Final Data Management Plan

DELIVERABLE 6.4

Author(s):	Rona Thompson
Date:	22-12-2025
Version:	1
Responsible Partner:	NILU
Deliverable due Date:	31-12-2025
Dissemination Level:	Public

Call:	HORIZON-CL5-2022-D1-02
Topic:	Climate Sciences and Responses
Project Type:	Research and Innovation Action
Lead Beneficiary:	NILU - Norsk Institutt for Luftforskning

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Document History

Version	Date	Comment	Modifications made by
0.1	09-12-2025	First draft of update	Rona Thompson (NILU)
0.2	19-12-2025	Internal review	Wilfried Winiwarter (IIASA)
1.0	22-12-2025	Final version	Rona Thompson (NILU)

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Summary

The Data Management Plan (DMP) is a key element for the overall management of the EYE-CLIMA project. It describes the data management procedures from the collection of input data, processing, and the data to be generated by EYE-CLIMA. This document describes how the EYE-CLIMA consortium intends to make their research data findable, accessible, interoperable and reusable (FAIR), to ensure it is soundly managed. This DMP describes the types of data that will be used and generated during the course of the project, how the data will be shared within the consortium, and how this data will be made openly available through a repository. This DMP also includes a description of the data policy and the data standards (including meta-data) that will be used. In this update to the DMP, considerations of the data formats (in order to facilitate easy inter-comparisons with other EU projects in the same field, i.e., PARIS and AVENGERS) are also included.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

TABLE OF CONTENTS

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

1. Introduction

The primary objective of the EYE-CLIMA project is to support the need for independent verification of national greenhouse gas inventories (NGHGIs) by developing top-down methods based on atmospheric inversion to a level of readiness where they can be used to determine emissions at national and sub-national scales and be incorporated into NGHGIs. As such, EYE-CLIMA partners will be preparing flux estimates using both process-based and data-driven models, analysing satellite observations of the land biosphere and atmospheric composition, running atmospheric transport models and estimating fluxes based on atmospheric inversions.

These activities will support the following specific project objectives:

1. Further develop the atmospheric inversion methodology to the point where it can be used to verify and support NGHGIs. This includes: i) improving the spatial resolution to a level where national emissions can be accurately estimated and details such as hotspots and sub-national distributions can be resolved, ii) developing methods to ensure the consistency of national, regional and global atmospheric inversions, and iii) improving quantification of uncertainty, and developing methods to identify and minimize systematic errors.
2. Address the need for a transparent and repeatable verification method for estimates of climate forcer emissions, especially at national scale, by establishing guidelines on how to perform atmospheric inversions and evaluate the results.
3. Better understand the needs of NGHGI compilers for verification and improvement of NGHGIs by working together with them in pilot projects. And through this collaboration, prepare guidelines on how to use top-down data products to verify inventories and to make these data products available to NGHGI compilers.
4. Further develop methodologies to detect and quantify large emissions of CH_4 using high resolution satellite data and integrate these emission estimates into national inventories.
5. Improve estimates of CO_2, CH_4 and N_2O land biosphere fluxes, and emissions and removals of CO_2 in the LULUCF sector, using combined atmospheric CO_2, CH_4 and N_2O observations, advanced land surface modelling and remote sensing data. This will allow emissions (e.g. harvest, fires) to be separated from removals at sub-national scale, as well as the impact of management versus environmental changes on the GHG balance of the EU.
6. Support the implementation and monitoring of EU policy through providing accurate estimates of GHG fluxes and BC emissions at national level over multiple years.
7. Disseminate the developments in methodology, the atmospheric inversion benchmark, and data products to a large group of users, in particular, NGHGI compilers, national and EU decision makers and researchers in the field of GHG emissions accounting.

2. Data Summary

2.1 Description of data to be used

The data to be used in EYE-CLIMA can be broadly grouped into the following categories:

1. Remote sensing data of the land biosphere and atmosphere
2. Meteorological data
3. Ground-based atmospheric and flux observations
4. Emission inventories and activity data

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

A detailed description for each of these data categories is provided below.

2.1.1 Remote sensing data

Remote sensing data of the land biosphere will be used to prepare maps of above ground biomass (AGB) and its changes to support data-driven modelling of CO_2 emissions and removals in the land-use land-use-change and forestry (LULUCF) sector. Satellite retrievals of total column CH_4 (XCH_4) will be used to detect and estimate large emission sources, which will be incorporated into emission inventories. In addition, these will be used in regional atmospheric inversions to determine fluxes. Retrievals of CO_2 total columns (XCO_2) will be used in atmospheric inversions to estimate CO_2 land-biosphere fluxes as a complimentary method to determine emissions and removals in the LULUCF sector. The specific remote sensing data sources that will be used are described in Table 1.

Table 1. Summary of remote sensing data (and derived data products) to be used.

Variable	**Instrument(s)/Product(s)**	**Source(s)**	**Comment**
XCO_2	OCO-2	NASA	This includes all required meta data and quality flags
XCH_4	GOSAT, GOSAT-2, TROPOMI Level 2 Data	JAXA, ESA CCI, Copernicus C3S, CAMS, SRON, IUP	This includes all required meta data and quality flags
XCH_4	EMIT, PRISMA, EnMAP, S2 hyper- and multi-spectral Level 1 Data	Mission Data portals	This includes all required meta data
Biomass and biomass change	SMOS, Sentinel-1, Sentinel-2, ALOS PALSAR	NASA, ESA CCI Biomass, University of Maryland	
Landcover and plant functional type	Copernicus LC or HILDA+	Copernicus, ESA world cover	
Above-ground biomass	SMOS L-VOD	ESA	
Tree height and forest biomass	High-resolution imagery	planet.com	High resolution images in four multispectral RGB and NIR bands from Dove satellites
Wetland extent	BAWLD, WAD2M	University of Alberta (from Arctic Data Center), University of Maryland (from Zenodo)	

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

2.1.2 Meteorological data

Meteorological data will be needed to drive the atmospheric transport models used in the inversions, and as input to process-based models of the land biosphere, which will be used to estimate land-biosphere fluxes of CO_2, CH_4 and N_2O. The datasets used are described in Table 2.

Table 2. Summary of meteorological data to be used.

Data product	Source	Comment
ERA5 meteorological reanalysis	ECMWF	Used to drive atmospheric transport models
ERA5-Land at 0.1° resolution and 3-hourly time steps	ECMWF	Bias correction using CRU. Variables needed: air temperature, precipitation, short and long wave radiation, wind speed, specific humidity, air pressure

2.1.3 Ground-based atmospheric observations

Observations of atmospheric mixing ratios (for CO_2, CH_4, N_2O, and the selected F-gases) and concentrations (for BC) from ground-based sites will be used in inversions. For CO_2 and CH_4, these will be in addition to remote sensing data (see Table 1). Ground-based observations will be used from established networks with an open access data policy. Networks and data products are chosen based on the data quality determined by the use of established measurement protocols and calibration scales. In addition, openly available flux datasets from open access publications will be used. The data are summarized in Table 3.

Table 3. Summary of ground-based atmospheric observations to be used.

Variable	Description	Source	Comment
CO_2	In-situ and flask measurements	ICOS, NOAA and ICOS ObsPacks, JR-STATION	ObsPack data will be used prior to start of ICOS data
CH_4	In-situ and flask measurements	ICOS, NOAA and ICOS ObsPacks, JR-STATION	ObsPack data will be used prior to start of ICOS data
N_2O	In-situ and flask measurements	ICOS, NOAA, WDCGG	Pre-ICOS data will be compiled and harmonized
F-gases	In-situ and flask measurements	AGAGE, WDCGG, NOAA, NASA, IAGOS-CARIBIC	
BC	Filter absorption photometer data (aethalometer) including absorption coefficient, and other meta data	ACTRIS, EMEP, FAN, and other publicly available on ebas.nilu.no	For conversion to equivalent black carbon (eBC) after source apportionment
N_2O flux	Eddy-covariance and chamber flux measurements at European sites	ETHZ, Fluxnet, and other publicly available sources based on literature	VERIFY data will be used prior to start of ICOS data

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

2.1.4 Emission inventory and other flux data

Inventories of anthropogenic emissions of CO_2 (specifically fossil fuel, bio-fuels, biomass burning and cement production), CH_4, and N_2O will be used in combination with the process-based model estimates as prior information in the inversions. A range of estimates from different datasets will be used to gauge the uncertainty in these estimates, in particular, the uncertainty per source sector, which will be propagated into the uncertainty per sector for the top-down emission estimates from the EYE-CLIMA project. Activity data on management (e.g. crops, harvest, fertilizer and manure application) and data on N-deposition will be used from various data sources. In addition, inland water and ocean flux estimates will be used as prior information for the inversions but will not be prepared in the framework of EYE-CLIMA.

Table 4. Summary of emission data, activity data, and other related data to be used.

Data product	Source	Comment
Crops, harvest, fertilizer use	FAO	
Agricultural activity data	CAPRI	Crop and livestock production, changes in land use etc
Land use data	LUH2	Harmonized set of land-use scenarios, agricultural management data
Crop sector inputs	ISIMIP3	Crop sector data from the Intersectoral impact model intercomparison project
N-deposition	JRC	
EDGAR	JRC	Anthropogenic emissions of greenhouse gases
TNO	TNO	Anthropogenic emissions of greenhouse gases
CEDS	CMIP6	Historical anthropogenic emissions data
Ocean fluxes of CO_2, CH_4 and N_2O	GCP	Estimates from recent GCP global budgets using ocean biogeochemistry models
Fire emissions	GFAS, GFED	GFAS from CAMS
Inland waters	VERIFY project	Lake and river GHG fluxes prepared in VERIFY by P. Regnier and his group
Geological seeps	Etiope et al. Earth Sys Sci Data (2019) https://doi.org/10.5194/essd-11-1-2019, 2019.	Annual climatological estimate for onshore seeps
Termites	CAMS	Climatological estimate for global termite emissions
Country masks	Python maps package	Create raster file of country masks to be used to calculate country total emissions

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

2.2 Description of data to be produced

Following the Description of Action, EYE-CLIMA will produce various data types, including gridded (or spatial) data, timeseries data, algorithms and code, as well as text data. These different types of data can be grouped into the following categories:

1. Maps of biomass and biomass change
2. Flux and emission data
3. Atmospheric mixing ratio and concentration data
4. Synthesized data products
5. Code and methodologies

The data produced for each of these categories are described below along with their intended applications.

2.2.1 Biomass maps

Maps of above ground biomass will be derived using remote sensing data (see Table 1) and will be used to derive maps of total biomass carbon using tree carbon content and below ground expansion factors. The resulting biomass carbon maps will be used in conjunction with a book-keeping model to derive maps of carbon losses (emissions) resulting from fire/clearing and carbon gains (removals) due to recovery from past disturbances.

Table 5. Summary of biomass and biomass change datasets.

Data description	Format	Size	Applications
Map of Above Ground Biomass (AGB) changes for Europe at 25 km resolution for 2010 - 2022	NetCDF	Few GB	Used to derive maps of above ground biomass carbon and using under-ground expansion factors to total biomass carbon
Map of AGB changes for Russia at 25 km resolution for 2010 - 2022	NetCDF	Few GB	
Spatial Sample of Drivers of Biomass Change	GeoTIFF	Few GB	Policy decisions

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

2.2.2 Flux and emission data

Gridded fluxes of CO_2, CH_4 and N_2O will be produced by process-based models for the whole globe but with a particular focus on Europe. In addition, gridded N_2O fluxes from agricultural land will be produced for Europe using a data-driven model. Gridded fluxes of CO_2, CH_4 and N_2O will be produced for Europe, and for CO_2 and CH_4 also for Russia, from atmospheric inversions. Emissions of BC and F-gases will be produced from atmospheric inversions globally, and for BC and SF_6 also at higher resolution for Europe. Lastly, a global inventory of point CH_4 emission sources will be produced using satellite remote sensing.

Table 6. Summary of flux and emissions datasets.

Data description	Format	Size	Applications
GAINS emission inventory for Europe at 0.1° spatial resolution and monthly for CH_4, N_2O, BC and annually for F-gases.	NetCDF	Few GB	Prior information for inversions and for sectorial composition of emissions
Land-biosphere fluxes of CO_2, CH_4 and N_2O from process-based models for Europe and northern latitudes	NetCDF	Few GB	Prior information for inversions. Understanding driving factors for fluxes
European agricultural fluxes of N_2O from data-driven model at 0.1° resolution	NetCDF	Few GB	Independent comparison with process-based model and inversion estimates
European fluxes of CO_2, CH_4, and N_2O from atmospheric inversions at 0.5° resolution and monthly from 2005 to 2023	NetCDF	Few GB	Support European policy by providing observation based emission trends
European fluxes of CO_2, CH_4, and N_2O from atmospheric inversions at 0.2° resolution and monthly from 2018 to 2023	NetCDF	Few GB	Support European NGHGIs and policy
Russian fluxes of CO_2 and CH_4 at 1.0° or 0.5° resolution and monthly from 2015 to 2021	NetCDF	Few GB	Support NGHGIs and policy
Global emissions of F-gases with at least 1.0° resolution from 2014 (for SF_6 from 2005) to at least 2021	NetCDF	Few GB	Support European NGHGIs and policy
European emissions of BC at 0.5° resolution from 2015 (joint deliverable with the PARIS project)	NetCDF	Few GB	Support policy
Global emissions of BC at 1.0° resolution from 2015	NetCDF	Few GB	Support policy
Global inventory of CH_4 point emission sources from remote sensing	NetCDF	Few GB	Support NGHGIs and policy
Compilation of NGHGI estimates (as reported to UNFCCC) and their uncertainties	ASCII	Few MB	Support verification of NGHGIs and comparisons with independent emission estimates

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

2.2.3 Atmospheric mixing ratio and concentration data

Atmospheric data will be produced from the atmospheric transport models used in the inversions. Although these data are not main data products of EYE-CLIMA, they are useful for supporting future modelling studies (e.g. the 4D mixing ratio fields) and for evaluating the quality of the bottom-up flux estimates (used in the prior information) and the fluxes estimated by the inversions.

Table 7. Summary of atmospheric datasets.

Data description	**Format**	**Size**	**Applications**
4D mixing ratio fields of CO_2, CH_4, N_2O and F-gases from nudged runs of FLEXPART-CTM	NetCDF	Few GB	Support regional atmospheric inversions
Simulated mixing ratios of CO_2, CH_4, N_2O and F-gases from atmospheric transport models	ASCII	Few MB	Support evaluation of bottom-up flux estimates and the atmospheric inversions
eBC concentrations from model simulations	ASCII	Few MB	Support evaluation of bottom-up flux estimates and the atmospheric inversions
eBC concentrations from observations attributed to solid fuel (e.g. biomass burning) and liquid fuel (e.g. fossil fuel) combustion sources	ASCII	Few MB	Used in inversions of BC and to support emission reduction policy

2.2.4 Synthesized data products

To support the objectives of EYE-CLIMA, a number of synthesised data products will be produced that are aimed directly at stakeholders, in particular, NGHGI compilers and policy makers.

Table 8. Summary of synthesised data products.

Data description	**Format**	**Size**	**Applications**
Emissions (and uncertainties) of CO_2, CH_4, N_2O and BC by sector and for EU27+3 countries based on atmospheric inversions with at least annual resolution	ASCII	Few GB	Support NGHGIs and policy
Russian emissions (and uncertainties) of CO_2 and CH_4 by sector based on atmospheric inversions with at least annual resolution	ASCII	Few GB	Independent estimates to compare with emissions reported to the UNFCCC
Map of managed land for Russia	GeoTIFF	Few GB	Support comparison of top-down estimates with estimates reported to the UNFCCC

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

2.2.5 Code and methodologies

An important objective of EYE-CLIMA is to further develop the methodology of atmospheric inversion and demonstrate the use of inversion data to support NGHGIs. In the context of this objective, EYE-CLIMA will produce code (and an associated report) on methodology for using top-down (inversion) estimates to derive emission variables that can be compared to what is reported in NGHGIs. In addition, EYE-CLIMA will produce good practice guidelines on atmospheric inversions to support NGHGIs in the form of a report. These guidelines will be produced in collaboration with the AVENGERS project. Lastly, EYE-CLIMA will contribute to the further development of the Community Inversion Framework (CIF).

Table 9. Summary of code and methodological products.

Description	Format	Application
Methodology for using top-down flux estimates to derive emissions comparable to what is required for reporting to the UNFCCC	Code and report	Support verification of NGHGIs
Good practice guidelines on atmospheric inversions for supporting NGHGIs (produced in collaboration with the AVENGERS project)	Report	Support verification of NGHGIs
Developments to the Community Inversion Framework (CIF)	Code	Atmospheric inversions

3. FAIR Data

3.1. Making data findable, including provisions for metadata

All data produced in EYE-CLIMA will be made discoverable with specific meta-data using standard qualifiers as used in the disciplines of atmospheric and climate sciences. The list of metadata associated with each dataset has evolved during the project following user feedback and experiences gained and are reflected in this final update to the DMP. A short data management protocol has also been prepared as a “quick reference” guide and as a stand-alone document (this is included in Appendix A).

The naming of data files will follow a specific convention that has been defined by the executive board of EYE-CLIMA and which contains the key meta-data information. Such naming convention also takes care of the version numbers, since during the course of the project more than one version of a dataset may be delivered.

Files will be named using the fields provided in Table 10 with the fields given in the following order:

Species_Variable_Sector_Region_Method_Timestep_FromTime_ToTime_Model_Institute_Version.
Filetype

Some examples are:

CO2_FLUX_ALL_EUR_INV_MONTH_20050101_20221231_CIF-CHIMERE_CNRS_V01.nc

N2O_FLUX_AGRIC_GBL_MOD_DAY_20200101_20201231_LPJGUESS_KIT_V01.nc

SF6_CONC_ALL_GBL_MOD_MONTH_20200101_20201231_FLEXPART-CTM_UNIVIE_V01.nc

This project has received funding from the European Union’s Horizon Europe
research and innovation programme under grant agreement No 101081395

Here, "dataset" is defined as a data file (or files) that result(s) from a given model or methodology, or a set of data files that otherwise have some logical connection. For datasets consisting of more than one species, the default will be to include only one species per file but include multiple variables per file if relevant (e.g. soil respiration, NPP, NBP). There may be datasets for which including multiple species (e.g. in the case of co-emitted species) in the same file may make sense, in which case deviations from the default will be useful.

Table 10. Meta-data to be included in the file names

Field	**Examples**	**Description**
Species	CO2, CH4, N2O, BC, HFC125	Species name
Variable	FLUX CONC	The main variable type in the file
Sector	ALL AGRIC FOSSIL LULUCF	Precision on the ecosystem or sector covered
Region	EUR: Europe RUS: Russia GBL: Global	Region covered
Method	DAT: data driven model MOD: model INV: inversion SYN: synthesis	Distinguishes the different methods used for the product derivation
Timestep	HOUR DAY MONTH YEAR	The time frequency of the product
FromTime	YYYYMMDD	Start date of dataset
ToTime	YYYYMMDD	End date of dataset
Model	CIF FLEXINVERT FLEXPART CHIMERE TM5MP ORCHIDEE LPJGUESS JSBACH-HIMMELI GAINS	A reference enabling people to uniquely identify the dataset, including the type of simulation that is produced. In some cases, names may need to be combined, e.g. CIF-FLEXPART.
Institute	Official short name	The institute who produced the data (using the official short names)
Version	VX	The version number
File type	.nc .tiff .txt .csv	The file format

Additional meta-data will be provided within the files following the Climate and Forecast meta-data convention (this is described further in Section 3.3. Making data interoperable).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

All public datasets from EYE-CLIMA will be assigned Digital Object Identifiers (DOIs) making them citable and trackable (this described further in Section 3.2 Making data accessible).

3.2. Making data accessible

All data produced in EYE-CLIMA will be stored in the Zenodo repository (https://zenodo.org/communities/eye-clima/) or another similar repository if there is a good reason to prefer that over Zenodo. Zenodo is a free repository funded by CERN and the European Union. Zenodo accepts all types of research data, including meta-data, without restriction on format. Uploaded datasets are versioned and provided with Digital Object Identifiers (DOIs) to make them citable and traceable. The only restriction in Zenodo is that for datasets exceeding 50 GB uploads need to be requested. This limit will not affect most datasets produced in EYE-CLIMA, and for any that do, a request will be made. On Zenodo we will use the "community" feature to create a collection of all data associated with the EYE-CLIMA project make the project's output more visible. We will also point to these datasets from the EYE-CLIMA website with a direct link to the data on Zenodo.

Although we will aim to make all final datasets publicly available as soon as possible, some especially novel datasets will be embargoed until publications related to the dataset have been accepted for publication. The embargo period, however, will not extend beyond the end of the project (31 December 2026). Zenodo allows datasets to be put under an embargo before being made open access.

Data will be released under the Creative Commons International license (CC BY 4.0) (https://creativecommons.org/licenses/by/4.0/), which allows material to be redistributed and adapted, but under the condition that appropriate credit is given and that no legal terms or technical measures are applied that restrict others using the material in a way that the license permits. A request will be made on each dataset submission that users acknowledge the data provider and the EYE-CLIMA project and cite the DOI in any publication that uses the dataset.

Zenodo allows for rich meta-data of each dataset according to the DataCite meta-data schema (https://schema.datacite.org). Meta-data are indexed and searchable directly in Zenodo's search engine immediately after publishing. (Further details on meta-data are given in section 3.3).

Data stored on Zenodo are backed-up and will be retained for the lifetime of the repository, which is currently defined for the next 20 years. Once datasets are made open access, it will not be possible to ascertain the identity of people accessing the data, however, Zenodo does provide statistics on the number of views and downloads for each dataset.

3.3. Making data interoperable

We will seek to make all publicly available datasets produced in EYE-CLIMA as inter-operable as possible. To do this, we will adopt international conventions for the data formats and metadata.

For spatial data, we will use the Network Common Data Format (NetCDF), which is especially suited to data with multiple dimensions. This data format is accessible across different operating systems and readable/writeable by numerous programming interfaces (e.g. Python, Matlab, R, Fortran, C++). We will follow as closely as possible the Climate and Forecast (CF) convention (http://cfconventions.org), which has been established to promote the processing and sharing of NetCDF files. The meta-data defined by the CF convention are included in the same file as the data making it "self-describing". The CF convention provides standard names for variables and standard units.

For simple timeseries data, we will use machine readable formats, such as .txt and .csv, which also allow the meta-data to be included in the same file as the data.

For map data, we will use the GeoTIFF data format, which allows geo-referencing information to be embedded within a TIFF file. The potential additional information includes map projection, coordinate

This project has received funding from the European Union's Horizon Europe
research and innovation programme under grant agreement No 101081395

systems, ellipsoids, datums, and everything else necessary to establish the exact spatial reference for the file.

Qualified references will be included in datasets produced in EYE-CLIMA that depend on other EYE-CLIMA datasets. This can be done by citing the DOIs of data already uploaded to the repository in the data description.

To further the facilitate the use of the data, within EYE-CLIMA, and for inter-comparisons with other related projects (e.g. PARIS and AVENGERS), and for the ease of use for other data-users, a specific dimension and variable convention will be used, especially for NetCDF files (see Appendix B).

Dimension variables

Table 11. Dimension variables for NetCDF files (as needed and wherever applicable).

Name	**Description**	**Units**
longitude	longitudinal mid-points	degrees east
latitude	latitudinal mid-points	degrees north
pressure	pressure	pascals
altitude	height above sea level	metres
time	time stamp (for the start of the time interval)	days since 1900-01-01 (or if hourly data used hours since start date)
time_index	for the time_bounds variable	none (index)
country_number	index of countries	none (index)

Dependent variables for inversion and other similar output

Specific points for variables

1. Sink/source convention: negative indicates flux from the atmosphere to the surface (sink), positive indicates flux from the surface to the atmosphere (source)
2. Include a “units” attribute: for units using mass include the molecular weight of the species to make it clear what mass was used (to avoid confusion between reporting e.g. mass of C versus mass of CO_2)
   (e.g. CO_2 = 44, CH_4 = 16, N_2O = 44)
3. Include a “long_name” attribute
4. Include a “missing_value” attribute

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101081395

Table 12. Mandatory dependent variables in inversion flux output as NetCDF files (see also Appendix B).

Name	**Description**	**Dimensions**	**Units**
time_bnds	Start and end time of each time interval of the reported variables	time, time_index	same as “time”
flux_total_prior	Total prior flux estimate	longitude, latitude, time	mol m^{-2} s^{-1} (positive = source; negative = sink)
flux_total_posterior	Total posterior flux estimate	longitude, latitude, time	mol m^{-2} s^{-1} (positive = source; negative = sink)
stdev_flux_total_prior	Total prior flux uncertainty as 1-sigma standard deviation	longitude, latitude, time	mol m^{-2} s^{-1}
stdev_flux_total_posterior	Total posterior flux uncertainty as 1-sigma standard deviation	longitude, latitude, time	mol m^{-2} s^{-1}
flux_[sector_name]_prior	Prior flux estimate for sector	longitude, latitude, time	mol m^{-2} s^{-1}
flux_[sector_name]_posterior	Posterior flux estimate for sector	longitude, latitude, time	mol m^{-2} s^{-1}
stdev_flux_[sector_name]_prior	Prior flux uncertainty for sector as 1-sigma standard deviation	longitude, latitude, time	mol m^{-2} s^{-1}
stdev_flux_[sector_name]_posterior	Posterior flux uncertainty for sector as 1-sigma standard deviation	longitude, latitude, time	mol m^{-2} s^{-1}
flux_total_prior_country	Total prior flux by country	country_number, time	kg yr^{-1}
flux_total_posterior_country	Total posterior flux by country	country_number, time	kg yr^{-1}
stdev_flux_total_prior_country	Pprior flux uncertainty by country	country_number, time	kg yr^{-1}
stdev_flux_total_posterior_country	Posterior flux uncertainty by country	country_number, time	kg yr^{-1}

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101081395

flux_[sector_name]_prior_country	Prior flux for sector and by country	country_number, time	kg yr^{-1}
flux_[sector_name]_posterior_country	Posterior flux for sector and by country	country_number, time	kg yr^{-1}
stdev_flux__[sector_name]_prior_country	Pprior flux uncertainty for sector by country	country_number, time	kg yr^{-1}
stdev_flux__[sector_name]_posterior_country	Posterior flux uncertainty for sector by country	country_number, time	kg yr^{-1}
country	Country abbreviation ISO_3166_1_alpha3	countrynumber, nchar	none
country_fraction	Fraction of grid cell associated to country	countrynumber, latitude, longitude	none
cell_area	Area of grid cells	longitude, latitude	m^2

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Table 13. Mandatory dependent variables in inversion mixing ratio output as NetCDF files (see also Appendix B)

Name	Description	Dimensions	Units
time	time of the mid-point of the observation interval	index	days since 1970-01-01 00:00:00
latitude	latitude of observation	index	degrees north
longitude	longitude of observation	index	degrees east
altitude	altitude of observation (surface elevation plus intake height)	index	meters above sea level
mf_observed	observed mole fraction in dry air	index	mol mol^{-1}
stdev_mf_model	uncertainty used in inversion	index	mol mol^{-1}
mf_prior	prior modelled mole fractions	index	mol mol^{-1}
mf_posterior	posterior modelled mole fractions	index	mol mol^{-1}
mf_bc_prior	prior simulated boundary condition mole fraction	index	mol mol^{-1}
mf_bc_posterior	posterior simulated boundary condition mole fraction	index	mol mol^{-1}
mf_outer_prior	prior modelled mole fraction from outside inversion domain (e.g. mf_prior = mf_prior_bc + mf_outer_prior + mf_inversiondomain_prior)	index	mol mol^{-1}
mf_outer_posterior	posterior modelled mole fraction from inside and outside inversion domain	index	mol mol^{-1}
identifier_of_platform	name of observing platform (e.g. site)	index, nchar	none

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Dependent variables for other output

For variable names please provide the fields in the following order: Variable_Species_[Sector]

Examples:
mass_concentration_black_carbon_total
flux_net_ecosystem_exchange
flux_nitrous_oxide_croplands

Table 14. Variable names for other output (and wherever applicable).

Field	**Examples**	**Description**
Variable	flux mole_fraction mass_concentration net_ecosystem_exchange	Physical variable
Species	carbon_dioxide sulfur_hexafluoride methane nitrous_oxide black_carbon hfc125	Species name
Sector	agriculture cropland lulucf	Additional description for unique variable names, or to distinguish fluxes from different sectors.

Table 14. Standard units for physical variables

Variable	**Units**
Flux	kg[species] m^{-2} s^{-1} (positive = source; negative = sink)
Mass concentration	kg[species] m^{-3}
Mole fraction	ppt, ppb, or ppm (as appropriate)

3.4. Increase data re-use

We will make all data open access by the end of the project. Data will be released using the Creative Commons International (CC BY 4.0) license, which allows data to be redistributed and modified, with the condition that credit is given to the data providers. Rich meta-data including a detailed data description and keywords will be provided with the datasets in Zenodo and will follow international conventions (e.g. CF convention). We will use file formats that allow for self-description of the data and are readable by a wide range of programming interfaces (see section 3.3).

We will also cite dataset DOIs in publications describing and/or using the datasets. This will increase the awareness of the data.

Quality of the data produced is ultimately the responsibility of the data provider, but we will, as far as possible, publish scientific articles that describe the methodology used to produce the data or results based on the data, which will be peer-reviewed. The data manager and the coordination team will check

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

compliance of all datasets with the DMP. In general, EYE-CLIMA partners will strive for the highest quality data and meta-data.

4. Other Research Outputs

Zenodo allows also for the upload of publications, posters, presentations, videos, and code. As for datasets, there is the possibility to associate rich meta-data with the material and it is given a DOI. For code, Zenodo also allows to preserve a GitHub repository.

An advantage of using the Zenodo repository also for outreach material (publications, posters etc.) is that these can be directly associated with the EYE-CLIMA collection along with datasets.

5. General Data Management Issues

5.1 Allocation of Resources

The cost of making the data FAIR will be shared across the data providers (each partner of EYE-CLIMA) and the coordinator as follows:

- The data providers will make sure that their data are compliant with this DMP and provide the necessary data description and metadata of their datasets.
- The coordinator (NILU) will ensure that the DMP is implemented and that the metadata description, key words and file naming complies with EYE-CLIMA's protocol. The coordination team will ensure that the data are uploaded to the Zenodo repository.

Since Zenodo is a free repository, there are no additional costs related to its use. Data stored on Zenodo are backed-up and will be retained for the lifetime of the repository, that is, at least 20 years from present.

5.2 Data Security

Zenodo is a secure repository. All data files are stored in CERN Data Centres, primarily Geneva, with replicas in Budapest. Data files are kept in multiple replicas in a distributed file system, which is backed up to tape on a nightly basis. Data are stored for the lifetime of the repository, which is guaranteed for the next 20 years at least.

5.3 Ethics

There are no specific legal or ethical issues in relation to the activities, the data used, or the data produced in EYE-CLIMA. Specifically, EYE-CLIMA will not be using personal or confidential data and will not produce any data of confidential nature.

In general, EYE-CLIMA's partners will comply with the ethical principles as set out in Article 14 of the Grant Agreement, which states that all activities must be carried out in compliance with:

- ethical principles, including the highest standards of research integrity
- applicable EU, international and national law

5.4 Other Issues

For the sharing of input data and intermediary results among the partners, EYE-CLIMA will use an SFTP server provided by the coordinator (NILU). The SFTP server will allow for secure (password protected) access to data and for the transfer of large data files. It is not intended, however, for long-term storage of data, and data there will not be backed-up.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Appendix A: EYE-CLIMA Data Management Protocol

Purpose of this protocol

The intention of this document is to detail the procedure for managing final datasets from the EYE-CLIMA project, i.e., the data products that should be submitted to a data repository. This protocol is based on the Data Management Plan (DMP) and describes the steps that should be followed once a data product has been produced. It also describes in detail the file naming convention to be used and the meta data to be provided. The data manager (DM) (Nalini Krishnankutty, NILU) will oversee and assist with the implementation of this protocol. Please note, that this protocol does not exclude submitting datasets to other repositories or servers, but the following is a minimum requirement.

I. Data management steps

1. Check the dataset conforms to the EYE-CLIMA conventions

Please check that the file and variable names conform to the conventions outlined in Section II and that sufficient meta-data is provided.

2. Inform the DM of the dataset

Pease send an email to coord-eyeclima@nilu.no to inform them that the dataset is ready and, if possible, provide a link to where the data are located (this could be by uploading the data to NILU's sftp server – see below for server details). If the dataset is too large to upload to a server within reasonable time limits, a description of the dataset including the file and variable names as well as the meta-data may be provided instead.

3. Upload the dataset to the repository

Once the dataset has been checked by the DM, please upload the data to the Zenodo repository: https://zenodo.org.

If you have not previously used Zenodo you will first need to create a user account. Once you have logged-in click on "New upload" to upload your dataset. This will open a new page where you can enter the meta-data and a place to drag-and-drop your dataset into. When uploading please remember to choose the EYE-CLIMA collection and to specify the funding source (see the box on the following page).

There is also a direct link to the EYE-CLIMA collection:

https://zenodo.org/communities/eye-clima/

You can also use this link to directly upload to the collection (log-in is still required).

4. Send the DOI to the DM

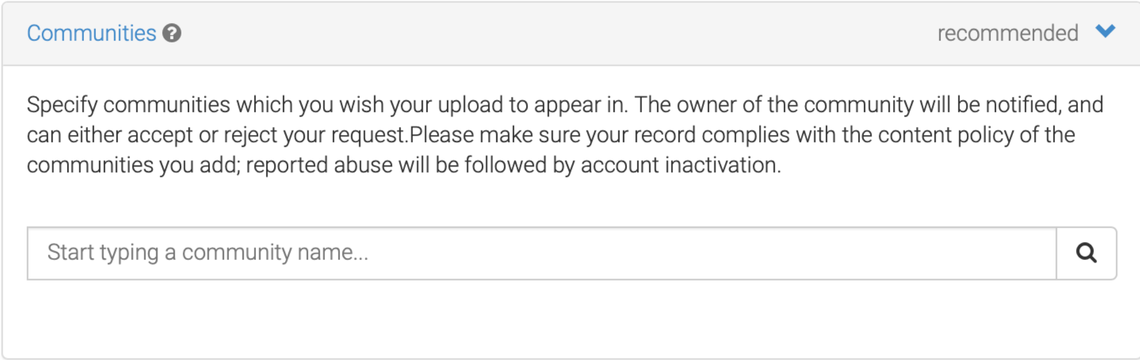
Once you have received the DOI of your dataset please add this to the Dataset list on the project sharepoint here: Datasets.xlsx. The DM will add the DOI and url link to the table to data products on the EYE-CLIMA website so that all datasets can be found via the website.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

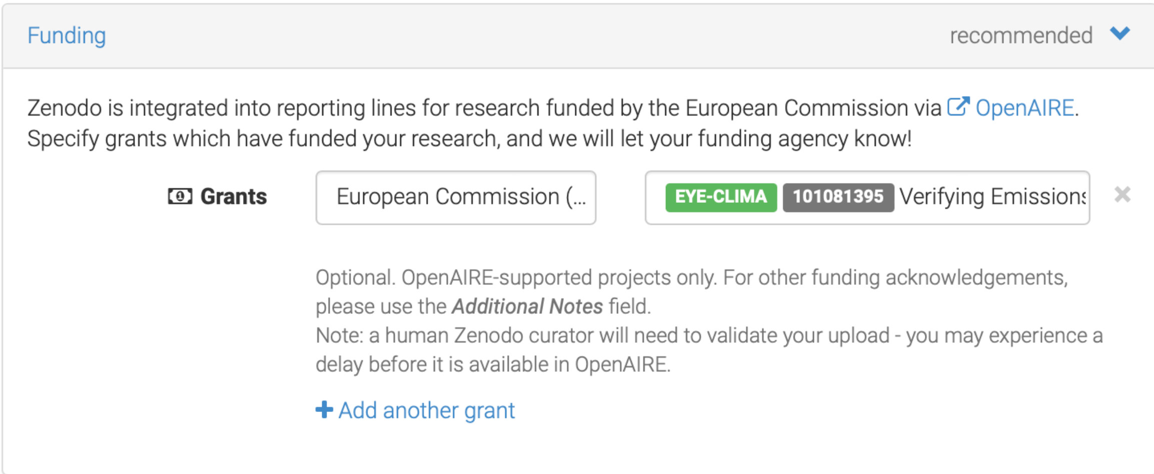
IMPORTANT

Please note the following:

i) in the field Communities enter "eye-clima"



ii) in the field Funding select "European Commission" and enter "EYE-CLIMA" (the project should automatically be found with the full name and grant number, if not enter the grant number: 101081395)



II. File and meta-data conventions

1. File formats and naming conventions

For spatial data, NetCDF files should be used, with the exception of map data, for which GeoTIFF files should be used. For timeseries data, we suggest you use ASCII formatted data (e.g. .txt and .csv).

Files should be named using the convention described below using the fields provided in Table 1. The fields should be given in the following order:

Species_Variable_Sector_Region_Method_Timestep_FromTime_ToTime_Model_Institute_Version.
Filetype

Some examples are:

CO2_FLUX_ALL_EUR_INV_MONTH_20050101_20221231_CIF-CHIMERE_CNRS_V01.nc

N2O_FLUX_AGRIC_GBL_MOD_DAY_20200101_20201231_LPJGUESS_KIT_V01.nc

SF6_CONC_ALL_GBL_MOD_MONTH_20200101_20201231_FLEXPART-CTM_UNIVIE_V01.nc

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Please include only one species per file, but include multiple variables per file if relevant (e.g. soil respiration, NPP, NBP).

Table 1. Overview of the different fields to be used in file names

Field	**Examples**	**Description**
Species	CO2, CH4, N2O, BC, HFC125	Species name
Variable	FLUX CONC	The main variable type in the file
Sector	ALL AGRIC FOSSIL LULUCF	Precision on the ecosystem or sector covered
Region	EUR: Europe RUS: Russia GBL: Global	Region covered
Method	DAT: data driven model MOD: model INV: inversion SYN: synthesis	Distinguishes the different methods used for the product derivation
Timestep	HOUR DAY MONTH YEAR	The time frequency of the product
FromTime	YYYYMMDD	Start date of dataset
ToTime	YYYYMMDD	End date of dataset
Model	CIF FLEXINVERT FLEXPART CHIMERE TM5MP ORCHIDEE LPJGUESS JSBACH-HIMMELI GAINS	A reference enabling people to uniquely identify the dataset, including the type of simulation that is produced. In some cases, names may need to be combined, e.g. CIF-FLEXPART.
Institute	Official short name	The institute who produced the data (using the official short names)
Version	VX	The version number
File type	.nc .tiff .txt .csv	The file format

2. Variable and dimension naming conventions

In NetCDF files (and wherever applicable), please use the dimensions, variables and names as outlined in Tables 2 to 5, and the units as outlined in Table 6. These tables are based on the Climate and Forecast conventions (http://cfconventions.org/Data/cf-standard-names/docs/guidelines.html).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Table 2. Dimension names for NetCDF files (and wherever applicable). Obviously only include the dimensions as applicable to your dataset.

Name	Description	Units
longitude	longitudinal mid-points	degrees east
latitude	latitudinal mid-points	degrees north
pressure	pressure	pascals
altitude	height above sea level	metres
time	time stamp (for the start of the time interval)	days since 1900-01-01 (or if hourly data used hours since start date)
time_index	for the time_bounds variable	none (index)
country_index	index of countries	none (index)

Specific points for variables:

1. Sink/source convention: negative indicates flux from the atmosphere to the surface (sink), positive indicates flux from the surface to the atmosphere (source)
2. Include a "units" attribute: for units using mass include the molecular weight of the species to make it clear what mass was used (to avoid confusion between reporting e.g. mass of C versus mass of CO_2)
   (e.g. CO2 = 44, CH4 = 16, N2O = 44)
3. Inlcude a "long_name" attribute
4. Include a "missing_value" attribute

Dependent variables for inversion flux output

For inversion output files, there are specific variables that should be provided and a specific format for these to facilitate inter-comparison with results from other related projects (PARIS and AVENGERS).

Please refer to Appendix B of the Final Data Management Plan for the specific format.

Dependent variables for other output

For variable names please provide the fields in the following order: Variable_Species_[Sector]

Examples:

mass_concentration_black_carbon_total
flux_net_ecosystem_exchange
flux_nitrous_oxide_croplands

Table 3. Variable names for NetCDF files (and wherever applicable).

Field	Examples	Description
Variable	flux mole_fraction mass_concentration net_ecosystem_exchange	Physical variable
Species	carbon_dioxide sulfur_hexafluoride methane nitrous_oxide black_carbon hfc125	Species name

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Sector	agriculture cropland lulucf	Additional description for unique variable names, or to distinguish fluxes from different sectors.

Table 4. Standard units for physical variables

Variable	**Units**
Flux	kg[species] m^{-2} s^{-1} (positive = source; negative = sink)
Mass concentration	kg[species] m^{-3}
Mole fraction	ppt, ppb, or ppm (as appropriate)

3. Meta-data conventions

Please provide at least the following meta-data:

i. Production date (YYYYMMDD)
ii. Author name(s)
iii. Keywords

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Appendix B: AVEYPA Inversion output data formats

To facilitate inter-comparisons between the three projects, we have defined a common data format for inversion output, which should be used by all projects. The format is designed to contain all information that may be needed to allow an in-depth analysis into why the results between projects may differ, and also to be useful for other data users. Note also that names should be given as character arrays and not strings as not all tools to read NetCDF files can cope with strings.

1. Fluxes

The format for flux output is described fully in Section 1. This is based on the following guidelines:

- separate variables for each source type (where inversion optimizes more than one source type)
- include a variable for the total source
- for gridded fluxes use units of molecules per area per time (not mass since for e.g. CO2 the mass is reported as either for C or for CO2)
- for country totals use the mass of species (CO2 = 44, CH4 = 16, N2O = 44) (variable attribute) which is what is used in national inventories
- include an additional variable for the area of grid cells
- inclusion of covariances is optional
- include uncertainty given as a standard deviation for inversions using Gaussian prior distributions, and percentiles for inversions using non-Gaussian distributions
- include source sector information in meta-data

2. Mole fractions

The format for mole fraction output is described fully in Section 2. This is based on the following guidelines:

- format inspired by Obspack format
- model simulations should be reported as dry-air mole fractions
- No dimension for stations
- optional variables for background, initial mixing ratio etc.
- Include time bounds (start and end time of observation time window)
- Use UTC time

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

1. Flux NetCDF file format

```
Note: comments are indicated by //
dimensions:
        longitude = nx ;
        latitude = ny;
        time = nt ;
        countrynumber = ncntry ;
        nchar = 3;
        nbnds = 2 ;
        sectornumber = nsector ;        // optional
        sectornchar = 20 ;        // conditional (mandatory if sectornumber present)
variables:
        double longitude(longitude) ;
                longitude:units = "degrees_east" ;
                longitude:long_name = "longitude of grid cell centre" ;
                longitude:standard_name = "longitude" ;
        double latitude(latitude) ;                // mandatory
                latitude:units = "degrees_north" ;
                latitude:long_name = "latitude of grid cell centre" ;
                latitude:standard_name = "latitude" ;
        double time(time) ;
                time:units = "days since 1970-01-01 00:00:00" ;
                time:long_name = "mid of flux interval in UTC" ;
                time:standard_name = "time" ;
                time:calendar = "proleptic_gregorian" ;
        double time_bnds(time, nbnds) ; // mandatory
                time_bnds:long_name = "start and end points of each time step" ;
                time_bnds:calendar = "proleptic_gregorian" ;
                time_bnds:units = "days since 1970-01-01 00:00:00" ;
//
// TOTAL FLUX ON SPATIAL GRID
//      'total' variables should be present even if fluxes by sector are given separately
        float flux_total_prior(time, latitude, longitude) ;
                flux_total_prior:units = "mol m-2 s-1" ;
                flux_total_prior:_FillValue = NaNf ;
                flux_total_prior:long_name = "prior total <species> fluxes" ;
                flux_total_prior:cell_methods = "time:mean area:mean" ;
        float flux_total_posterior(time, latitude, longitude) ;
                flux_total_posterior:units = "mol m-2 s-1" ;
                flux_total_posterior:_FillValue = NaNf ;
                flux_total_posterior:long_name = "posterior total <species> fluxes" ;
                flux_total_posterior:cell_methods = "time:mean area:mean" ;
//
// UNCERTAINTY OF TOTAL FLUX ON SPATIAL GRID
//      by default, uncertainty should be reported as standard deviation (1-sigma)
        float stdev_flux_total_prior(time, latitude, longitude) ;
                stdev_flux_total_prior:units = "mol m-2 s-1" ;
                stdev_flux_total_prior:_FillValue = NaNf ;
                stdev_flux_total_prior:long_name = "standard deviation of prior total <species>
                fluxes" ;
                stdev_flux_total_prior:cell_methods = "time:mean area:mean" ;
```


This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

```
        float stdev_flux_total_posterior(time, latitude, longitude) ;
                stdev_flux_total_posterior:units = "mol m-2 s-1" ;
                stdev_flux_total_posterior:_FillValue = NaNf ;
                stdev_flux_total_posterior:long_name = "standard deviation of posterior total
                <species> fluxes" ;
                stdev_flux_total_posterior:cell_methods = "time:mean area:mean" ;
        // 'percentile' variables are optional for systems that report non-Gaussian uncertainty
        //band do not report 'stdev' variables
//
// FLUX ON SPATIAL GRID BY SECTOR (one set of variables per sector)
//      all sector variables are optional. However, if a sector is given the same variables as for the
        total are mandatory.
//      'sector' variables are optional for systems that give fluxes by sector ;
//      sector names should be given in the sector_names variable: typical names may be
'agriculture', 'energy' 'waste', 'industry'
//      'sector' needs to be replaced by the individual names
        float flux_sector_name_prior(time, latitude, longitude) ;
                flux_sector_name_prior:units = "mol m-2 s-1" ;
                flux_sector_name_prior:_FillValue = NaNf ;
                flux_sector_name_prior:long_name = "prior <species> fluxes from sector_name";
                flux_sector_name_prior:cell_methods = "time:mean area:mean" ;
        float flux_sector_name_posterior(time, latitude, longitude) ;
                flux_sector_name_posterior:units = "mol m-2 s-1" ;
                flux_sector_name_posterior:_FillValue = NaNf ;
                flux_sector_name_posterior:long_name = "posterior <species> fluxes from
                <sector_name>" ;
                flux_sector_name_posterior:cell_methods = "time:mean area:mean" ;
//
// UNCERTAINTY OF FLUX ON SPATIAL GRID BY SECTOR
//      by default, uncertainty should be reported as standard deviation (1-sigma) around the
        mean, alternatively percentiles can be given in 'percentile' variables
        float stdev_flux_sector_name_prior(time, latitude, longitude) ;
                stdev_flux_sector_name_prior:units = "mol m-2 s-1" ;
                stdev_flux_sector_name_prior:_FillValue = NaNf ;
                stdev_flux_sector_name_prior:long_name = "standard deviation prior <species>
                fluxes from <sector_name>" ;
                stdev_flux_sector_name_prior:cell_methods = "time:mean area:mean" ;
        float stdev_flux_sector_name_posterior(time, latitude, longitude) ;
                stdev_flux_sector_name_posterior:units = "mol m-2 s-1" ;
                stdev_flux_sector_name_posterior:_FillValue = NaNf ;
                stdev_flux_sector_name_posterior:long_name = "standard deviation posterior
                <species> fluxes from <sector_name>" ;
                stdev_flux_sector_name_posterior:cell_methods = "time:mean area:mean" ;
//
//      TOTAL FLUX BY COUNTRY
//
        float flux_total_prior_country(time, countrynumber) ;
                flux_total_prior_country:units = "kg yr-1" ;
                flux_total_prior_country:_FillValue = NaNf ;
                flux_total_prior_country:long_name = "country-total prior <species> fluxes" ;
                flux_total_prior_country:cell_methods = "time:mean countrynumber:point" ;
```


This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

```
        float flux_total_posterior_country(time, countrynumber) ;
                flux_total_posterior_country:units = "kg yr-1" ;
                flux_total_posterior_country:_FillValue = NaNf ;
                flux_total_posterior_country:long_name = "country-total posterior <species> fluxes"
                ;
                flux_total_posterior_country:cell_methods = "time:mean countrynumber:point";
        float stdev_flux_total_prior_country(time, countrynumber) ;
                stdev_flux_total_prior_country:units = "kg yr-1" ;
                stdev_flux_total_prior_country:_FillValue = NaNf ;
                stdev_flux_total_prior_country:long_name = "standard deviation of country-total
                prior <species> fluxes" ;
                stdev_flux_total_prior_country:cell_methods = "time:mean countrynumber:point" ;
        float stdev_flux_total_posterior_country(time, countrynumber) ;
                stdev_flux_total_posterior_country:units = "kg yr-1" ;
                stdev_flux_total_posterior_country:_FillValue = NaNf ;
                stdev_flux_total_posterior_country:long_name = "standard deviation of country-
                total posterior <species> fluxes" ;
                stdev_flux_total_posterior_country:cell_methods = "time:mean
                countrynumber:point" ;
        // covariance between countries
        float covariance_flux_total_posterior_country(time, countrynumber, countrynumber) ;
                covariance_flux_total_posterior_country:units = "kg2 yr-2" ;
                covariance_flux_total_posterior_country:_FillValue = NaNf ;
                covariance_flux_total_posterior_country:long_name = "covariance of country-total
                posterior <species> fluxes" ;
                covariance_flux_total_posterior_country:cell_methods = "time:mean
                countrynumber:point countrynumber:point" ;
//
//      FLUX BY SECTOR AND COUNTRY
//
        float flux_sector_name_prior_country(time, countrynumber) ; // optional
                flux_sector_name_prior_country:units = "kg yr-1" ;
                flux_sector_name_prior_country:_FillValue = NaNf ;
                flux_sector_name_prior_country:long_name = "country-<sector_name> prior
                <species> fluxes" ;
                flux_sector_name_prior_country:cell_methods = "time:mean countrynumber:point"
                ;
        float flux_sector_name_posterior_country(time, countrynumber) ; // optional
                flux_sector_name_posterior_country:units = "kg yr-1" ;
                flux_sector_name_posterior_country:_FillValue = NaNf ;
                flux_sector_name_posterior_country:long_name = "country-<sector_name>
                posterior <species> fluxes" ;
                flux_sector_name_posterior_country:cell_methods = "time:mean
                countrynumber:point" ;
        float stdev_flux_sector_name_prior_country(time, countrynumber) ; // optional
                stdev_flux_sector_name_prior_country:units = "kg yr-1" ;
                stdev_flux_sector_name_prior_country:_FillValue = NaNf ;
                stdev_flux_sector_name_prior_country:long_name = "standard deviation of
                country-<sector_name> prior <species> fluxes" ;
                stdev_flux_sector_name_prior_country:cell_methods = "time:mean
                countrynumber:point" ;
```


This project has received funding from the European Union's Horizon Europe
research and innovation programme under grant agreement No 101081395

```
        float stdev_flux_sector_name_posterior_country(time, countrynumber) ;
                stdev_flux_sector_name_posterior_country:units = "kg yr-1" ;
                stdev_flux_sector_name_posterior_country:_FillValue = NaNf ;
                stdev_flux_sector_name_posterior_country:long_name = "standard deviation of
                country-<sector_name> posterior <species> fluxes" ;
                stdev_flux_sector_name_posterior_country:cell_methods = "time:mean
                countrynumber:point" ;
// covariance within each sector between countries
        float covariance_flux_sector_name_posterior_country(time, countrynumber,
                countrynumber) ; // optional
                covariance_flux_sector_name_posterior_country:units = "kg2 yr-2" ;
                covariance_flux_sector_name_posterior_country:_FillValue = NaNf ;
                covariance_flux_sector_name_posterior_country:long_name = "covariance of
                posterior <species> fluxes of <sector_name> between countries" ;
                covariance_flux_sector_name_posterior_country:cell_methods = "time:mean
                countrynumber:point countrynumber:point" ;
//
//      AUXILIARY VARIABLES
        // Country abbreviation
        // please store as character array not as string; some older interfaces cannot deal with
        netcdf string type
        char country(countrynumber, nchar) ;    // mandatory
                country:long_name = "country_ISO_3166_1_alpha3" ;
        // sector names should only contain lower case letters (no separator characters), since they
        should be used in variable names
        // optional when reporting separate fluxes by sector
        char sector_names(sectornumber, sectornchar) ;          // optional
                country:long_name = "short name of flux sector" ;
        //
        float country_fraction(countrynumber, latitude, longitude) ;      // mandatory
                country_fraction:units = "1" ;
                country_fraction:_FillValue = NaNf ;
                country_fraction:long_name = "fraction of grid cell associated to country" ;
                country_fraction:standard_name = "area_fraction" ;
        //
        float cell_area(latitude, longitude) ;       // mandatory
                cell_area:units = "m2" ;
                cell_area:_FillValue = NaNf ;
                cell_area:long_name = "surface area of grid cell" ;
                cell_area:standard_name = "cell_area" ;

global attributes:
        :title = "GHG flux distribution and country totals" ;
        :institution = <institute name> ;
        :source = "Estimated flux from trace gas observations transport simulations and inversion
        code." ;
        :creator = <name> ;
        :creation_date = <date> ;
        :contact = <email> ;
        :frequency = <time resolution, e.g. monthly> ;
        :transport_model = <model name, e.g. CHIMERE> ;
```


This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

```
        :transport_model_version = "" ;
        :inversion_system = <inversion framework name, e.g. CIF>;
        :inversion_system_version= <version number>;
        :experiment = <experiment name, e.g. GAINS prior>;
        :project = "EYE-CLIMA" ;
```

2. Mole fractions NetCDF file format

Note: comments are indicated by //

```
dimensions:
        index = unlimited;
        percentile = 2 ;
        nbnds = 2 ;
        nchar = 3 ;
variables:
// characterising observation
        double time(index) ;
                time:units = "days since 1970-01-01 00:00:00" ;
                time:long_name = "time of mid of observation interval; UTC" ;
                time:calendar = "proleptic_gregorian" ;
                double time_bnds(index, nbnds) ;
                time_bnds:long_name = "start and end points of each time step" ;
                time_bnds:calendar = "proleptic_gregorian" ;
                time_bnds:units = "days since 1970-01-01 00:00:00" ;
        double longitude(index) ;
                 longitude:_FillValue = NaNf ;
                longitude:long_name = "sample_longitude_in_decimal_degrees" ;
                longitude:units = "degrees_east" ;
                longitude:comment = "Longitude at which air sample was collected." ;
                longitude:standard_name = "longitude" ;
        double latitude(index) ;
                latitude:_FillValue = NaNf ;
                latitude:long_name = "sample_latitude_in_decimal_degrees" ;
                latitude:units = "degrees_north" ;
                latitude:comment = "Latitude at which air sample was collected." ;
                latitude:standard_name = "latitude" ;
        float altitude(index) ;
                altitude:_FillValue = NaNf ;
                altitude:long_name = "sample_altitude_in_meters_above_sea_level" ;
                altitude:units = "m" ;
                altitude:comment = "Altitude (surface elevation plus sample intake height) at which
                air sample was collected" ;
                altitude:standard_name = "altitude" ;
        short number_of_identifier(index) ;
                number_of_identifier:_FillValue = -9 ;
                number_of_identifier:long_name = "Index of identifier of observing platform" ;
                number_of_identifier:units = "1" ;

// observation and uncertainty
        float mf_observed(index) ;
                mf_observed:units = "mol mol-1" ;
```


This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

```
                mf_observed:_FillValue = NaNf ;
                mf_observed:long_name = "observed mole fraction of <species> in dry air" ;
        float stdev_mf_model(index) ;                                        // optional
                stdev_mf_model:units = "mol mol-1" ;
                stdev_mf_model:_FillValue = NaNf ;
                stdev_mf_model:long_name = "model uncertainty of simulated mole fraction" ;
        float stdev_mf_total(index) ;
                stdev_mf_total:units = "mol mol-1" ;
                stdev_mf_total:_FillValue = NaNf ;
                stdev_mf_total:long_name = "total model observation uncertainty applied in
                inversion" ;
// simulated mole fractions
        // mf_prior and mf_posterior contain the complete simulated concentration, i.e. sum of
        // the regional contribution within the transport domain
        // and a boundary (baseline) concentration (given as mf_prior_bc and mf_posterior_bc),
        // i.e. mf_posterior = mf_posterior_regional + mf_posterior_bc.
        float mf_prior(index) ;
                mf_prior:units = "mol mol-1" ;
                mf_prior:_FillValue = NaNf ;
                mf_prior:long_name = "prior simulated mole fraction of <species> in dry air" ;
        float mf_posterior(index) ;
                mf_posterior:units = "mol mol-1" ;
                mf_posterior:_FillValue = NaNf ;
                mf_posterior:long_name = "posterior simulated mole fraction of <species> in dry
                air" ;

//  individual sector contributions        (optional); the sum should add up to
// mf_posterior_regional, which is not included in file
        float mf_sector_name_prior(index) ;                              // optional
                mf_sector_name_prior:units = "mol mol-1" ;
                mf_sector_name_prior:_FillValue = NaNf ;
                mf_sector_name_prior:long_name = "prior simulated mole fraction of <species> in
                dry air from <sector_name>" ;
        float mf_sector_name_posterior(index) ;                          // optional
                mf_sector_name_posterior:units = "mol mol-1" ;
                mf_sector_name_posterior:_FillValue = NaNf ;
                mf_sector_name_posterior:long_name = "posterior simulated mole fraction of
                <species> in dry air from <sector_name>" ;

// mf_bc_prior and mf_bc_posterior should contain boundary condition (baseline)
        float mf_bc_prior(index) ;
                mf_bc_prior:units = "mol mol-1" ;
                mf_bc_prior:_FillValue = NaNf ;
                mf_bc_prior:long_name = "prior simulated boundary condition mole fraction" ;
        float mf_bc_posterior(index) ;
                mf_bc_posterior:units = "mol mol-1" ;
                mf_bc_posterior:_FillValue = NaNf ;
                mf_bc_posterior:long_name = "posterior simulated boundary condition mole
                fraction" ;
        // mf_prior_outer and mf_posterior_outer are part of mf_prior_regional
        // and mf_posterior_regional, respectively,
```


This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

```
        // i.e. mf_posterior_regional = mf_posterior_outer + mf_posterior_inversionDomain (the //
        latter not given in file)
        float mf_outer_prior(index) ;                    // optional
                mf_outer_prior:units = "mol mol-1" ;
                mf_outer_prior:_FillValue = NaNf ;
                mf_outer_prior:long_name = "prior simulated mole fraction contribution from
                distant regions" ;
        float mf_outer_posterior(index) ;                // optional
                mf_outer_posterior:units = "mol mol-1" ;
                mf_outer_posterior:_FillValue = NaNf ;
                mf_outer_posterior:long_name = "posterior simulated mole fraction contribution
                from distant regions" ;

// AUXILIARY VARIABLES
        // please use 2D character array instead of netcdf string type, because some older
        // interfaces do not know how to deal with strings
        char identifier_of_platform(index, nchar) ;
                identifier_of_platform:long_name = "identifier of observing platform" ;

// global attributes:
                :Conventions = "CF-1.8" ;
                :title = "In-situ mole fractions at sites: observed and simulated" ;
                :institution = <institution name> ;
                :source = "Trace gas mole fractions from observations and transport simulations /
                inverse estimation." ;
                :creator = <name> ;
                :creation_date = <yyyy-mm-dd> ;
                :contact = <email>;
                :transport_model = <transport model> ;
                :transport_model_version = "" ;
                :inversion_system = <inversion framework e.g. CIF> ;
                :inversion_system_version = "" ;
                :experiment = <experiment name, e.g. GAINS prior> ;
                :project = "EYE-CLIMA" ;
                :references = "" ;
```

iv. Description (what data are provided and how it was generated including any references)

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

https://eyeclima.eu

BRUSSELS, 22 12 2025

The publication reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395