over the inversion period ($10+0.3\ \mathrm{CH_4^{emis}}$ instead of $10+0.3\ \overline{\mathrm{CH_4^{emis}}}$).

– Case 14 tests the ability of the system to capture temporally varying emissions. For this case, similar to the temporal varying background scaling factors in case 11, a second set of "true" scaling factors is selected for the emissions in different categories and regions, again randomly drawn from a normal distribution with the same variances as given in

385 the a priori uncertainties. These true emissions linearly move with every 10 d window from the first set (shown in Fig.

Table 2. Overview of the synthetic simulations. In the column "mdm" the formula used for calculating the model-data mismatch is specified. The column "variation" indicates if the emissions varied temporarily in the simulation that was used to produce the pseudo-observations. The column "bg" indicates if the background mole fractions were varied in the simulation producing the pseudo-observations and optimized in the inversion. Boldface numbers or text signify changes compared to the reference inversion.

Case	members	mdm [ppb]	localization	variation	bg	perturbation	remark
1	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	None	spatially correlated	reference
2	**50**	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	None	spatially correlated	
3	**300**	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	None	spatially correlated	
4	192	**2**	σ=600 km	None	None	spatially correlated	
5	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	**None**	None	None	spatially correlated	
6	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=**1200 km**	None	None	spatially correlated	
7	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	None	**by region**	
8	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	None	spatially correlated	**full state propagation**
9	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	**yes**	spatially correlated	bg uniformly scaled
10	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	**yes**	spatially correlated	8 λ_{bg}
11	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	**yes**	spatially correlated	8 λ_{bg}(t)
12	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	None	spatially correlated	**4 additional stations**
13	192	$10 + 0.3\ \mathrm{CH}_4^{\mathrm{emis}}$	σ=600 km	None	None	spatially correlated	**mdm from instant. values**
14	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	**yes**	None	spatially correlated	
15	192	$10 + 0.3\ \overline{\mathrm{CH}_4^{\mathrm{emis}}}$	σ=600 km	None	None	spatially correlated	**3 emission categories**

6, left column) to this second set such that the second set represents the true emissions at the end of the inversion period (03.03.2018).

– Finally, case 15 tests the capability of the system to optimize the agricultural emissions separately from the remainder of the anthropogenic emissions. In this case, the emissions from agriculture represented the category with the largest domain total flux (15.5 Tg/yr), followed by non-agricultural anthropogenic emissions (14.2 Tg/yr) and natural sources (7.9 Tg/yr).

390
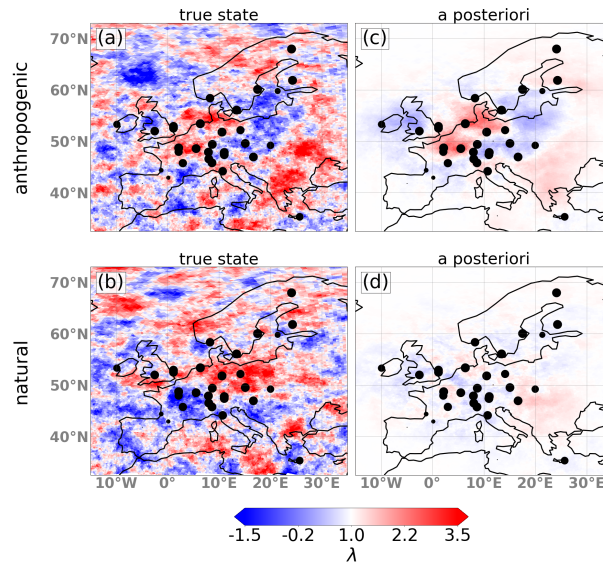
Figure 6. Map showing the "true" scaling factors (a and b) and the 2-month mean a posteriori scaling factors (c and d) for the two optimized emission categories for the reference inversion. Overlaid are the measurement locations used in the inversion with the size of the circle proportional to the number of observations available.

3 Results and discussion

3.1 Idealized Setup

3.1.1 Reference Inversion

Figure 6 shows the "true" (left) and the mean (from 12.01.2018 - 03.03.2018) a posteriori scaling factors (right) for the two optimized emission categories. It can be seen that the optimized state is close to the true state for the anthropogenic emissions in regions with high emissions (see Fig. 3), large deviations of λ from 1 in the true state and/or good coverage with observation sites, which is the case in central Europe. The a posteriori scaling factors for the natural emissions stay close to 1 and do not agree well with the true state, which means that the natural fluxes cannot be constrained independently from the larger anthropogenic fluxes. It is important to note that the performance of the optimization is strongly depending on the choices of the true state. Especially for the lower natural emissions, a different set of true scaling factors may yield different results. In this setup, the inability to reproduce the true state of the natural emissions may come from the fact that the perturbations are relatively small in areas with large emissions (e.g. in Scandinavia) and that the density of observation stations in Scandinavia is low.

Figure 7 shows the total emissions (a) as well as the a priori (b) and a posteriori (c) error of the total emissions summed over both categories compared to the true state. The improvement (prior - posterior error) in terms of total flux is shown in (d).
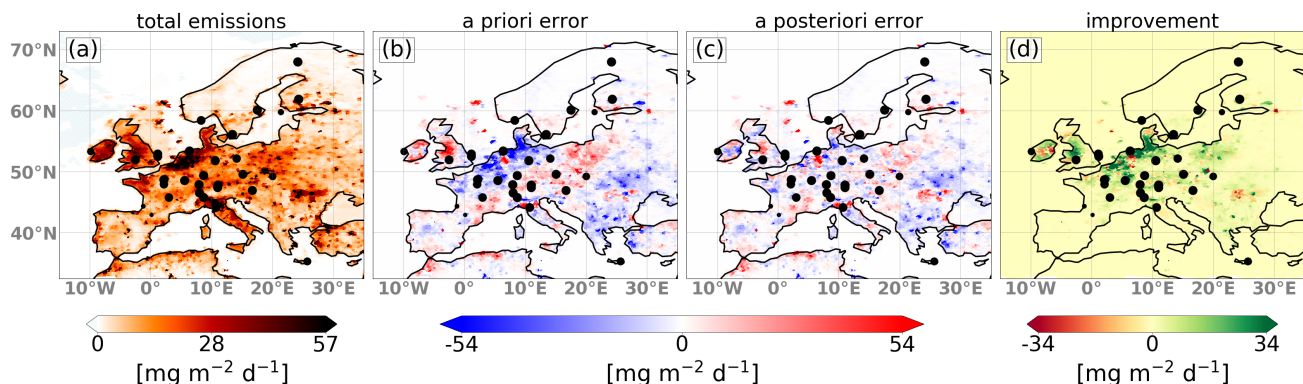
Figure 7. Maps showing the total emissions (a), the a priori (b) and a posteriori (c) deviations of the fluxes and the total improvement (d) for the reference inversion.

The results were computed from the mean fluxes for 12.01.2018 - 03.03.2018. The first optimization window (02.01.2018 - 12.01.2018) is discarded as this first window is only optimized once by the smoother (see Fig . 2). Almost every grid cell with deviations from the true state improves clearly in the a posteriori field, especially in the regions with a high station density in central Europe and on the British Isles. The mean absolute error of the2-month mean total fluxes in this setup is reduced by 18.8% in the a posteriori compared to the a priori emissions.

The system is thus capable of optimizing the total anthropogenic emissions (and, as shown in case 15, also agricultural emissions) of Europe. However, the system is not able to optimize the natural CH_4 emissions individually, which have comparatively small fluxes.

3.1.2 Sensitivity Simulations

For each of the 15 sensitivity inversions (see Table 2) we quantify the reduction of mean absolute error in the a posteriori compared to the a priori emissions expressed as a percentage. The results are summarized in Table 3. They show the error reduction for the total emissions based on the 10 d windows $(1 - E_a^e/E_b^e)$ as well as for the 2 emission categories (anth,natural) individually (for case 15 also the agricultural emission error reduction is shown). For the 3 inversions with background optimization, the reduction in the error of the 8 corresponding state vector parameters $\left(1 - E_a^{s,bg}/E_b^{s,bg}\right)$ is shown additionally. The reduction of the a posteriori mean absolute error of the mean total emissions fluxes for the considered period (i.e, of the temporally averaged fluxes), is shown in the column $(1 - E_a^{e,mean}/E_b^{e,mean})$. All results are computed for a 2-month period (12.01.2018 - 03.03.2018).

In the reference inversion (case 1), the error in the total a priori fluxes is reduced by 18.1%, while the error of the 2-month mean total fluxes is reduced by 18.8%. The anthropogenic emissions show an error reduction of 21.0%, whereas the error for the natural emissions increased by 2.7%.

The performance for the total fluxes as well as for the individual categories decreases as expected when reducing the number of ensemble members (case 2) or switching off the localization of the Kalman Gain matrix (case 5), highlighting the importance of using the loacalization techinque. Other sensitivity inversions show similar performance as the reference inversion, such as inversions with 300 instead of 192 ensemble members (case 3), which justifies the use of 192 ensemble members in our setup. Also a longer localization length scale (case 6) or adding 4 additional stations in central Europe, where the observation network is already dense (case 11), don't change the performance significantly. The use of a model-data mismatch based on instantaneous concentration values (case 13) instead of the average concentrations also doesn't change the performance. As explained in Sect. 2.2.9, we expect the use of a model-data mismatch based on mean instead of instantaneous concentrations to have advantages when assimilating real data where we have a real transport error between simulated and real tracer transport. This situation is not mimicked in this idealized setup. Propagating the state vector by 100% (case 8) improves the error reduction significantly. Such an improvement is to be expected for an idealized setup with constant emissions, perfect transport model and without any measurement noise. The same is true for case 3 with a very small model-data mismatch (2 ppb), as a lower model-data mismatch gives more weight to the unbiased pseudo-observations and hence pulls the a posteriori scaling factors closer to the "true" scaling factors. A significantly better performance (especially for the natural fluxes) is reached in case 7, where a different type of "true state" (11 scaled European regions) is used. This produces a more distinct flux signal, which is easier to optimize. Disturbing the background with a constant factor (0.98) and optimizing the background mole fractions with 8 additional state vector parameters (case 10) influences the optimization of the emissions only minimally. This is due to the fact that the error in the 8 state vector parameters representing the background mole fractions is reduced by 91.1%. Similarly, using 8 different (but still temporarily constant) scaling factors for the background mole fractions (case 10) reduces the performance of the emission optimization very little. The error in the 8 state vector parameters representing the background mole fractions is again reduced strongly by 88.2%. As for the emissions, also the performance of the background optimization is highly dependent on the chosen "true" scaling factors. Due to the dominating westerly winds, the 2 scaling factors at the western border of the domain can be constrained the best. Using 8 different scaling factors for the background mole fractions which additionally change in time (case 11), shows performance of the emission optimization similar to cases 9 and 10, but shows a worse performance in optimizing the 8 background state vector parameters (which are now time-varying). The reason for this is one single background parameter (representing the inflow region 1, see Fig. 4). As there is almost no inflow from this region and the distance from this region to the next observation stations is very large, this region cannot be constrained. When only looking at the inflow regions 2-7, the error in the scaling factors is reduced by 84.7%, which is a similar perfomance as in the cases 9 and 10. In case 14, where the scaling factors of the true state are time-dependent, i.e., where they gradually move to a second set of scaling factors, the improvement of the total flux is even slightly better than in the reference inversion, which shows that the system is also capable of capturing slowly varying emission strengths. Case 15 finally shows that agricultural emissions individually can be optimized well (17.8%) but then the error reduction of the remaining anthropogenic sources without the agricultural emissions significantly drops from 21.0% to 6.8%.

In all cases, the total error reduction is based on the error reduction of the anthropogenic emissions (or the agriculture emissions in case 15) while the error in the natural emissions almost always increase, especially for the cases that perform

Table 3. Statistical results of the sensitivity simulations for the period 02.01.2018 - 03.03.2018. $1-E^e_a/E^e_b$: the a posteriori mean absolute error of the total emissions fluxes (E^e_a) compared to the a priori situation at the beginning of the inversion (E^e_b), expressed as a percentage reduction on the prior. The statistics are computed considering every single flux component (every single grid cell and every single assimilation window). $1-E^{e,anth}_a/E^{e,anth}_b$ and $1-E^{e,natural}_a E^{e,natural}_b$ show the same as $1-E^e_a/E^e_b$, but for the 2 categories individually. $1-E^{s,bg}_a/E^{s,bg}_b$ denotes the reduction of MAE for the 8 state vector parameters optimizing the background CH_4 mole fractions. $1-E^{e,mean}_a/E^{e,mean}_b$ finally shows the same as $1-E^e_a/E^e_b$, but for the mean fluxes during the period instead of considering the individual assimilation windows. Boldface numbers signify the reference inversion.

Case	Sensitivity	$1-\frac{E^e_a}{E^e_b}$	$1-\frac{E^{e,anth}_a}{E^{e,anth}_b}$	$1-\frac{E^{e,natural}_a}{E^{e,natural}_b}$	$1-\frac{E^{e,agri}_a}{E^{e,agri}_b}$	$1-\frac{E^{s,bg}_a}{E^{s,bg}_b}$	$1-\frac{E^{e,mean}_a}{E^{e,mean}_b}$
1	**reference**	**18.1%**	**21.0%**	**-2.7%**	**-**	**-**	**18.8%**
2	50 member	14.1%	16.9%	-6.4%	-	-	14.9%
3	300 members	18.6%	21.3%	-1.8%	-	-	19.2%
4	small mdm	23.4%	25.2%	2.0%	-	-	28.0%
5	no localization	15.7%	20.0%	-10.3%	-	-	16.4%
6	localization w. L=1200 km	18.8%	22.4%	-5.7%	-	-	19.4%
7	different true state	30.8%	29.8%	16.6%	-	-	32.6%
8	full state propagation	20.2%	23.4%	-2.6%	-	-	20.8%
9	bg uniformly scaled	17.9%	21.0%	-3.6%	-	91.1%	18.5%
10	8 λ_{bg}	17.6%	20.8%	-4.0%	-	88.2%	18.3%
11	8 λ_{bg}(t)	18.0%	21.3%	-3.1%	-	49.7%	18.7%
12	4 more stations	18.5%	21.3%	-2.6%	-	-	19.1%
13	mdm from instant. values	17.6%	20.1%	-2.0%	-	-	18.1%
14	temp. varying true state	20.3%	21.1%	-1.9%	-	-	23.4%
15	3 emission categories	16.2%	6.8%	3.0%	17.8%	-	16.6%

less well in general (e.g., when only using 50 ensemble members or switching off localization). This suggests that the natural emissions cannot be estimated independently from the anthropogenic. This is different for case 7 with a different type of true state, which highlights the impact of the chosen true state on the performance of the inversions.

465   The sensitivity runs allow to draw conclusions for the application with real observations. They emphasize the significance of the localization technique for optimal performance, suggesting that a σ of 600 km is a favorable selection. Additionally, the findings indicate that 192 ensemble members are sufficient. As a "perfect transport model" without transport error is used in this idealized setup, these inversions show the upper limit of what we can expect from the application with real observations: the total fluxes and anthropogenic emissions can be well optimized, this is especially true for central and western Europe

470   with good observational coverage. However, categories with small fluxes (such as natural emissions in Central Europe) can hardly be optimized independently. It is also to be expected that the improvement in regions with few measurements (e.g.,
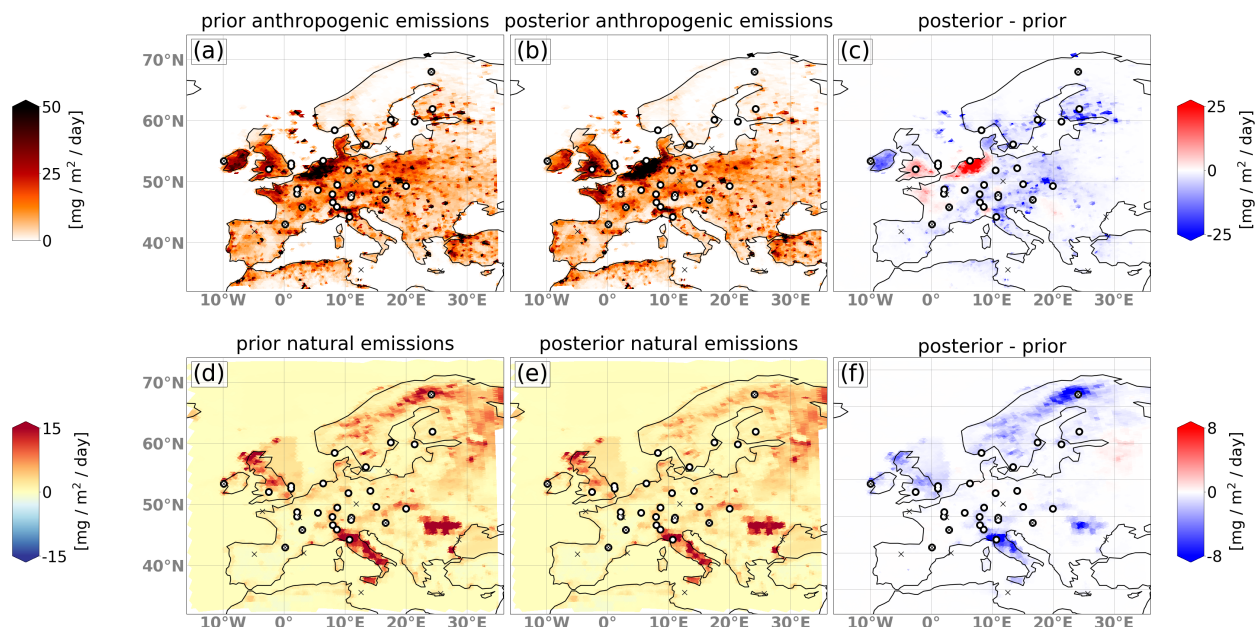
Figure 8. Map showing the mean a priori (a, d) and a posteriori (b, e) fluxes for the year 2018 as well as the difference (c, f) for the anthropogenic (a-c) and natural sources (d-f). The results for 2008 and 2013 are shown in the appendix B.

South-Western and Eastern Europe) will be very small. We further see that the system can correct slowly varying large-scale background biases of around 1% very well.

3.2 Real Data Application

3.2.1 Posterior Fluxes and Background Mole Fractions.

Figure 8 shows maps of annual mean CH_4 emissions for 2018 for the a priori (left), a posteriori (middle) and the difference between the two (right) for the anthropogenic (top) and natural (bottom) fluxes. The anthropogenic emissions show a strong upward correction of up to 25 $\mathrm{mg\,m^{-2}\,day^{-1}}$ in the Benelux countries and a moderate upward correction in northwestern France and southern England. In the rest of the domain, the anthropogenic emissions are corrected downwards, especially over Ireland, but also in the Alpine region and in southern Finland. In the annual mean, the natural fluxes are corrected downward almost everywhere, especially over Italy (up to -22%), the British Isles (-15%), Romania/Moldova (-10%) and Scandinavia (-10%).

Over the Iberian Peninsula, North Africa, and Russia, the increments are generally very low in both categories, as those regions are poorly or not at all constrained by observations. The downward correction of the natural fluxes in Italy and Romania/Moldova is most likely due to the very high a priori geological emissions in these regions. The downward correction of
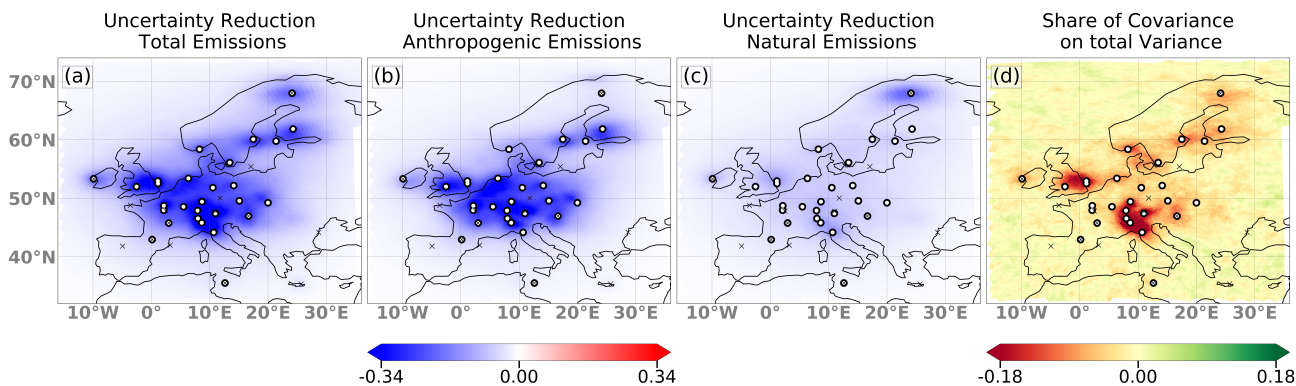
Figure 9. Maps showing the differences in the a posteriori uncertainties compared to the a priori uncertainties for the total emissions (a), anthropogenic emissions (b) and natural emissions (c). The fraction of the covariance on the total variance is shown in (d). Optimized fluxes are based on an inversion with real observations.

the Scandinavian natural fluxes is possibly due to the unusually hot summer 2018, though the study of Thompson et al. (2022) did not find a clear anomaly of CH_4 fluxes in this region in 2018. The a priori wetland and peatland emissions provided by JSBACH-HIMMELI for the year 2018 are a repetition from the year 2017 and hence do not account for the lower CH_4 emissions expected for these sources during a hot and dry summer caused by low water table levels (e.g. Bridgham et al., 2013).

490   For the years 2008 and 2013, the system corrects the natural fluxes in Northern Europe much less downward than in 2018. This can be seen in the appendix A in Fig. A1. It shows the time series of the a priori and a posteriori fluxes for the two categories and the three years 2008, 2013 and 2018 for the EU28 countries as well as for northern, western, eastern, and southern Europe. It also shows less downward correction of anthropogenic fluxes in Northern Europe for the years 2008 and 2013 which could be an effect of cross-sector covariances (see Fig. 9d).

495   Figure 9 presents maps of yearly mean uncertainty reductions (negative values mean a reduction in uncertainty) for the total emissions (a), anthropogenic emissions (b) and natural emissions (c). The fraction of the covariance on the total variance is shown in (d). The total uncertainty is computed from the variances of the two categories as well as the covariances between them, for each gridbox. Since the covariances are typically negative, the total uncertainty reduction is larger than the sum of the uncertainty reduction in the two categories. This is mainly the case in southwestern England, the Alpine region and to a

500   lesser extent also around the Scandinavian measurement locations. Negative covariances mean that in those regions the system cannot distinguish well between the two fluxes, i.e., an improvement match to observed mole fractions could be achieved by increments of either anthropogenic or natural fluxes in these regions.

As described in Sect. 2.2.6, the system not only optimizes emissions but also background mole fractions in 8 different boundary regions. The effect of this optimization on the mole fractions of the background field in the domain is shown in

505   Fig. 10, which displays the mean correction of the background field in the lowest model level for 2018, spatially, and as a time series of the domain-wide total. The yearly mean mole fraction difference shows only very small corrections between 0 and
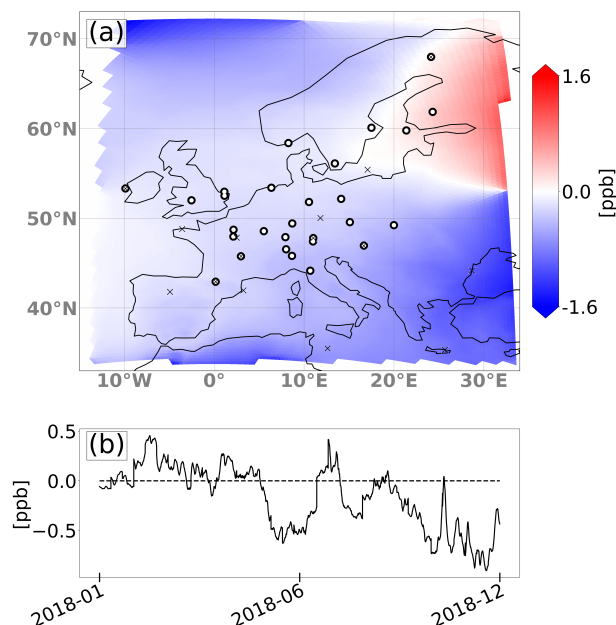
Figure 10. The difference of the a posteriori to a priori background dry air mole fractions in the lowest model layer shown as a yearly mean on a map (a) and as a time series of of the domain-wide total (b).

1.6 ppb of the background field, mainly downwards, except for the northeasterly inflow region. The time series shows that the domain-total corrections were slightly upward in spring, while the corrections were downwards during early summer and at the end of the year. The minor adjustments of the background mole fractions indicate a very small bias in the used reanalysis product CAMSv19r1, significantly smaller than what we assumed in the sensitivity simulations (cases 10-12).

3.2.2 Time Series at Measurement Sites

Figure 11 shows examples of time series for the period 30 March - 13 June 2018 at the stations Lutjewad (NL), Hyytiälä (FI) and Heidelberg (DE). Simulated dry air mole fractions (from a priori emissions in red and a posteriori emissions in blue) are compared to observations (in black). Additionally, the assimilated afternoon measurements (black dots) and the simulated background tracer dry air mole fractions are shown, in dark green for the a posteriori and light green for the a priori. The light green line is mostly overlaid by the dark green line. At all three stations, the simulation with the a posteriori emissions is closer to the observations. Furthermore, the a priori background mole fractions match the lowest measured mole fractions in this period, indicating that the bias of the background field is very low.

Table 4 summarizes the statistics of the model performance at the in-situ measurement stations assimilated in the inversion (first part of the table). It shows the root-mean squared error (E) and Pearson correlation (r^2) both for dry air mole fractions at hours where observations were assimilated (i.e., afternoon values at non-mountain sites and night-time values at mountain sites)
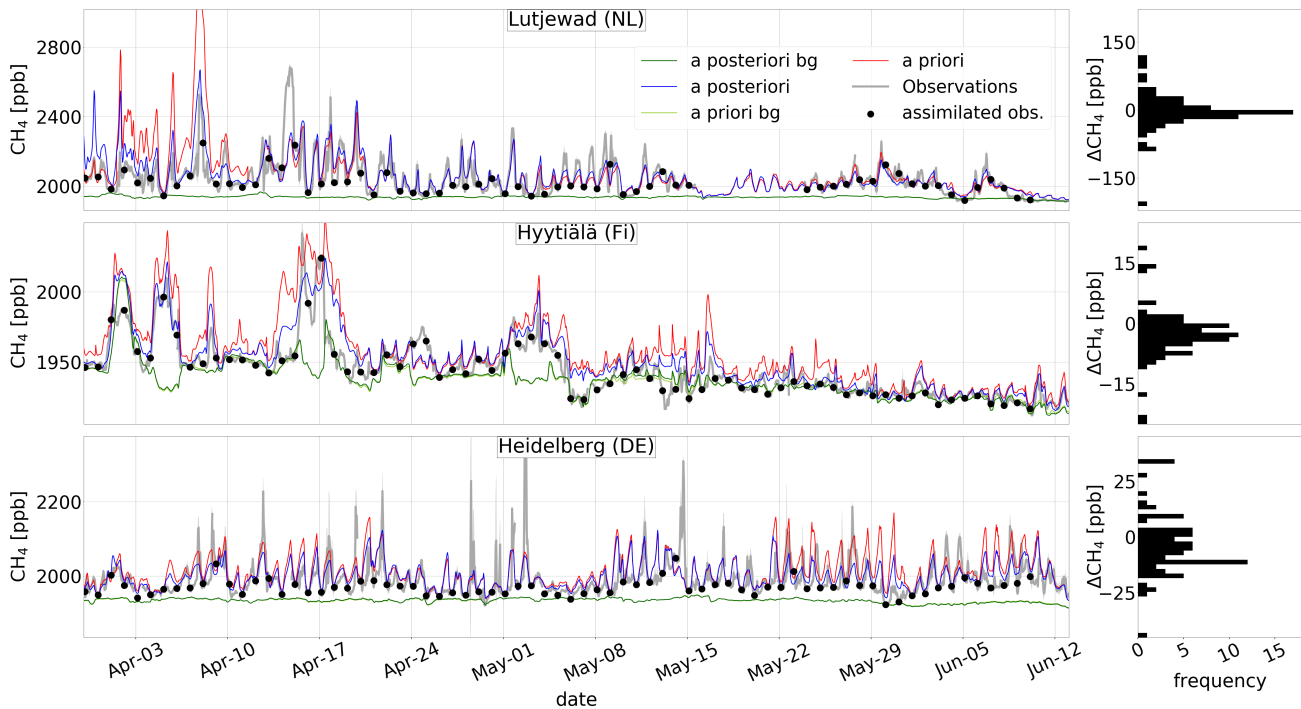
Figure 11. Time series of the simulated dry air mole fractions with a priori (red) and a posteriori (blue) emissions. Additionally, the background dry air mole fractions are indicated in dark and light green for the a posteriori and a priori, respectively, as well as the observations (black) with the reported measurement uncertainty (gray shaded, often too small to see) for the period 30 March - 13 June at the stations Lutjewad (NL), Hyytiälä (Fi) and Heidelberg (DE). The assimilated afternoon measurements are shown as black dots. The right column shows the residuals of the assimilated observations in the a posteriori simulation for the same period.

and for all 24 hours per day. For the assimilated stations, the error always decreases and the correlation always increases for the hours that were assimilated. Also, the error almost always decreases and the correlation mostly increases for the observations during the entire day. The a posteriori correlation coefficients (r^2) for the assimilated observations range from a minimum of 0.59 at the station Ispra (Italy) to 0.93 at the station Ridge Hill (UK). They are above 0.7 at 25 out of 28 stations (89%) and above 0.8 at 19 stations (68%). These values are similar to those in the high-resolution inversion study of Bergamaschi et al. (2022) and suggest an excellent model performance typically explaining 70-90% of the observed variance.

The bottom of Table 4 shows the same statistics for the validation sites, which were not assimilated in the inversion. At these validation sites, the error always decreases and the correlation always increases or stays the same for the afternoon hours (night-time hours for the mountain-site Sonnenblick (labelled with SNB)). When considering all 24 hourly averages per day, the error mostly decreases and the correlation mostly increases.

Table 4. Statistical results at measurement stations assimilated in the inversion and used for validation (bottom). The root-mean squared error (E) and Pearson correlation (r^2) are given once for dry air mole fractions only at hours where observations were assimilated (denoted with 'assim.') and for all 24 hours per day (denoted with 'all obs.').

Station	E_{prior} assim. [ppb]	E_{post} assim. [ppb]	r^2_{prior} assim.	r^2_{post} assim.	E_{prior} all obs. [ppb]	E_{post} all obs. [ppb]	r^2_{prior} all obs.	r^2_{post} all obs.
BIR	19.2	11.9	0.92	0.95	21.3	13.9	0.89	0.91
CMN	17.1	13.6	0.78	0.80	21.2	17.6	0.69	0.73
HEI	27.4	23.0	0.85	0.87	62.0	49.7	0.66	0.68
HPB	25.3	19.3	0.80	0.89	29.9	25.8	0.75	0.81
HTM	18.5	13.6	0.93	0.93	20.3	15.7	0.91	0.92
HUN	37.2	32.3	0.66	0.73	41.3	38.9	0.59	0.65
IPR	71.2	55.4	0.59	0.75	137.2	134.8	0.48	0.44
JFJ	11.4	10.6	0.85	0.87	13.7	12.7	0.79	0.82
KAS	22.1	17.4	0.65	0.74	26.6	23.3	0.57	0.63
KRE	18.6	14.7	0.87	0.91	22.7	20.2	0.81	0.83
LIN	22.3	16.1	0.89	0.94	25.1	22.2	0.86	0.87
LUT	78.8	56.1	0.75	0.86	106.5	87.5	0.76	0.79
MHD	13.6	8.7	0.82	0.88	16.8	10.4	0.76	0.84
NOR	21.3	11.0	0.90	0.93	23.9	12.2	0.87	0.91
OPE	19.3	16.2	0.88	0.91	22.1	10.5	0.85	0.86
PAL	17.6	9.4	0.79	0.91	19.5	11.7	0.73	0.86
PDM	7.5	7.4	0.95	0.95	8.39	8.37	0.92	0.92
PUY	17.2	16.5	0.78	0.82	18.5	17.9	0.75	0.78
RGL	15.3	14.2	0.93	0.94	20.1	20.9	0.88	0.87
SAC	28.7	21.7	0.86	0.91	35.2	30.3	0.83	0.84
SMR	27.1	13.6	0.86	0.91	30.0	15.1	0.83	0.90
SSL	19.4	15.6	0.78	0.83	19.6	17.9	0.76	0.78
TAC	20.8	19.3	0.90	0.91	25.0	26.7	0.85	0.84
TOH	16.5	14.3	0.82	0.85	18.3	17.8	0.78	0.79

3.2.3 National-Scale Emissions

Figure 12 compares 2018 a priori and a posteriori country total emissions for both anthropogenic (left) and natural sources (right) together with their uncertainties for the 15 largest countries in Europe and the Benelux countries. Anthropogenic emis-

Station	E_{prior} assim. [ppb]	E_{post} assim. [ppb]	r^2_{prior} assim.	r^2_{post} assim.	E_{prior} all obs. [ppb]	E_{post} all obs. [ppb]	r^2_{prior} all obs.	r^2_{post} all obs.
TRN	20.5	17.2	0.87	0.91	24.6	23.4	0.83	0.85
UTO	27.0	13.8	0.81	0.92	26.7	14.2	0.81	0.91
WAO	28.1	20.1	0.84	0.92	31.3	31.2	0.83	0.83
ZSF	15.3	12.1	0.77	0.82	16.3	13.2	0.75	0.80
GAT	16.59	16.49	0.88	0.88	21.07	20.53	0.82	0.83
KIT	25.40	20.02	0.89	0.90	30.08	27.47	0.84	0.82
SNB	15.09	12.99	0.76	0.80	16.11	13.85	0.75	0.78
SVB	17.16	11.20	0.89	0.92	17.76	11.53	0.88	0.91

sions are additionally compared to the values reported to UNFCCC (including the reported uncertainty range) for 2018 (year of report 2021). The a posteriori fluxes in the Benelux countries as well as Germany, France, Norway and Finland are higher than reported. On the other hand, the a posteriori fluxes are lower in the remaining countries. The strong downward correction of anthropogenic emissions in Italy is likely an effect of the very high natural (geological) emissions. Although they are corrected downwards strongly in the inversion, the a posteriori natural emissions in Italy still seem unrealistically high, potentially leading to a misattribution of anthropogenic and natural emissions in the a posteriori.

For the sum of all EU27 countries + United Kingdom, the inversion reduces the a priori emissions from 19.9 $\mathrm{Tg\,yr^{-1}}$ to 17.4 $\mathrm{tg\,yr^{-1}}$ in the a posteriori. The reported value for the EU27 + UK for the year 2018 is 17.8 $\mathrm{Tg\,yr^{-1}}$. We performed two additional inversions for the years 2008 and 2013. In these years, the a posteriori fluxes are also lower than the a priori, but still slightly higher than the reported values (2008: a priori: 21.8 $\mathrm{Tg\,yr^{-1}}$, a posteriori: 21.3 $\mathrm{Tg\,yr^{-1}}$, reported: 20.5 $\mathrm{Tg\,yr^{-1}}$; 2013: a priori: 20.5 $\mathrm{Tg\,yr^{-1}}$, a posteriori: 18.9 $\mathrm{Tg\,yr^{-1}}$, reported: 18.6 $\mathrm{Tg\,yr^{-1}}$).

4 Conclusions

We developed a new inverse modeling system combining the atmospheric transport model ICON-ART with the ensemble Kalman smoother data assimilation system CTDAS, and evaluated its performance in idealized and real CH_4 inversions over Europe. For this purpose, we extended ICON-ART with modules for efficient handling of emissions and online generation of the ensemble of perturbed fluxes and with a nudging scheme to keep the simulations close to analyzed meteorology. We showed that with this system, we can optimize total anthropogenic European CH_4 fluxes on a national-scale using pseudo-observations in an idealized setup using a realistic network of measurement stations. However, from the subcategories anthropogenic and natural, the observations can only successfully constrain the larger source of the anthropogenic emissions (and in one sensitivity
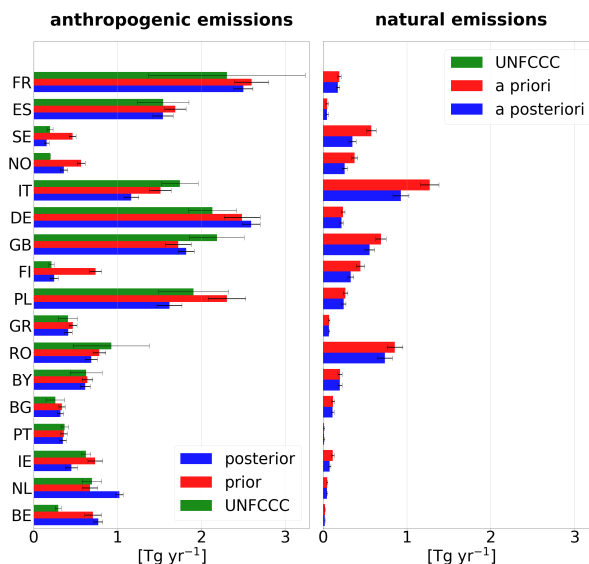
Figure 12. Country total emissions for the year 2018 in the a priori (red) and a posteriori (blue) for anthropogenic and natural sources separately as well as the reported anthropogenic emissions to UNFCCC (green, if available). The black bar indicates the total uncertainty, as derived from the error covariance matrices (for the a priori and a posteriori emissions) or as reported to UNFCCC by the countries. The results for 2008 and 2013 are shown in the appendix B.

inversion also the agricultural emissions). The natural emissions in turn, cannot be constrained independently. Furthermore, we
555 have investigated the sensitivities towards different parameters of the inversions setup with 15 sensitivity runs in the idealized setup. We then applied the system to real in-situ observations from 28 European stations. We used a priori anthropogenic fluxes from the EDGARv6 inventory and a priori natural fluxes from various sources (peatlands, mineral soils, inland water, termites, ocean, biofuels and biomass burning as well as geology).

Our results show that the anthropogenic emissions are significantly underestimated in EDGARv6 for the year 2018 in
560 the Benelux countries (by ca. 25%) and to a weaker extent in northwestern France and southern England. In the rest of the domain, the anthropogenic fluxes are corrected downwards by the inversion. The natural fluxes are corrected downwards almost everywhere, especially over Italy and Romania/Moldova where both regions have very high a priori geological emissions in the data set from Etiope et al. (2019) (scaled to a global total of 15 Tg) as well as in England and Scandinavia (during the hot and dry summer 2018). For most countries, the a posteriori country-total emissions are closer to the values that were
565 independently determined and reported to UNFCCC than the a priori emissions. The total anthropogenic fluxes for the EU27 + UK is corrected downwards from 19.9 $\mathrm{Tg\,yr^{-1}}$ to 17.4 $\mathrm{Tg\,yr^{-1}}$. The emissions reported to UNFCCC (inventory year 2021) are 17.8 $\mathrm{Tg\,yr^{-1}}$.

The a posteriori anthropogenic emissions in our study are lower than in most other regional inversions (Bergamaschi et al., 2018, 2022; Petrescu et al., 2023). It is important to emphasize that in the various inversions partially other in-situ observations

570 or satellite measurements were assimilated. While the a posteriori anthropogenic emissions in our study are lower, the pattern of the emission increments are comparable to Bergamaschi et al. (2022). However, our results for anthropogenic emissions are comparable to the results for Europe from global inversions (Deng et al., 2022; Petrescu et al., 2023), both when these assimilate in-situ measurements or satellite observations. It is noticeable that the inversion strongly corrects the overall European emissions downwards such that the independently determined emissions that are reported to the UNFCCC are better matched.

575 However, the large spread in the various inversions of recent studies shows that there is still substantial uncertainty with inverse emission estimation of European emissions and care needs to be taken to assign realistic a posteriori uncertainties. The most prominent pattern in our results is the increase over the Benelux countries. This increase of emissions is also visible in other inversion studies, at least in parts of Benelux (e.g. Bergamaschi et al., 2022). This may indicate higher emissions than reported for 2018 in this region due to agricultural emissions.

580 This overall downward correction is more pronounced for 2018 than for the other two years where we applied our system (2008 and 2013). The reason is most likely the unusually hot and dry summer 2018, which was not taken into account in the a priori peatland emissions and which influences the a posteriori anthropogenic emissions due to cross-sector covariances.

This study presents the new CTDAS-ICON system and shows its application for CH_4 in an idealized setup and with in-situ measurements. In future applications, the system could be extended to assimilate additionally satellite observations or

585 estimate N_2O or CO_2 emissions, the latter requiring an extension of ICON-ART with a biosphere-atmosphere exchange flux model. The large uncertainties associated with inverse modelling could be addressed by extending the system with a flow-dependent model-data mismatch, where the emission flux ensemble is coupled to meteorological ensemble for a more realistic representation of the model transport error.

Appendix A: Further Results 2018
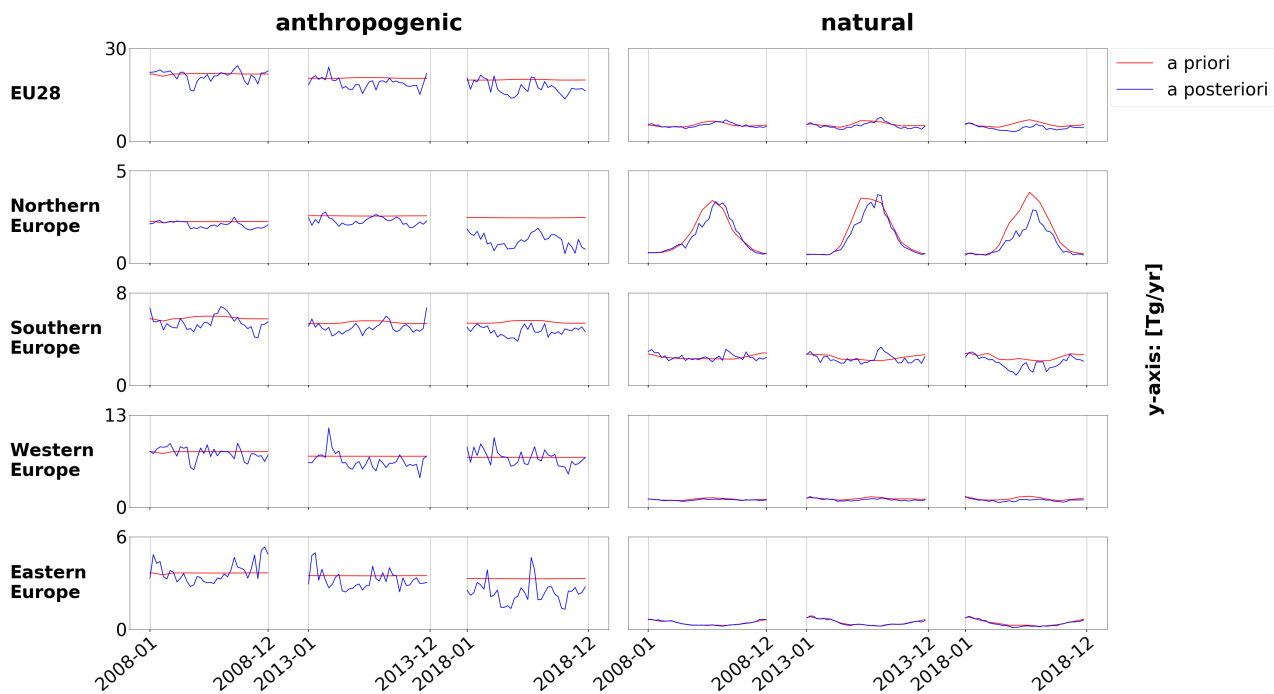
Figure A1. Time series for the years 2008, 2013 and 2018 of a priori (red) and a posteriori (blue) CH_4 emissions for the EU28 countries as well as for northern (Norway, Sweden, Finland, Denmark, Estonia, Latvia, Lithuania), western (United Kingdom, Ireland, Netherlands, Belgium, Luxembourg, France, Georgia, Switzerland, Austria), eastern (Poland, Czech Republic, Slovakia, Hungary), and southern (Portugal, Spain, Italy, Slovenia, Croatia, Greece, Romania, Bulgaria) Europe.
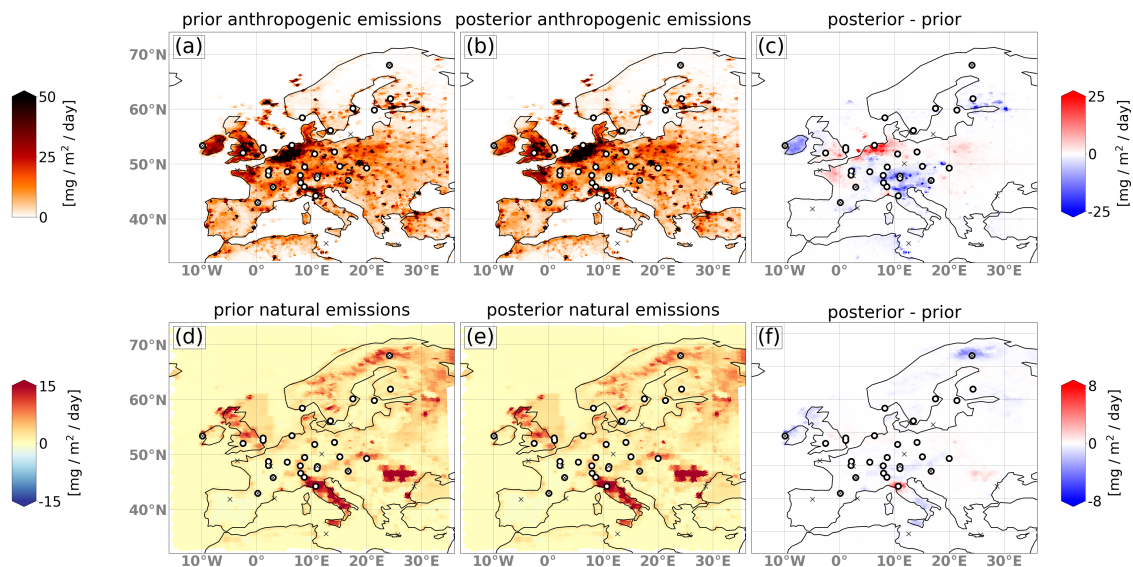
Figure B1. As Fig. 8 but for the year 2008.
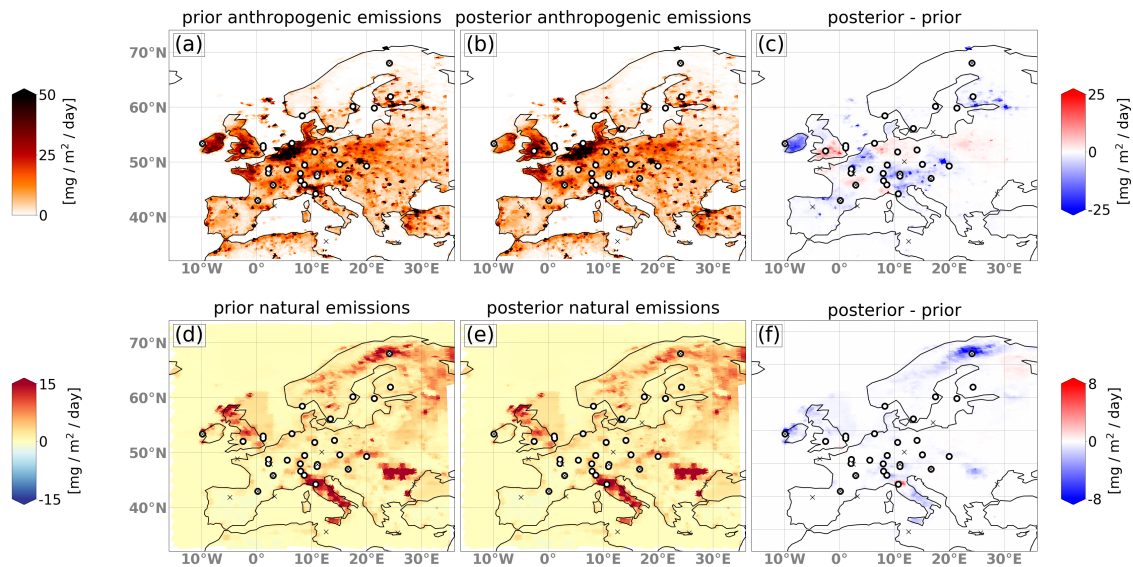
Figure B2. As Fig. 8 but for the year 2013.
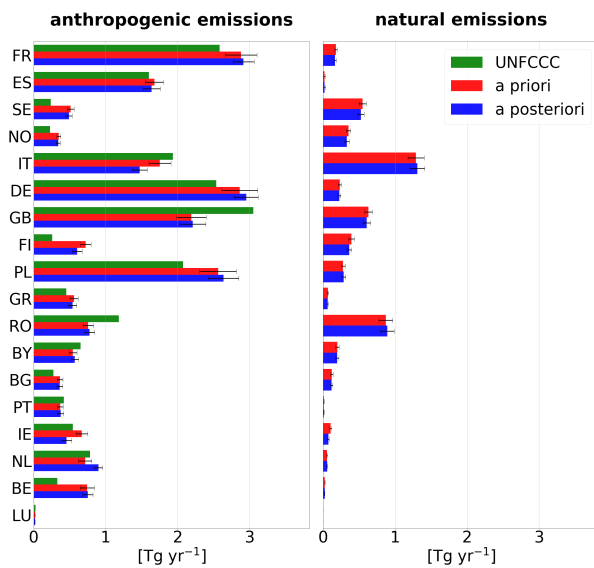
Figure B3. As Fig. 12 but for the year 2008.
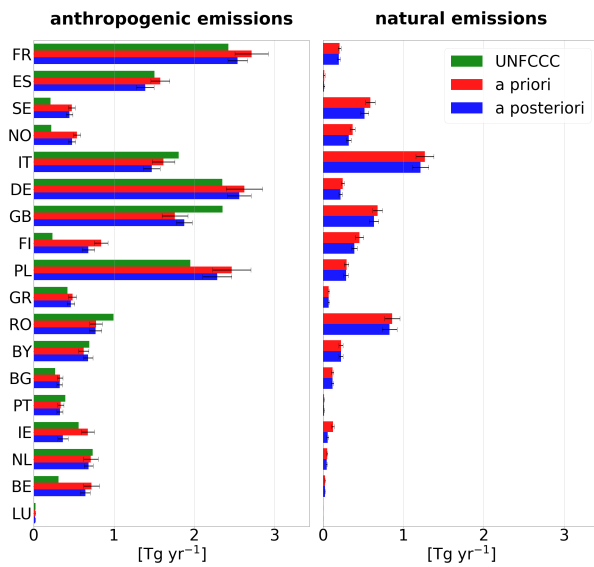
Figure B4. As Fig. 12 but for the year 2013.

Code and data availability. Access to the git repository containing the official CTDAS code can be obtained by contacting the main developers. The code version used for this study is available in a separate branch of the repository. The ICON code can be obtained from the Max Planck Institute for Meteorology (MPI-M) and the German Weather Service (DWD), who manage the license. For ART, please contact Bernhard Vogel (bernhard.vogel@kit.edu). Our implementations in the ART code are available in the official ART repository. The atmospheric observations used in the inversions are available at the ICOS carbon portal (https://www.icos-cp.eu), GAW WDCGG data center (https://gaw.kishou.go.jp), NOAA ESRL data (https://www.esrl.noaa.gov/gmd/), and EBAS data center (http://ebas.nilu.no). Information about the pre-processing applied on the observations can be requested from Rona Thompson (rlt@nilu.no).

Author contributions. MS and DB initiated the project. MS made the implementations in ICON, coupled ICON with CTDAS, performed and analyzed the simulations and inversion under the supervision of DB. MS, DB and SHe interpreted the results. WP and IL helped with the coupling of ICON to CTDAS. MS wrote the paper with substantial contribution of DB, SHe, WP and IL. HC and SHa provided the data of the measurement stations Lutjewad and Heidelberg and revised the manuscript.

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

Acknowledgements. We thank Ivan Mammarella, Petri Keronen and Pasi Kolari for providing the atmospheric observations for the Hyytiälä station. We thank the Environment Agency Austria for providing the atmospheric observations for the Sonnblick station. We thank Cathrine Lund Myhre, Paolo Cristofanelli, Dagmar Kubistin, Jennifer Müller-Williams, Matthias Lindauer, Tobias Biermann, László Haszpra, Giovanni Manca, Lukas Emmenegger, Lukasz Chmura, Michal Marek, Morgan Lopez, Irene Lehner, Michel Ramonet, Juha Hatakka, Marc Delmotte, François Gheusi, Aurélie Colomb, Simon O'Doherty, Dickon Young, Joe Pitt, Kieran Stanley, Martina Schmidt, Paul Smith, Juha Hatakka, Grant Forster and Cédric Couret for providing the atmospheric observations from the other stations. The Heidelberg measurements were financed as part of the ICOS CRL pilot stations. We thank Lisbeth Florentie and Sander Houweling for the coordination of the intercomparison exercise within TRANSCOM and setting up the protocol. TRANSCOM received support from the EU-funded CoCO2 project. Furthermore we thank Rona Thompson for the harmonization and pre-processing of the atmospheric observation within the VERIFY project. The Center for Climate Systems Modeling (C2SM) at ETH Zurich is acknowledged for providing technical and scientific support. ICON-ART simulations were performed at the Swiss National Supercomputing Cetre (CSCS) under grants s1152 and s1091.

References

615 Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S., and Zhu, Z.: TransCom 3 inversion inter-comparison: Impact of transport model errors on the interannual variability of regional CO2 fluxes, 1988–2003, Global Biogeochemical Cycles, 20, https://doi.org/https://doi.org/10.1029/2004GB002439, 2006.

Baker, D. F., Bösch, H., Doney, S. C., O'Brien, D., and Schimel, D. S.: Carbon source/sink information provided by column CO_2 measure-
620 ments from the Orbiting Carbon Observatory, Atmospheric Chemistry and Physics, 10, 4145–4165, https://doi.org/10.5194/acp-10-4145-2010, 2010.

Bergamaschi, P., Danila, A., Weiss, R., Ciais, P., Thompson, R., Brunner, D., Levin, I., Meijer, Y., Chevallier, F., Janssens-Maenhout, G., Bovensmann, H., Crisp, D., Basu, S., Dlugokencky, E., Engelen, R., Gerbig, C., Günther, D., Hammer, S., Henne, S., and Krol, M.: Atmospheric monitoring and inverse modelling for verification of greenhouse gas inventories, https://doi.org/10.2760/759928, 2018.

625 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 novel FLEXPART-COSMO TM5 4DVAR inverse modelling system, Atmospheric Chemistry and Physics Discussions, 2022, 1–37, https://doi.org/10.5194/acp-2022-118, 2022.

630 Brasseur, G. P. and Jacob, D. J.: Modeling of Atmospheric Chemistry, Cambridge University Press, https://doi.org/10.1017/9781316544754, 2017.

Bridgham, S. D., Cadillo-Quiroz, H., Keller, J. K., and Zhuang, Q.: Methane emissions from wetlands: biogeochemical, microbial, and mod-eling perspectives from local to global scales, Global Change Biology, 19, 1325–1346, https://doi.org/https://doi.org/10.1111/gcb.12131, 2013.

635 Bruhwiler, L., Dlugokencky, E., Masarie, K., Ishizawa, M., Andrews, A., Miller, J., Sweeney, C., Tans, P., and Worthy, D.: CarbonTracker-CH_4: an assimilation system for estimating emissions of atmospheric methane, Atmospheric Chemistry and Physics, 14, 8269–8293, https://doi.org/10.5194/acp-14-8269-2014, 2014.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P.: An improved Kalman Smoother for atmospheric inversions, Atmospheric Chemistry and Physics, 5, 2691–2702, https://doi.org/10.5194/acp-5-2691-2005, 2005.

640 Brunner, D., Henne, S., Keller, C. A., Reimann, S., Vollmer, M. K., O'Doherty, S., and Maione, M.: An extended Kalman-filter for regional scale inverse emission estimation, Atmospheric Chemistry and Physics, 12, 3455–3478, https://doi.org/10.5194/acp-12-3455-2012, 2012.

Burgers, G., van Leeuwen, P. J., and Evensen, G.: Analysis Scheme in the Ensemble Kalman Filter, Monthly Weather Review, 126, 1719 – 1724, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2, 1998.

Butler, M. P., DAVIS, K. J., DENNING, A. S., and KAWA, S. R.: Using continental observations in global atmospheric inversions of CO2:
645 North American carbon sources and sinks, Tellus B, 62, 550–572, https://doi.org/https://doi.org/10.1111/j.1600-0889.2010.00501.x, 2010.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/https://doi.org/10.1029/2005JD006390, 2005.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Lo Vullo, E., Solazzo, E., Monforti-Ferrario, F., Olivier, J., and Vignati, E.:

650   EDGAR v6.0 Greenhouse Gas Emissions., European Commission, Joint Research Centre (JRC) [Dataset], http://data.europa.eu/89h/
       97a67d67-c62e-4826-b873-9d972c4f670b, 2021.

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.,

655   and Chevallier, F.: Comparing national greenhouse gas budgets reported in UNFCCC inventories against atmospheric inversions, Earth
       System Science Data, 14, 1639–1675, https://doi.org/10.5194/essd-14-1639-2022, 2022.

Eggleston, H. S., Buendia, L., Miwa, K., Ngara, T., and Tanabe, K.: 2006 IPCC Guidelines for National Greenhouse Gas Inventories, 2006.

Etiope, G., Ciotoli, G., Schwietzke, S., and Schoell, M.: Gridded maps of geological methane emissions and their isotopic signature, Earth
       System Science Data, 11, 1–22, https://doi.org/10.5194/essd-11-1-2019, 2019.

660   Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics,
       Journal of Geophysical Research: Oceans, 99, 10 143–10 162, https://doi.org/https://doi.org/10.1029/94JC00572, 1994.

Evensen, G.: Geir Evensen The Ensemble Kalman Filter : theoretical formulation and practical implementation, vol. 53, pp. 1616–7228,
       https://doi.org/https://doi.org/10.1007/s10236-003-0036-9, 2003.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological

665   Society, 125, 723–757, https://doi.org/https://doi.org/10.1002/qj.49712555417, 1999.

Giorgetta, M. A., Brokopf, R., Crueger, T., Esch, M., Fiedler, S., Helmert, J., Hohenegger, C., Kornblueh, L., Köhler, M., Manzini, E.,
       Mauritsen, T., Nam, C., Raddatz, T., Rast, S., Reinert, D., Sakradzija, M., Schmidt, H., Schneck, R., Schnur, R., Silvers, L., Wan, H.,
       Zängl, G., and Stevens, B.: ICON-A, the Atmosphere Component of the ICON Earth System Model: I. Model Description, Journal of
       Advances in Modeling Earth Systems, 10, 1613–1637, https://doi.org/https://doi.org/10.1029/2017MS001242, 2018.

670   Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J.,
       Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO_2 exchange: inter-comparison of
       modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, https://doi.org/10.5194/bg-9-457-
       2012, 2012.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I., Gloor,

675   M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B., Randerson, J., Sarmiento,
       J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: Towards robust regional estimates of CO2 sources and sinks using atmospheric transport
       models, Nature, 415, 626–630, https://doi.org/10.1038/415626a, 2002.

He, W., van der Velde, I. R., Andrews, A. E., Sweeney, C., Miller, J., Tans, P., van der Laan-Luijkx, I. T., Nehrkorn, T., Mountain, M., Ju, W.,
       Peters, W., and Chen, H.: CTDAS-Lagrange v1.0: a high-resolution data assimilation system for regional carbon dioxide observations,

680   Geoscientific Model Development, 11, 3515–3536, https://doi.org/10.5194/gmd-11-3515-2018, 2018.

Heiskanen, J., Brümmer, C., Buchmann, N., Calfapietra, C., Chen, H., Gielen, B., Gkritzalis, T., Hammer, S., Hartman, S., Herbst, M.,
       Janssens, I. A., Jordan, A., Juurola, E., Karstens, U., Kasurinen, V., Kruijt, B., Lankreijer, H., Levin, I., Linderson, M.-L., Loustau, D., Mer-
       bold, L., Myhre, C. L., Papale, D., Pavelka, M., Pilegaard, K., Ramonet, M., Rebmann, C., Rinne, J., Rivier, L., Saltikoff, E., Sanders, R.,
       Steinbacher, M., Steinhoff, T., Watson, A., Vermeulen, A. T., Vesala, T., Vítková, G., and Kutsch, W.: The Integrated Carbon Observation

685   System in Europe, Bulletin of the American Meteorological Society, 103, E855 – E872, https://doi.org/https://doi.org/10.1175/BAMS-D-
       19-0364.1, 2022.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J.,
Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Houtekamer, P. L. and Mitchell, H. L.: Data Assimilation Using an Ensemble Kalman Filter Technique, Monthly Weather Review, 126, 796 – 811, https://doi.org/https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2, 1998.

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, Geoscientific Model Development, 13, 2379–2392, https://doi.org/10.5194/gmd-13-2379-2020, 2020.

Kalnay, E.: Ensemble Kalman Filter: Current Status and Potential, pp. 69–92, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-74703-1_4, 2010.

Lauritzen, P. H., Erath, C., and Mittal, R.: On simplifying 'incremental remap'–based transport schemes, Journal of Computational Physics, 230, 7957–7963, https://doi.org/https://doi.org/10.1016/j.jcp.2011.06.030, 2011.

Leuenberger, D., Koller, M., Fuhrer, O., and Schär, C.: A Generalization of the SLEVE Vertical Coordinate, Monthly Weather Review - MON WEATHER REV, 138, 3683–3689, https://doi.org/10.1175/2010MWR3307.1, 2010.

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, Atmospheric Chemistry and Physics, 8, 6341–6353, https://doi.org/10.5194/acp-8-6341-2008, 2008.

Michalak, A. M., Bruhwiler, L., and Tans, P. P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/https://doi.org/10.1029/2003JD004422, 2004.

Miura, H.: An Upwind-Biased Conservative Advection Scheme for Spherical Hexagonal–Pentagonal Grids, Monthly Weather Review, 135, 4038 – 4044, https://doi.org/10.1175/2007MWR2101.1, 2007.

Nisbet, E. and Weiss, R.: Top-Down Versus Bottom-Up, Science, 328, 1241–1243, https://doi.org/10.1126/science.1189936, 2010.

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO2 surface fluxes from atmospheric trace gas observations, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/https://doi.org/10.1029/2005JD006157, 2005.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proceedings of the National Academy of Sciences, 104, 18 925–18 930, https://doi.org/10.1073/pnas.0708986104, 2007.

Peters, W., Krol, M. C., Van Der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., Di Sarra, A. G., Picentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Ström, J., Haszpra, L., Meijer, H. A. J., Var Der Laan, S., Neubert, R. E. M., Jordan, A., Rodó, X., Morguí, J.-A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M., and Tans, P. P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, Global Change Biology, 16, 1317–1337, https://doi.org/https://doi.org/10.1111/j.1365-2486.2009.02078.x, 2010.

725 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.,
730 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 System Science Data, 15, 1197–1268, https://doi.org/10.5194/essd-15-1197-2023, 2023.

Pham, T. V., Steger, C., Rockel, B., Keuler, K., Kirchner, I., Mertens, M., Rieger, D., Zängl, G., and Früh, B.: ICON in Climate Limited-area Mode (ICON release version 2.6.1): a new regional climate model, Geoscientific Model Development, 14, 985–1005,
735 https://doi.org/10.5194/gmd-14-985-2021, 2021.

Raivonen, M., Smolander, S., Backman, L., Susiluoto, J., Aalto, T., Markkanen, T., Mäkelä, J., Rinne, J., Peltola, O., Aurela, M., Lohila, A., Tomasic, M., Li, X., Larmola, T., Juutinen, S., Tuittila, E.-S., Heimann, M., Sevanto, S., Kleinen, T., Brovkin, V., and Vesala, T.: HIMMELI v1.0: HelsinkI Model of MEthane buiLd-up and emIssion for peatlands, Geoscientific Model Development, 10, 4665–4691, https://doi.org/10.5194/gmd-10-4665-2017, 2017.

740 Reick, C. H., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, Journal of Advances in Modeling Earth Systems, 5, 459–482, https://doi.org/https://doi.org/10.1002/jame.20022, 2013.

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, Geoscientific Model Development, 8, 1659–1676, https://doi.org/10.5194/gmd-8-1659-2015, 2015.

745 Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton,
750 J. R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000–2017, Earth System Science Data, 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020,
755 2020.

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, Geoscientific Model Development, 11, 4043–4068, https://doi.org/10.5194/gmd-11-4043-2018, 2018.

760 Smith, N. E., Kooijmans, L. M. J., Koren, G., van Schaik, E., van der Woude, A. M., Wanders, N., Ramonet, M., Xueref-Remy, I., Siebicke, L., Manca, G., Brümmer, C., Baker, I. T., Haynes, K. D., Luijkx, I. T., and Peters, W.: Spring enhancement and summer reduction in

carbon uptake during the 2018 drought in northwestern Europe, Philosophical Transactions of the Royal Society B: Biological Sciences, 375, 20190 509, https://doi.org/10.1098/rstb.2019.0509, 2020.

Thompson, R. L., Groot Zwaaftink, C. D., Brunner, D., Tsuruta, A., Aalto, T., Raivonen, M., Crippa, M., Solazzo, E., Guizzardi, D., Regnier, P., and Maisonnier, M.: Effects of extreme meteorological conditions in 2018 on European methane emissions estimated using atmospheric inversions, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 380, 20200 443, https://doi.org/10.1098/rsta.2020.0443, 2022.

Tsuruta, A., Aalto, T., Backman, L., Hakkarainen, J., van der Laan-Luijkx, I. T., Krol, M. C., Spahni, R., Houweling, S., Laine, M., Dlugokencky, E., Gomez-Pelaez, A. J., van der Schoot, M., Langenfelds, R., Ellul, R., Arduini, J., Apadula, F., Gerbig, C., Feist, D. G., Kivi, R., Yoshida, Y., and Peters, W.: Global methane emission estimates for 2000–2012 from CarbonTracker Europe-CH$_4$ v1.0, Geoscientific Model Development, 10, 1261–1289, https://doi.org/10.5194/gmd-10-1261-2017, 2017.

Tsuruta, A., Aalto, T., Backman, L., Krol, M. C., Peters, W., Lienert, S., Joos, F., Miller, P. A., Zhang, W., Laurila, T., Hatakka, J., Leskinen, A., Lehtinen, K. E. J., Peltola, O., Vesala, T., Levula, J., Dlugokencky, E., Heimann, M., Kozlova, E., Aurela, M., Lohila, A., Kauhaniemi, M., and Gomez-Pelaez, A. J.: Methane budget estimates in Finland from the CarbonTracker Europe-CH4 data assimilation system, Tellus B: Chemical and Physical Meteorology, 71, 1565 030, https://doi.org/10.1080/16000889.2018.1565030, 2019.

van der Laan-Luijkx, I. T., van der Velde, I. R., Krol, M. C., Gatti, L. V., Domingues, L. G., Correia, C. S. C., Miller, J. B., Gloor, M., van Leeuwen, T. T., Kaiser, J. W., Wiedinmyer, C., Basu, S., Clerbaux, C., and Peters, W.: Response of the Amazon carbon balance to the 2010 drought derived with CarbonTracker South America, Global Biogeochemical Cycles, 29, 1092–1108, https://doi.org/https://doi.org/10.1002/2014GB005082, 2015.

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, Geoscientific Model Development, 10, 2785–2800, https://doi.org/10.5194/gmd-10-2785-2017, 2017.

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth System Science Data, 9, 697–720, https://doi.org/10.5194/essd-9-697-2017, 2017.

Wan, H., Giorgetta, M. A., Zängl, G., Restelli, M., Majewski, D., Bonaventura, L., Fröhlich, K., Reinert, D., Rípodas, P., Kornblueh, L., and Förstner, J.: The ICON-1.2 hydrostatic atmospheric dynamical core on triangular grids – Part 1: Formulation and performance of the baseline version, Geoscientific Model Development, 6, 735–763, https://doi.org/10.5194/gmd-6-735-2013, 2013.

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

Weimer, M., Schröter, J., Eckstein, J., Deetz, K., Neumaier, M., Fischbeck, G., Hu, L., Millet, D. B., Rieger, D., Vogel, H., Vogel, B., Reddmann, T., Kirner, O., Ruhnke, R., and Braesicke, P.: An emission module for ICON-ART 2.0: implementation and simulations of acetone, Geoscientific Model Development, 10, 2471–2494, https://doi.org/10.5194/gmd-10-2471-2017, 2017.

Whitaker, J. S. and Hamill, T. M.: Ensemble Data Assimilation without Perturbed Observations, Monthly Weather Review, 130, 1913 – 1924, https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2, 2002.

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, Quarterly Journal of the Royal Meteorological Society, 141, 563–579, https://doi.org/https://doi.org/10.1002/qj.2378, 2015.