D3 - ALGORITHM THEORICAL BASIS DOCUMENT (V3)

WORLD EMISSION

Prepared by:	LSCE-ULB-KAYRROS-MPIC-GMV
Approved by:	Gregoire Broquet
Authorized by:	Beatriz Revilla-Romero

Code:	D3-ATBD-V3
Version:	v2.0
Date:	27/10/2023
Internal code:	GMV 24675/23 V2/23

DOCUMENT STATUS SHEET

Version	Date	Pages	Changes
v1.0	27/10/2023	71	This deliverable D3v3 (v1.0) is an update of D3v2, which has been submitted at the end of phase 1 of World Emission. The updates mainly consist in: • a slight update of the references (Table 2-2: updating the status of some publications, adding new ones, and adding references to the project deliverables) • the detailed description of the data and processing for the quantification of SO_2 point sources (which had just started at the time of deliverable D3v2) in section 3.3.4.3 • a significant extension of the detailed description of the data and updated processing for the quantification of NO_x emissions from megacities and conurbations (which was still at an initial stage at the time of deliverable D3v2) in section 3.3.5. • an update of section 3.4.3 corresponding to the update of the regional NO_x, CO and CO_2 products • an extension of section 3.5.3.2 detailing new analysis of the uncertainties in the global CH_4 product • an update of section 3.5.4 to describe the new mass-balance approach to generate the global NO_x, SO_2 and NH_3 products
v2.0	27/10/2023	71	Comments from ESA implemented: • Updates in Section 3.3.4.3 (SO2 point source) • Updates in Section 3.3.5 (NOx megacities) • Updates in Section 3.4.3.2 (NOX, CO AND CO2 regional) • Updates in Section 3.4.3.3 (NOX, CO AND CO2 regional) • References and formatting

TABLE OF CONTENTS

1. INTRODUCTION 8
1.1. PURPOSE 8
1.2. SCOPE 8
1.3. ACRONYMS 8
2. REFERENCES 11
2.1. APPLICABLE DOCUMENTS 11
2.2. REFERENCE DOCUMENTS 11
3. ALGORITHM THEORETICAL BASIS 20
3.1. INTRODUCTION 20
3.2. ASSESSMENT OF SATELLITE DATASETS 20
3.3. ASSESSMENT OF POINT SOURCE TOP-DOWN INVERSION MODELS AND DESCRIPTION OF THE ALGORITHMS USED IN WORLD EMISSION 24
3.3.1. Introduction 24
3.3.2. Methane (CH4) point sources inventory 24
3.3.2.1. Point sources inversion 24
3.3.2.2. Error Analysis 26
3.3.2.3. Alternative methods and inventories 27
3.3.3. Carbon Dioxide (CO2) point sources inventory 27
3.3.3.1. Point source inversion 27
3.3.3.2. Error Analysis 29
3.3.3.3. Alternative methods 29
3.3.4. Nitrogen Oxides (NOx) and Sulphur Dioxide (SO2) point sources inventories 29
3.3.4.1. NOx 29
3.3.4.2. Error Analysis 31
3.3.4.3. SO_2 32
3.3.4.4. Alternative methods and inventories 33
3.3.5. NOx from megacities and conurbations 33
3.3.5.1. Datasets 34
3.3.5.2. Data processing 34
3.3.5.3. Lifetime and emission fit 37
3.3.5.4. Selection of fit results 37
3.3.6. Ammonia (NH3), Acetylene (C2H2), Ethylene (C2H4) and Methanol (CH3OH) point sources inventories 37
3.3.6.1. Point sources inventories 38
3.3.6.2. Error Analysis 42
3.3.6.3. Alternative method 42
3.4. ASSESSMENT OF REGIONAL TOP-DOWN INVERSION MODELS AND DESCRIPTION OF THE ALGORITHMS USED IN WORLD EMISSION 42
3.4.1. Introduction 42
3.4.2. Methane (CH4) basin-scale emissions 43
3.4.2.1. Regional inversion 43
3.4.2.2. Error Analysis 44
3.4.3. Nitrogren Oxides (NOx), Carbon Monoxide (CO), Carbon Dioxide (CO2) regional emissions 45
3.4.3.1. State of the art 45

3.4.3.2. General overview of the NOx, CO and CO2 regional products and of their derivation in World Emission 45
3.4.3.3. Components of the NO_x, CO and CO_2 regional inverse modeling system 46
3.4.3.4. Error Analysis 53
3.4.4. Ammonia (NH3) regional emissions 53
3.5. ASSESSMENT OF GLOBAL TOP-DOWN INVERSION MODELS AND DESCRIPTION OF THE ALGORITHMS USED IN WORLD EMISSION 54
3.5.1. Introduction 54
3.5.2. The LMDz-INCA model 54
3.5.3. Carbon Monoxide (CO) and Methane (CH4) 54
3.5.3.1. Description of the inversion 55
3.5.3.2. Uncertainty analysis 56
3.5.4. Nitrogen oxides (NOx), Sulphur Dioxide (SO2), Ammonia (NH3) 58
3.5.4.1. Satellite observations 58
3.5.4.2. Data processing 59
3.5.4.3. Description of the inversion 59
3.5.4.4. Estimated emissions 61
3.5.4.5. Uncertainty analysis 62
3.5.5. Isoprene (C5H8) 63
3.5.6. Particulate Matter (PM) 64
3.5.7. Water Vapor (H2O) 64
3.6. PRODUCT GENERATION WORKFLOW 65

LIST OF TABLES AND FIGURES

Table 1-1 Acronyms....8

Table 2-1 Applicable Documents....11

Table 2-2 Reference Documents....11

Table 3-1. Overview of the satellite datasets for emission inventories used in the project and their availability....21

Table 3-2: Methane plume detection and quantification....26

Table 3-3: CO2 point sources detection and quantification....28

Table 3-4: NO_X and SO_2 point sources detection and quantification....31

Table 3-5: NH3, C2H2, C2H4 and CH3OH point sources detection and quantification....42

Table 3-6: Assumptions on the uncertainty in the NO_x and CO emission products detailed in section 3.5 that are used to derive the covariance matrix of the prior uncertainty B over Europe and over Eastern China....52

Figure 3-1: Methane plume detection algorithm. Top: Sentinel-5P Level-2 methane mixing ratio bias corrected data. Bottom-left: mask of the methane anomaly detected at step (2) of the detection algorithm. Bottom-right: 2 methane plumes are disentangled after step (3) of the detection algorithm. Source: [RD.1]....25

Figure 3-2: Methane plume detection and quantification. Left: Sentinel-5P Level-2 methane mixing ratio bias corrected data. Center: polygon of two methane plumes detected is overlaid on the TROPOMI image. Right: corresponding HYSPLIT simulations from which flux rates are determined. Source: [RD.1]....26

Figure 3-3: Relative uncertainty on flux rate estimates. Median relative uncertainty is 45%....27

Figure 3-4: track selection for CO2 flux rate quantification. (a) and (b) display the XCO2 data product retrieved by OCO-2. (c) shows the XCO2 value along-track after quality filtering. (d) displays XCO2 enhancements after background removal. (e) shows XCO2 values as modelled by the Gaussian framework. Source: [RD.10]....28

Figure 3-5: Sample maps of the temporal mean advection, corrected for topography, for (a) the first candidate classified as point source, i.e. the Secunda coal liquefier (South Africa), (b) Niederaußem/Neurath power plants (Germany), (c) the Navajo power plant (USA), and (d) the candidate with lowest derived emissions, i.e. the Al Yamama cement factory (Saudi Arabia). Results of the candidate classification are indicated by triangles for point sources and circles for area sources. The large dashed circle reflects the 15 km radius used for candidate classification procedure as well as for spatial integration. Note the different color scales. Source: [RD.16]....31

Figure 3-6 Mean NOx distribution (summer) for Paris depending on wind direction. The central panel displays the distribution for calm conditions (< 2 m/s). The surrounding panels show the respective patterns for the 8 different wind sectors....36

Figure 3-7 Mean NOx line densities (summer) for Paris depending on wind direction. The panels show the different wind axis. For each axis, the line densities for calm (blue), forward (green) and backward (purple) wind directions are displayed as straight lines. If the lifetime/emission fit succeeds, the corresponding fitted line densities are shown as dashed lines. The fitted emission density is displayed in red. Fitted lifetime and emissions are provided as text in each panel....36

Figure 3-8 Averaging techniques applied to C2H4 IASI measurements over Mengxi Park (Inner Mongolia, China). a, Binned average on a 0.15° × 0.15° spatial resolution grid, b, oversampled average, c, wind-rotated oversampling, and d, wind-rotated supersampling of the IASI C2H4 HRI at a 0.01° × 0.01° spatial resolution, around the hotspot of Mengxi Park (Inner Mongolia, China). At the latitude of the hotspot (39.9° N), this corresponds to a grid resolution of 12.8 × 16.6 km (lon × lat) for a, and of 0.9 × 1.1 km for b-d. The coordinates of the presumed emitter, marked by a white square, is used as the wind rotation centre. Here, the winds are realigned to the east (in the x direction). The area delimited by the black dashed line is used to calculate the averaged downwind value. This figure is taken from [RD.36]....39

Figure 3-9: Oversampling and supersampling-based point sources map. Panel a shows the output of applying oversampling over Canada. Panel b displays the output of wind-rotated supersampling. Black circles represent identified point sources. Sources: [RD.32] 40

Figure 3-10: IASI hyperfine resolution distribution, with hotspots and point-sources of ethylene. The central panel shows a zoom-in of the 13-year IASI average of C2H4 on satellite visible imagery, over an industrial valley of the Shanxi province, China. Hotspots are indicated with black squares. The side panels give examples of close-up views on point-source emitters. Sources: [RD.36] 41

Figure 3-11: Example of C2H4 flux calculation over Secunda (South Africa). The panel a shows the IASI C2H4 column distribution produced by the wind-rotated supersampling, with the point-source (Secunda, South Africa) delimited in white and the area used to calculate the C2H4 background level shaded in semi-transparent. The panel b shows the IASI distribution after subtracting the background column. Sources: [RD.36]. 41

Figure 3-12: Example of TROPOMI enhancement, reconstruction based on prior emissions, and reconstruction based on inversion, in Shanxi (China). The panel a shows the TROPOMI CH4 enhancement (y), panel b shows prior-based methane enhancement reconstruction (Kxp), panel c shows posterior methane enhancement reconstruction (Kx). Source: [RD.41] 43

Figure 3-13 (left) observed composite methane column enhancements over a coal extracting basin in Australia. (right) inversion model fit to these enhancements after optimization of emitting coal mines sources across the entire basin.....44

Figure 3-14 Domain of our CHIMERE configuration and for the inversions in Europe with the averages of the CO concentrations a) observed by MOPITT-v8J and b) simulated by CHIMERE using the Carbon Monitor-CEDS anthropogenic emission estimate where and when MOPITT-v8J "surface" super observations are available, in ppbv, at the 0.5°x0.5° CHIMERE grid-cell resolution, in January 2019. 48

Figure 3-15 Averages of the NO_2 tropospheric columns in Europe a) observed by TROPOMI-PAL and b) simulated by CHIMERE using the Carbon Monitor-CEDS anthropogenic emission estimate where and when TROPOMI-PAL super observations are available, in molec.cm^{-2}, at the 0.5°x0.5° CHIMERE grid-cell resolution, in January 2019. 49

Figure 3-16 Domain of our CHIMERE configuration and for the inversions in Eastern China with the averages of the CO concentrations a) observed by MOPITT-v8J and b) simulated by CHIMERE using the Carbon Monitor-CEDS anthropogenic emission estimate where and when MOPITT-v8J "surface" super observations are available, in ppbv, at the 0.5°x0.5° CHIMERE grid-cell resolution, in January 2019. 49

Figure 3-17 Averages of the NO2 tropospheric columns in Europe a) observed by TROPOMI-PAL and b) simulated by CHIMERE using the Carbon Monitor-CEDS anthropogenic emission estimate where and when TROPOMI-PAL super observations are available, in molec.cm-2, at the 0.5°x0.5° CHIMERE grid- 50

Figure 3-18 Annual NO_x emissions for the year 2019 as estimated by Carbon Monitor-CEDS, in ktNO_2/year, at the 0.5°x0.5° CHIMERE grid-cell resolution. 51

Figure 3-19 Spatial distribution of the global CO budget and 2000–2017 trends. Annual average CO total source and sink during 2000–2017 are shown at the spatial resolution of 3.75° longitude × 1.9° latitude in (a) and (c), respectively, and linear trends of each grid cell are shown in (b) and (d), which are estimated using the linear least squares fitting method based on annual time series.....55

Figure 3-20 Spatial patterns of CH4 emission anomaly between 2020 and 2019 from top-down inversions (left panel) and bottom-up estimates including wetland, fire and anthropogenic emissions (right panel). 56

Figure 3-21 Anomalies of CH4 emissions in 2020 relative to 2019 derived from the ensemble of six inversions. a-f, Spatial patterns of CH4 emission anomalies derived from each member of the ensemble. The global net emission change is given for each panel on the bottom right. g, Spatial pattern of CH4 emission anomalies averaged over the six inversions of the ensemble. The shaded areas indicate that posterior fluxes from all six inversions have the same changing direction. h, Coefficient of variation in the CH4 emission anomalies from the six inversions of the ensemble. For each model grid, the coefficient of variation is defined as the standard deviation (SD) of emission anomalies from six inversions divided by the absolute value of their mean. Darker colors indicate better agreement among inversions. 57

Figure 3-22: Annual averages of the NH3 total columns observed by IASI (left) projected on the model grids, simulated by LMDZ-INCA (middle), and the difference between the model and IASI columns (right) for 2019. 59

Figure 3-23: Spatial distribution of monthly mean β values for April-2019 estimated from LMDZ-INCA model with 40% perturb anthropogenic NH3 emissions. 60

Figure 3-24: Comparison of monthly estimated from IASI observations (right) and bottom-up (CEDS) (left) NH3 emissions of April 2020 using (a) the method described in[RD.94], and (b) the mass-balance approach. 61

Figure 3-25: Illustration of the steps for the global inversion of NO2, SO2, NH3 62

Figure 3-26 Relationship between isoprene emissions and HCHO concentrations over different regions of the US [RD.99]. 63

Figure 3-27: Overall system flow 66

Figure 3-28: Detailed analysis flow for NOx, NH3, SO2. 68

Figure 3-29: Detailed analysis flow for CH4, CO. 69

Figure 3-30: Detailed analysis flow for CO2. 70

1. INTRODUCTION

1.1. PURPOSE

This document evaluates the suitability of all the potential satellite datasets for World Emission, reviews existing inventories of point sources for the considered species, and existing regional and global inversion systems for their performances and applications to retrieve the emissions of the species of interest. We assess the models chosen in World Emission, keeping in mind the fact that they all have a rather unique position in their domain, and we detail the corresponding algorithm for the production of emission estimates.

The objective of the Algorithm Theoretical Basis Document (ATBD) is to develop, verify and detail the necessary algorithms to generate high quality satellite-based products that can be used for the monitoring of emissions. This includes:

- An analysis of existing satellite data in view of their value for verification or improvement of existing emission inventories or the generation of new inventories;
- An analysis of existing top-down inventories including the quality of current algorithms;
- A detailed description of each product in form of an Algorithm Theoretical Basis Document (ATBD), including supporting documentation, corresponding metadata, supporting ancillary information;
- An assessment of the alternative methodological approaches and algorithms for each product;
- A documented justification of the selected methodology, algorithms, and a product generation workflow (input data, end-to-end-processes, output);

1.2. SCOPE

The inversion systems considered here map emissions at various time scales from spaceborne sensors. The species considered are CO_2, CO, CH_4, NO_x, SO_2, NH_3., CH_3OH, C_2H_2, C_2H_4, C_5H_8, PM and H_2O.

This document is structured according to the following sections:

- In section 3.2, the assessment of existing satellite datasets in view of their value for verification or improvement of existing emission inventories or the generation of new inventories is presented,
- In section 3.3, the assessment of point source top-down inversion models and the technical details of the ones selected for the project is presented,
- In section 3.4, the assessment of regional top-down inversion models and the technical details of the ones selected for the project is presented,
- In section 3.5, the assessment of global top-down inversion models and the technical details of the ones selected for the project is presented,
- In section 3.6, the product generation workflow is presented.

1.3. ACRONYMS

Acronyms used in this document and needing a definition are included in the following table:

Table 1-1 Acronyms

Acronym	Definition
ACOS	Atmospheric CO2 Observations from Space
AERONET	Aerosols Robotic NETwork
AMC	Air Mass Corrected
AMF	Air Max Factor
AOD	Aerosol Optical Depth
ASO	August, September, October
BAMS	Bulletin of the American Meteorological Society

Acronym	Definition
BC	Black carbon
BVOC	Biogenic volatile organic compounds
CAMS	Copernicus Atmospheric Monitoring Service
CEDS	Community Emissions Data System
CHIMERE	CHIMERE chemistry-transport model
CHOCHO	Glyoxal
CIF	Community Inversion Framework
CMA	Center of Mass Altitude
COBRA	Covariance-Based Retrieval Algorithm
CSF	Cross-Sectional Flux
CTBT	Comprehensive Nuclear-Test-Ban Treaty
CTM	chemistry transport model
DECSO	Daily Emission estimates Constrained by Satellite Observations
DOAS	Differential Optical Absorption Spectroscopy
ECMWF	European Centre for Medium-Range Weather Forecasts
EDGAR	Emissions Database for Global Atmospheric Research
EMEP	European Monitoring and Evaluation Programme
EMG	Exponentially Modified Gaussian
EOS	Earth Observation System
enKF	ensemble Kalman filter
FFCO2	fossil fuel combustion emissions of CO2
FMA	February, March, Abril
FORLI	Fast Optimal Retrievals on Layer for IASI
FTP	File transfer protocol
GCM	General Circulation Model
GDAL	Geospatial Data Abstraction Library
GDAS	Global Data Assimilation System
GFED	Global Fire Emissions Database
GFS	NCEP's Global Forecast System
GMES	Global Monitoring for Environment and Security
GNFR	Gridding Nomenclature for Reporting
GOME	Global Ozone Monitoring Experiment
GOSAT	Greenhouse gases Observing SATellite
GPPD	Global Public Procurement Database
GPS	Global Positioning System
GSD	Ground Sampling Distance
HCHO	Formaldehyde
HYSPLIT	HYbrid Single-Particle Lagrangian Integrated Trajectory
IASI	Infrared Atmospheric Sounding Interferometer
IBP	Iterative Back-Projection
IME	Integrated Mass Enhancement
INCA	INteraction with Chemistry and Aerosols
KNMI	Royal Netherlands Meteorological Institute
LMDZ	Atmospheric Global Circulation Model of the Laboratoire de Météorologie Dynamique with a zoom capability
MEIC	Multi-resolution Emission Inventory for China
MJJ	May, Jun, July

Acronym	Definition
MODIS	Moderate Resolution Imaging Spectroradiometer
MOPITT	Measurement of Pollution in the Troposphere
NCEP	National Centers for Environmental Prediction
NDJ	November, December, January
NEI	National Emissions Inventory
NetCDF	Network Common Data Form
NIR	Near-infrared
NMHC	Non-Methane HydroCarbons
NMVOC	Non-Methane Volatile Organic Compound
NOX	Nitric Oxide
OC	Organic carbon
OCO-2	Orbiting Carbon Observatory 2
OCO-3	Orbiting Carbon Observatory 3
OMI	Ozone Monitoring Instrument
ORCHIDEE	ORganizing Carbon and Hydrology in Dynamic Ecosystems
PBL	Planetary Boundary Layer
PM	Particulate Matter
POM	Particulate Organic Matter
PSC	Polar Stratospheric Cloud
REAS	Regional Emission inventory in Asia
S5P	Copernicus Sentinel-5 Precursor
SAA	South Atlantic Anomaly
SACS	Simplified Atmospheric Chemistry Scheme
SZA	Solar Zenith Angle
TCCON	Total Carbon Column Observing Network
TES	Technology Experiment Satellite
TIR	Thermal-Infrared
TNO	Netherlands Organisation for Applied Scientific Research
TOVS	TIROS Operational Vertical Sounder
TROPOMI	TROPOspheric Monitoring Instrument
TVCD	Tropospheric Vertical Column Density
UTC	Coordinated Universal Time
VCD	Vertical Column Density
VIIRS	Visible Infrared Imaging Radiometer Suite
VOC	Volatile Organic Compound
VPD	Vapour-pressure Deficit
VZA	Viewing Zenithal Angle
WCD	World City Database
WRF	Weather Research & Forecasting Model

2. REFERENCES

2.1. APPLICABLE DOCUMENTS

The following documents, of the exact issue shown, form part of this document to the extent specified herein. Applicable documents are those referenced in the Contract or approved by the Approval Authority. They are referenced in this document in the form [AD.x]:

Table 2-1 Applicable Documents

Ref.	Title	Code	Version	Date
[AD.1]	Statement of Work. ESA Express Procurement Plus - [EXPRO+] World Emission	EOP-SD-SOW-2019-0101 Tender Reference ESA AO/1-10482/21/I-EF	1	15/04/2021
[AD.2]	World Emission Proposal	GMV 12175/21	2	7/12/2021
[AD.3]	ESA Contract No. 4000137291/22/I-EF	4000137291/22/I-EF	1	15/02/2022

2.2. REFERENCE DOCUMENTS

The following documents, although not part of this document, amplify or clarify its contents. Reference documents are those not applicable and referenced within this document. They are referenced in this document in the form [RD.x]:

Table 2-2 Reference Documents

Ref.	Title	Code	Version	Date
[RD.1]	Lauvaux, T., Giron, C., Mazzolini, M., D'Aspremont, A., Duren, R., Cusworth, D., Shindell, D. &, Ciais, P. (2022). Global assessment of oil and gas methane ultra-emitters, Science, 10.1126/science.abj4351	-	-	2022
[RD.2]	GDAL/OGR contributors. Geospatial Data Abstraction software Library. https://gdal.org (2021).	-	-	2021
[RD.3]	Stein, A.F., Draxler, R.R, Rolph, G.D., Stunder, B.J.B., Cohen, M.D., and Ngan, F., (2015). NOAA's HYSPLIT atmospheric transport and dispersion modelling system, Bull. Amer. Meteor. Soc., 96, 2059-2077, http://dx.doi.org/10.1175/BAMS-D-14-00110.1	-	-	2015
[RD.4]	Pandey, S., Ritesh Gautam, Sander Houweling, Hugo Denier van der Gon, Pankaj Sadavarte, Tobias Borsdorff, Otto Hasekamp, Jochen Landgraf, Paul Tol, Tim van Kempen, Ruud Hoogeveen, Richard van Hees, Steven P. Hamburg, Joannes D. Maasakkers, Ilse Aben: Satellite observations reveal extreme methane leakage from a natural gas well blowout, Proceedings of the National Academy of Sciences Dec 2019, 116 (52) 26376-26381; DOI:10.1073/pnas.1908712116.	-	-	2019
[RD.5]	Varon, D. J., Jacob, D. J., McKeever, J., Jervis, D., Durak, B. O. A., Xia, Y., & Huang, Y. (2018). Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. Atmospheric Measurement Techniques, 11, 5673–5686. https://doi.org/10.5194/amt-11-5673-2018	-	-	2018
[RD.6]	Ehret, T., et al. (2021). Global Tracking and Quantification of Oil and Gas Methane Emissions from Recurrent Sentinel-2 Imagery, https://doi.org/10.48550/arXiv.2110.11832	-	-	2021
[RD.7]	Irakulis-Loitxate, I., et al. (2022). Satellite-based survey of extreme methane emissions in the Permian basin, Science Advances, 10.1126/sciadv.abf4507	-	-	2022
[RD.8]	Barré, J., et al. (2021). Systematic detection of local CH4 emissions anomalies combining satellite measurements and high-resolution forecasts, Atmos. Chem. Phys, https://doi.org/10.5194/acp-2020-550	-	-	2021
[RD.9]	Chevallier, F., B. Zheng, et al. (2020), Local anomalies in the column-averaged dry air mole fractions of carbon dioxide across the globe during the first months of the coronavirus recession, Geophys. Res. Lett, doi:10.1029/2020GL090244	-	-	2020

Ref.	Title	Code	Version	Date
[RD.10]	Zheng, B., Chevallier, F., Ciais, P., Broquet, G., Wang, Y., Lian, J., and Zhao, Y.: Observing carbon dioxide emissions over China's cities and industrial areas with the Orbiting Carbon Observatory-2, Atmos. Chem. Phys., 20, 8501–8510, doi:10.5194/acp-20-8501-2020, 2020	-	-	2020
[RD.11]	Chevallier, F., Broquet, G., Zheng, B., Ciais, P., & Eldering, A. (2022). Large CO2 emitters as seen from satellite: Comparison to a gridded global emission inventory. Geophysical Research Letters, 49, e2021GL097540. https://doi.org/10.1029/2021GL097540	-	-	2022
[RD.12]	Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., & Crisp, D. (2017). Quantifying CO_2 emissions from individual power plants from space. Geophysical Research Letters, 44, 10045–10053. https://doi.org/10.1002/2017gl074702	-	-	2017
[RD.13]	Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., et al. (2019). Towards monitoring localized CO_2 emissions from space: Co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites. Atmospheric Chemistry and Physics, 19, 9371–9383. https://doi.org/10.5194/acp-19-9371-2019	-	-	2019
[RD.14]	Beirle, S. et al. Pinpointing nitrogen oxide emissions from space. Sci. Adv. 5, 1–7 (2019).	-	-	2019
[RD.15]	Beirle, S., Borger, C., Dörner, S., Eskes, H., Kumar, V., de Laat, A., and Wagner, T.: Catalog of NOx emissions from point sources as derived from the divergence of the NO2 flux for TROPOMI, Earth Syst. Sci. Data, 13, 2995–3012, https://doi.org/10.5194/essd-13-2995-2021, 2021	-	-	2021
[RD.16]	Beirle, S., Borger, C., Jost, A., Wagner, T.: Improved catalog of NO_x point source emissions (version 2), Earth Syst. Sci. Data, 15, 3051–3073, https://doi.org/10.5194/essd-15-3051-2023, 2023	-	-	2023
[RD.17]	Eskes, H., van Geffen, J., Sneep, M., Veefkind, P., Niemeijer, S., and Zehner, C.: S5P Nitrogen Dioxide v02.03.01 intermediate reprocessing on the S5P-PAL system: Readme file, [online] Available from: https://data-portal.s5p-pal.com/product-docs/NO2/PAL_reprocessing_NO2_v02.03.01_20211215.pdf	-	-	2021
[RD.18]	Jöckel, P., Tost, H., Pozzer, A., Kunze, M., Kirner, O., Brenninkmeijer, C. A. M., Brinkop, S., Cai, D. S., Dyroff, C., Eckstein, J., Frank, F., Garny, H., Gottschaldt, K.-D., Graf, P., Grewe, V., Kerkweg, A., Kern, B., Matthes, S., Mertens, M., Meul, S., Neumaier, M., Nützel, M., Oberländer-Hayn, S., Ruhnke, R., Runde, T., Sander, R., Scharffe, D. and Zahn, A.: Earth System Chemistry integrated Modelling (ESCiMo) with the Modular Earth Submodel System (MESSy) version 2.51, Geoscientific Model Development, 9(3), 1153–1200, Doi: HTTPs://doi.org/10.5194/gmd-9-1153-2016	-	-	2016
[RD.19]	Byers, L., Friedrich, J., Hennig, R., Kressig, A., Li, X., McCormick, C., and Malaguzzi Valeri, L.: A Global Database of Power Plants, World Resources Institute, Washington, DC. [online] Available from: https://datasets.wri.org/dataset/globalpowerplantdatabase	-	-	2019
[RD.20]	World Cities Database (WCD), provided on simplemaps.com/data/world-cities.	-	-	
[RD.21]	American Meteorological Society: "Advection". Glossary of Meteorology, [online] Available from: http://glossary.ametsoc.org/wiki/Advection	-	-	2012
[RD.22]	de Foy, B. and Schauer, J.: An improved understanding of NOx emissions in South Asian megacities using TROPOMI NO2 retrievals. Environ. Res. Letters 17, https://doi.org/10.1088/1748-9326/ac48b4	-	-	2022
[RD.23]	Sun, K.: Derivation of Emissions from Satellite-Observed Column Amounts and Its Application to TROPOMI NO2 and CO Observations, Geophys. Res. Letters, 49, e2022GL101102, https://doi.org/10.1029/2022GL101102	-	-	2022
[RD.24]	Lange, K., Richter, A., and Burrows, J. P.: Variability of nitrogen oxide emission fluxes and lifetimes estimated from Sentinel-5P TROPOMI observations, Atmos. Chem. Phys., 22, 2745–2767, https://doi.org/10.5194/acp-22-2745-2022	-	-	2022
[RD.25]	Wagner, T., Warnach, S., Beirle, S., Bobrowski, N., Jost, A., Puķīte, J., and Theys, N.: Investigation of 3D-effects for UV/vis satellite and ground based observations of volcanic plumes, Atmos. Meas. Tech. Discuss. [preprint], https://doi.org/10.5194/amt-2022-253, in review	-	-	2022

Ref.	Title	Code	Version	Date
[RD.26]	Theys, N., Fioletov, V., Li, C., De Smedt, I., Lerot, C., McLinden, C., Krotkov, N., Griffin, D., Clarisse, L., Hedelt, P., Loyola, D., Wagner, T., Kumar, V., Innes, A., Ribas, R., Hendrick, F., Vlietinck, J., Brenot, H., and Van Roozendael, M.: A sulfur dioxide Covariance-Based Retrieval Algorithm (COBRA): application to TROPOMI reveals new emission sources, Atmos. Chem. Phys., 21, 16727–16744, https://doi.org/10.5194/acp-21-16727-2021	-	-	2021
[RD.27]	Fioletov, V., McLinden, C. A., Griffin, D., Theys, N., Loyola, D. G., Hedelt, P., Krotkov, N. A., and Li, C.: Anthropogenic and volcanic point source SO2 emissions derived from TROPOMI on board Sentinel-5 Precursor: first results, Atmos. Chem. Phys., 20, 5591–5607, https://doi.org/10.5194/acp-20-5591-2020, 2020	-	-	2020
[RD.28]	Fioletov, V. E., McLinden, C. A., Krotkov, N., Li, C., Joiner, J., Theys, N., Carn, S., and Moran, M. D.: A global catalogue of large SO2 sources and emissions derived from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 16, 11497–11519, https://doi.org/10.5194/acp-16-11497-2016, 2016	-	-	2016
[RD.29]	Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G. and Wagner, T.: Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space, Science, 333(6050), 1737–1739, doi:10.1126/science.1207824, 2011.	-	-	2011
[RD.30]	Liu, F., Beirle, S., Zhang, Q., Dörner, S., He, K. and Wagner, T.: NOx lifetimes and emissions of cities and power plants in polluted background estimated by satellite observations, Atmos. Chem. Phys., 16(8), 5283–5298, doi:10.5194/acp-16-5283-2016, 2016.	-	-	2016
[RD.31]	Liu, F., Tao, Z., Beirle, S., Joiner, J., Yoshida, Y., Smith, S. J., Knowland, K. E., and Wagner, T.: A new method for inferring city emissions and lifetimes of nitrogen oxides from high-resolution nitrogen dioxide observations: a model study, Atmos. Chem. Phys., 22, 1333–1349, https://doi.org/10.5194/acp-22-1333-2022, 2022.	-	-	2022
[RD.32]	Clarisse, L., Van Damme, M., Clerbaux, C., Coheur, P.-F. Tracking down global NH3 point sources with wind-adjusted superresolution. Atmospheric Measurement Techniques 12, 5457-5473,, doi: 10.5194/amt-12-5457-2019, 2019	-	-	2019
[RD.33]	Van Damme, M., Clarisse, L., Whitburn, S., Hadji-Lazaro, J., Hurtmans, D., Clerbaux, C., Coheur, P.-F. Industrial and agricultural ammonia point sources exposed. Nature 564, 99-103, doi: 10.1038/s41586-018-0747-1, 2018	-	-	2018
[RD.34]	Dammers, E., McLinden, C. A., Griffin, D., Shephard, M. W., Van Der Graaf, S., Lutsch, E., Schaap, M., Gainairu-Matz, Y., Fioletov, V., Van Damme, M., Whitburn, S., Clarisse, L., Cady-Pereira, K., Clerbaux, C., Coheur, P. F., and Erisman, J. W.: NH_3 emissions from large point sources derived from CrIS and IASI satellite observations, Atmos. Chem. Phys., 19, 12261–12293, https://doi.org/10.5194/acp-19-12261-2019, 2019.	-	-	2019
[RD.35]	Irani, M. and Peleg, S.: Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency, J. Vis. Commun. Image R., 4, 324–335, https://doi.org/10.1006/jvci.1993.1030, 1993.	-	-	1993
[RD.36]	Franco, B., Clarisse, L., Van Damme, M., Hadji-Lazaro, J., Clerbaux, C., and Coheur, P.: Ethylene industrial emitters seen from space, Nat. Comm., 13, 6452, https://doi.org/10.1038/s41467-022-34098-8, 2022.	-	-	2022
[RD.37]	de Foy, B., Wilkins, J. L., Lu, Z., Streets, D. G. & Duncan, B. N. Model evaluation of methods for estimating surface emissions and chemical lifetimes from satellite data. Atmos. Environ. 98, 66–77 (2014).	-	-	2014
[RD.38]	Zhang, Y.; et al.,, Attribution of the accelerating increase in atmospheric methane during 2010–2018 by inverse analysis of GOSAT observations. Atmospheric Chemistry and Physics 2021, 21, 3643-3666.	-	-	2021
[RD.39]	Shen, L., et al., (2021) Unravelling a large methane emission discrepancy in Mexico using satellite observations, Remote Sensing of Environment, Volume 260, 112461, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2021.112461	-	-	2021

Ref.	Title	Code	Version	Date
[RD.40]	Heald, C. L., et al., (2004), Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, J. Geophys. Res., 109, D23306, doi:10.1029/2004JD005185.	-	-	2004
[RD.41]	Peng, S., et al., High resolution assessment of coal mining methane emissions by satellite in Shanxi, China, EarthArxiv, https://eartharxiv.org/repository/view/3401/ (2022)	-	-	2022
[RD.42]	Stein, A. et al. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. Bulletin of the American Meteorological Society 96, 2059–2077 (2015).	-	-	2015
[RD.43]	Shen, L., et al., Satellite quantification of oil and natural gas methane emissions in the US and Canada including contributions from individual basins, Atmospheric Chemistry and Physics Discussions, 1-31 (2022).	-	-	2022
[RD.44]	Burrows, J. P., et al. : The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results, J. Atmos. Sci., 56(2), 151 – 175, 1999.	-	-	1999
[RD.45]	Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview, Atmos. Meas. Tech., 9, 1279–1301, https://doi.org/10.5194/amt-9-1279-2016, 2016.	-	-	2016
[RD.46]	Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., et al. (1999). SCIAMACHY: Mission objectives and measurement modes. Journal of the Atmospheric Sciences, 56(2), 127–150.	-	-	1999
[RD.47]	Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., van der A, R., McLinden, C., Fioletov, V., Carn, S., de Laat, J., DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., González Abad, G., Arola, A., Boersma, F., Chan Miller, C., Chance, K., de Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H., and Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, Atmos. Chem. Phys., 18, 5699–5745, https://doi.org/10.5194/acp-18-5699-2018, 2018.	-	-	2018
[RD.48]	Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Mao, D., Martínez-Alonso, S., Worden, H. M., Ziskin, D., and Andreae, M. O.: Radiance-based retrieval bias mitigation for the MOPITT instrument: the version 8 product, Atmos. Meas. Tech., 12, 4561–4580, https://doi.org/10.5194/amt-12-4561-2019, 2019.	-	-	2019
[RD.49]	Visser, A. J., Boersma, K. F., Ganzeveld, L. N., and Krol, M. C.: European NOx emissions in WRF-Chem derived from OMI: impacts on summertime surface ozone, Atmospheric Chemistry and Physics, 19, 11 821–11 841, https://doi.org/10.5194/acp-19-11821-2019, 2019	-	-	2019
[RD.50]	Stavrakou, T., Müller, J.-F., Boersma, K. F., van der A, R. J., Kurokawa, J., Ohara, T., and Zhang, Q.: Key chemical NOx sink uncertainties and how they influence top-down emissions of nitrogen oxides, Atmospheric Chemistry and Physics, 13, 9057–9082, https://doi.org/10.5194/acp-13-9057-2013, 2013.	-	-	2013
[RD.51]	Ding, J., van der A, R. J., Mijling, B., and Levelt, P. F.: Space-based NOx emission estimates over remote regions improved in DECSO, Atmos. Meas. Tech., 10, 925–938, https://doi.org/10.5194/amt-10-925-2017, 2017.	-	-	2017
[RD.52]	Konovalov, I. B., Berezin, E. V., Ciais, P., Broquet, G., Zhuravlev, R. V., and Janssens-Maenhout, G.: Estimation of fossil-fuel CO2 emissions using satellite measurements of "proxy" species, Atmos. Chem. Phys., 16, 13509–13540, https://doi.org/10.5194/acp-16-13509-2016, 2016.	-	-	2016
[RD.53]	Fortems-Cheiney, A., and Broquet, G.: Final re-analysis of the national scale CO2 anthropogenic emissions over 2005-2015, VERIFY public deliverable D2.12, https://verify.lsce.ipsl.fr/index.php/repository/public-deliverables/wp2-verification-methods-for-fossil-co2-emissions/d2-12-final-re-analysis-of-the-national-scale-co2-anthropogenic-emissions-over-2005-2015	-	-	2022

Ref.	Title	Code	Version	Date
[RD.54]	Berchet, A., Sollum, E., Thompson, R. L., Pison, I., Thanwerdas, J., Broquet, G., Chevallier, F., Aalto, T., Berchet, A., Bergamaschi, P., Brunner, D., Engelen, R., Fortems-Cheiney, A., Gerbig, C., Groot Zwaaftink, C. D., Haussaire, J.-M., Henne, S., Houweling, S., Karstens, U., Kutsch, W. L., Luijkx, I. T., Monteil, G., Palmer, P. I., van Peet, J. C. A., Peters, W., Peylin, P., Potier, E., Rödenbeck, C., Saunois, M., Scholze, M., Tsuruta, A., and Zhao, Y.: The Community Inversion Framework v1.0: a unified system for atmospheric inversion studies, Geosci. Model Dev., 14, 5331–5354, https://doi.org/10.5194/gmd-14-5331-2021, 2021.	-	-	2021
[RD.55]	Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G., Turquety, S., Valari, M., Vautard, R., and Vivanco, M. G.: CHIMERE 2013: a model for regional atmospheric composition modelling, Geosci. Model Dev., 6, 981–1028, https://doi.org/10.5194/gmd-6-981-2013, 2013.	-	-	2013
[RD.56]	Mailler, S., Menut, L., Khvorostyanov, D., Valari, M., Couvidat, F., Siour, G., Turquety, S., Briant, R., Tuccella, P., Bessagnet, B., Colette, A., Létinois, L., Markakis, K., and Meleux, F.: CHIMERE-2017: from urban to hemispheric chemistry-transport modeling, Geosci. Model Dev., 10, 2397–2423, https://doi.org/10.5194/gmd-10-2397-2017, 2017.	-	-	2017
[RD.57]	Fortems-Cheiney, A., Pison, I., Broquet, G., Dufour, G., Berchet, A., Potier, E., Coman, A., Siour, G., and Costantino, L.: Variational regional inverse modeling of reactive species emissions with PYVAR-CHIMERE-v2019, Geosci. Model Dev., 14, 2939–2957, https://doi.org/10.5194/gmd-14-2939-2021, 2021.	-	-	2021
[RD.58]	Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, 2006. https://doi.org/10.5194/acp-6-3181-2006.	-	-	2006
[RD.59]	Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q. L., Rozemeijer, N. C., Veefkind, J. P., and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, Atmos. Meas. Tech., 10, 1957–1986, https://doi.org/10.5194/amt-10-1957-2017, 2017.	-	-	2017
[RD.60]	Boersma, K. F., Eskes, H., Richter, A., De Smedt, I., Lorente, A., Beirle, S., Van Geffen, J., Peters, E., Van Roozendael, M. and Wagner, T.: QA4ECV NO2 tropospheric and stratospheric vertical column data from OMI (Version 1.1) [Data set], Royal Netherlands Meteorological Institute (KNMI), 2017. http://doi.org/10.21944/qa4ecv-no2-omi-v1.1	-	-	2017
[RD.61]	Veefkind, J., Aben, I., McMullan, K., Forster, H., de Vries, J., Otter, G., Levelt, P.: TROPOMI on the esa sentinel-5 precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. Remote Sensing of Environment, 120, 70-83, 2012. https://doi.org/10.1016/j.rse.2011.09.027	-	-	2012
[RD.62]	van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., Maasakkers, J. D., and Veefkind, J. P.: TROPOMI ATBD of the total and tropospheric NO2 data products, Report S5P-KNMI-L2-0005-RP, version 2.1.0, to be released, KNMI, De Bilt, the Netherlands, available at: http://www.tropomi.eu/documents/atbd/ (last access: 4 December 2019), 2020.	-	-	2020
[RD.63]	Cholakian, A., Beekmann, M., Coll, I., Ciarelli, G., and Colette, A.: Biogenic secondary organic aerosol sensitivity to organic aerosol simulation schemes in climate projections, Atmos. Chem. Phys., 19, 13209-13226, 2019. https://doi.org/10.5194/acp-19-13209-2019	-	-	2019
[RD.64]	Ciarelli, G., Theobald, M. R., Vivanco, M. G., Beekmann, M., Aas, W., Andersson, C., Bergstrom, R., Manders-Groot, A., Couvidat, F., Mircea, M., Tsyro, S., Fagerli, H., Mar, K., Raffort, V., Roustan, Y., Pay, M.-T., Schaap, M., Kranenburg, R., Adani, M., Briganti, G., Cappelletti, A., D'Isidoro, M., Cuvelier, C., Cholakian, A., Bessagnet, B., Wind, P., and Colette, A.: Trends of inorganic and organic aerosols and precursor gases in Europe: insights from the EURODELTA multi-model experiment over the 1990-2010 period, Geosci. Model Dev., 12, 4923-4954, 2019. https://doi.org/10.5194/gmd-12-4923-2019	-	-	2019
[RD.65]	Menut, L., Bessanet, B., Siour, G., Mailler, S., Pennel, R. and Cholakian, A. : Impact of lockdown measures to combat Covid-19 on air quality over western Europe, Science of The Total Environment, 741, 2020. https://doi.org/10.1016/j.scitotenv.2020.140426	-	-	2020

Ref.	Title	Code	Version	Date
[RD.66]	Savas, D.; Dufour, G.; Coman, A.; Siour, G.; Fortems-Cheiney, A.; Broquet, G.; Pison, I.; Berchet, A.; Bessagnet, B.: Anthropogenic NOx Emission Estimations over East China for 2015 and 2019 Using OMI Satellite Observations and the New Inverse Modeling System CIF-CHIMERE. Atmosphere, 14, 154. https://doi.org/10.3390/atmos14010154, 2023.	-	-	2023
[RD.67]	Liu, Z., Ciais, P., Deng, Z. et al.: Sci Data 7, 392, https://doi.org/10.1038/s41597-020-00708-7, 2020.	-	-	2020
[RD.68]	Super, I., Dellaert, S. N. C., Visschedijk, A. J. H., and Denier van der Gon, H. A. C.: Uncertainty analysis of a European high-resolution emission inventory of CO2 and CO to support inverse modelling and network design, Atmos. Chem. Phys., 20, 1795–1816, 2020. https://doi.org/10.5194/acp-20-1795-2020	-	-	2020
[RD.69]	Gilbert, J. C., and C. Lemaréchal: Some numerical experiments with variable-storage quasi-Newton algorithms, Math. Program., 45, 407–435, 1989. https://link.springer.com/article/10.1007/BF01589113	-	-	1989
[RD.70]	Zhu, L. et al. Constraining U.S. ammonia emissions using TES remote sensing observations and the GEOS-Chem adjoint model. J. Geophys. Res. Atmos. 118, 3355–3368 (2013).	-	-	2013
[RD.71]	Whitburn, S., Van Damme, M., Kaiser, J.W., van der Werf, G.R., Turquety, S., Hurtmans, D., Clarisse, L., Clerbaux, C. and Coheur, P.F., 2015. Ammonia emissions in tropical biomass burning regions: Comparison between satellite-derived emissions and bottom-up fire inventories. Atmospheric Environment, 121, pp.42-54.	-	-	2015
[RD.72]	Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., et al. (2020). LMDZ6A: The atmospheric component of the IPSL climate model with improved and better tuned physics. Journal of Advances in Modeling Earth Systems, 12, e2019MS001892. https://doi.org/10.1029/2019MS001892	-	-	2020
[RD.73]	Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., & Bastrikov, V., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. Journal of Advances in Modeling Earth Systems, 12, e2019MS002010. https://doi.org/10.1029/2019MS002010	-	-	2020
[RD.74]	Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M.-A., Walters, S., Lamarque, J.-F., and Holland, E. A. (2004), Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: Description and background tropospheric chemistry evaluation, J. Geophys. Res., 109, D04314, doi:10.1029/2003JD003957.	-	-	2004
[RD.75]	Hauglustaine, D. A., Balkanski, Y., & Schulz, M. (2014). Simulation of present and future nitrate aerosols and their direct radiative forcing of climate. Atmospheric Chemistry and Physics, 14, 11,031–11,063. https://doi.org/10.5194/acp-14-11031-2014	-	-	2014
[RD.76]	Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C. (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cycles, 19, GB1015, doi:10.1029/2003GB002199.	-	-	2005
[RD.77]	Hourdin, F., and J. P. Issartel (2000), Sub-surface nuclear tests monotoring through the CTBT xenon network, Geophys. Res. Lett., 27, 2245–2248.	-	-	2000
[RD.78]	Folberth, G. A., Hauglustaine, D. A., Lathière, J., and Brocheton, F.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: model description and impact analysis of biogenic hydrocarbons on tropospheric chemistry, Atmos. Chem. Phys., 6, 2273–2319, https://doi.org/10.5194/acp-6-2273-2006, 2006.	-	-	2006
[RD.79]	Bauer, S. E., Y. Balkanski, M. Schulz, D. Hauglustaine, and F. Dentener (2004), Global modeling of heterogeneous chemistry on mineral aerosol surfaces: Influence on tropospheric ozone chemistry and comparison to observations, J. Geophys. Res., 109, D02304, doi:10.1029/2003JD003868.	-	-	2004
[RD.80]	Lefèvre, F., G. P. Brasseur, I. Folkins, A. K. Smith, and P. Simon, Chemistry of the 1991-1992 stratospheric winter: threedimensional model simulations, J. Geophys. Res., 99, 8183, 1994.	-	-	1994

Ref.	Title	Code	Version	Date
[RD.81]	McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., Zheng, B., Crippa, M., Brauer, M., and Martin, R. V.: A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS), Earth Syst. Sci. Data, 12, 3413–3442, https://doi.org/10.5194/essd-12-3413-2020, 2020.	-	-	2020
[RD.82]	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 Syst. Sci. Data, 9, 697–720, https://doi.org/10.5194/essd-9-697-2017, 2017.	-	-	2017
[RD.83]	Lathière, J., D. A. Hauglustaine, N. De Noblet-Ducoudré, G. Krinner, et G. A. Folberth, Past and future changes in biogenic volatile organic compound emissions simulated with a global dynamic vegetation model, Geophys. Res. Lett., 32, L20818, doi:10.1029/2005GL024164, 2005	-	-	2005
[RD.84]	Jourdain, L. et D. A. Hauglustaine, The global distribution of lightning NOx simulated on-line in a general circulation model, Phys. Chem. Earth, 26, 585-591, 2001	-	-	2001
[RD.85]	Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P. (2005), Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res., 110, D24309, doi:10.1029/2005JD006390.	-	-	2005
[RD.86]	Pison, I., P. Bousquet, F. Chevallier, S. Szopa and D. A. Hauglustaine, 2009: Multi-species inversion of CH4, CO and H2 emissions from surface measurements, Atmos. Chem. Phys., 9, 5281-5297, doi:10.5194/acp-9-5281-2009	-	-	2009
[RD.87]	Fortems-Cheiney, A., F. Chevallier, I. Pison, P. Bousquet, M. Saunois, S. Szopa, C. Cressot, T.P. Kurosu, K. Chance and A. Fried, 2012: The formaldehyde budget as seen by a global-scale multi-constraint and multi-species inversion system. Atmos. Chem. Phys., 12, 6699-6721, doi:10.5194/acp-12-6699-2012	-	-	2012
[RD.88]	Yin, Y., F. Chevallier, P. Ciais, G. Broquet, A. Fortems-Cheiney, I. Pison, and M. Saunois, 2015: Decadal trends in global CO emissions as seen by MOPITT. Atmos. Chem. Phys., 15, 13433-13451, doi:10.5194/acp-15-13433-2015.	-	-	2015
[RD.89]	Zheng, B., Chevallier, F., Yin, Y., Ciais, P., Fortems-Cheiney, A., Deeter, M. N., Parker, R. J., Wang, Y., Worden, H. M., and Zhao, Y., 2019: Global atmospheric carbon monoxide budget 2000–2017 inferred from multi-species atmospheric inversions, Earth Syst. Sci. Data, 11, 1411–1436, doi:10.5194/essd-11-1411-2019.	-	-	2019
[RD.90]	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, 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 Syst. Sci. Data, 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020, 2020.	-	-	2020
[RD.91]	Peng, S., Lin, X., Thompson, R.L. et al. Wetland emission and atmospheric sink changes explain methane growth in 2020. Nature612, 477–482 (2022). https://doi.org/10.1038/s41586-022-05447-w	-	-	2022

Ref.	Title	Code	Version	Date
[RD.92]	Zheng, B., Chevallier, F., Ciais, P., Yin, Y., Deeter, M., Worden, H., Wang, Y. L., Zhang, Q., and He, K. B.: Rapid decline in carbon monoxide emissions and export from East Asia between years 2005 and 2016, Environ. Res. Lett., 13, 044007, https://doi.org/10.1088/1748-9326/aab2b3, 2018a.	-	-	2018
[RD.93]	Zheng, B., Chevallier, F., Ciais, P., Yin, Y., and Wang, Y.: On the role of the flaming to smoldering transition in the seasonal cycle of African fire emissions. Geophys. Res. Lett., 45, 11,998–12,007. https://doi.org/10.1029/ 2018GL079092, 2018b	-	-	2018
[RD.94]	Zheng, B., Geng, G., Ciais, P., Davis, S. J., Martin, R. v., Meng, J., Wu, N., Chevallier, F., Broquet, G., Boersma, F., van Der, R. A., Lin, J., Guan, D., Lei, Y., He, K., & Zhang, Q. (2020). Satellite-based estimates of decline and rebound in China's CO2 emissions during COVID-19 pandemic. Science Advances, 6(49). https://doi.org/10.1126/SCIADV.ABD4998/SUPPL_FILE/ABD4998_SM.PDF	-	-	2020
[RD.95]	Lamsal, L. N., Martin, R. V., Padmanabhan, A., Van Donkelaar, A., Zhang, Q., Sioris, C. E., Chance, K., Kurosu, T. P., & Newchurch, M. J. (2011). Application of satellite observations for timely updates to global anthropogenic NOx emission inventories. *Geophysical Research Letters*, *38*(5). https://doi.org/10.1029/2010GL046476	-	-	2011
[RD.96]	Cooper, M., Martin, R. V., Padmanabhan, A., & Henze, D. K. (2017). Comparing mass balance and adjoint methods for inverse modeling of nitrogen dioxide columns for global nitrogen oxide emissions. *Journal of Geophysical Research: Atmospheres*, *122*(8), 4718–4734. https://doi.org/10.1002/2016JD025985	-	-	2017
[RD.97]	Evangeliou, N., Balkanski, Y., Eckhardt, S., Cozic, A., Van Damme, M., Coheur, P.F., Clarisse, L., Shephard, M.W., Cady-Pereira, K.E. and Hauglustaine, D., 2021. 10-year satellite-constrained fluxes of ammonia improve performance of chemistry transport models. Atmospheric Chemistry and Physics, 21(6), pp.4431-4451. https://doi.org/10.5194/acp-21-4431-2021	-	-	2021
[RD.98]	Hoesly, R., O'Rourke, P., Braun, C., Feng, L., Smith, S. J., Pitkanen, T., Seibert, J. J., Vu, L., Presley, M., Bolt, R., Goldstein, B., & Kholod, N. (2019). Community Emissions Data System. https://doi.org/10.5281/ZENODO.3592073	-	-	2019
[RD.99]	Palmer, Paul I., Daniel J. Jacob, Arlene M. Fiore, Randall V. Martin, Kelly Chance, and Thomas P. Kurosu. 2003. "Mapping Isoprene Emissions over North America Using Formaldehyde Column Observations from Space." Journal of Geophysical Research 108 (D6). doi:10.1029/2002jd002153.	-	-	2003
[RD.100]	Messina, P., Lathière, J., Sindelarova, K., Vuichard, N., Granier, C., Ghattas, J., Cozic, A., and Hauglustaine, D. A.: Global biogenic volatile organic compound emissions in the ORCHIDEE and MEGAN models and sensitivity to key parameters, Atmos. Chem. Phys., 16, 14169–14202, https://doi.org/10.5194/acp-16-14169-2016, 2016.	-	-	2016
[RD.101]	Kok, J., Ridley, D., Zhou, Q. et al. Smaller desert dust cooling effect estimated from analysis of dust size and abundance. Nature Geosci 10, 274–278 (2017). https://doi.org/10.1038/ngeo2912	-	-	2017
[RD.102]	Torres, O., Jethva, H., Ahn, C., Jaross, G., and Loyola, D. G.: TROPOMI aerosol products: evaluation and observations of synoptic-scale carbonaceous aerosol plumes during 2018–2020, Atmos. Meas. Tech., 13, 6789–6806, https://doi.org/10.5194/amt-13-6789-2020, 2020.	-	-	2020
[RD.103]	Zhao S, Yu Y, Yin D, He J, Liu N, Qu J, Xiao J. Annual and diurnal variations of gaseous and particulate pollutants in 31 provincial capital cities based on in situ air quality monitoring data from China National Environmental Monitoring Center. Environ Int. 2016 Jan; 86:92-106. doi: 10.1016/j.envint.2015.11.003. Epub 2015 Nov 9. PMID: 26562560.	-	-	2015
[RD.104]	Huang, RJ., Zhang, Y., Bozzetti, C. et al. High secondary aerosol contribution to particulate pollution during haze events in China. Nature 514, 218–222 (2014). https://doi.org/10.1038/nature13774	-	-	2014
[RD.105]	Holben, B. N., et al. (1998), AERONET–A federated instrument networkand data archive for aerosol characterization, Remote Sens. Environ.,66,1–16, doi:10.1016/S0034-4257(98)00031-5	-	-	1998

Ref.	Title	Code	Version	Date
[RD.106]	Sinyuk, A., Dubovik, O., Holben, B. N., Eck, T. F., Breon, F. M., Martonchik, J., Kahn, R., Diner, D. J., Vermote, E. F., Roger, J. C., Lapyonok, T., and Slutsker, I.: Simultaneous retrieval of aerosol and surface properties from combination of AERONERT and satellite data, Remote Sens. Environ., 107, 90–108, 2007.	-	-	2007
[RD.107]	Borger, C., Beirle, S., and Wagner, T.: A 16-year global climate data record of total column water vapour generated from OMI observations in the visible blue spectral range, Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-2021-319, in review, 2021.	-	-	2021
[RD.108]	Küchler, T., Noël, S., Bovensmann, H., Burrows, J. P., Wagner, T., Borger, C., Borsdorff, T., and Schneider, A.: Total water vapour columns derived from Sentinel 5P using the AMC-DOAS method, Atmos. Meas. Tech., 15, 297–320, https://doi.org/10.5194/amt-15-297-2022, 2022.	-	-	2022
[RD.109]	Alhajeri, H.M.; Almutairi, A.; Alenezi, A.; Alshammari, F. Energy Demand in the State of Kuwait During the Covid-19 Pandemic: Technical, Economic, and Environmental Perspectives. Energies 2020, 13, 4370. https://doi.org/10.3390/en13174370	-	-	2020
[RD.110]	Mora, C., Dousset, B., Caldwell, I. et al. Global risk of deadly heat. Nature Clim Change 7, 501–506 (2017). https://doi.org/10.1038/nclimate3322	-	-	2017
[RD.111]	Yuan, W. et al. Increased atmospheric vapor pressure deficit reduces global vegetation growth. Sci. Adv. 5, eaax1396 (2019).	-	-	2019
[RD.112]	Green, J. K., Ballantyne, A., Abramoff, R., Gentine, P., Makowski, D., & Ciais, P. (2022). Surface temperatures reveal the patterns of vegetation water stress and their environmental drivers across the tropical Americas. Global Change Biology, 28, 2940– 2955. https://doi.org/10.1111/gcb.16139	-	-	2022
[RD.113]	D3 - Algorithm Theoretical Basis Document v2	GMV 23087 /22 V1/22	3	17/04/2023
[RD.114]	Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, Atmos. Chem. Phys., 15, 11411–11432, https://doi.org/10.5194/acp-15-11411-2015, 2015.	-	-	2015

3. ALGORITHM THEORETICAL BASIS

3.1. INTRODUCTION

This document consolidates a reliable design for the backend resources of the World Emission service. It was written after assessing the current state-of-the-art and doing a comprehensive review of the currently available datasets and models. It is structured into five parts that have been written in parallel by the corresponding partners and then have been aggregated together. The first part briefly lists the satellite datasets to be used by the service, the second part describes the point source inversion models. The regional top-down inversion models and the global ones are presented in the third and fourth parts, respectively. Finally, the fifth part describe how the generation of the different products is integrated within a single workflow. This version 3 of the ATBD is the update of version 2 (last update in April 2023 [RD.113]) and the final one of the project.

3.2. ASSESSMENT OF SATELLITE DATASETS

[Minor update with respect to previous version]

The list of existing satellite datasets to be used by World Emission in view of their value for verification or improvement of existing emission inventories or the generation of new inventories has been reviewed. These satellite datasets are listed in Table 3-1 (below), including the type of data (e.g., data unit, temporal resolution), the time period and spatial coverage that are currently available, where the data can be found or downloaded, and the main references describing the datasets.

All the satellite datasets listed in Table 3-1 are Level-2 data, either publicly accessible or already available to the partners of the World Emission consortium. These datasets consist of multi-annual measurements made by satellite instruments achieving a (near-)global coverage on a daily basis. The data are available as total or tropospheric columns, and as vertical profiles for some instruments. Hence, these satellite datasets are suitable for inversion of gas emissions and allow fulfilling the goals of World Emission.

Specifically, the comprehensive TROPOMI CH_4, OCO-2 and OCO-3 CO_2, OMI and TROPOMI NO_2, and IASI NH_3 satellite products are used to generate top-down emission fluxes at three different scales, i.e., for point sources and hotspots, regions of interest, and at the global scale. The recent TROPOMI SO_2 product will serve to inverse fluxes globally and for a large suite of hotspots, whereas the CO products from the TROPOMI, IASI and MOPITT instruments will be assimilated by models to produce global and regional emissions. Regarding the NMVOC species, top-down emissions of anthropogenic point sources (mainly in Asia) will be derived from the IASI CH_3OH, C_2H_2 and C_2H_4 products. Moreover, the IASI CH_3OH, and TROPOMI HCHO and CHOCHO data will be used to evaluate model results and could provide additional constraints on isoprene emissions. The different methods that will be implemented to calculate the top-down fluxes and assimilate the satellite data are described in the next sections (3.3, 3.4 and 3.5).

Table 3-1. Overview of the satellite datasets for emission inventories used in the project and their availability.

Species	Satellite instrument	Data level	Data Product	Available time period	Spatial coverage	Temporal resolution	Data type / units	Available at / from	Reference
CO_2	OCO-2	2	ACOS bias-corrected L2 retrievals	2014-ongoing	global	daily	total column	https://oco2.gesdisc.eosdis.nasa.gov/	https://docserver.gesdisc.eosdis.nasa.gov/public/project/OCO/OCO2_OCO3_B10_DUG.pdf
	OCO-3	2	ACOS bias-corrected L2 retrievals	2019-ongoing	global	daily	total column	https://oco2.gesdisc.eosdis.nasa.gov/	https://docserver.gesdisc.eosdis.nasa.gov/public/project/OCO/OCO2_OCO3_B10_DUG.pdf
CH4	TROPOMI	2	Level 2 Offline Methane mixing ratio bias corrected	2019-ongoing	global (limited quality over water and tropics)	daily	concentrations (total column ppb)	C. Giron, Kayrros (c.giron@kayrros.com)	Lauvaux et al. (2022), Science, 375 (6580), 557-561, https://doi.org/10.1126/science.abj4351
CO	IASI	2	FORLI-CO v20151001	2007-ongoing	global	daily	total column and profiles	Through the AERIS portal at https://iasi.aeris-data.fr/co/	Hurtmans et al. (2012), J. Quant. Spectrosc. Ra., 113, 1391–1408, https://doi.org/10.1016/j.jqsrt.2012.02.036 George et al. (2015), Atmos. Meas. Tech., 8, 4313-4328, https://doi.org/10.5194/amt-8-4313-2015
		2	MOPITT-v8	2000-Dec 2021	global	2-3 days	total column and profiles	Through NASA's EarthData portal at https://earthdata.nasa.gov/	Deeter et al. (2019), Atmos. Meas. Tech., 12, 4561–4580, https://doi.org/10.5194/amt-12-4561-2019
		2	MOPITT-v9	2000-ongoing	global	2-3 days	total column and profiles	Through NASA's EarthData portal at https://earthdata.nasa.gov/	Deeter et al. (2022), Atmos. Meas. Tech., 15, 2325–2344, https://doi.org/10.5194/amt-15-2325-2022, 2022.
	TROPOMI	2	To be decided in the course of the project One test may be done by end of phase 2 but no inversion based on TROPOMI CO	2018-ongoing	global	daily	total column	Copernicus Open Access Hub at https://scihub.copernicus.eu/	Veefkind et al. (2012), Remote Sen. Environ., 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027 Landgraf et al. (2016), Atmos. Meas. Tech., 9, 4955–4975, https://doi.org/10.5194/amt-9-4955-2016

Species	Satellite instrument	Data level	Data Product	Available time period	Spatial coverage	Temporal resolution	Data type / units	Available at / from	Reference
NO_x	OMI	2	OMI-QA4ECV	2004-March 2021	global	daily	tropospheric column	Through the QA4ECV portal at http://www.qa4ecv.eu/ecv/no2-pre/data (the QA4ECV product has been interrupted intentionally in March 2021. The next OMI product targeted for the regional inversions (OMI "collection 4") is not available yet but its distribution is currently planned within few months.	Levelt et al. (2006), IEEE T. Geosci. Remote, https://doi.org/10.1109/TGRS.2006.872333 Boersma et al. (2017), [Data set], KNMI, http://doi.org/10.21944/qa4ecv-no2-omi-v1.1 Zara et al. (2018), Atmos. Meas. Tech., 11, 4033–4058, https://doi.org/10.5194/amt-11-4033-2018
	TROPOMI	2	OFFL v02	Jul 2021-ongoing	global	daily	tropospheric column	Copernicus Open Access Hub at https://scihub.copernicus.eu/	Veefkind et al. (2012), Remote Sen. Environ., 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027 Copernicus Sentinel-5P (processed by ESA), 2021, TROPOMI Level 2 Nitrogen Dioxide total column products. Version 02. European Space Agency. https://doi.org/10.5270/S5P-9bnp8q8
			PAL	2018-Nov 2021	global	daily	tropospheric column	https://data-portal.s5p-pal.com	Veefkind et al. (2012), Remote Sen. Environ., 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027 Eskes et al., https://data-portal.s5p-pal.com/product-docs/no2/PAL_reprocessing_NO2_v02.03.01_20211215.pdf
			v2.4	2019-ongoing	global	daily	tropospheric column	Copernicus Open Access Hub at https://scihub.copernicus.eu/	Van Geffen, J., Eskes, H., K.F., B., and Veefkind, J.: TROPOMI ATBD of the total and tropospheric NO2 data products, Tech. rep., https://sentinel.esa.int/documents/247904/2476257/sentinel-5p-tropomi-atbd-no2-data-products, 2022
SO_2	TROPOMI	2	COBRA	2018-ongoing	global	daily	total column	COBRA SO_2 data provided by N. Theys (nicolas.theys@aeronomie.be); included in the S5P PAL since summer 2022	Theys et al. (2021), Atmos. Chem. Phys., 21, 16727–16744, https://doi.org/10.5194/acp-21-16727-2021

Species	Satellite instrument	Data level	Data Product	Available time period	Spatial coverage	Temporal resolution	Data type / units	Available at / from	Reference
NH_3	IASI	2	Level 2, reanalyzed with the ERA-5 meteorological fields	2007-ongoing	global	daily	total column	M. Van Damme, ULB (Martin.Van.Damme@ulb.be) L. Clarisse, ULB (lieven.clarisse@ulb.be)	Van Damme et al. (2021), Environ. Res. Lett., 16, 055017, https://doi.org/10.1088/1748-9326/abd5e0
CH_3OH	IASI	2	Level 2, reanalyzed with the ERA-5 meteorological fields	2007-ongoing	global	daily	total column	B. Franco, ULB (bruno.franco@ulb.be) L. Clarisse, ULB (lieven.clarisse@ulb.be)	Franco et al. (2018), J. Geophys. Res. - Atmos., 123, 13963-84, https://doi.org/10.1029/2018JD029633
C_2H_2	IASI	2	Level 2, reanalyzed with the ERA-5 meteorological fields	2007-ongoing	global (limited quality over water)	daily	total column	B. Franco, ULB (bruno.franco@ulb.be) L. Clarisse, ULB (lieven.clarisse@ulb.be)	/
C_2H_4	IASI	2	Level 2, reanalyzed with the ERA-5 meteorological fields	2007-ongoing	global (limited quality over water)	daily	total column	B. Franco, ULB (bruno.franco@ulb.be) L. Clarisse, ULB (lieven.clarisse@ulb.be)	Franco et al. (2022), Nat. Comm., 13, 6452, https://www.nature.com/articles/s41467-022-34098-8
HCHO	OMI	2	QA4ECV	2004-ongoing	global	daily	tropospheric column	Through the QA4ECV portal at http://www.qa4ecv.eu/ecv/hcho-p/data	De Smedt et al. (2017), [Data set], QA4ECV, https://doi.org/10.18758/71021031
	TROPOMI	2	To be decided in the course of the project One test may be done by end of phase 2 but no inversion of isoprene emissions based on TROPOMI HCHO	2018-ongoing	global	daily	tropospheric column	Copernicus Open Access Hub at https://scihub.copernicus.eu/	https://sentinels.copernicus.eu/web/sentinel/data-products/-/asset_publisher/fp37fc19FN8F/content/sentinel-5-precursor-level-2-formaldehyde

3.3. ASSESSMENT OF POINT SOURCE TOP-DOWN INVERSION MODELS AND DESCRIPTION OF THE ALGORITHMS USED IN WORLD EMISSION

3.3.1. INTRODUCTION

The purpose of this section is to review existing inventories of point sources for the considered species. The methods for the point source inventories to be included in the World Emission platform are detailed and, when alternative inventories are available, a comparative analysis is conducted.

The inventories considered in this comparative analysis are point source inventories derived from spaceborne sensors. The species considered are CO_2, CH_4, NO_x, SO_2, NH_3, C_2H_2, C_2H_4 and CH_3OH.

3.3.2. METHANE (CH4) POINT SOURCES INVENTORY

The methodology used to build the World Emission methane point sources is detailed in the peer-reviewed research article [RD.1]. The World Emission methane point sources inventory is based on the TROPOMI data. TROPOMI is a hyperspectral sensor onboard the Sentinel-5P satellite which orbits the earth on a near-polar, sun-synchronous trajectory and senses most of the Earth surface on a daily basis by means of its 2600-kilometer-wide swath. Its ground sampling distance is 5.5 × 7 km.

3.3.2.1. POINT SOURCES INVERSION

The method consists of a detection phase and a quantification step. In the detection phase, each TROPOMI Level-2 methane mixing ratio bias corrected data product is processed in a detection pipeline relying on classical computer vision techniques. The images are first reprojected on a regular grid in GPS coordinates using the standard GDAL library [RD.2]. Reprojected image pixels are filtered to only take into account the measures that results from interpolation of high-quality pixels (QA-value above 75). The computed-vision operations hereafter applied to detect methane plumes are the following:

1. the image is denoised using gaussian filtering;
2. local methane anomalies are identified and delineated using patch-based anomaly detection. On each patch, an anomaly map is defined as $AnomalyMap = Patch - background - k \times StandardDeviation$ and anomalies at defined as contiguous pixels flagged by an *a contrario* anomaly detection;
3. contiguous but distinct methane plumes are disentangled using sharpening kernel based saddle point detection. The saddle points are emphasized by applying to the image a convolutional kernel of size (3, 3) for which sum of coefficients is 0, the center pixel is negative and the surrounding pixels are equal and positive. The convoluted image will have higher values at local minima and saddle points. Local minima are by definition not part of methane plumes, and saddle point are possible delineations between two contiguous plumes (see Figure 3-2);

Methane plumes automatically detected are subjected to human check. The purpose of this human check is to remove spurious detections, by checking that the plumes is consistent with the wind direction at acquisition time (from ECMWF-ERA5 meteorological data - computing angle between plume direction and wind direction, and asserting that it is below some threshold) and that there is no substantial correlation between the methane plume detected and the albedo measure retrieved as an auxiliary information in the TROPOMI Level-2 data product. The detection phase is illustrated in Figure 3-1.

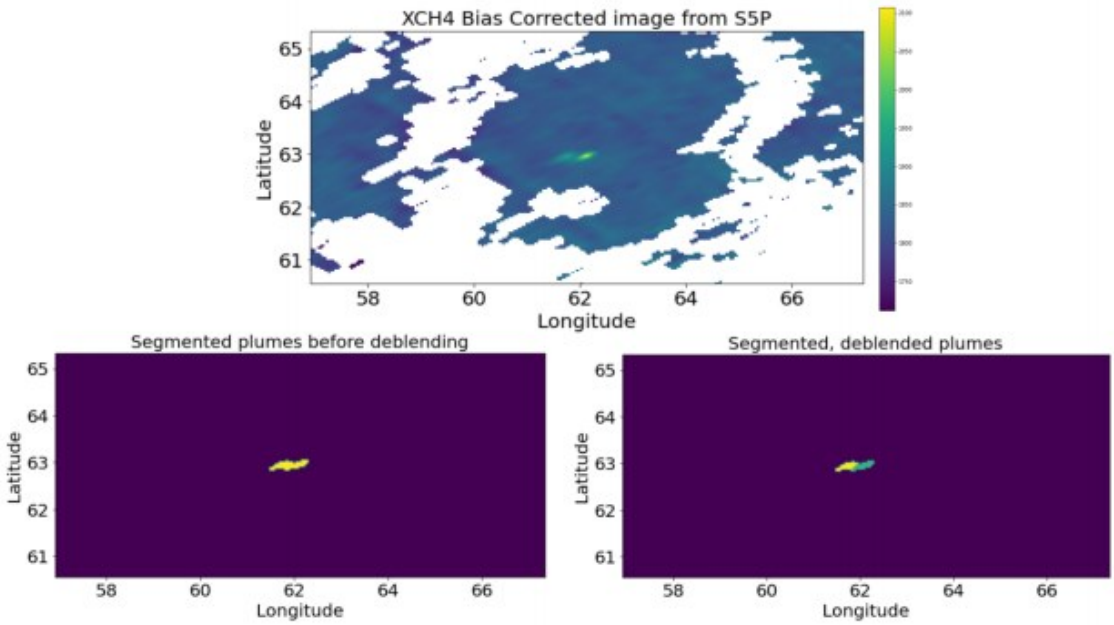


Figure 3-1: Methane plume detection algorithm. Top: Sentinel-5P Level-2 methane mixing ratio bias corrected data. Bottom-left: mask of the methane anomaly detected at step (2) of the detection algorithm. Bottom-right: 2 methane plumes are disentangled after step (3) of the detection algorithm. Source: [RD.1].

Each plume validated after the detection phase is tentatively quantified. For that purpose, the following processing pipeline is applied:

1. An approximate source location is inferred by following the upwind direction from the centroid of the detected plume polygon to one of the edges;
2. HYSPLIT [RD.3] forward point source simulations are run and emission flux rates are estimated using the mass balance method [RD.4]. The mass balance estimate is given by $Q = Q\omega \times \frac{X}{X\omega}$, where Q, Qω, X and Xω are respectively the observed flow rate, the flow rate used in the simulation, the observed methane enhancement within the plumes mask and the simulated methane concentration within the plume mask;
3. Human check is performed to assess the validity of the mass balance method (quantification is considered spurious if wind speed is too high or low, if simulated and observed plumes are not collinear, or if the methane anomaly detected is likely to have originated from distinct and distant point sources).

At step 2., the HYSPLIT model is run in concentration mode on a 0.01x0.01 degree grid. Gaussian puffs (in the horizontal direction) representing a fixed amount of CH_4 are released continuously. The Planetary Boundary Layer (PBL) in which these puffs are mixing is provided by the meteorological input fields. 2500 are released continuously during each simulated hour of emission. Releases start 7 hours before sensing time, which has empirically been found sufficient to simulate plumes in steady state. If the observed plume extends beyond the simulated plume, new simulations are performed with earlier release times until the plume length matches the observed one. The particles are released at 10 meters above ground level to account for high-pressure injection heights. This is also consistent with the height of devices found to be responsible of point-source emissions observed on TROPOMI images (e.g. flaring stacks, pipeline compressor station, coal mine ventilation air-methane devices). Meteorological data used as input for the HYSPLIT simulations are retrieved from the Global Forecast System (GFS) by the National Centers for Environmental Prediction (NCEP) at 0.25-degree and hourly resolutions. When GFS is not available on the NOAA FTP server, we rely on the Global Data Assimilation System (GDAS) meteorological data from NCEP at 1-degree and hourly resolutions.

Further details on the transport model parameters and human check process are given in [RD.1]. Figure 3-2 illustrates the detection and quantification process.

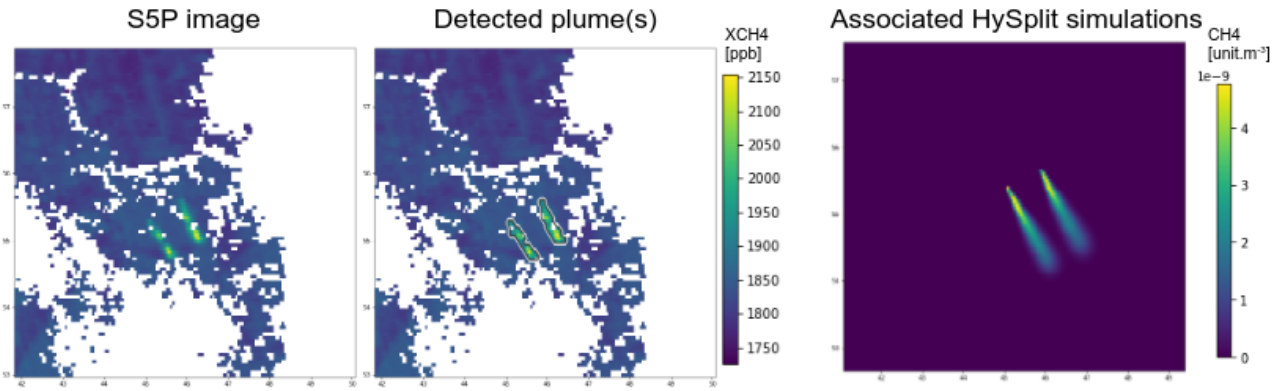


Figure 3-2: Methane plume detection and quantification. Left: Sentinel-5P Level-2 methane mixing ratio bias corrected data. Center: polygon of two methane plumes detected is overlaid on the TROPOMI image. Right: corresponding HYSPLIT simulations from which flux rates are determined. Source: [RD.1].

The results of the flux rate inversions have been shown consistent with reported data and scientific literature (see [RD.1] for more details).

The following table summarises the features of the output of the methane point source inversion.

Table 3-2: Methane plume detection and quantification

Feature	Value
Satellite data used	TROPOMI Level-2 methane mixing ratio bias corrected data product
Coverage	Global, onshore, mid-latitudes regions ([20, 55] degrees and [-40, -20] degrees latitude)
Revisit	Daily
Detection threshold	5 to 25 metric tons of methane per hour (subject to location and wind speed)
Uncertainty on flux rate estimates	+/- 50% relative uncertainty on average

3.3.2.2. ERROR ANALYSIS

Uncertainty in source rate estimation mainly stems from uncertainty in the model input parameters. For a set of parameters responsible for flow rate estimates uncertainty, we build an ensemble of estimates with independent variations in those parameters. We apply this ensemble of estimates to a random set of 200 plumes to derive sensitivity intervals. Assuming that these parameters affect independently the flux rate estimates, the standard deviation of the flow rates estimates computed with respect to each varying parameters are then summed in quadrature to derive the sensitivity intervals:

$$\sigma_{total} = \sqrt{\sigma^2_{location} + \sigma^2_{measure} + \sigma^2_{weather} + \sigma^2_{offset\ hour} + \sigma^2_{background}}$$

.

The parameters for which ensembles of estimates are computed are the following:

- longitude and latitude of the source location varies on a 0.1 * 0.1 degrees grid to account for the uncertainty in estimating the point source location;
- meteorological data: to account for the uncertainties in meteorological data, simulations are performed using both the GDAS-1 and the GFS-0.25 data sources;
- HYSPLIT simulations start time is offset by +/- 2 hours with hourly sampling;
- Background methane concentration is computed using various methods (mean and median, with and without and exclusion window);

TROPOMI measurement error, taken from the data product *methane_mixing_ratio_precision*, is propagated through the quantification procedure to estimate the sensor measurement error.

The results of this ensembling approach are shown by Figure 3-3. Details on the procedure are given in [RD.1].

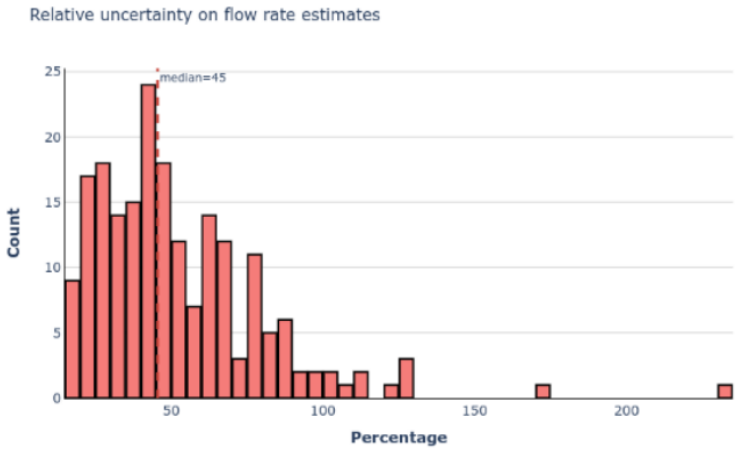


Figure 3-3: Relative uncertainty on flux rate estimates. Median relative uncertainty is 45%.

3.3.2.3. ALTERNATIVE METHODS AND INVENTORIES

No top-down point source inventory combining the same features is available in the literature. However, some methods and projects tackle similar issues.

- **Flux rate quantification method**: the mass balance method could have been replaced by the Integrated Mass Enhancement (IME) method of the Cross-Sectional Flux (CSF) methane, as developed in [RD.5]. Those two methods have the benefit of relying on wind data only, whereas the mass balance method relies on a transport model (in this case, HYSPLIT), hence being heavier to run. The mass balance method is however robust when part of the methane plume is missing, which is frequently observed on TROPOMI data (see Figure 3-2 for example). The IME would not apply in such situations, whereas the CSF would require proper wind conditions.
- **High-resolution methane point-sources inventories**: [RD.6] and [RD.7] provide high-resolution spaceborne methane plumes inventories, respectively based on multispectral and hyperspectral satellites. Due to their revisit periods (few days at best, pushbroom acquisition for multispectral satellites, tasking for hyperspectral), such satellites cannot provide a comprehensive, global inventory with daily revisit. However, they provide facility-level attribution of the emissions (against more than 15 km of uncertainty of source with TROPOMI inversions) and their detection threshold can be as low as 1 to 10 metric tons per hour for multispectral satellites and a few hundreds of kilograms per hour for hyperspectral satellites.
- **Local anomaly detection**: [RD.8] provides local methane anomaly detections based on the difference between methane concentrations simulated by the CAMS model and methane concentrations retrieved from TROPOMI data products. This approach uses time-averaged TROPOMI observation whereas the World Emission approach works on single images, hence providing daily refresh [RD.8] does not focus on quantification and indistinguishably detects anomalies that can be due to point source or diffuse surface emissions.

3.3.3. CARBON DIOXIDE (CO2) POINT SOURCES INVENTORY

Methods used in World Emission to estimate CO_2 emissions from cities and industrial complexes are detailed in the peer-reviewed research articles ([RD.9], [RD.10], [RD.11]). These inventories rely on quality-filtered XCO2 bias-corrected data from the OCO-2 satellite and the OCO-3 sensor. OCO-2 orbits Earth in polar, sun-synchronous orbits and has a narrow, 2-10 kilometer-wide swath and a 16-day orbital cycle, and its GSD is below 3 km. OCO-3 is a similar sensor onboard the International Space Station.

3.3.3.1. POINT SOURCE INVERSION

The World Emission data-driven approach to provide flux-rate estimates of urban and industrial areas without relying on 3D transport models is developed in [RD.9][RD.10][RD.11]. This method includes 4

stages. The first and second stages consist in selecting 200-kilometer-wide segments of OCO tracks on which point source inversion can be performed. The third stage restricts the number of processed segments to the vicinity of industrial and urban emitters. The last stage quantifies the emission flux rate.

In the first step of the inversion process, OCO-2 and OCO-3 quality-filtered retrievals containing substantial XCO2 anomalies are selected. This step is executed within 200-kilometer-long segments by flagging XCO_2 retrievals significantly different from the median value.

In the second step, the hypothesis that the emission plume can be shaped with a bell curve on top of a linear background is tested. The soundings are represented by their XCO_2 retrievals and by their position along the satellite track: their across-track position is ignored, the satellite narrow swath being used only to damp retrieval noise. A restriction to the soundings of the dominant surface type (land or water) mainly skips the artificial discontinuities in the retrievals linked to the corresponding changes in the surface albedo. The adjustable parameters are the standard deviation and the height of the Gaussian, and the slope and offset of the line. The centre of the peak is positioned in the middle of the 200 km moving window and is not adjusted. This procedure was implemented by [RD.10], and then refined and automated in [RD.9] and [RD.11], including a series of quality control criteria related to the values of the adjusted parameters, the quality of the fit and the data density. Figure 3-4 illustrates the track selection and background removal processes.

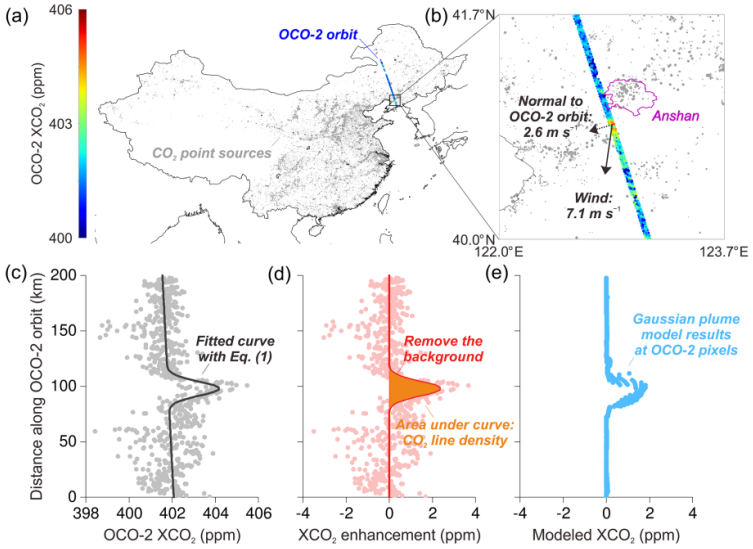


Figure 3-4: track selection for CO2 flux rate quantification. (a) and (b) display the XCO2 data product retrieved by OCO-2. (c) shows the XCO2 value along-track after quality filtering. (d) displays XCO2 enhancements after background removal. (e) shows XCO2 values as modelled by the Gaussian framework. Source: [RD.10]

In a third step, relevant use cases are selected based on their proximity to industrial emitters so that only fresh plumes are kept. The fourth step consists of quantifying the emission flux rate. This is performed by applying CSF ([RD.5], [RD.8], [RD.10]), by assuming steady wind conditions and using the ECMWF-ERA5 meteorological data.

Table 3-3: CO2 point sources detection and quantification

Feature	Value
Satellite data used	OCO-2 and OCO-3 quality-filtered XCO_2 bias-corrected retrievals
Coverage	Global, but the spatial density is biased towards China and India
Revisit	16-day orbital cycles but eligible tracks are subject to restrictive data selection. Empirically, 313 OCO-2 enhancements and 38 OCO-3 enhancements have been selected in respectively 7 years (2014-2021) and 2 years (2019-2021) [RD.11].

Feature	Value
Detection threshold	1 ktCO2/h per 0.1 deg grid cell
Uncertainty on flux rate estimates	Mostly random, 50% standard deviation

3.3.3.2. ERROR ANALYSIS

The quantification process has been compared, in Chinese cities, to the MEIC emission inventory and shows consistency up to a factor of two [RD.10]. In [RD.11], estimated flux rates are attributed and compared to EDGAR 6.0 grid cells for which emissions exceed 1000 metric tons per hour; a determination coefficient about 0.4 is found for flux rate estimates from both sensors. The study estimates that the emission retrieval typically has a 1-sigma relative uncertainty of 30-60%.

3.3.3.3. ALTERNATIVE METHODS

Some pioneer papers have focussed on the plumes from a few specific emission hotspots, also with a cross-sectional approach ([RD.12] and [RD.13]), but we are not aware of any other work at the global scale based on real data.

3.3.4. NITROGEN OXIDES (NOX) AND SULPHUR DIOXIDE (SO2) POINT SOURCES INVENTORIES

3.3.4.1. NOX

The methodology used to identify and quantify NO_x point sources is based on the continuity equation: the divergence, i.e. spatial derivative, of the horizontal flux, directly yields local sources and sinks. The basic idea was introduced and applied to some selected regions in [RD.14], and a global catalog of NO_x point source emissions (v1) was presented in [RD.15].

Within the World Emission project, an updated catalog (v2) of NO_x point source emissions was compiled. The update includes several improvements of the algorithm, resulting in higher (and more realistic) NO_x emissions compared to [RD.15]. Below we provide a summary of the input data and the processing steps for the NO_x point source catalog v2. Further details can be found in a dedicated description of the catalog v2 data product [RD.16].

Datasets

The NO_x point source catalog v2 is based on the following input data:

- Vertical Column Density (VCD): TROPOMI nitrogen dioxide tropospheric columns are taken from the PAL product [RD.17], a consistently reprocessed dataset covering the period May 2018-November 2021. Due to the improved cloud product used in the PAL processor versions, NO_2 columns are about 10%-40% higher than for the offline product versions used in [RD.14].
- Wind fields: Horizontal wind fields are taken from ECMWF.
- Ozone (needed for upscaling NO_2 to NO_x): An Ozone climatology is compiled from a global chemistry model [RD.18].
- Power Plants: The Global Power Plant Database (GPPD) [RD.19] is used in order to automatically look for matches of the detected point sources.
- Cities: A world city database (WCD) [RD.20] is used in order to automatically match the detected point sources with cities.

Data processing

Data processing consists of the following steps:

- Air Mass Factor (AMF) correction: As the spatial derivative is sensitive for the *excess* column added by the point source, while the local background is intrinsically removed, the operational VCD is corrected according to the averaging kernel (AK) at the assumed plume height of 500 m above ground. For NO_2, this correction is 1.61 on average.

- NO_2 to NO_x: The tropospheric NO_2 VCD is upscaled to NO_x assuming photo stationary state based on (a) photolysis frequency parameterized by the SZA, (b) reaction rate constant for $[NO]+[O_3]$ parameterized by temperature, and (c) Ozone mixing ratio taken from a model climatology [RD.18]. The NO_x/NO_2 ratio has been found as 1.38 on average.
- Advection: In [RD.14], the divergence $D:=\nabla\cdot(\boldsymbol{w}V)$ was calculated, with horizontal wind field $\boldsymbol{w}$ and VCD V. According to product rule, this equals $D = (\nabla\cdot\boldsymbol{w})V + \boldsymbol{w}\cdot(\nabla V)$. The first term represents the divergence of the wind field multiplied with the VCD. As we are interested in detecting changes caused by emissions rather than wind divergence, however, we skip this term in v2 of the catalog and only consider the second term $A:= \boldsymbol{w}\cdot(\nabla V)$, i.e. the scalar product of horizontal wind fields and the spatial gradient of the VCD. This term is denoted as "advection" in meteorology (in the meaning of "the rate of change of the value of the advected property", [RD.21]). Note that this has almost no impact on the resulting emission estimate, as the temporal mean divergence of wind fields (at 500 m above ground) is negligible.
- Derivative on TROPOMI grid: As proposed in [RD.22], the spatial derivative is now calculated on the original TROPOMI grid (along-track x across-track) which avoids steps in the gridding process which caused spikes in the derivative in v1 [RD.15].
- Topographic correction: As shown by [RD.23], the simplification of assuming horizontal transport only (2D perspective) is not appropriate in case of inhomogeneous terrain. [RD.23] derives a topographic correction term that takes into account spatial gradients of surface elevation: $C_{topo} = V/H\ \boldsymbol{w}\cdot(\nabla z_0)$, with H being the NO_x scale height. We applied this correction for the catalog v2, which improves the quality of the resulting maps and the respective emission estimates over mountains.
- Gridding & averaging: Advection (corrected for topography) is calculated per orbit, gridded on a regular 0.025° grid, and averaged over time (monthly, annual, and full period mean).
- Point source identification: Based on the temporal mean advection map (full period), point sources are identified in an automated iterative process. Local advection maxima are considered as point source candidates, whereby locations with spatial gaps, large negative advection around, or spatially extended peaks are rejected due to insufficient data, artifacts not consistent with the assumptions made, or area sources, respectively. 50000 iterations (i.e. candidates) have been processed for v2 of the catalog.
- Emission estimate: Point source emissions are derived by integrating the mean advection map spatially within 15 km. This approach has proven to be far more robust than the Gaussian peak fitting applied in [RD.15], and allows to quantify also annual or even monthly emissions for point sources with good statistics (i.e., cloud free conditions).
- Lifetime correction: An explicit correction of the chemical loss is applied based on the integrated loss during the residence time of the point source plume within the considered 15 km radius. For this purpose, NO_x lifetimes were parameterized as function of latitude as reported in [RD.24]. The lifetime correction factor has been found to be 1.40 on average.

The catalogue of NO_x emissions v2 [RD.16] lists 1139 "point sources" which include power plants, cement plants, industrial complexes, or small cities.

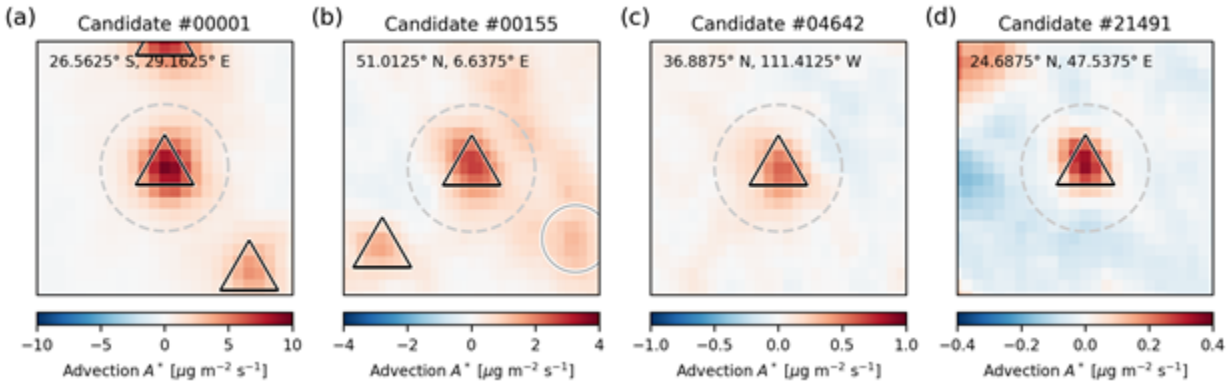


Figure 3-5: Sample maps of the temporal mean advection, corrected for topography, for (a) the first candidate classified as point source, i.e. the Secunda coal liquefier (South Africa), (b) Niederaußem/Neurath power plants (Germany), (c) the Navajo power plant (USA), and (d) the candidate with lowest derived emissions, i.e. the Al Yamama cement factory (Saudi Arabia). Results of the candidate classification are indicated by triangles for point sources and circles for area sources. The large dashed circle reflects the 15 km radius used for candidate classification procedure as well as for spatial integration. Note the different color scales. Source: [RD.16]

Table 3-4: NO_x and SO_2 point sources detection and quantification

Feature	NO_x	SO_2
Satellite data used	TROPOMI Level-2 nitrogen dioxide tropospheric column data product (PAL)	TROPOMI Level-2 sulfur dioxide total vertical column data product (COBRA)
Coverage	Global, excluding polar regions (50°S to 72°N, SZA < 65°), subject to cloud coverage and adverse atmospheric conditions.	Global, excluding polar regions (50°S to 72°N, SZA < 65°), subject to cloud coverage and adverse atmospheric conditions
Temporal averages	Full period, annually (default), monthly (good conditions only)	
Detection threshold	3 kt / year (1 kt / year for good conditions) [RD.14]	[RD.1]
Uncertainty on flux rate estimates	+/- 50% relative for individual point sources; +/- 20% on average [RD.14]	+/- 50% relative on average [RD.27][RD.28]

3.3.4.2. ERROR ANALYSIS

In the NO_x catalog v2 [RD.16], the uncertainties of the individual processing steps are considered, and a total error is calculated by error propagation and included in the catalog. Typical relative errors are about 20%-40%.

In addition, there are potentially systematic errors which are hard to quantify (and are thus not included in the provided error):

- uncertainty of wind fields: as discussed in [RD.14], statistical as well as systematic errors in wind direction both result in a systematic low bias, as the actual transport in the true wind direction is underestimated. This effect was estimated to be about 3% for Riyadh, and was thus ignored in [RD.16]. However, the impact might be stronger in particular for point sources where wind speeds are generally low.

- mountains: In addition to higher uncertainties in wind fields, also 3D effects of transport come into play as soon as the terrain has spatial gradients. The consideration of the topographic advection term [RD.23] improves the advection maps over mountains significantly. However, there are still some artifacts (both positive and negative) remaining. For further improvements, wind fields with better spatial and/or temporal resolution should be used in future.

- 3D effects of radiative transfer for power plant plumes: AMFs are usually calculated for a-priori trace gas profiles without the consideration of horizontal gradients and applying the independent pixel

approximation. With TROPOMI, however, pixel size becomes so small that 3D effects of radiative transfer matters. As shown in [RD.25], horizontal light paths lead to a smearing out of the satellite observations of a confined plume: VCDs of plume pixels are generally biased low when derived with a 1D AMF, while neighbouring pixels are biased high. Note, however, that this effect is slightly dampened by the spatial integration within 15 km applied in the catalog v2. For an accurate quantitative estimate and potential correction, further studies are required that take the specific geometry of power plant plumes into account.

- Spatial integration: It has to be kept in mind that the reported emissions are derived by spatial integration within 15 km radius. Any other NO_x source within this radius will thus contribute to the emission estimate.

Overall, we estimate the uncertainty of the updated emissions to about 50% for individual point sources, and about 20% on average.

3.3.4.3. SO_2

[Major update with respect to previous version]

The identification and quantification of SO_2 point sources uses the same methodology as for NO_x. The workflow was slightly modified where necessary and some processing steps (e.g., the upscaling of NO_2 to NO_x) are obsolete for SO_2. The detailed workflow settings are described below.

Datasets

The SO_2 point source catalog is based on the following input data:

- Vertical Column Density (VCD): TROPOMI sulfur dioxide total vertical columns are taken from the COBRA product [RD.26], a consistently reprocessed dataset covering the period May 2018-July 2022. Due to an improved SO_2 retrieval scheme used in the COBRA product, noise and biases in the SO_2 columns are reduced significantly. The data is available on request and was kindly provided by Nicolas Theys.

- Auxiliary data: Additional variables needed for the data processing workflow are taken from the operational TROPOMI SO_2 product (e.g. averaging kernels, surface altitude, and 10 m wind components (for topographic correction)).

- Wind fields: High resolution (0.25°) horizontal wind fields are taken from ECMWF and interpolated to a height of 500 m (default) and 300 m (error analysis).

- Power Plants: The Global Power Plant Database (GPPD) [RD.19] is used in order to automatically look for matches of the detected point sources.

- Cities: A world city database (WCD) [RD.20] is used in order to automatically match the detected point sources with cities.

Data processing

Data processing consists of the following steps:

Data selection: To keep the data of the best quality, only TROPOMI pixels associated with a quality assurance value above 0.5 are used. This includes e.g., a solar zenith angle restriction to below 60°, a radiometric cloud fraction lower than 30%, and excluding the 50 outermost rows at the swath edge.

- SO_2 artefact removal: in cases where the SO_2 signal is very strong (e.g., volcanic SO_2) the covariance matrix used in the COBRA algorithm for the derivation of SO_2 VCDs is corrupted and may lead to stripes in the data. In order to account for this, the median of the along-track VCDs is calculated for each cross-track position. If a threshold value of 2.85 DU is exceeded, the affected ground pixel is omitted for the whole orbit.

- Air Mass Factor (AMF) correction: see section 3.3.4.1 The AMFs are taken from the COBRA product while the averaging kernels (AKs) from the operational SO_2 product are used, since this information was not provided in the COBRA dataset.

- Advection: see section 3.3.4.1
- Derivative on TROPOMI grid: see section 3.3.4.1
- Topographic correction: see section 3.3.4.1 The surface altitude (z0) and 10-meter wind components (combined in the variable w) are taken from the operational product.
- Gridding & averaging: see section 3.3.4.1
- Point source identification: Based on a 2-dimensional median smoothed temporal mean advection map (full period), point sources are identified in an automated iterative process analog to section 3.3.4.1 The median smoothing is necessary for SO_2 as the SO_2 VCDs are generally noisier than for NO_2, therefore the averaged advection is noisier as well. The SO_2 algorithm would run much longer and miss out potential point sources. Instead of 50000, 100000 iterations (i.e., candidates) have been processed for the catalog to increase the number of potential point sources even further.
- Emission estimate: see section 3.3.4.1
- Lifetime correction: An explicit correction of the chemical loss is applied based on the integrated loss during the residence time of the point source plume within the considered 15 km radius. For this, a SO_2 lifetime of 6 hours is assumed [RD.28].

The catalogue of SO_2 emissions lists 130 "point sources" which include power plants, smelters, oil and gas industry, or industrial complexes.

In the SO_2 catalog, just as for NO_x, the uncertainties of the individual processing steps are considered, and a total error is calculated by error propagation and included in the catalog. Typical relative errors are about 33%-50%.

The systematic errors of the NO_2 and SO_2 catalog are discussed in 3.3.4.2.

3.3.4.4. ALTERNATIVE METHODS AND INVENTORIES

On NO_x, some methods and projects tackle similar issues and deal differently with the sink term: In [RD.13] and [RD.14], the chemical loss due to the reaction of NO_2 with OH is described by a first-order time constant and is estimated directly from the NO_2 measured column itself. In [RD.13] the first-order time constant is derived empirically from the NO_2 downwind plume evolution whereas in [RD.14], NO_2, NO, and OH fields from The Copernicus Atmospheric Monitoring Service (CAMS) are used in the computation of the first-order time constant. In [RD.24], the chemical loss of NOx is determined empirically from the mean downwind decay under particular consideration of potentially interfering NOx sources.

On SO_2, a point source inventory is provided in [RD.11]: For potential point sources, all SO_2 quality-filtered observations are rotated according to ECMWF wind fields, and temporally averaged. An exponentially modified Gaussian is then fitted to the mean rotated SO_2 observations after background correction.

In addition, a method has been developed in [RD.15] to quantify SO_2 emissions from MT. Etna volcano using back-trajectory modelling with the PlumeTraj transport model instead of an empirical plume model.

3.3.5. NOX FROM MEGACITIES AND CONURBATIONS

[Major update with respect to previous version]

For the quantification of NOx emissions from megacities and conurbations, the advection method is not as sensitive as for point sources, as the emissions are distributed over a larger area, and spatial gradients are not as strong. In addition, effects of the chemical loss of NOx are more important, as longer distances have to be considered.

Thus, a different approach is used for determining both the effective NOx lifetime and NOx emissions from megacities, which is based on the works of [RD.29], [RD.30] and [RD.31]. The key idea is to (a)

sort the observations by wind direction and (b) derive the NOx lifetime and NOx emissions from the downwind evolution of the NO2 plume.

In [RD.29], megacities are considered as single source with a spatial distribution that can be described by a Gaussian function. This is only applicable for isolated megacities without interfering emissions around; [RD.29] could derive emissions for 8 megacities and conurbations. In [RD.30] a modified approach is proposed that works also for complex spatial distributions of multiple sources: The observed pattern of NO2 columns at calm winds is used as proxy for the spatial distribution of emissions. From the comparison of the respective patterns for windy conditions, the effective NOx lifetime can then be derived (Figure 3-6). As shown in [RD.29], the emissions of 53 cities in the US and China could be derived, with very good agreement to bottom-up inventories (9%±49% mean±standard deviation). Recently, the algorithm was refined further [RD.31] such that lifetime and emissions are derived in a single step instead of the two step scheme in [RD.29].

Within World Emission, NOx emissions from megacities and conurbations were planned to be derived based on the algorithm described in [RD.31][RD.31][RD.31]. During processing, however, we noticed some shortcomings of this approach:

- The background level of NO2 VCDs is included as one fit parameter in [RD.30] and [RD.31]. However, we noticed that the background itself depends on wind direction in many cases, and cannot generally be assumed to be the same for calm vs. windy conditions.
- The method requires sufficient observations for calm conditions; otherwise, no proxy for the emission distribution is available. Note that the sorting of TROPOMI into 8 wind direction bins plus calm is demanding with respect to sample size, even when multiple years of data are considered.
- Even if only calm observations are considered, which still includes wind speeds up to 2 m/s, column density patterns are smeared out compared to emission patterns, limiting the reachable agreement between observation and forward model.

In order to account for these shortcomings, a modified procedure was developed. The basic idea is to consider the spatial distribution of emissions also as fit parameters; in order to have sufficient observations for a well constrained fit, the fit parameters (distribution of emissions, background and lifetime) are derived from the combined observations for calm as well as two opposite wind directions. Details of this approach are detailed in "Lifetime and emission fit" below.

3.3.5.1. DATASETS

The NO_x hotspot emission catalog is based on the following input data:

- Vertical Column Density (VCD): TROPOMI nitrogen dioxide tropospheric columns are taken from the PAL product [RD.17].
- Wind fields: Horizontal wind fields are taken from ECMWF.
- Cities: The world city database (WCD) [RD.20] is used in order to generate a list of hotspot candidates. The NO_x hotspot emission algorithm is applied to all cities with more than one million inhabitants.

3.3.5.2. DATA PROCESSING

Data processing consists of the following steps:

- NO2 column densities (PAL product) are restricted to SZA<65°, VZA<56°, qa>0.75.
- NO_2 is upscaled to NO_x the same way as for the point source catalog.
- NOx column densities are sorted into "calm" (wind speed below 2 m/s) and 8 wind direction *sectors* (45° steps) based on ECMWF wind fields at 500 m above ground.
- All TROPOMI pixels are re- gridded on a regular lat-lon grid with 0.05° resolution.
- Temporal means are calculated. Since the NO_x lifetime is one fit parameter, which is expected to depend on season, the analysis is performed on seasonal basis. As the main impact of season is expected to be that on the SZA, seasons are defined with respect to solstice, i.e., winter (NDJ), spring (FMA), summer (MJJ) and autumn (ASO).

The temporal mean NO_x column density for the different wind sectors is shown exemplarily in Figure 3-6 for Paris in summer.

- For each *axis* (west to east, southwest to northeast, south to north, and southeast to northwest), the mean line density is calculated for calm, forward and backward wind direction, by integrating the seasonal mean NO_x column density maps in across-wind direction (+/- 50 km), yielding NO_x amount per length unit. Note that the VCD for calm is the same for each axis, but due to the integration in across-wind direction, the line densities are different for calm for the 4 axes. The resulting line densities for Paris in summer are displayed in Figure 3-7.

Note that in contrast to the point source emission catalog, no additional AMF correction is applied, since for point sources the altitude of the excess plume is rather well constrained, which is not the case for area sources. For megacities, on the other hand, it can be generally assumed that they are included in emission inventories, and CTMs reproduce the NO_2 profiles reasonably well, and the calculated AK in the PAL product are appropriate in most cases.

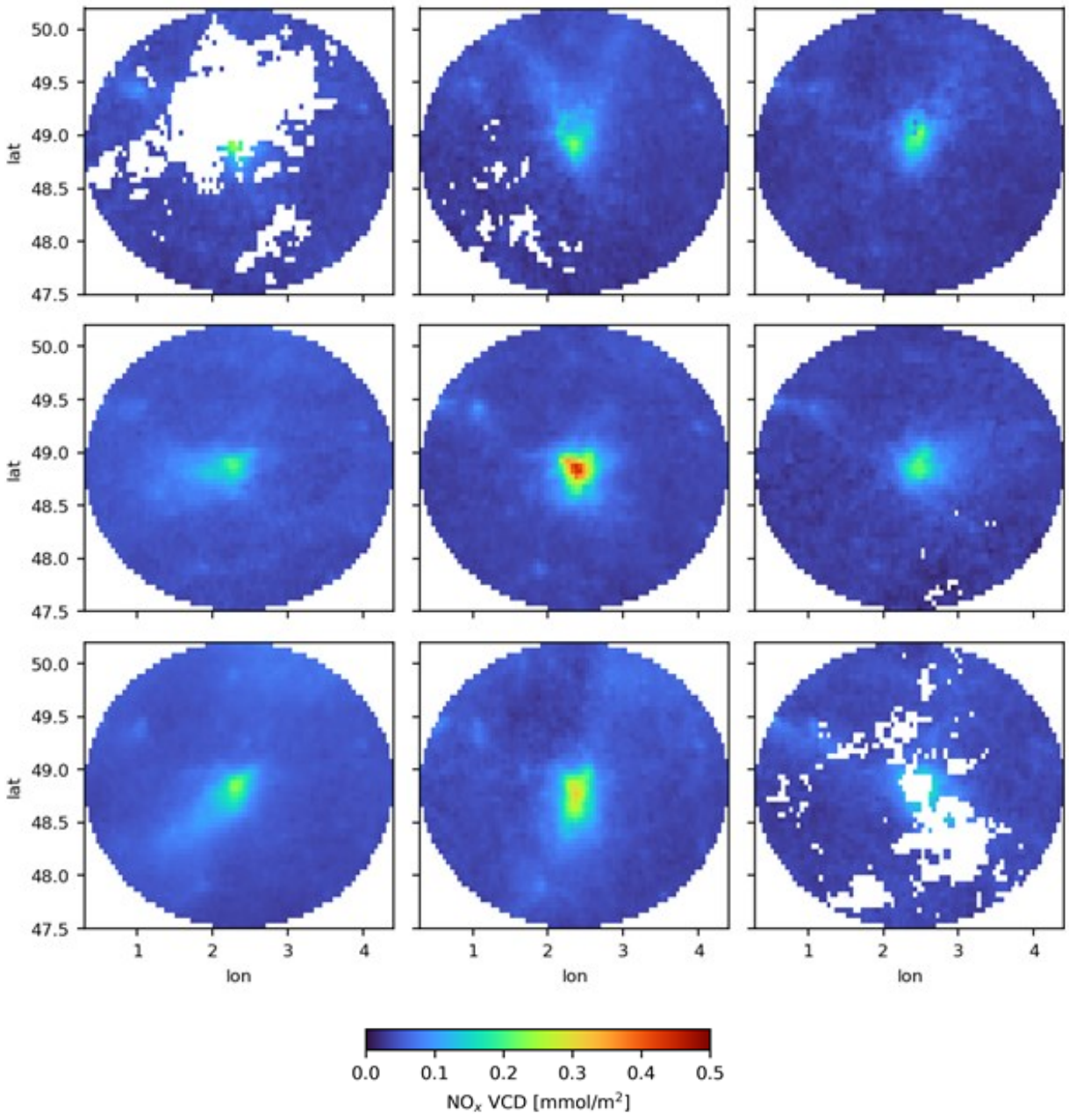


Figure 3-6 Mean NOx distribution (summer) for Paris depending on wind direction. The central panel displays the distribution for calm conditions (< 2 m/s). The surrounding panels show the respective patterns for the 8 different wind sectors.

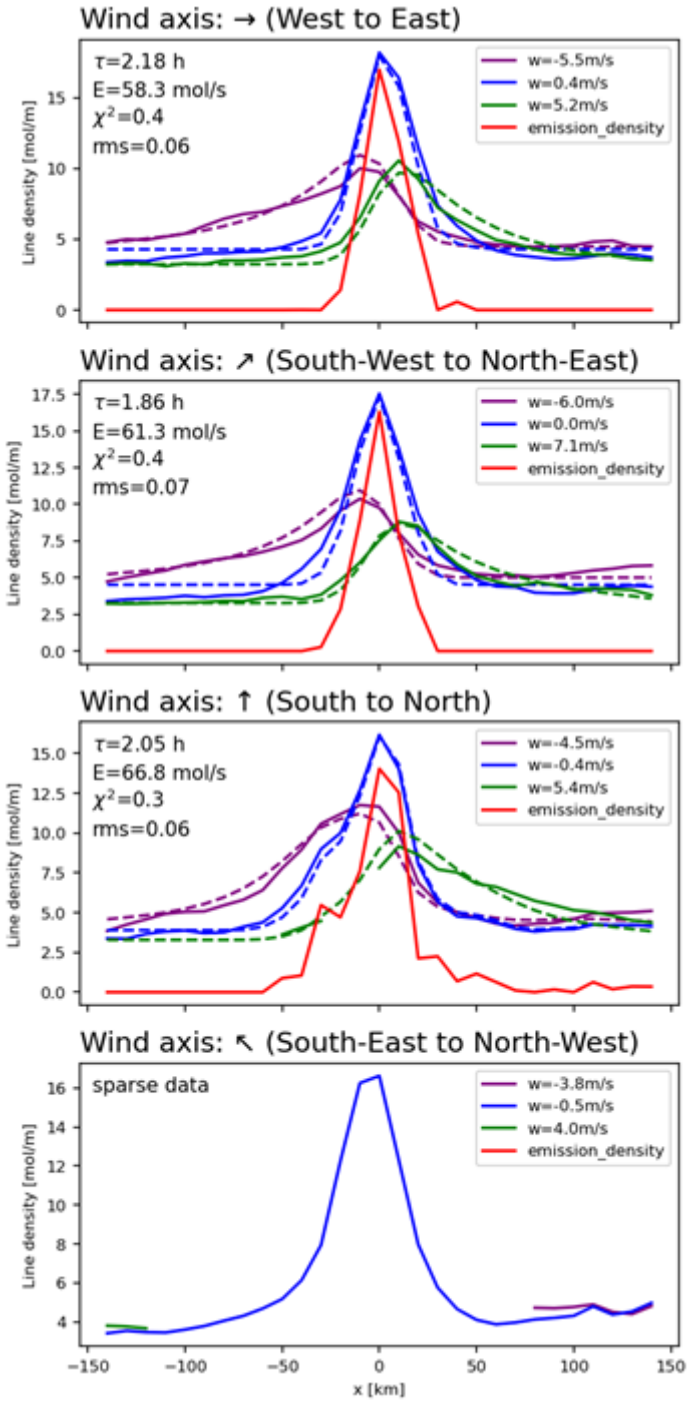


Figure 3-7 Mean NOx line densities (summer) for Paris depending on wind direction. The panels show the different wind axis. For each axis, the line densities for calm (blue), forward (green) and backward (purple) wind directions are displayed as straight lines. If the lifetime/emission fit succeeds, the

corresponding fitted line densities are shown as dashed lines. The fitted emission density is displayed in red. Fitted lifetime and emissions are provided as text in each panel.

3.3.5.3. LIFETIME AND EMISSION FIT

The lifetime and emission fit are performed for each axis separately based on the following forward model:

$L_i(x) = E(x) * \exp(-x/(w_i \tau)) * \exp(-x^2/(2\sigma^2)) + b_i$

The index i refers to the wind direction (calm, forward, backward).

$L_i(x)$ denotes the line density, i.e. the column integrated in across-wind direction.

E(x) represents the spatial density of emissions. It is considered to be the same for all 3 wind directions. E(x) has the same unit as L(x) (mol per m) and corresponds to the line density that would be observed if not wind transport would occur.

The symbol * indicates mathematical convolution. The first convolution term (to be truncated to x>0) describes the downwind decay of the emitted NO_x with the e-folding lifetime τ which is converted to an e-folding distance by multiplication with the mean wind speed w_i.

The second convolution represents a simple Gaussian smoothing that accounts for smoothing effects like e.g. temporal variations of wind speed. σ is fixed to 7 km.

b_i is the respective NO_x background line density which can be different for different wind directions and calm.

Based on the observed line densities, the distribution of emission densities E(x), lifetime τ and background b_i are fitted *simultaneously*.

Figure 3-7 displays the measured (straight) and fitted (dashed) line densities for Paris in summer. The fitted emission density is shown in red.

The emission rate of the considered hot spot is then derived by spatial integration of E(x) from -50 km to +50 km (yielding total amount), divided by τ (yielding emission rates in amount per time unit).

Finally, the results for the different axis are averaged, weighted by the number of contributing directions for each axis (the more observations, the higher the weight) as well as the fit performance (the lower the χ^2, the higher the weight).

3.3.5.4. SELECTION OF FIT RESULTS

The forward model described above allow to quantify the emission distribution and total emissions around hotspots. However, for robust fit results, sufficient observations are necessary. As fit results for insufficient data were often found to be dubious, the fit results are selected according to several strict criteria:

- For individual axis, a fit is considered only if at least 2 directions have sufficient data (less then 10% spatial gaps in the seasonal mean column) and the difference in wind speed is sufficiently large (4 m/s between calm and windy or 8 m/s between for- and backward wind).
- The emission results for the hotspot candidates are kept if no interfering emissions were found, and the fit worked for at least one season with either one sector where all 3 directions are available, or 2 sectors where 2 directions are available.

With these strict selection criteria, emissions were derived for 100 hotspots out of the original list of 700 cities with > 1 million inhabitants. Note that for most cities, emission estimates could only be derived for 1 or 2 seasons.

3.3.6. AMMONIA (NH3), ACETYLENE (C2H2), ETHYLENE (C2H4) AND METHANOL (CH3OH) POINT SOURCES INVENTORIES

Methods used to build existing NH_3 point source inventories are detailed in peer-reviewed research articles ([RD.32], [RD.33] and [RD.34]). Specifically, [RD.32] and [RD.33] present a NH_3 point source detection and quantification method that can be straightforwardly applied to C_2H_2, C_2H_4 and CH_3OH. The

World Emission NH_3, C_2H_2, C_2H_4 and CH_3OH point source inventories are built on IASI data products. IASI is an infrared interferometer sensor onboard the Metop-A (2007-2021), Metop-B (2013-ongoing), Metop-C satellite (2019-ongoing), which orbits the Earth on sun-synchronous, polar trajectories.

3.3.6.1. POINT SOURCES INVENTORIES

[Minor update with respect to previous version]

The point source inventories of NH_3, C_2H_2, C_2H_4 and CH_3OH have been established in two major steps, one of identification and categorization (where the emission sector is attributed) and one of quantifications, where the emission flux is estimated. Their application over the course of this project is explained below.

The point source identification relies on oversampling [RD.33] and supersampling techniques [RD.32]. Relying on the availability of a large number of satellite observations, these techniques allow producing averages at a much higher spatial resolution than the native resolution of the sounder by exploiting the varying location, shape and orientation of the IASI pixels from successive overpasses. Indeed, after successive satellite overpasses, the varying footprints on the ground of the satellite pixels overlap, and sub-pixel information becomes exploitable. Therefore, for each target species, a large dataset of satellite measurements is built that consists of the 2008-2020 IASI/Metop-A, the 2013-2020 IASI/Metop-B, and the 2019-2020 IASI/Metop-C observations. Only the IASI data from the morning overpasses are used here to take advantage of the larger sensitivity of the satellite measurements to the lower tropospheric layers (thanks to a larger thermal contrast). Moreover, the observations affected by >10% cloud coverage are discarded from the IASI datasets. We also excluded all the measurements outside the 60° S-70° N latitudinal band, as those are typically affected by lower sensitivity, and because anthropogenic gas emissions are limited in polar regions.

Oversampling consists in constructing a high-resolution sub-grid and assigning to each pixel of this sub-grid the average of the IASI image values re-projected on this same sub-grid. The improvement brought by the oversampling over a regular gridded average for the representation of a hotspot of a short-lived pollutant from space, is highlighted in Figure 3-8 a and b. Supersampling is a refinement of the oversampling technique, since oversampling usually tends to smooth observations and blurs the output. The supersampling requires first the application of the wind rotation technique. This latter consists in rotating each measurement of the satellite time series around the presumed source of the target species, taking into account the daily horizontal wind direction. The result is a distribution of the measurements in which the winds always blow in the same direction from the presumed gas emitter (i.e., the rotation centre). This distribution is hereafter called the downwind average. An example of wind rotation applied to an oversampled average of IASI measurements over a hotspot is presented in c. Compared to the regular oversampling, the hotspot exhibits enhanced magnitude and a more concentrated plume of the target species from the presumed source. The ECMWF ERA5 daily horizontal wind fields, collocated with the IASI measurements and interpolated vertically at the mid-height of the planetary boundary layer, are used for the application of the wind rotation. The supersampling consists in repeating the wind-rotated oversampling procedure and in correcting, at each iteration, the resulting oversampled average according to the differences with the satellite observations. The application of the wind rotation guarantees the homogeneity of the observation scene by removing most of the variability around the point source [RD.32]. The supersampling is computed using an iterative algorithm (iterative back-projection, IBP, [RD.35]) which partially corrects the smoothing effect generated by the oversampling. In the first iteration, the supersampled image SS_1 is equal to the oversampled image OS_1. After each supersampling iteration, simulations of what the sensor would observe if SS_i were the ground truth, are computed (M_i). In the *k*th iteration, the supersampled image is equal to the supersampled image computed at iteration *k-1* plus the oversampling of M_1 minus the oversampling of M_{k-1}. Full methodology is detailed in [RD.32]. As illustrated in Figure 3-8d, the supersampling is more effective at smoothing the background noise and at reproducing the real strength of the hotspot.

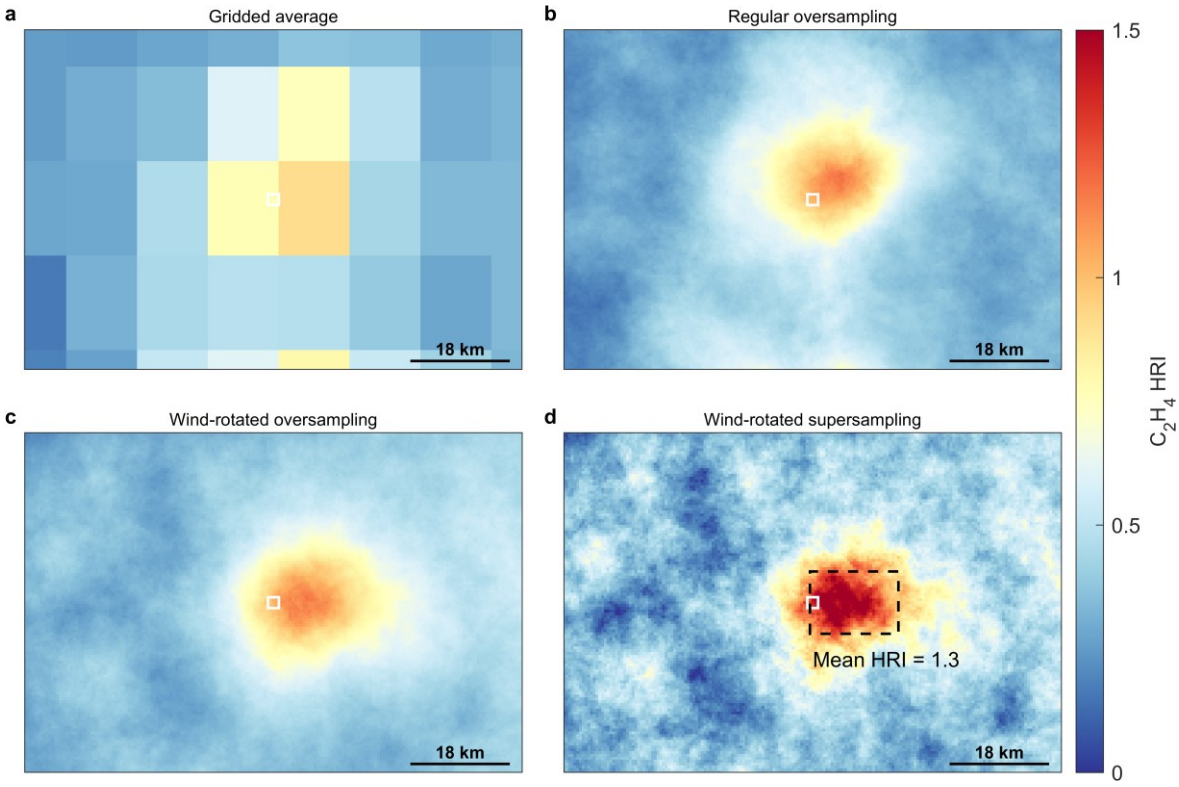


Figure 3-8 Averaging techniques applied to C2H4 IASI measurements over Mengxi Park (Inner Mongolia, China). a, Binned average on a 0.15° × 0.15° spatial resolution grid, b, oversampled average, c, wind-rotated oversampling, and d, wind-rotated supersampling of the IASI C2H4 HRI at a 0.01° × 0.01° spatial resolution, around the hotspot of Mengxi Park (Inner Mongolia, China). At the latitude of the hotspot (39.9° N), this corresponds to a grid resolution of 12.8 × 16.6 km (lon × lat) for a, and of 0.9 × 1.1 km for b-d. The coordinates of the presumed emitter, marked by a white square, is used as the wind rotation centre. Here, the winds are realigned to the east (in the x direction). The area delimited by the black dashed line is used to calculate the averaged downwind value. This figure is taken from [RD.36]

To provide a global mapping of the gas emitters with the supersampling technique, we assume that each location on Earth is a potential point source. Therefore, the wind-rotated supersampling is applied successively to each grid cell of a high-resolution world map (typically 0.01°×0.01°) [RD.32]. For each grid cell, the averaged value is calculated from the supersampled distribution, over an area downwind of the emitter, by considering all the observations included in a box extending from 0 to 20 km in the downwind direction and from -5 to +5 km in the orthogonal direction. Such box is illustrated by the black dashed line in Figure 3-8d. This accounts for the gas transport and benefits both from the supersampling/oversampling and from the wind rotation heuristics. In particular, supersampling is more efficient than oversampling to reduce background noise and pinpoint emitting locations. As it is able to resolve much finer spatial features, it also allows the discovery of small hotspots that would be difficult to detect with the oversampling alone. This is illustrated in Figure 3-9, which present the NH_3 hotspots map on one location in North America. Compared to the first inventory based on oversampling [RD.33], supersampling allowed doubling the number of NH_3 point sources identified with IASI (now over 500 points sources have been identified and categorised).

Through a careful analysis of the supersampled high-resolution distribution of each species, local enhancements of the target species, typically of 20-50km spatial extent, can be detected, which correspond to hotspots. As an example, a zoom-in over an industrial valley in the Shanxi Province, China (Figure 3-10), reveals a series of C_2H_4 hotspots associated with the presence of heavy and coal-related industries. These hotspot maps are combined with visible airborne and spaceborne imagery, public data sources such as infrastructure or emission inventories, to identify the emission point sources and classify them in categories (see [RD.32] and [RD.36]). For each target species, a catalogue is built that includes all the point sources detected with the IASI measurements and the category(ies) of emitters with which they are associated.

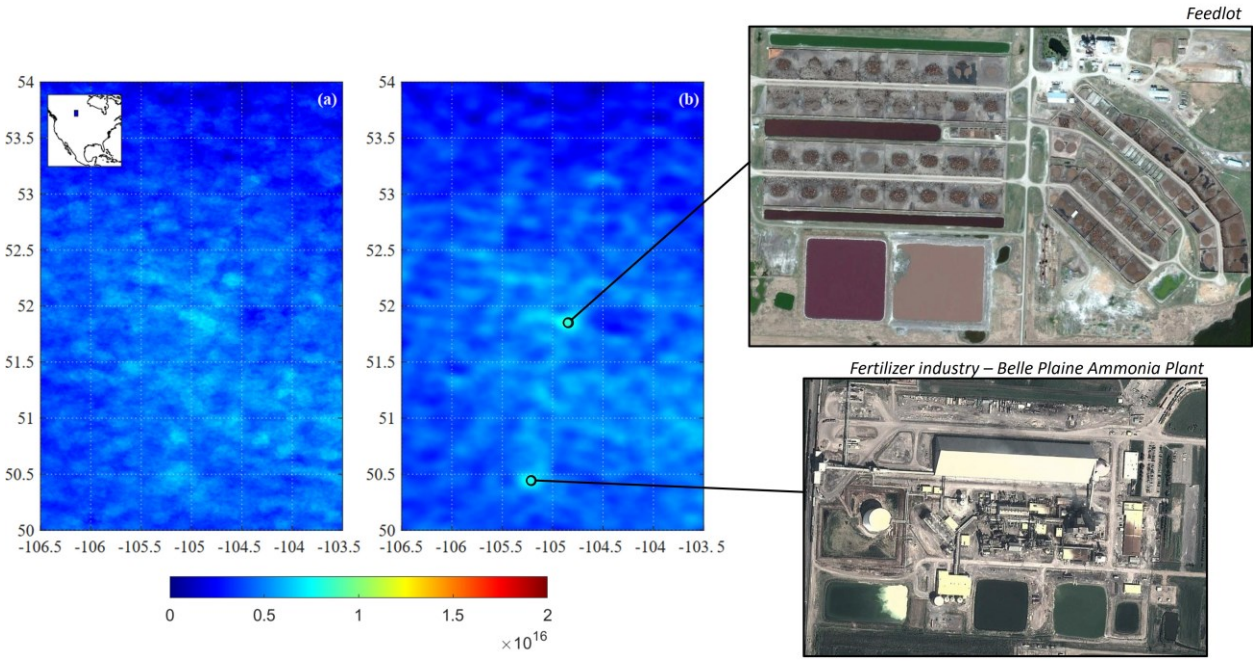


Figure 3-9: Oversampling and supersampling-based point sources map. Panel a shows the output of applying oversampling over Canada. Panel b displays the output of wind-rotated supersampling. Black circles represent identified point sources. Sources: [RD.32]

Once the point sources are detected, their flow rates *E* are quantified using the box model $E = M/T$, with *M* the amount of gas (total mass) contained in the box and *T* the effective lifetime of the target species ([RD.37]; see [RD.33] and [RD.36] for details on its application to IASI NH_3 and C_2H_4 data, respectively). With this model, a steady state and first-order loss terms are assumed. The size of the box around the point source is considered large enough to encompass most of the plume and to avoid the transport of the target species out of the area. *M* is calculated directly from the 0.01° x 0.01° distribution of IASI column measurements produced by the wind-rotated supersampling, using the coordinates of the identified point source as the rotation point [RD.36]. The wind rotation allows concentrating the bulk of gas by aligning the wind fields, while preserving the distance of the satellite measurements to the identified emitter [RD.32]. To account for an ambient concentration of the target species in the area, a column background is subtracted from the gas column distribution. As illustrated for a C_2H_4 hotspot in Figure 3-11a, this background is calculated as the gas column averaged over a side band of the box that is not affected by the transport plume. *M* is obtained from the resulting gas distribution (e.g., Figure 3-11b) by summing up the gas masses in each 0.01° x 0.01° grid cell. The effective lifetime *T* of each target species is evaluated, based on the literature.

A prerequisite of this mass-balance approach for inferring the top-down emission fluxes from a point source is that the representation of the hotspot, and in particular of the plume of pollutant, with the IASI observations must be realistic enough. Therefore, depending on the target species, the inference of top-down fluxes might not be doable for all the identified point sources. This is especially the case for the NMVOC species, namely CH_3OH, C_2H_4 and C_2H_2, which are weak absorbers in the thermal infrared, and for which the signal in the IASI observations is typically weaker than the signal of NH_3 (e.g., [RD.36]). For instance, in [RD.36] the top-down fluxes of C_2H_4 were calculated only for 57 global hotspots out of the >300 identified by IASI, i.e., for the hotspots with the highest detected signal and for those presenting the largest contrasts relative to the surrounding background.

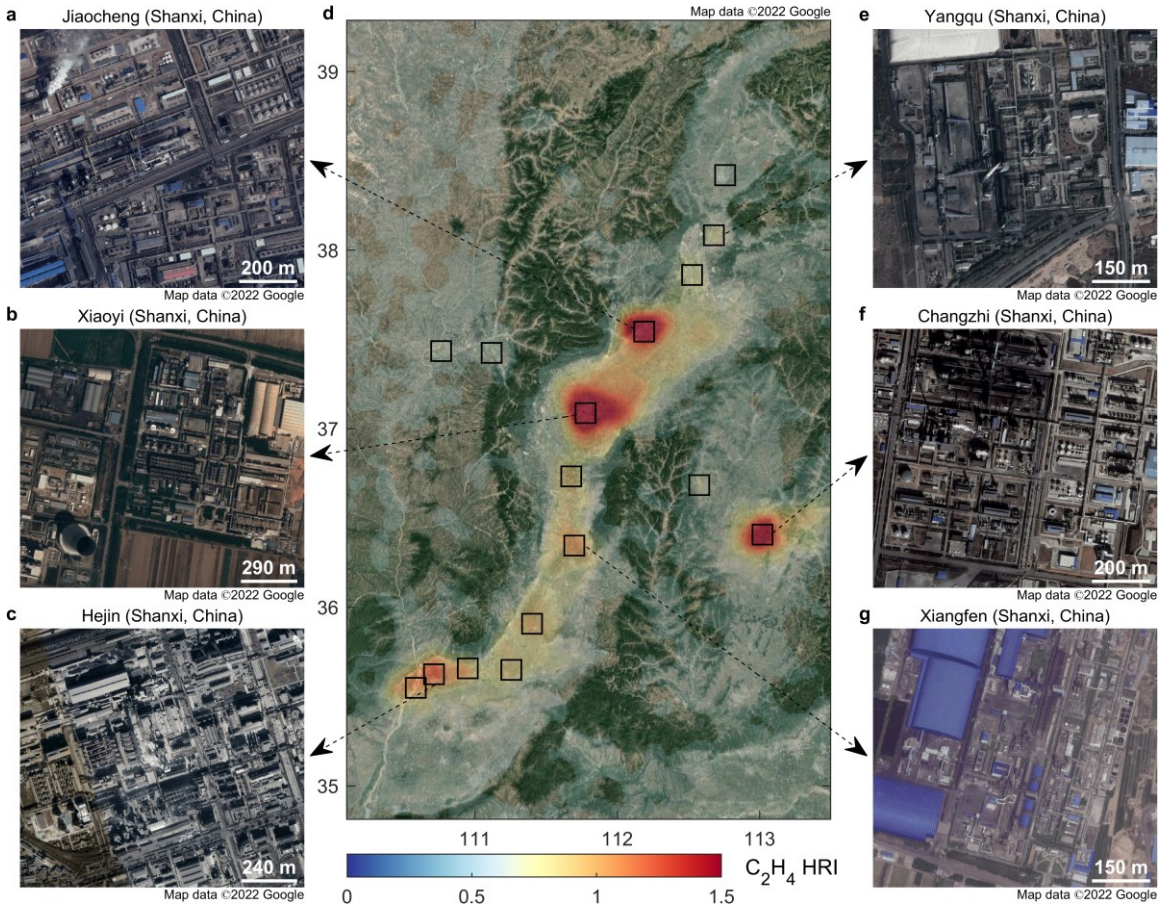


Figure 3-10: IASI hyperfine resolution distribution, with hotspots and point-sources of ethylene. The central panel shows a zoom-in of the 13-year IASI average of C2H4 on satellite visible imagery, over an industrial valley of the Shanxi province, China. Hotspots are indicated with black squares. The side panels give examples of close-up views on point-source emitters. Sources: [RD.36]

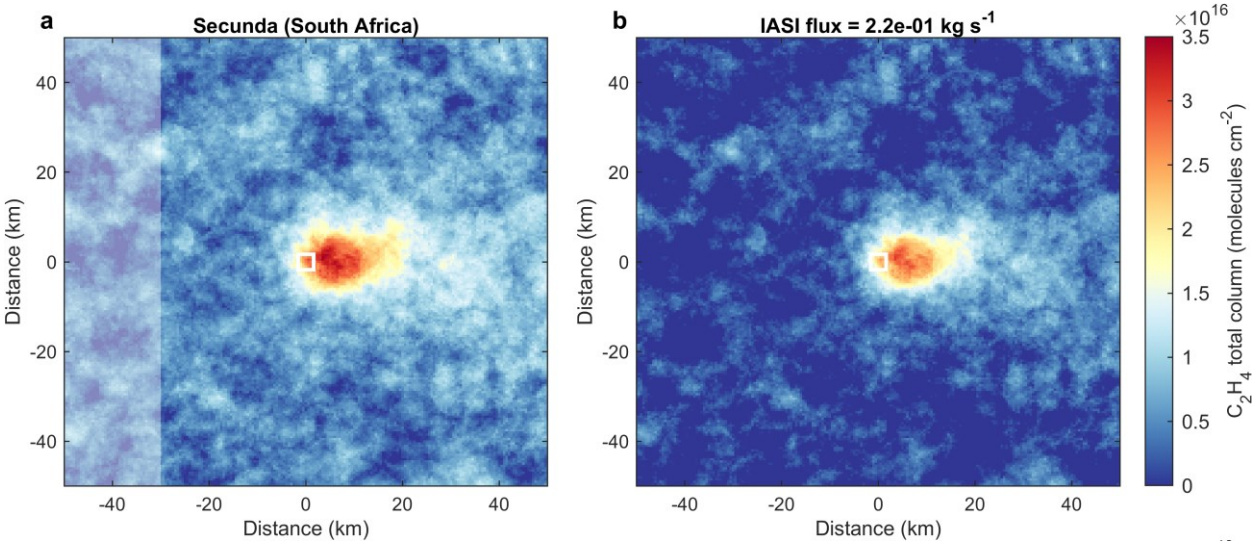


Figure 3-11: Example of C2H4 flux calculation over Secunda (South Africa). The panel a shows the IASI C2H4 column distribution produced by the wind-rotated supersampling, with the point-source (Secunda, South Africa) delimited in white and the area used to calculate the C2H4 background level shaded in semi-transparent. The panel b shows the IASI distribution after subtracting the background column. Sources: [RD.36].

3.3.6.2. ERROR ANALYSIS

The uncertainty on the estimated gas emission fluxes from the point sources depends primarily on (1) the errors and potential biases affecting the satellite measurements, and on (2) the validity of the assumptions made by the box model.

1. The uncertainties on the retrieved columns from IASI are quantified during the retrieval process. Considering the amount of IASI observations that are averaged over the same scene by the oversampling/supersampling, these uncertainties are significantly reduced. Potential systematic biases are evaluated via comparisons of the IASI retrieved columns with independent measurements.
2. The effective lifetime of the target gas that is assumed by the box model to derive the emission fluxes from the point sources represents an important source of uncertainty in the flux estimates. Therefore, the gas emissions are also calculated assuming the upper- and lower-bound estimates of the gas effective lifetime found in the literature, as done in [RD.33] and [RD.36] for NH_3 and C_2H_4, respectively. This provides a range of emission fluxes for each point source.

Another assumption made by the box model is the steady state. The uncertainty on the flux rate estimates are highly variable, depending on space (e.g., satellite measurements at high latitudes are usually affected by poorer sensitivity), time (flux rates on shorter periods will be affected by higher uncertainties due to less satellite measurements available), and on the target gas (the errors on the VOC retrieved columns are typically larger than those on NH_3).

3.3.6.3. ALTERNATIVE METHOD

One can also estimate the lifetime T directly from the measurements, following techniques that have been developed for NO_2 and SO_2 ([RD.28], [RD.29]). In particular, the exponentially modified Gaussian (EMG) method technique has been applied on NH_3 emissions in [RD.34], where in addition to emissions, lifetimes ($T = 1/\lambda$, with λ the decay rate) and plume widths (σ) from 228 large agricultural and industrial NH_3 point sources were estimated [RD.34].

Table 3-5: NH3, C2H2, C2H4 and CH3OH point sources detection and quantification

Feature	Value
Satellite data used	IASI data products
Coverage	NH3, C2H4: Global, provided for ~50 point sources C2H2, CH3OH: Mainly in Asia. Provided for ~15 largest point sources
Temporal averages	Over the full IASI dataset timeframe (2008-2021) for all species, and annually for the main NH3 point sources
Detection threshold	NH_3: 10^{-2} kg.s^{-1} in [RD.33] (depending on thermal contrast, lifetime and vertical distribution) C_2H_2, C_2H_4 and CH_3OH: not applicable (too highly variable, depending on thermal contrast and vertical profile)
Uncertainty on flux rate estimates	Highly variable – depending on space and time

3.4. ASSESSMENT OF REGIONAL TOP-DOWN INVERSION MODELS AND DESCRIPTION OF THE ALGORITHMS USED IN WORLD EMISSION

3.4.1. INTRODUCTION

A usual strategy for providing weekly gridded emission products over regions of interest relies on top-down models with a high spatial resolution, with either Lagrangian or Eulerian mesoscale transport models. By definition, regional models have a limited area domain. This regional concentration of the computation effort allows fine-scale simulation of the atmospheric species of interest, and thus a higher resolution for the top-down approach. The World Emission inversion models infer emission details over 1) key basins emitting CH_4 from the extraction of oil, gas or coal, 2) Western Europe and Eastern China,

two densely populated and industrialized regions which are among the top global emission hotspots of greenhouse gases and pollutants, and, potentially, 3) the Eastern Amazon, a region with megacities, intense agricultural emissions, and large deforestation emissions associated with biomass burning in the dry season.

The project does not include specific regional inversion of SO_2 emissions but the global scale inversions of SO_2 emissions (see section 3.5.4) should be downscaled using high resolution inventories.

3.4.2. METHANE (CH4) BASIN-SCALE EMISSIONS

3.4.2.1. REGIONAL INVERSION

The regional inverse modelling (basin inversion) of CH_4 emissions in World Emission provides monthly total emissions estimates in selected extraction basins with numerous potential emitting sources. The assimilated data are bias-corrected CH_4 mixing ratio column data product from TROPOMI, providing global daily imagery at a horizontal spatial resolution of 5.5 × 7 km, with a sensitivity of 5 to 10 nanomole per mole (abbreviated as ppb, for parts per billion). The regional inversion model also relies on very detailed inventory data to assess the potential emitting sources within an extraction basin. This comprises coal mines from public databases, well completions data in the US from KAYRROS' Operations Tracker product, pipeline compressor station, flaring data collected from the VIIRS sensor, and bottom-up gridded emissions inventory data from EDGAR v5.0.

Basin-scale inversions are generally performed using classical Bayesian frameworks ([RD.38], [RD.39]), which are also widely used in global inversions. In these Bayesian frameworks, the optimization program

$$\mathrm{Min}_{\mathbf{x}}\,[J(\mathbf{x})] = (\mathbf{y} - \mathbf{K}\mathbf{x})^T\mathbf{S}_O^{-1}(\mathbf{y}-\mathbf{K}\mathbf{x}) + (\mathbf{x}-\mathbf{x}_P)^T\mathbf{S}_p^{-1}(\mathbf{x}-\mathbf{x}_P)$$

is solved in closed form (or using variational methods) to estimate the gridded emission rates $\mathbf{x}$. $\mathbf{x}_P$ stands for the prior emissions, $\mathbf{y}$ is the XCH_4 level retrieved from TROPOMI, $\mathbf{S}_O$ is the covariance matrix for observational error, $\mathbf{S}_p$ is the covariance matrix for prior error [RD.40], and $\mathbf{K}$ stands for the Jacobian matrix of the regional dispersion model with respect to emission rates ($\mathbf{x}$).

The regional inverse modeling (basin inversion) of CH_4 emissions used in this project introduces key changes to the classical Bayesian approach, to get rid of the prior dependency to emission inventories. The method, detailed in [RD.41], solves the following quadratic program:

$$\mathrm{Min}_{\mathbf{x}}\,[J(\mathbf{x})] = (\mathbf{y} - \mathbf{K}\mathbf{x})^T\mathbf{S}_O^{-1}(\mathbf{y}-\mathbf{K}\mathbf{x}) + \lambda\,\mathbf{x}^T\mathbf{S}_p^{-1}\mathbf{x}$$

under the constrain $\mathbf{x} \geq 0$

where $\mathbf{K}$ stands for CH_4 dispersion footprints simulated by the Lagrangian Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT; [RD.42]), $\mathbf{y}$ are methane enhancements and λ is a scaling factor meant to avoid model overfitting by a L2-penalization while keeping the quadratic program well-conditioned [RD.41].

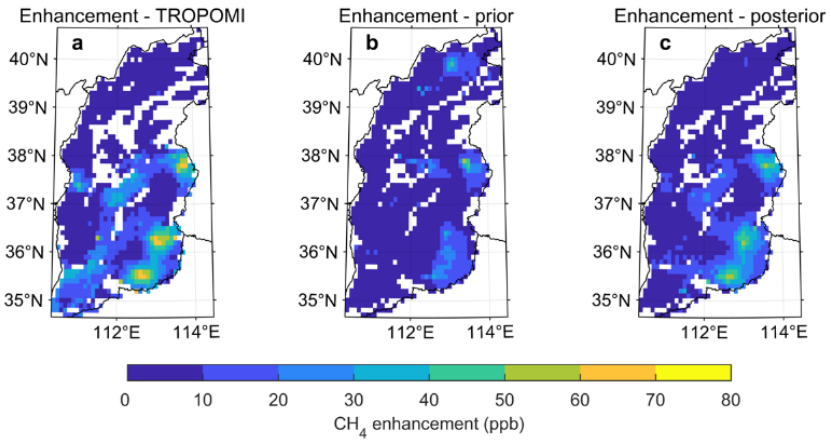


Figure 3-12: Example of TROPOMI enhancement, reconstruction based on prior emissions, and reconstruction based on inversion, in Shanxi (China). The panel a shows the TROPOMI CH4 enhancement (y), panel b shows prior-based methane enhancement reconstruction (Kxp), panel c shows posterior methane enhancement reconstruction (Kx). Source: [RD.41]

The full method can be summarized as follows: first identify all potential emitting sources; select then TROPOMI images that are suited for a basin inversion. The criteria for selecting TROPOMI images are the completeness of the images (percentage of valid pixels above some threshold) and the quality of the methane retrieval (variable "qa_value" of the soundings above some threshold). TROPOMI images are re-projected (typically, on a 0.1 × 0.1 square degree grid), and the prior (emitting sources) is re-projected on the same geometry. Then we compute methane enhancement images by removing the methane background from the TROPOMI methane mixing ratio bias corrected images. For every image, for every pixel containing a potential methane emitting source, a dispersion model (HYSPLIT using wind data from GFS) is run to obtain the footprint of methane emissions originating from this potential methane emitter, and finally the emissions are quantified by fitting the modeled to the observed methane images by solving the quadratic optimization program discussed in the previous paragraph.

This basin-level inversion is applied over key coal, oil, and gas production basins in the USA, Algeria, South Africa, Middle-East, Central Asia and Australia. The results are very consistent with [RD.38][RD.35] over the Permian shale oil and gas basin. [RD.39] estimated methane emissions from Mexico, and [RD.43] quantified methane emissions in most of the key oil, gas, and coal production basins in North America.

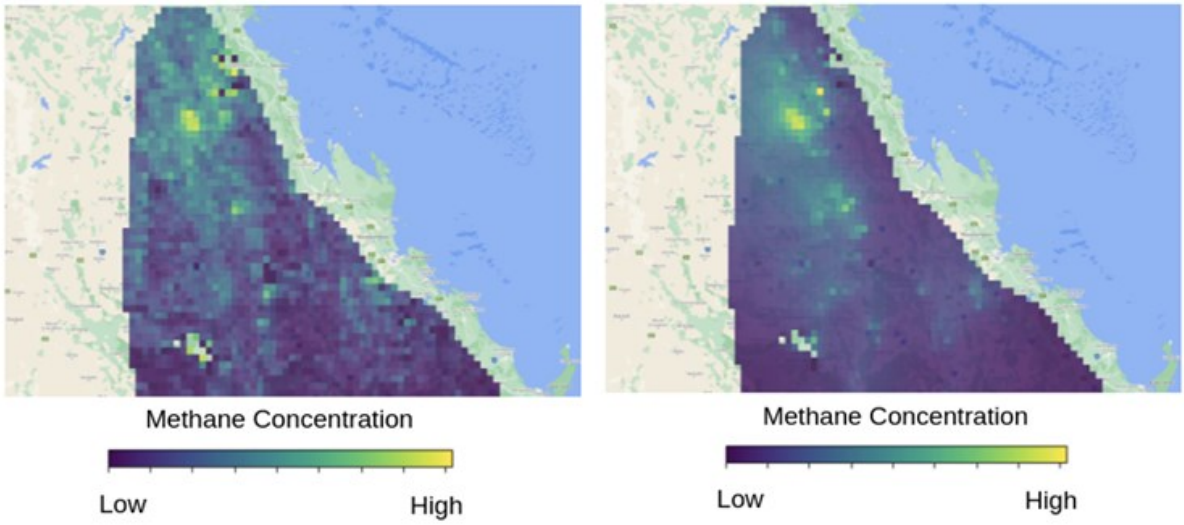


Figure 3-13 (left) observed composite methane column enhancements over a coal extracting basin in Australia. (right) inversion model fit to these enhancements after optimization of emitting coal mines sources across the entire basin

3.4.2.2. ERROR ANALYSIS

Uncertainty in the monthly methane emissions estimates mainly stems from uncertainty in the background methane concentration estimates, uncertainty in the meteorological data, TROPOMI measurement error, and uncertainty associated with the HYSPLIT simulations. To account for these uncertainties, we run simulations and estimates on an ensemble of parameters. Assuming these sources of uncertainty are independent, 1-σ sensitivity intervals are then derived by summing in quadrature the uncertainty associated with each source of uncertainty [RD.38]. In addition, uncertainty associated with the sampling of the usable TROPOMI images is taken into account by adopting a polling methods [RD.41].

Relative uncertainties of monthly methane emissions estimates are typically 20-30%. The uncertainty can vary greatly from one basin to another; these discrepancies are mainly driven by differences in TROPOMI coverage.

3.4.3. NITROGREN OXIDES (NOX), CARBON MONOXIDE (CO), CARBON DIOXIDE (CO2) REGIONAL EMISSIONS

[Major update with respect to previous version]

3.4.3.1. STATE OF THE ART

Since the 2000s, the NO_2 atmospheric mixing ratios have been monitored around the globe with space-borne instruments, such as the Global Ozone Monitoring Experiment GOME [RD.44] and GOME-2 [RD.45], the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY SCIAMACHY [RD.46], the Ozone Monitoring Instrument OMI [RD.47], and, since 2017, the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Copernicus Sentinel-5 Precursor (S5P). The CO concentrations have also been monitored with satellites, for example with the Measurement of Pollution in the Troposphere (MOPITT) instrument [RD.48]. In this context of growing availability and reliability of satellite observations, attempts have been made to develop atmospheric transport inverse modelling methods, complementary to bottom-up inventories, to derive NO_x or CO emissions from the NO_2 or CO satellite data. Global scale approaches have been considered. However, there are scientific and societal needs for the quantification of pollutant emissions at relatively high-spatial resolution, which requires the use of regional-scale systems.

Although regional inversion systems are suited to the simulation and analysis of the very large spatial and temporal variability of the CO and NO_x concentrations at fine scales, they have hardly been used to quantify CO and NO_x emissions based on satellite data. Mass-balance approaches have been performed at the regional scale [RD.49], accounting for non-linear relationships between NO_x emission changes and NO_2 tropospheric columns via reactions with hydroxyl radicals (OH) but with simple scaling factors. A more suitable account for the complex NO_x chemistry is probably required for the accurate derivation of NO_x emissions from NO_2 satellite data [RD.50]. More elaborated inversion approaches, based on the ensemble Kalman filter technique (enKF) [RD.51] have been used to infer NO_x emissions.

Variational inversion systems allow for solving high-dimensional problems, typically solving for the fluxes at high spatial and temporal resolution, which can be critical to fully exploit satellite images. If relying on a chemistry transport model whose adjoint code is available and fully accounts for the chemistry scheme used to simulate NO_2 concentrations, variational frameworks are suited to account for the non-linearities of the chemistry. In World Emission, we perform regional inversions of the NO_x and CO emissions based on such a variational atmospheric inversion framework and on such a chemistry transport model.

Due to the lack of CO_2 satellite data with dense coverage that can be used directly for the regional scale inversion of fossil fuel combustion emissions of CO_2 ($FFCO_2$ emissions), we derive these emissions from our inversions of the NO_x and CO emissions based on the NO_2 and CO satellite data. Attempts at exploiting NO_2 and CO satellite data assimilation to infer $FFCO_2$ emissions have been documented by studies like [RD.52] but this activity keeps on being exploratory. In particular, using anthropogenic CO-to-$FFCO_2$ or NOx-to-$FFCO_2$ emission ratios at coarse resolution limits the potential of the approaches due to the variations of these ratios across the sectors of activity, or in space and time. In World Emission, we attempt at supporting such an inference by strengthening the derivation and the spatial and temporal resolution of the NOx and CO emissions.

3.4.3.2. GENERAL OVERVIEW OF THE NOX, CO AND CO2 REGIONAL PRODUCTS AND OF THEIR DERIVATION IN WORLD EMISSION

Monthly maps at 0.5° resolution and national-scale budgets of the NOx, CO and fossil fuel CO_2 ($FFCO_2$) emissions have been derived for Europe, from 2019 to 2021 during the phase 1 and 2 and are now derived for Eastern China in the frame of phase 2. A potential attempt at deriving emissions in South America was initially stated as a best effort during the phase 2, but it now appears that it cannot be achieved in the frame of this phase. The corresponding sequence of computations over each region has been implemented by [RD.53] and adapted to the specific requirements in the World Emission project. It consists in:

1) the atmospheric inversions of maps of the NO_x or CO emissions during 2019-2022 at 1-day and 0.5° resolution. These inversions are based on the coupling between the variational mode of the advanced and recently developed Community Inversion Framework (CIF, [RD.54]), the CHIMERE regional atmospheric chemistry transport model ([RD.55], [RD.56]) and the adjoint code of this model ([RD.57]). This CIF-CHIMERE regional mesoscale variational inversion framework is close to the system described

by [RD.57]. The NO_x and CO inversions assimilate respectively atmospheric NO_2 and CO products from spaceborne instruments: NO_2 satellite data from OMI and TROPOMI and CO data from MOPITT. They apply 0.5°/1-day resolution corrections to the maps of anthropogenic emissions from a gridded version of the Carbon Monitor-CEDS inventory of the NO_x, CO and CO_2 emissions and to the maps of biogenic emissions of NO_x from the MEGAN global model ([RD.58]) to better fit these assimilated observations. This calculation is performed by LSCE in collaboration with LISA, a partner laboratory of the IPSL.

2) for each country in Europe (for the different administrative areas in Eastern China): the conversion of the monthly maps of NO_x or CO anthropogenic emissions from these inversions into monthly maps at 0.5° resolution and national (regional) scale estimates of the fossil fuel CO_2 emissions for five large groups of sectors of emitting activities. This conversion relies on the sectoral maps of emissions of the three species and, implicitly, on the emission ratios between the species for each sector, country (or administrative region) and month from the Carbon Monitor-CEDS inventory.

The following section details the different components of this general scheme. Results from the inversions are derived at a 1-day resolution but aggregated at 1-month resolution before the conversion into CO_2 emission estimates to ensure good confidence levels. The initial protocol relied on the analysis and conversion of the results from the inversions at 1-week resolution. This temporal resolution was coarsened to 1-month in order to decrease the computational cost of the processing chain. However, the objective is to refine it back to 1-week temporal resolution on the near-term. Furthermore, the initial target was to produce 3-monthly estimates of the monthly emission maps with a six-month (phase 1) to three-month (phase 2) latency between the emission and their estimates due to the need for accessing and processing the satellite data. However, the current lag-time associated to the distribution of the suitable satellite datasets prevented from achieving such a short latency in Phase 1 (in January 2023, the OMI-QA4ECV product was not extended beyond March 2021, and the TROPOMI-PAL product was not extended beyond November 2021). This latency will be made as short as possible by the end of the project, with the aim to reach a more realistic target during the year 2023 of 6-monthly estimates of the monthly emission maps with a six-month latency. Finally, the TROPOMI CO data were initially identified as one of the main potential source of information for the CO inversions. However, the analysis of the current TROPOMI CO products and their comparison to the CO simulations with CHIMERE raise concerns regarding the capability to exploit them to derive estimates of the CO anthropogenic emissions in Europe. On-going analysis are currently conducted to support the use of these observations in the CIF-CHIMERE CO inversions.

Coupling the results from the inversions at 1-month and 0.5° resolution for the two co-emitted species and information on the sectoral, temporal and spatial distribution of the emissions from the Carbon Monitor-CEDS inventory enables some distribution per sectors of activity, and the strengthening of the results in the focus regions. In this process, the co-emission of NO_x, CO and CO_2 by the fossil fuel combustion is exploited by combining the CO and NO_x inversions with the emission factors and ratios from the Carbon Monitor-CEDS inventory to derive fossil fuel CO_2 emission estimates. Regional emissions are compared with bottom-up regional (TNO for Europe; REAS, MEIC for Eastern China) and/or global (EDGAR, CAMS) inventories.

3.4.3.3. COMPONENTS OF THE NO_X, CO AND CO_2 REGIONAL INVERSE MODELING SYSTEM

Satellite observations

The NO_x inversions have assimilated NO_2 Tropospheric Vertical Column Density (TVCD) products: first the OMI-QA4ECV (www.qa4ecv.eu and http://temis.nl/qa4ecv/no2.html) and the TROPOMI reprocessed PAL (https://data-portal.s5p-pal.com/products/no2.html) products. The OMI-QA4ECV product was not available after March 2021 and, therefore, it has not been used in Phase 2. The TROPOMI-PAL observations have been updated and we now use the TROPOMI-v2.4 data. The CO inversions assimilate the MOPITTv8 product from 2019 to December 2021 (https://l5ftl01.larc.nasa.gov/) and the MOPITTv9 product from December 2021 to December 2022. The use of the TROPOMI CO products is currently postponed due to the concerns regarding the ability to constraint the CO anthropogenic emissions in Europe with these data, raised by the analysis of the current versions of the TROPOMI CO observations and their comparisons to the CO simulations with CHIMERE.

OMI-QA4ECV-v1.1 NO2

The Ozone Monitoring Instrument (OMI) is an ultraviolet-visible (UV-Vis) instrument launched in July 2004 onboard the Earth Observation System (EOS) Aura satellite, which flies on a 705 km sun-synchronous orbit that crosses the Equator at 13:40 LT. OMI provides measurements for the derivation of NO2 TVCD observations. Compared with other UV-Vis instruments providing a long archive of NO2 observations, it has the highest spatial resolution and least degradation ([RD.59], [RD.47]), allowing the study of long-term variations for various regions of the world. We use the OMI-QA4ECV-v1.1 NO2 TVCD level-2 product ([RD.60]). The data selection follows the criteria of the data quality statement from [RD.60]:

- the processing error flag equals 0 for a pixel
- the solar zenith angle is lower than 80°,
- the snow ice flag is lower than 10 or equal to 255,
- the ratio of tropospheric air mass factor (AMF) over geometric AMF is higher than 0.2 to avoid situations in which the retrieval is based on very low (relative) tropospheric air mass factors,
- the cloud fraction is lower than 0.5.

We add an additional criterion: the error associated to the retrieval must be lower than 100%.

TROPOMI-PAL NO_2

The Tropospheric Monitoring Instrument (TROPOMI, [RD.61]) was launched onboard the Copernicus Sentinel-5 Precursor (S5P) satellite in October 2017. It flies on a 824 km altitude sun-synchronous orbit that crosses the Equator at 13:30 LT. This imaging spectrometer covers a UV-Vis band supporting the derivation of NO2 TVCD observations. In the phase 1 of the project, we used the Product Algorithm Laboratory (PAL) version 2.3 NO_2 TVCD level-2 product ([RD.17]). The TROPOMI data have been updated and we now use the TROPOMI-v2.4 data.

Our TROPOMI data selection follows the criteria of [RD.62]. We only select observations with a quality assurance (qa) value higher than 0.75 and a cloud radiance fraction lower than 0.5. As for OMI, we only select observations when the error associated to the retrieval is lower than 100%.

MOPITT CO

We use CO observations from the MOPITT instrument ([RD.48]). The MOPITT Infrared (IR) spectrometer has been flown onboard the NASA EOS-Terra satellite, on a low sun-synchronous orbit that crosses the Equator at 10:30 and 22:30 LST since 1999. The spatial resolution of its measurements is about 22 ×22 km2 at nadir. It has been operated nearly continuously since March 2000. MOPITT CO products are available in three variants: thermal-infrared (TIR) only, near-infrared (NIR) only and the multispectral TIR-NIR product, all containing total columns and retrieved profiles (expressed on a 10-level grid from the surface to 100 hPa). In Phase 1, we choose to use the MOPITT "surface" multispectral MOPITTv8-NIR-TIR product (also called MOPITT-v8J), as it provides the highest number of observations, with a good evaluation against in situ data from NOAA stations ([RD.48]) for the period 2019 to November 2021. However, the release of the CO regional inversions for the period December 2021-December 2022 in Phase 2 is based on the most recent version of the MOPITT CO products: the MOPITT-v9 product.

Configurations of CHIMERE and of its adjoint model for the simulation of NO2 and CO concentrations

Configuration for the simulations over Europe

CHIMERE is widely used to model air quality at regional scale ([RD.63], [RD.64], [RD.65]). It is one of the seven state-of-the-art chemistry transport models (CTMs) behind the operational ensemble simulations of the Copernicus Atmosphere Monitoring Service (CAMS) regional services. Here, CHIMERE and its adjoint code are driven by the CIF to simulate NO_2 and CO atmospheric concentrations over Europe and their sensitivity to the surface emissions. It is run over a 0.5°x0.5° regular horizontal grid with 17 vertical layers, from the surface to 200 hPa, with 8 layers within the first two kilometres. The domain covers 15.25°W-35.75°E; 31.75°N-74.25°N (see Figure 3-14) and therefore includes 101 (longitude) x 85 (latitude) x 17 (vertical levels) grid-cells.

The chemical scheme used in CHIMERE and its adjoint model is MELCHIOR-2, with more than 100 reactions, including 24 for inorganic chemistry. CHIMERE is driven by the European Centre for Medium-Range Weather Forecasts (ECMWF) operational meteorological forecasts. The emissions maps and

boundary conditions used for the simulations are discussed in the next sections. Due to the need for a compromise between the robustness of the simulation of the chemistry in the model and the computational cost with a complex chemical scheme, the aerosols modules of CHIMERE have not been included in its adjoint code yet and are therefore not activated in the CHIMERE forward simulations.

Configuration for Eastern China

The configuration of CHIMERE and of its adjoint for the simulations over Eastern China also relies on meteorological fields from ECMWF, and on a 0.5° × 0.5° horizontal resolution. It is based on that of [RD.66]. In particular, the domain covers 18° N–50° N; 102° E–132° E and thus Eastern and Southern China, and Korean Peninsula (see Figure 3-16).

Comparison between the satellite observations and CHIMERE

As the spatial resolution of the satellite data is finer than that of the CHIMERE grid, the selected OMI and TROPOMI NO_2 TVCDs, and MOPITT surface concentrations are aggregated into "super-observations" in each model grid-cell. In the Phase 1 of the project, in order to associate the OMI, TROPOMI and MOPITT super-observations to a real averaging kernel, these super-observations were taken as the observations (with their averaging kernel) corresponding to the median of the observed concentrations within the 0.5°×0.5° model grid-cell and within the CHIMERE physical time step of about 5-10 minutes. In the Phase 2 of the project, this technique is still used with the MOPITT observations. However, since the spatial resolution and number of the TROPOMI observations is much higher than for the other instruments, the TROPOMI super-observations are now taken as the observations (with their averaging kernel) corresponding to the value closest to the mean of the observed concentrations within the CHIMERE grid-cell and physical time step.

The CHIMERE simulation of the CO concentrations and NO_2 TVCDs corresponding to these super-observations are derived from the CO and NO_2 simulated 3D-fields by applying vertical integrations with the averaging kernels associated to the super-observations (Figure 3-14, Figure 3-15 and Figure 3-15Figure 3-16, Figure 3-17 respectively).

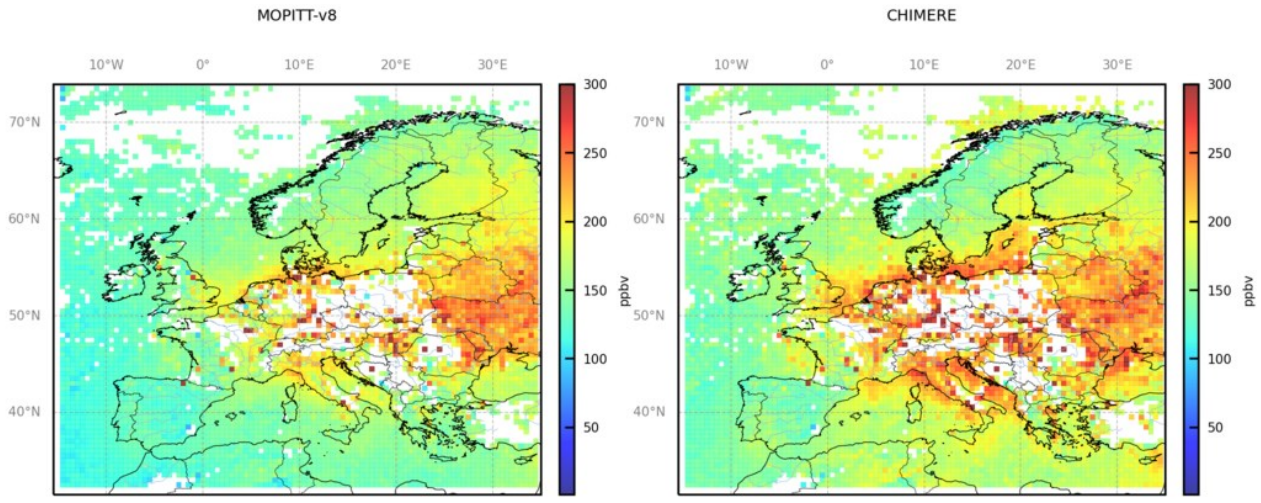


Figure 3-14 Domain of our CHIMERE configuration and for the inversions in Europe with the averages of the CO concentrations a) observed by MOPITT-v8J and b) simulated by CHIMERE using the Carbon Monitor-CEDS anthropogenic emission estimate where and when MOPITT-v8J "surface" super observations are available, in ppbv, at the 0.5°x0.5° CHIMERE grid-cell resolution, in January 2019.

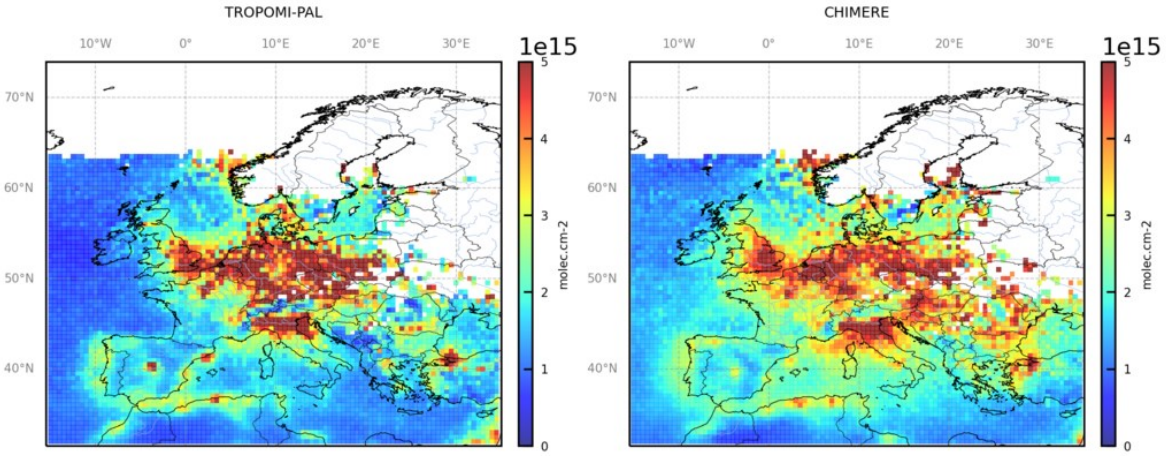


Figure 3-15 Averages of the NO_2 tropospheric columns in Europe a) observed by TROPOMI-PAL and b) simulated by CHIMERE using the Carbon Monitor-CEDS anthropogenic emission estimate where and when TROPOMI-PAL super observations are available, in molec.cm^{-2}, at the 0.5°x0.5° CHIMERE grid-cell resolution, in January 2019.

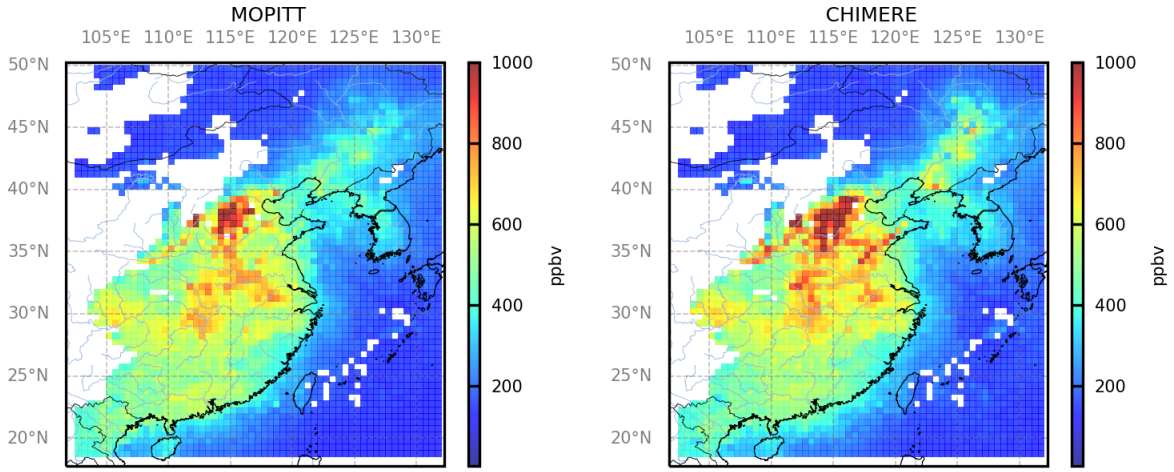


Figure 3-16 Domain of our CHIMERE configuration and for the inversions in Eastern China with the averages of the CO concentrations a) observed by MOPITT-v8J and b) simulated by CHIMERE using the Carbon Monitor-CEDS anthropogenic emission estimate where and when MOPITT-v8J "surface" super observations are available, in ppbv, at the 0.5°x0.5° CHIMERE grid-cell resolution, in January 2019.

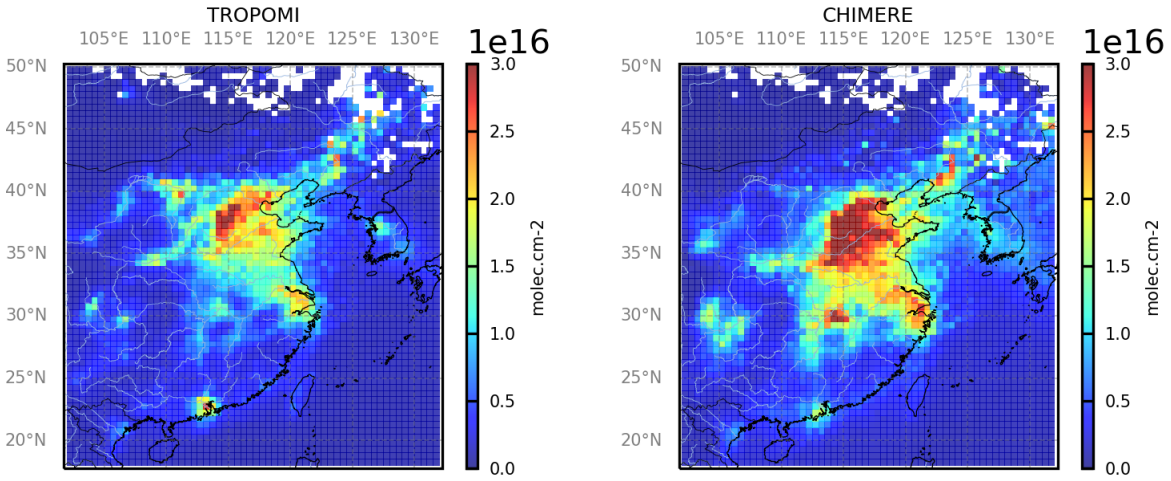


Figure 3-17 Averages of the NO2 tropospheric columns in Europe a) observed by TROPOMI-PAL and b) simulated by CHIMERE using the Carbon Monitor-CEDS anthropogenic emission estimate where and when TROPOMI-PAL super observations are available, in molec.cm-2, at the 0.5°x0.5° CHIMERE grid-

Prior estimate of the emissions and of the boundary conditions

The Bayesian framework of the atmospheric inversion underlies the use of a prior estimate of the NO_x and CO emissions in the NO_x and CO emission inversions. This prior estimate feeds the initial simulation with CHIMERE in the iterative process of the variational inversions. The inversion system controls the initial and lateral and top boundary conditions for the concentrations in the model together with the surface emissions, but as for the latter, it requires a prior estimate of these boundary conditions. Actually, here, the inversion rescales some of these "prior" estimates for the emissions and boundary conditions, so that, strictly speaking, the corresponding emission products detailed below belong to the observation operator of the inversions. Finally, the simulations of the atmospheric chemistry require estimates of the regional emissions of VOCs, which are not modified by the inversion.

For both regional configurations (Europe and Eastern China), the prior anthropogenic NO_x, CO and $FFCO_2$ emission estimates from 2019 to 2021 are taken from the Carbon Monitor-CEDS gridded inventory ([RD.67]), mapped at a 0.5° horizontal resolution (see). The 11 energy-related and cement production sectors of EDGAR are aggregated into six main sectors, including power, industry, ground transportation, residential (public, commercial buildings and households), domestic and international aviation. The temporal disaggregation of the daily into hourly emissions is based on temporal profiles provided per GNFR sector code with typical diurnal variations from the TNO inventory ([RD.68]) for Europe, and from the MEIC inventory for Eastern China. For this study, the international aviation is not considered.

The prior biogenic NO_x emissions are derived from the MEGAN model ([RD.58]). The fixed VOCs biogenic emission estimates also come from the MEGAN model. CO biogenic emissions are assumed to be negligible and are not considered. For Europe, the fixed anthropogenic emissions for VOCs are obtained from the EMEP inventory (https://www.ceip.at). For Eastern China, the fixed anthropogenic emissions for VOCs are obtained from the HTAP inventory [RD.114].

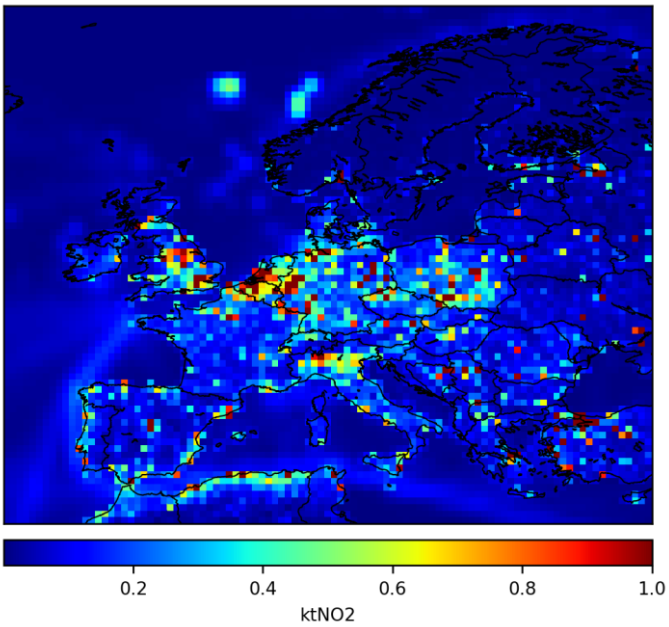


Figure 3-18 Annual NO_x emissions for the year 2019 as estimated by Carbon Monitor-CEDS, in $ktNO_2$/year, at the 0.5°x0.5° CHIMERE grid-cell resolution.

Climatological values from the LMDZ-INCA global model are used to prescribe concentrations at the lateral and top boundaries and the initial atmospheric composition in the regional domains.

Variational inversion of the NO_x and CO emissions

The inversions of the NO_x or CO emissions follow a Bayesian theoretical framework. They consist in correcting the prior estimate of these emissions and of the model initial and boundary conditions to improve the fit between CHIMERE and the satellite NO_2 or CO data, respectively. [RD.57] provides details on the principle and configuration for these variational inversions. The optimal (posterior) estimate of the emissions in a statistical sense is found by iteratively minimizing the following cost function J(x):

$$J(\mathbf{x})=(\mathbf{x}-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)+(H(\mathbf{x})-\mathbf{y})^T\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{y})$$

where

- $\mathbf{x}$ is the control vector and $\mathbf{x}^b$ is its prior value, which are detailed in the following

- the observation vector $\mathbf{y}$ gathers the super-observations from the OMI, TROPOMI or MOPITT datasets presented in sections 3.2 and 3.4

- the observation operator H projecting the control vector to simulations of concentrations corresponding to these super-observation combines the emission and boundary conditions presented in section 3.5, the CHIMERE model presented in section 3.3 and the sampling and vertical interpolation corresponding to the observations as explained in section 3.4.

- the prior and observation error covariance matrices $\mathbf{B}$ and $\mathbf{R}$ characterize the uncertainties in $\mathbf{x}^b$ and in $\mathbf{y}$ and H respectively, under the assumption that these uncertainties are unbiased and follow a normal distribution

Series of inversions over successive 1-day windows are conducted for NO_x and then combined to provide an estimate of the NO_x emissions over the whole period of analysis. The control vector $\mathbf{x}$ in the NO_x inversions contains the variables to be optimized by the NO_x inversion over Europe and Eastern China (the dimensions of the control vector provided below correspond to the configuration for Europe):

- the logarithm of the NO and NO_2 anthropogenic emissions at a 1-day temporal resolution, at a 0.5° × 0.5° (longitude, latitude) horizontal resolution and over the first 8 vertical levels of CHIMERE i.e, for each of the corresponding grid cells,
- the logarithm of the NO biogenic emissions at a 1-day temporal resolution, at a 0.5° × 0.5° (longitude, latitude) resolution and at the surface (over 1 vertical level only), i.e. for each of the corresponding grid cells ,

- factors scaling the daily NO and NO_2 3D initial conditions at 0:00 UTC, at a 0.5° × 0.5° (longitude, latitude) resolution and over the 17 vertical levels of CHIMERE.

For CO, series of inversions over successive 1-month windows are conducted. The control vector **x** for the CO inversions is similar to that for the NO_x inversions. However, due to the relatively long lifetime of CO, we have to account for the uncertainties in the CO lateral boundary conditions at the borders of the CHIMERE domain. Furthermore, the chemical production of CO by VOCs is fixed and not controlled here by the inversion. Therefore, the control vector for the CO inversions over Europe and Eastern China (the dimensions of the control vector provided below correspond to the configuration for Europe) contains the following scaling factors:

- factors scaling the CO anthropogenic emissions at a 1-day temporal resolution, at a 0.5° ×0.5° (longitude, latitude) horizontal resolution and over the first 8 vertical levels of CHIMEREfactors scaling the CO lateral boundary conditions at 1-day temporal resolution, at a 0.5° × 0.5° (longitude, latitude) horizontal resolutionfactors scaling the CO 3D initial conditions at 0:00 UTC the first day of the month, at a 0.5° × 0.5° (longitude, latitude) horizontal resolution and over the 17 vertical levels of CHIMERE.

The prior estimate $\mathbf{x}^b$ of the control vectors for the inversions consist in either the products presented above in the section "prior estimate of the emissions and of the boundary conditions" (for the NOx emissions) or sets of unitary scaling factors (for the scaling factors to be applied to initial and boundary conditions, or to the CO emissions), which implicitly fixes the prior knowledge of the emissions and boundary conditions to the products detailed in the section "prior estimate of the emissions and of the boundary conditions" in all cases. The assumptions made to define the **B** and **R** matrices for Europe and for Eastern China have been detailed in [RD.57]. The main features for the parts of **B** corresponding to the surface emissions are recalled in Table 3-6. However, the configuration of the **R** matrix has changed compared to Phase 1: in Phase 2, the covariance matrix **R** only takes into account the estimates of observation errors reported in the TROPOMI data sets. Indeed, the errors associated to the observation operator (in particular those associated to the chemistry-transport modelling with CHIMERE) are ignored since they are assumed to be much smaller than those associated to the TROPOMI data.

Table 3-6: Assumptions on the uncertainty in the NO_x and CO emission products detailed in section 3.5 that are used to derive the covariance matrix of the prior uncertainty B over Europe and over Eastern China.

	Anthropogenic emissions	Biogenic emissions
NO_x	50% at 0.5° × 0.5° and daily scale Spatial correlations of 50km No day to day temporal correlation	100% at 0.5° × 0.5° and daily scale Spatial correlations of 50km
CO	100% at 0.5° × 0.5° and daily scale Spatial correlations of 50km No day to day temporal correlation	

The inversions are conducted using the variational mode of the CIF with the M1QN3 algorithm for the minimization of the cost function J [RD.69]. At each iteration of this minimization, the CIF uses a CHIMERE simulation to compute J for a new estimate of **x** and the adjoint code of CHIMERE to compute the gradient of J for this new estimate of **x** ([RD.57], [RD.54]).

In the frame of Phase 1, for Europe, three inversions had been conducted:

- the inversion of NO_x emissions based on the OMI data over 2019-March 2021

- the inversion of NO_x emissions based on the TROPOMI data over 2019-November 2021

- the inversion of CO emissions based on the MOPITT data over 2019-2021

In the frame of Phase 2, for Europe, two new inversions have been conducted:

- the inversion of NO_x emissions based on the TROPOMI data over 2019-2022

- the inversion of CO emissions based on the MOPITT data over 2019-2022

Two years of inversions of NO_x emissions for Eastern China based on the TROPOMI data have also been performed.

The posterior estimate of the anthropogenic emissions from these different NO_x or CO inversions are used for the subsequent analysis.

Derivation of the FFCO2 emissions from the NO_x or CO inversions

The derivations of the 1-month/0.5° resolution maps and national scale (for Europe; regional scale for Eastern China) budgets of the $FFCO_2$ emissions with a sectoral distribution (across five large sectors of activity) relies on the conversion of the estimate of the anthropogenic emissions from the different NO_x or CO inversions into estimates of 1-month, national/regional and sectoral scaling factors to be applied to the Carbon Monitor-CEDS inventory in order to better fit the satellite data.

For this purpose, the monthly averages of the maps of national/regional and sectoral NO_x or CO emissions provided by the Carbon Monitor-CEDS inventory are compared to the monthly maps of total anthropogenic emissions from the inversions at 0.5° resolution. The sectoral distribution is defined here by five large groups of sectors of anthropogenic activity: energy, industry, residential, road transport and the rest of the sectors.

For each country in Europe (or for the different administrative region in Eastern China) and each month, optimal national/regional and sectoral scaling factors to be applied to the Carbon Monitor-CEDS NO_x and CO emissions maps are derived to fit the 0.5° resolution national/regional and monthly maps of NO_x and CO emission from the inversions. A simple analytical inversion scheme is used to infer the optimal set of five sectoral scaling factors for each country/region and month. These optimal scaling factors are used to derive new estimates of the CO_2 fossil fuel emissions at 0.5° and 1-month resolution.

On the near-term, this conversion for the derivation of $FFCO_2$ emissions should be conducted at a refined temporal resolution of 1-week instead of 1-month.

3.4.3.4. ERROR ANALYSIS

We cannot provide an error analysis for these products at this stage. The comparison of these regional products to independent inventories and experiments with pseudo-data (with synthetic true emissions, emission ratios and concentrations) are used to assess the uncertainties in these products (due to uncertainties in the assimilated observations, in the atmospheric chemistry transport model and thus in the variational inversions, in the emissions ratios from the Carbon Monitor-CEDS inventory, etc.).

3.4.4. AMMONIA (NH3) REGIONAL EMISSIONS

Next to traditional inverse modelling efforts for NH_3 [RD.70], the simpler mass-balance approach pioneered by [RD.33] for the dense IASI data and presented in Section 3.3.6.1 will be implemented. It allows calculating regional emissions of NH_3 as yearly emissions and as a weekly climatology for coherent emission regions (e.g., Po Valley, Central US). These regional NH_3 emissions will be compared with bottom-up regional (such as REAS, NEI or EMEP) and global (such as EDGAR or CEDS) inventories to make an overall assessment of their accuracy. The error analysis is similar to the one detailed for the NH_3 point sources (see Section 3.3.6.2).

An important, but poorly constrained source of NH_3 is emissions from fires. Expanding on the work performed in [RD.71], we will use the full IASI dataset to constrain NH_3 emissions over the Amazon region, and possibly extend this approach to the world's largest biomass burning regions. The product will be compared against the GFED inventory.

Whatever the responsible sector, seasonal variations of NH_3 emissions are currently poorly known and almost absent in inventories. For example, some of the most advanced models use a single seasonal cycle of emission for the entire Europe, while it is in fact only representative for the Netherlands. This causes model failure in reproducing local and regional pollution episodes, such as spring-time hazes that occur after spreading manure intensively on the agricultural fields. The IASI dataset of daily NH_3 global distributions gives a unique opportunity to examine the temporal variability of regional NH_3 emissions under a range of climates and practices (agricultural, industrial, societal) for the target regions and other high emission regions of the world. We will derive seasonal profiles of NH_3 for the main hotspot regions.

3.5. ASSESSMENT OF GLOBAL TOP-DOWN INVERSION MODELS AND DESCRIPTION OF THE ALGORITHMS USED IN WORLD EMISSION

3.5.1. INTRODUCTION

The World Emission strategy to deliver global gridded emission products for the target species mostly relies on the global PYVAR-LMDZ-SACS variational inversion system. This system includes the LMDZ transport model guided by meteorological wind reanalyses and a full chemistry model (INCA) (section 3.5.2). The application to the various target species is discussed in the following sections (3.5.3 - 3.5.7). Ancillary data are the global prior emission inventories (described in each section) and ECMWF wind fields (described in LMDZ INCA section 3.5.2).

3.5.2. THE LMDZ-INCA MODEL

The LMDZ-INCA global chemistry-aerosol-climate model couples on-line the LMDZ (version 6) General Circulation Model ([RD.72], [RD.73]) and the INCA (INteraction with Chemistry and Aerosols) (version 5) model ([RD.74], [RD.75]). The interaction between the atmosphere and the land surface is ensured through the coupling of LMDZ with the ORCHIDEE (version 1.9) dynamical vegetation model [RD.76]. The model now includes 79 hybrid vertical levels extending up to 70 km. Two horizontal resolutions are available. The development version has a resolution of 1.25° in latitude and 2.5° in longitude and the production version uses a 1°×1° resolution. The meteorological data from the ECMWF ERA-5 reanalysis are used to the nudge the GCM winds. The relaxation of the GCM winds towards ECMWF meteorology is performed by applying at each time step a correction term to the GCM u and v wind components with a relaxation time of 2.5 hours ([RD.74], [RD.77]). The ECMWF fields are provided every 6 hours and interpolated onto the LMDZ grid.

INCA initially included a state-of-the-art CH_4-NOx-CO-NMHC-O_3 tropospheric photochemistry ([RD.74], [RD.78]). The tropospheric photochemistry and aerosols scheme used in this model version is described through a total of 123 tracers including 22 tracers to represent aerosols. The model includes 234 homogeneous chemical reactions, 43 photolytic reactions and 30 heterogeneous reactions. For tropospheric aerosols, the INCA model simulates the distribution of aerosols with anthropogenic sources such as sulfates, nitrates, black carbon (BC), organic carbon (OC), as well as natural aerosols such as sea salt and dust. The heterogeneous reactions on both natural and anthropogenic tropospheric aerosols are included in the model ([RD.79], [RD.74], [RD.75]). Chemical species and reactions specific to the middle atmosphere are included in this model version. A total of 31 species were added to the standard chemical scheme, mostly belonging to the chlorine and bromine chemistry, and 66 gas phase reactions and 26 photolytic reactions. Heterogeneous processes on Polar Stratospheric Clouds (PSCs) and stratospheric aerosols are parameterized in INCA following the scheme implemented in [RD.80].

For the baseline simulations, the a-priori anthropogenic emissions compiled by [RD.81] and the biomass burning emissions from GFED4 [RD.82] are added to the natural fluxes used in the INCA model. The ORCHIDEE vegetation model has been used to calculate off-line the biogenic surface fluxes of isoprene, terpenes, acetone, and methanol as described by [RD.83]. Natural emissions of dust and sea salt are computed using the 10-m wind components from the ECMWF reanalysis. The lightning NOx emissions are parameterized in the model based on convective cloud heights as described in [RD.84]. Based on this parameterization, the total lightning NOx emissions for the baseline simulation is 5.5 TgN/yr.

The INCA model daily and monthly fields for key species are made available as NetCDF files.

3.5.3. CARBON MONOXIDE (CO) AND METHANE (CH4)

PYVAR-LMDZ-SACS is a multi-tracer variational inverse system with the capacity to assimilate observations of the CH4-HCHO-CO oxidation chain to better constrain the sources and sinks of these species at individual model grid cell ([RD.85], [RD.86], [RD.87], [RD.88], [RD.89]). Given observed changes in temporal and spatial variations of all the three species, the system optimizes simultaneously (i) methane emissions, (ii) CO emissions, (iii) HCHO sources (surface emissions + chemical productions from volatile organic compound oxidation), and (iv) OH concentrations. These terms are optimized at a weekly temporal resolution and a 1.9° by 3.75° spatial resolution. Besides, initial concentrations of all the four species are also optimized at individual horizontal model grid. By design, the multispecies inversion system benefits from more observational information than a single species system to distinguish between primary and secondary sources and optimize different species involved in the same oxidation chain in a synergistic way. The chemical feedback between different species (e.g., increased

CO concentrations – lower OH – increased CH_4 lifetime) can also be addressed much better than a single species system. It exploits the wide availability of satellite observations of CH_4, HCHO and CO total or partial columns or vertical profiles, e.g., reaching 20 years for MOPITT CO retrievals. Emissions can also be disaggregated into source sectors from inversion-based total fluxes based on the heterogeneity of spatial-temporal distributions of different sources.

3.5.3.1. DESCRIPTION OF THE INVERSION

First, we assimilate the latest MOPITT version 9 CO retrievals[2], which have benefited from improved cloud detection (resulting in substantially improved retrieval coverage over land) and mapping of highly polluted scenes compared with previous MOPITT retrieval versions. Specifically, we use the Level 2 MOPITT version 9 "multispectral" retrieval product which simultaneously exploits MOPITT's thermal-infrared (TIR) and near-infrared (NIR) measurements. Second, the prior anthropogenic fossil fuel and biofuel fluxes 2000–current year is derived from the latest Community Emissions Data System (CEDS) emission inventory, which corrects an overestimation of global CO emissions in the previous CEDS version, especially in China and India. Third, to provide prior fluxes for 2020 and 2021, while the CEDS data ends in 2019, we used daily country- and sector-level CO_2 emission growth rates from the Carbon Monitor project and combined them with the CEDS emissions data in 2019. The integration of MOPITT CO retrievals and fast-track Carbon Monitor emission growth rates provides the near real-time update of observational constraints and prior anthropogenic fluxes and helps better attribute the observed enhanced CO plume signals over burning regions to fire emission sources. The fire emission flux priors were derived from the fourth version of the Global Fire Emissions Database (GFED 4.1s) from 2000 to current year, and the 2000–2020 average flux was used as the prior flux for 2021.

The inversion-based global CO emissions are further processed to estimate fire and anthropogenic CO and CO2 emissions. Fire CO emissions are estimated using our previously developed approach, which combines inversion total CO emissions with percentages of fire emissions in each model grid cell based on the prior information. Fires dominate CO emissions in burning grid cells during fire seasons because fires concentrate over regions distinct from human settlements and present large seasonal cycles due to local environmental conditions; therefore, fires exhibit contrasting spatial–seasonal distribution patterns than anthropogenic fossil fuel and biofuel CO sources. The estimated fire CO emissions are converted to CO2 emissions based on a dynamic conversion factor ratio map of CO2 to CO, which are estimated through inversion-based modified combustion efficiency and based on field measurement datasets4. We estimate global fire CO2 emissions from 2000 to 2020 using this method and rely on the monthly average CO2-to-CO emission ratio maps between 2000 and 2020 to infer fire CO2 emissions in the current year. The multiannual average of CO2 to CO emission ratio maps helps constrain the random errors and reduce the influence of potential uncertainties on the estimate of fire emissions.

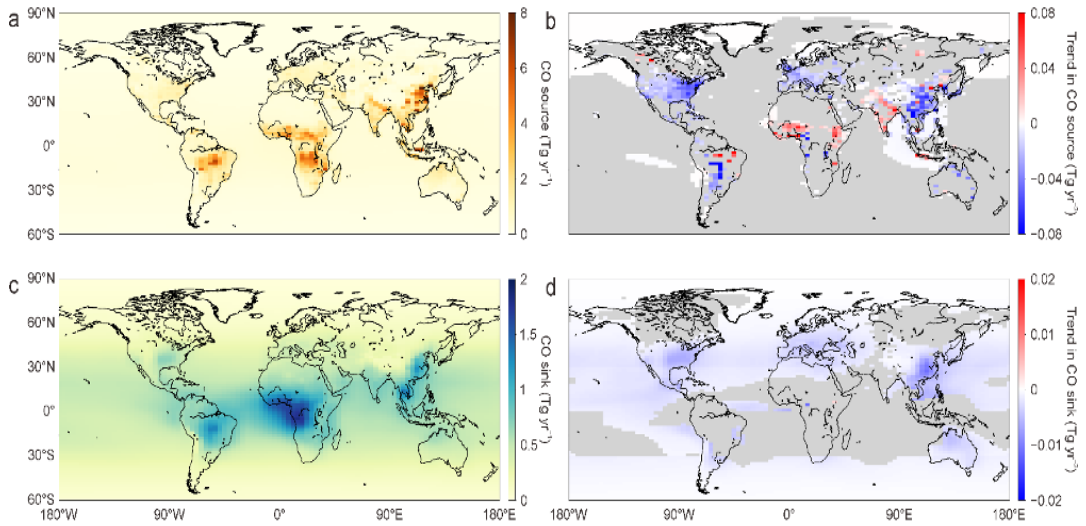


Figure 3-19 Spatial distribution of the global CO budget and 2000–2017 trends. Annual average CO total source and sink during 2000–2017 are shown at the spatial resolution of 3.75° longitude × 1.9° latitude in (a) and (c), respectively, and linear trends of each grid cell are shown in (b) and (d), which are estimated using the linear least squares fitting method based on annual time series

For CH_4 inversions, the prior fluxes were compiled from existing bottom-up estimates for different anthropogenic and natural sources. The dataset incorporates recent development of emission

inventories and current understanding of various CH_4 sources and sinks ([RD.90]), and therefore has been proposed for use as the prior for top-down CH_4 inversions contributing to the next phase of the global methane budget assessment. The OH and O(^{1}D) fields were prescribed from model output of the chemistry-climate model LMDZ-INCA with a full tropospheric photochemistry scheme, using anthropogenic emissions from CEDS inventory up to 2019 and were updated to 2020 and 2021 based on the sectoral CO_2 emission changes estimated by Carbon Monitor. Observations used to constrain CH_4 fluxes were obtained from either ground-based greenhouse gas monitoring networks or GOSAT XCH_4 retrievals. More details of model configurations were documented in [RD.91]. The exceptionally high atmospheric CH_4 growth rates in 2020 and 2021 (15.2 and 18.1 ppb yr^{-1} respectively, highest since 1980s when ground-observations started) have been studied using the global CH_4 inversion system described here, together with near-real-time bottom-up estimates to identify regions and drivers that explain the phenomenon (Figure 3-20, [RD.91]). The inversion results were evaluated against independent observations from various networks and platforms including TCCON XCH4 observations and vertical profiles from aircraft and AirCore measurements. The evaluation demonstrated overall good performance of our global CH_4 inversions in terms of representing large-scale atmospheric mixing and separating latitudinal fluxes [RD.91].

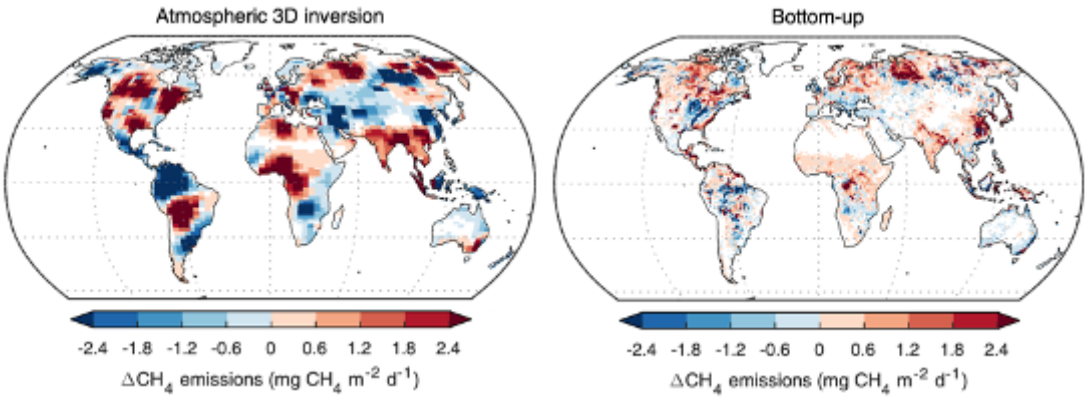


Figure 3-20 Spatial patterns of CH4 emission anomaly between 2020 and 2019 from top-down inversions (left panel) and bottom-up estimates including wetland, fire and anthropogenic emissions (right panel).

3.5.3.2. UNCERTAINTY ANALYSIS

[Major update with respect to previous version]

The error analysis is performed as in [RD.87] using a Bayesian approach. The core of atmospheric Bayesian inversion is the minimization of the following cost function:

$$J(\mathbf{x}) = \left(\mathbf{x}-\mathbf{x}^{\mathbf{b}}\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^{\mathbf{b}}\right) + \left(H(\mathbf{x})-\mathbf{y}\right)^T \mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}\right)$$

x is the control vector that gathers the target variables we seek to optimize, and xb is a prior guess of these variables assuming unbiased Gaussian error statistics represented by a covariance matrix B. y is the observation vector containing all the observation data assimilated to constrain the inverse problem; their error statistics are assumed to be unbiased and Gaussian with a covariance matrix R. H is the forward model (the combination of the LMDz-SACS model, a sampling operator, and an averaging kernel operator) that calculates the equivalent of the observation data in y based on the control vector x. The forward model error and the representation error caused by the mismatch between model and observation resolutions are also included in R, making R represent a combination of measurement, forward model, and representation errors. Configurations of the variables and vectors in the Equation above are summarized in [RD.89], most of which have already been described in previous papers ([RD.92], [RD.93]).
For CH4, we also perform an ex-post error analysis by taking the spread of 6 different inversions using 3 observation datasets (ground based, GOSAT U Leicester GSUOL, NIES full physics retrievals GSNIES) times two transport model settings with two different vertical transport (convection) parameterizations in Figure 3-21. The two GOSAT retrievals have been widely used in global or regional methane inverse

modeling to study recent trends and interannual variabilities. The ensemble of six inversions shows global increases in surface CH4 emissions by an average of 20.3±9.9 Tg CH4 yr-1 and 24.8±3.1 Tg CH4 yr-1 in 2020 and 2021 compared to 2019, the error being from the spread of inversions suggesting rather robust assessment of changes, although with a spread of ≈ 50% of the change from 2020 minus 2019

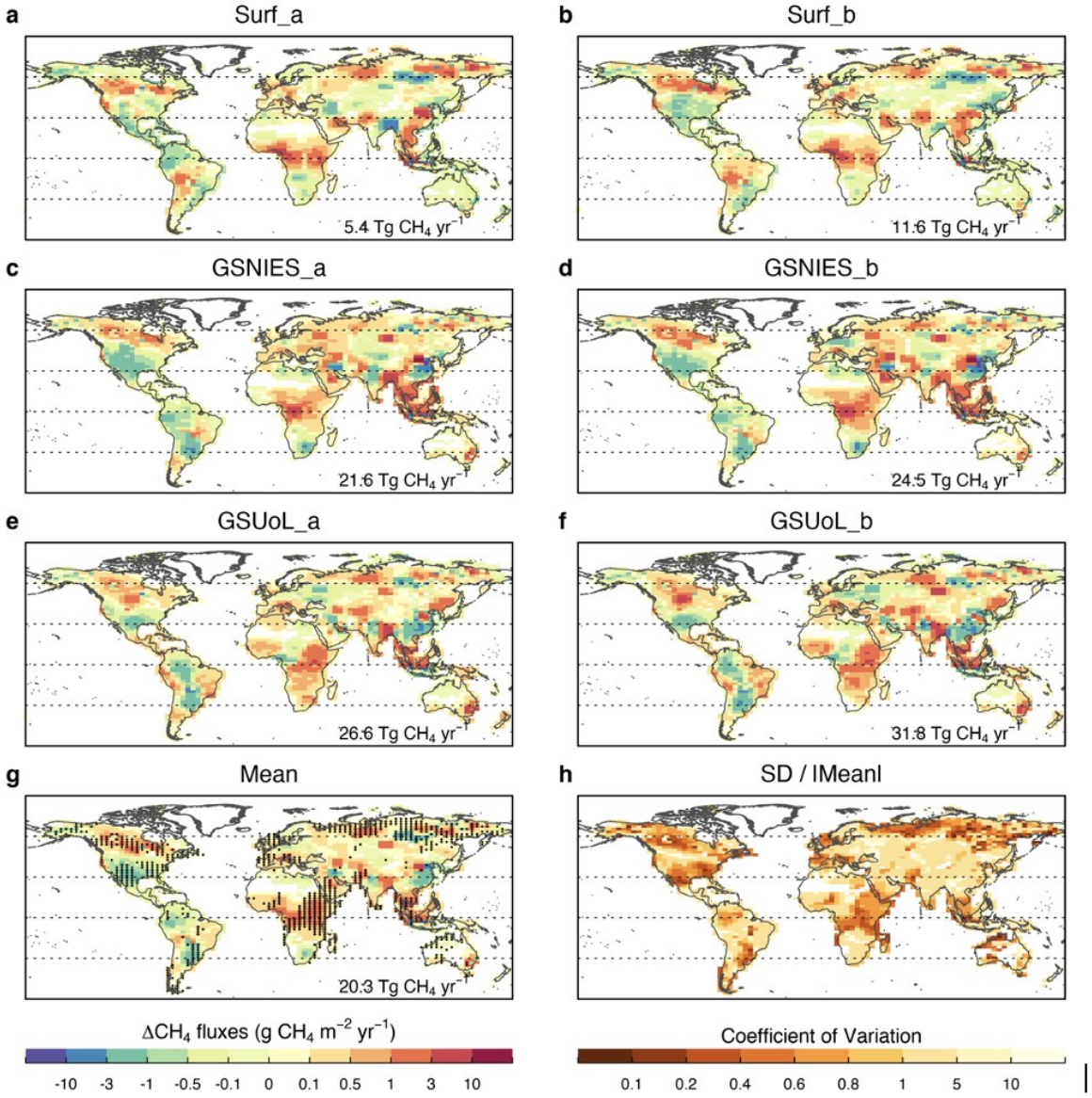


Figure 3-21 Anomalies of CH4 emissions in 2020 relative to 2019 derived from the ensemble of six inversions. a-f, Spatial patterns of CH4 emission anomalies derived from each member of the ensemble. The global net emission change is given for each panel on the bottom right. g, Spatial pattern of CH4 emission anomalies averaged over the six inversions of the ensemble. The shaded areas indicate that posterior fluxes from all six inversions have the same changing direction. h, Coefficient of variation in the CH4 emission anomalies from the six inversions of the ensemble. For each model grid, the coefficient of variation is defined as the standard deviation (SD) of emission anomalies from six inversions divided by the absolute value of their mean. Darker colors indicate better agreement among inversions.

3.5.4. NITROGEN OXIDES (NOX), SULPHUR DIOXIDE (SO2), AMMONIA (NH3)

[Major update with respect to previous version]

For reactive species NOx, SO_2 and NH_3 which have relatively short lifetimes of hours to days, atmospheric chemistry processes need to be explicitly modelled. World Emission uses LMDZ-INCA to invert weekly emissions in each grid cell over the globe at a resolution of 1.27° by 1°. The approach was initially (during the phase 1 of the project) based on [RD.94] where the local sensitivity of pollutant concentrations to underlying emissions change is computed from simulations with perturbed emissions starting from a prior inventory, to quantify the emissions of the three reactive species. We have now also implemented a mass-balance method as described in [RD.95] and [RD.96] to estimate the total emissions of the three reactive species. The main innovation to provide rapid availability of realistic top-down emissions to users is to use the near real time Carbon Monitor inventory providing daily emissions maps from activity data for six sectors: aviation, ships, industry, power generation, ground transportation and residential. The Carbon Monitor near real time system produces national daily CO_2 emissions that will be further downscaled into prior emission maps at 10 x 10 km spatial horizontal resolution and converted to prior emissions for NOx and SO_2, co-emitted with fossil CO_2 by combustion processes, using sector and fuel specific emission ratios. To extend the inversion over a longer period and for emissions we will combine the Carbon Monitor daily inventory with the EDGAR spatial patterns of activities and NOx - SO2 to CO2 emission ratios. This work will be performed in collaboration with the EDGAR emission team at the Joint Research Center of the European Commission. For NH_3 emitted by industrial facilities and agricultural sources, we directly use the existing CEDS inventory [RD.81]. For NH_3, another possible and complementary approach will be to determine the NH_3 lifetime from the global model and then apply a local mass balance to derive its emissions based on satellite data (i.e., IASI). This method was recently applied by [RD.97] to estimate the global NH_3 emissions over the period 2008-2017.

The inversion framework for NOx, SO2, and NH3 global source inversions is based on a coarse (1.27°×2.5°) spatial resolution version of the global climate-aerosol-chemistry transport model LMDZ-INCA. LMDZ-INCA, which is used for simulating atmospheric NO_2 concentrations, is a coupled model based on the general circulation model LMDZ (Laboratoire de Météorologie Dynamique, V6; [RD.72]) and the chemistry model INCA (INteraction with Chemistry and Aerosol, V5; [RD.74], [RD.75]). The model uses a monotonic finite-volume second order parameterization to calculate large-scale advection of water vapor, liquid and solid water, and tracers. The simulations of global atmospheric NO_2 are forced with monthly global anthropogenic emission inventories from open-source Community Emissions Data System (CEDS) [RD.98] with a horizontal resolution of 0.5°×0.5°. The CEDS inventories offer the NO_x, SO_2, NMVOCs, NH_3, CO, OC, and BC emission from eleven anthropogenic sectors which are agriculture, energy, on-road, non-road transportation, residential, commercial, other, waste solvents, international shipping. In the present configuration, the horizontal resolution of LMDZ-INCA is 1.27°×2.5° (model spatial grid lat × lon: 143 × 144), and the vertical dimension is divided into 79 hybrid σ-θ pressure levels which stretches up to 80 km. The global simulations of these species consist in a spin-up from 2010 to 2018 and then on simulations from 2019, which are evaluated with the satellite observations.

3.5.4.1. SATELLITE OBSERVATIONS

NOx: We utilize the reprocessed PAL data product (S5P_PAL_L2__NO2) of the tropospheric total column NO_2 from TROPOMI for evaluating the model simulation in 2019 and for the global source inversions of anthropogenic NO_2 emissions in 2020-2021. Following the recommendations joined to the TROPOMI product of NO_2, we select the observations according to the threshold value of qa_value greater than 75 (i.e. high quality observations which exclude the cloud effect, snow cover, and any other problematic pixel from retrievals).

SO2: We use satellite SO2 total vertical column densities obtained by the TROPOMI and the OMI SO2 product (OMSO2.003) from OMI. As the TROPOMI operational offline L2 SO2 data product has high noise levels and large-scale variable biases, we use the TROPOMI COBRA SO2 data product [RD.26] in this study which has comparatively smaller noise. We select only measurements from both TROPOMI-COBRA and OMI and use filtering criteria for to select high quality observations from the inversions. For OMI, we select data with total column values between -10 and 5 DU (1 Dobson Unit (DU) = 2.69 x 10^{16} molecules cm^{-2}), with solar zenith angles (SZA) less than 65°, cloud radiance fraction less than 0.3, Air Mass Factor (AMF) greater than 0.2. We only use 6th-55th pixels on the cross-track dimensions which are less affected by the row anomaly and data in the regions with South Atlantic Anomaly (SAA) are also excluded. For TROPOMI-COBRA, we select high quality observations with QA_flag_filter = 1, discard

first and last 25 pixels on the cross-section dimensions due to row anomaly which has large radiance shot noise due to a difference in detector signal binning at the swath edges [RD.26], and apply the same filter for AMF, cloud radiance fraction as for the OMI dataset with SZA less than 60°, and the total column values between -5 and 5 DU. As we focused mainly on anthropogenic emissions, in order to minimize the influence of volcanic SO2 plumes on satellite SO2 columns, we detected and filtered volcanic SO2 pixels from TROPOMI & OMI SO2 products. We detected volcanic SO2 columns based on an algorithm presented in TROPOMI L2 product. In this algorithm for TROPOMI, a SO2 column is assigned corresponding to a volcanic plume if the SO2 VCD assuming a box profile at 15 km altitude (i.e. SO2 VCD_box) is greater than a SO2 threshold (= 0.3 DU). For OMI, we used SO2 VCD assuming upper tropospheric SO2 profile with a center of mass altitude (CMA) of 13 km (i.e. ColumnAmountSO2_TRU) for this detection. SO2 VCDs are removed if the volcanic pixel is detected within a radius of 300 km to a known volcano whose locations are taken for a catalogue from Fioletov et al. (2023). For explosive volcanos, SO2 plumes are tracked manually based on daily TROPOMI SO2 VCD_box and the SO2 VCD is removed if SO2 VCD_box is greater than the SO2 threshold value.

NH3: We use NH3 total columns from the IASI V3R.1.0 NH3 product for inversions of NH3 anthropogenic emissions. We select only land measurements from the morning overpass as IASI is more sensitive at this time to the boundary layer, owing to more favourable thermal conditions. We use the observations only with cloud coverage lower than 10 %.

3.5.4.2. DATA PROCESSING

We map LMDZ-INCA model hourly profile of a species on each high-quality satellite pixel to calculate the total/tropospheric model column correspond to each satellite observations. Further, we aggregate the selected model and satellite observations over model spatial grid for each day. Only TROPOMI-PAL NO2 product provides the total column averaging kernel and thus we could use the averaging kernel for NO2 only to process the model NO2 data for the atmospheric inversions of NOx emissions. In order to calculate the model column NO_2 in the troposphere corresponding to high quality filtered TROPOMI observations, we use the total column averaging kernel provided in the TROPOMI dataset, construct the tropospheric averaging kernel from the total air-mass factor (AMF), tropospheric AMF, and tropopause layer index and combine the tropospheric averaging kernel with the modelled NO_2 distribution interpolated over satellite pressure levels. For other species SO2 and NH3, the current accessible versions of the respective data products do not provide averaging kernel and we used the total column values from the LMDZ-INCA without rescaling to the effective satellite pressure and without convolution with the averaging kernel. For NH3, the total columns in a model grid are computed as a weighted average of the individual observations weighted with the relative error. Figure 3-22 shows an example of the simulated annual averaged NH3 total columns from LMDZ-INCA compared with the IASI observations for 2019. It shows that the model is underestimating the total columns compared to the IASI observations over most of the regions in Africa, North America, South America, and Europe, except some regions in India and China.

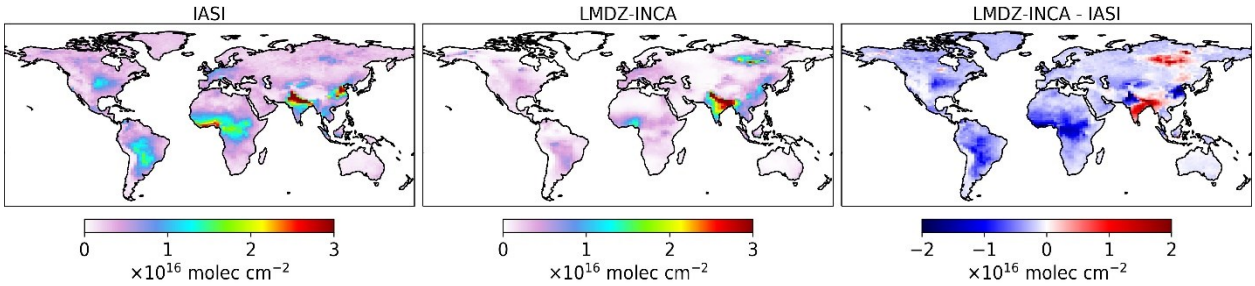


Figure 3-22: Annual averages of the NH3 total columns observed by IASI (left) projected on the model grids, simulated by LMDZ-INCA (middle), and the difference between the model and IASI columns (right) for 2019.

3.5.4.3. DESCRIPTION OF THE INVERSION

Method described in [RD.94]

A specific system has been developed for the inversions of global anthropogenic NO_x, SO2, and NH3 emissions based on LMDZ-INCA and on the satellite observations. It follows a scheme described in [RD.94]. We use daily averaged (as 10-day running mean) simulated and observed columns for the inversions to estimate daily fluxes. The inversion uses the local sensitivity and year-to-year observation changes. The estimated fluxes are derived by calculating the gridded local sensitivity of concentrations

to emission changes (β) and relative observation change of a year of interest (e.g., for 2020 or 2021) compared to the observations of 2019. Furthermore, LMDZ-INCA simulations of a species using fluxes of 2019 and meteorology of an inversion year (e.g., 2020) are employed for computing relative changes of the observations.

The inversion algorithm is based on the following steps:

Step 1. We compute β based on differences of modelled columns between LMDZ-INCA standard simulations ($\Omega^{2019}_{\text{base}}$) with bottom-up emissions ($\text{E}^{2019}_{\text{bottom-up}}$) for a reference year 2019 and the modelled columns ($\Omega^{2019}_{\text{perturb}}$) from the LMDZ-INCA simulations using 40% reduced anthropogenic fluxes ($\text{E}^{2019}_{\text{perturb}}$)

$$\beta = \frac{\Delta E^{2019}_{bottom-up} / E^{2019}_{bottom-up}}{(\Omega^{2019}_{perturb} - \Omega^{2019}_{base}) / \Omega^{2019}_{base}} \quad (1)$$

Step 2: Some filters are applied on β, on observed and/or modelled columns, and/or on bottom-up emissions to select the grids corresponding to the anthropogenic emissions and also to avoid any extreme or negative estimates of the emissions from the inversions. We select grids only for (i) $0 \leq \beta \leq 10$, (ii) $\beta \frac{\Delta\Omega}{\Omega}\Big|_{\text{anth}} \geq -1$, and (iii) Ω_{sat} is greater than 1 x 10^{15} molecules cm^{-2} for NO2 and NH3 and greater than 0.02 DU for SO2. Figure 3-23 shows an example of the distribution of monthly mean values of β computed for NH3 for April 2020. The values of β are less than 1.5 over most of the regions in the world.

Step 3: The fluxes ($\text{E}^{2019}_{\text{sat}}$) for an inversion year (e.g., 2020) using the satellite observations of 2019 ($\Omega^{2019}_{\text{sat}}$) and 2020 ($\Omega^{2020}_{\text{sat}}$) are estimated from bottom-up fluxes ($\text{E}^{2019}_{\text{bottom-up}}$) of 2019 using the following equation

$$E^{2020}_{sat} = \left(1 + \beta \frac{\Delta\Omega}{\Omega}\Big|_{anth}\right) E^{2019}_{bottom-up} \quad (2)$$

Where:

$$\frac{\Delta\Omega}{\Omega}\Big|_{anth} = \frac{\Omega^{2020}_{sat}}{\Omega^{2019}_{sat}} - \frac{\Omega^{2020}_{mod, E^{2019}_{bottom-up}}}{\Omega^{2019}_{mod}} \quad (3)$$

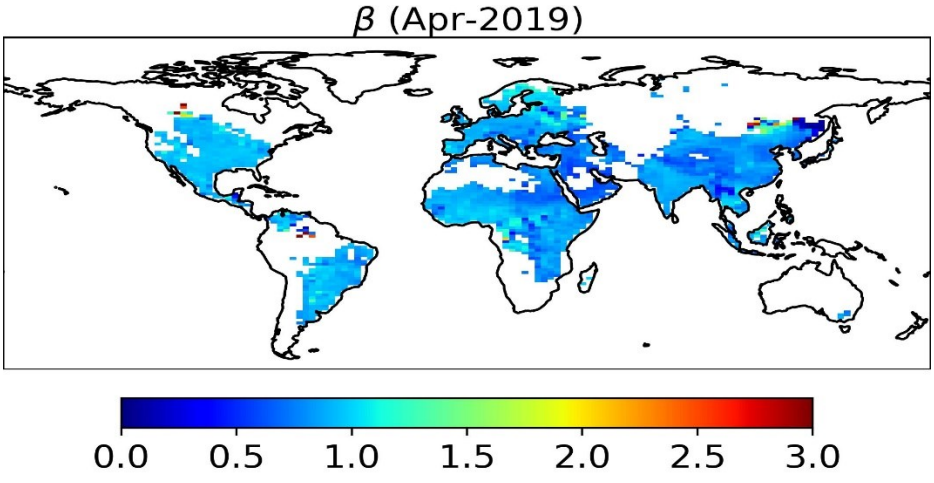


Figure 3-23: Spatial distribution of monthly mean β values for April-2019 estimated from LMDZ-INCA model with 40% perturb anthropogenic NH3 emissions.

Mass-balance approach

We have also implemented a finite difference mass-balance approach [RD.96] to estimate total emission for our study. This method utilizes a unit-less local sensitivity (β) as described above and the relative difference of tropospheric NO_2, SO_2 or NH_3 VCD

Where β is unit-less local sensitivity as describe above but will be calculated for year of interest using following equation

$$E_{post} = E_{prior}\left(1 + \beta\frac{\Delta\Omega}{\Omega}\right)$$

where β of a year of interest is calculated using below equation

$$\beta_{year} = \frac{E_{perturb,year} - E_{prior,year}}{E_{prior,year}} \Big/ \frac{\Omega_{perturb,year} - \Omega_{prior,year}}{\Omega_{prior,year}}$$

and

$$\frac{\Delta\Omega}{\Omega} = \frac{\Omega_{obs} - \Omega_{mod}}{\Omega_{mod}}$$

Ω_{obs} is the NO2 tropospheric column from the satellite and Ω_{mod} is the model NO2 tropospheric column.

3.5.4.4. ESTIMATED EMISSIONS

Figure 3-24 shows an example of a map of the estimated monthly emissions of NH3 using IASI observations for April 2020 and its comparison with the bottom-up CEDS inventories emissions used for the inversions. The estimated emissions using the method described in [RD.94] are approximately 29% higher compared to the bottom-up CEDS emissions for this month. Whereas the estimated emissions using mass-balance approach for the same month is approximately 76% higher than the bottom-up CEDS emissions. The spatial distribution of the estimates over the globe using method described in [RD.94] is similar to the bottom-up emissions and differ from the estimated emissions distribution from mass-balance approach. We conducted similar inversions for NO2 and SO2. For NO2, we also analysed the impact of covid lockdown periods on the relative changes of the estimated emissions compared to the previous year in different regions over the world. We observed reduction of estimated NOx fluxes during lockdown over East China and India which changes ~-21.74% and ~-17.24%, during respective lockdown period. Trend of estimated East US flux is similar to the bottom-up NOx emissions but seasonal amplitude of the estimated emissions is observed higher than the bottom-up anthropogenic emissions. For SO2, the estimated fluxes using the method described in [RD.94] are mostly higher compared to the bottom-up emissions.

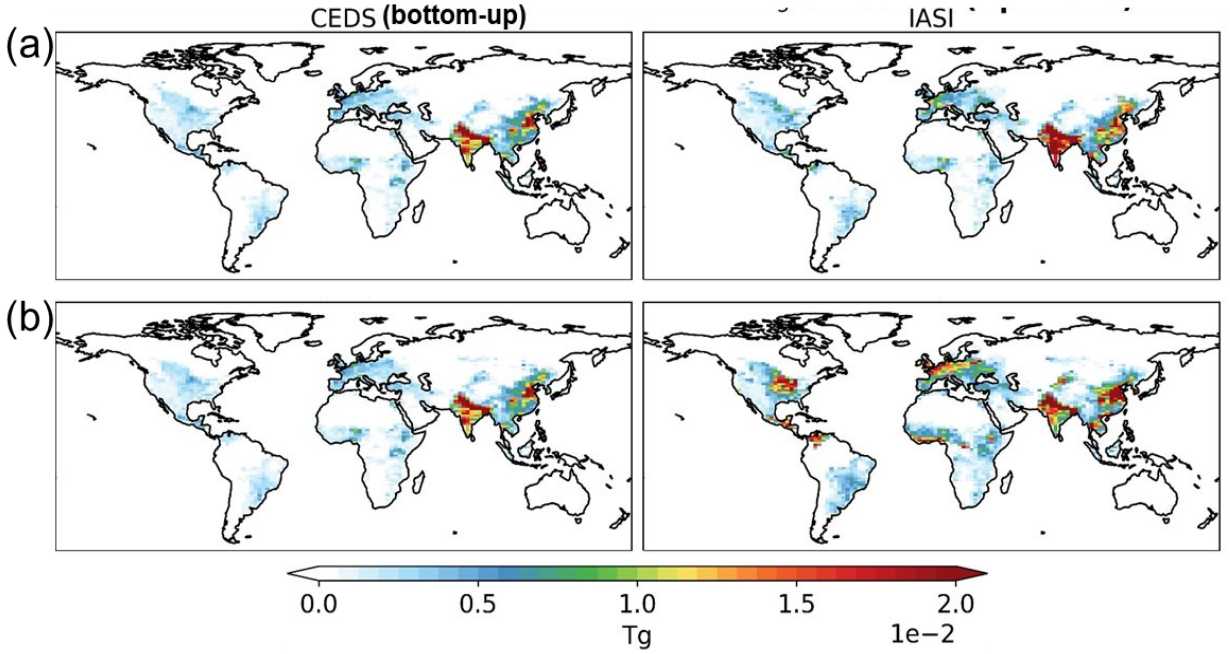


Figure 3-24: Comparison of monthly estimated from IASI observations (right) and bottom-up (CEDS) (left) NH3 emissions of April 2020 using (a) the method described in[RD.94], and (b) the mass-balance approach.

3.5.4.5. UNCERTAINTY ANALYSIS

This approach is subject to a number of uncertainties and limitations. The uncertainty in different steps of our analysis is discussed below.

First, the satellite retrievals of NO2 TVCDs suffer from uncertainties in spectral fitting and in a priori and ancillary data utilized for calculating the stratospheric NO2 background and the air mass factors. For example, the TROPOMI single-pixel errors are typically ~40-60% in the wintertime (19). However, we use spatial and temporal averaging in our analysis to reduce the random errors. Meanwhile, relative differences between 2019 and 2020 are used to derive NOx emission changes, which could cancel out some of the systematic errors.

Second, the values simulated by the LMDZ INCA model reflect the feedback of NOx emissions on NOx chemistry, which could be affected by many factors, such as the model representation of chemical and physical processes (e.g. transport) and changes in emissions of other species involved in the NOx chemistry. We evaluate the model simulation against satellite data to prove the model's ability to capture the characteristics of NO2 columns. We also conduct several sensitivity scenarios to show the robustness of value, as discussed in previous sections.

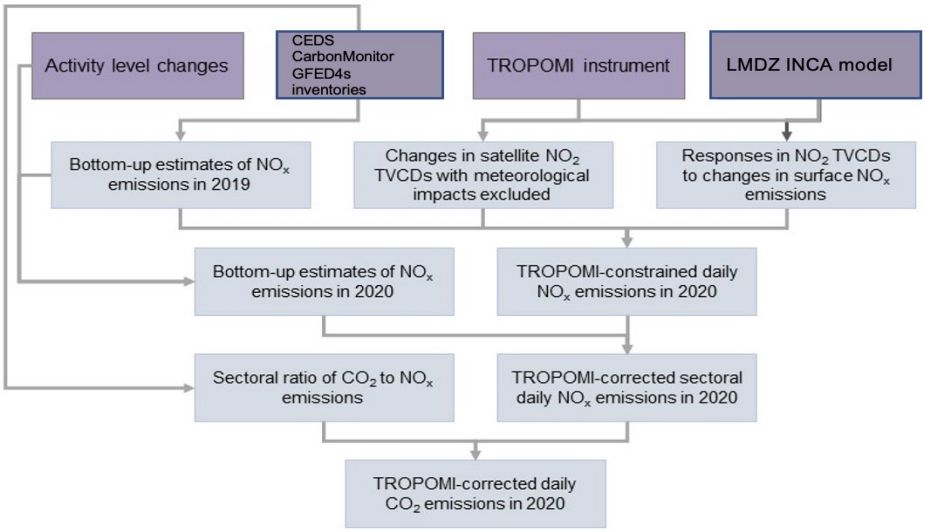


Figure 3-25: Illustration of the steps for the global inversion of NO2, SO2, NH3

3.5.5. ISOPRENE (C5H8)

HCHO is a Volatile Organic Compounds (VOC) not only emitted at the surface but also produced in the atmosphere through the oxidation of anthropogenic and biogenic hydrocarbons. The biogenic hydrocarbon isoprene is the main precursor of HCHO over specific regions such as the tropics or the South Eastern United-States.

Isoprene emissions by the vegetation vary as a function of temperature, water stress and plant functional types. Usually, these emissions are simulated by models like MEGAN or CAMEO (the ORCHIDEE land surface model modified with emissions parameterisations similar to MEGAN and coupled to the LMDZ-INCA model [RD.100]). In Word Emission, we will optimize prior emissions from CAMEO using satellite column measurements of HCHO from GOME and empirical relationships between daily modelled isoprene emissions and daily HCHO observations. These relationships will be established over mainly vegetated areas to reduce the impact of anthropogenic VOC emissions on HCHO. This work will allow ups to produce optimized weekly to monthly isoprene emissions based on HCHO observations. This approach is inspired from [RD.99] who applied it over the USA in the GESOSCHEM model. Figure 3-26 shows such relationships for different sub regions of USA. The regression slope will be determined for different continents and different time scales, including sub-divisions of continents into distinct biomes. A simulation without isoprene emissions will be performed with the LMDZ INCA to isolate the impact of emissions over the period 2019-2021. Based on the results, isoprene emissions will then be updated regularly from satellite observations of HCHO.

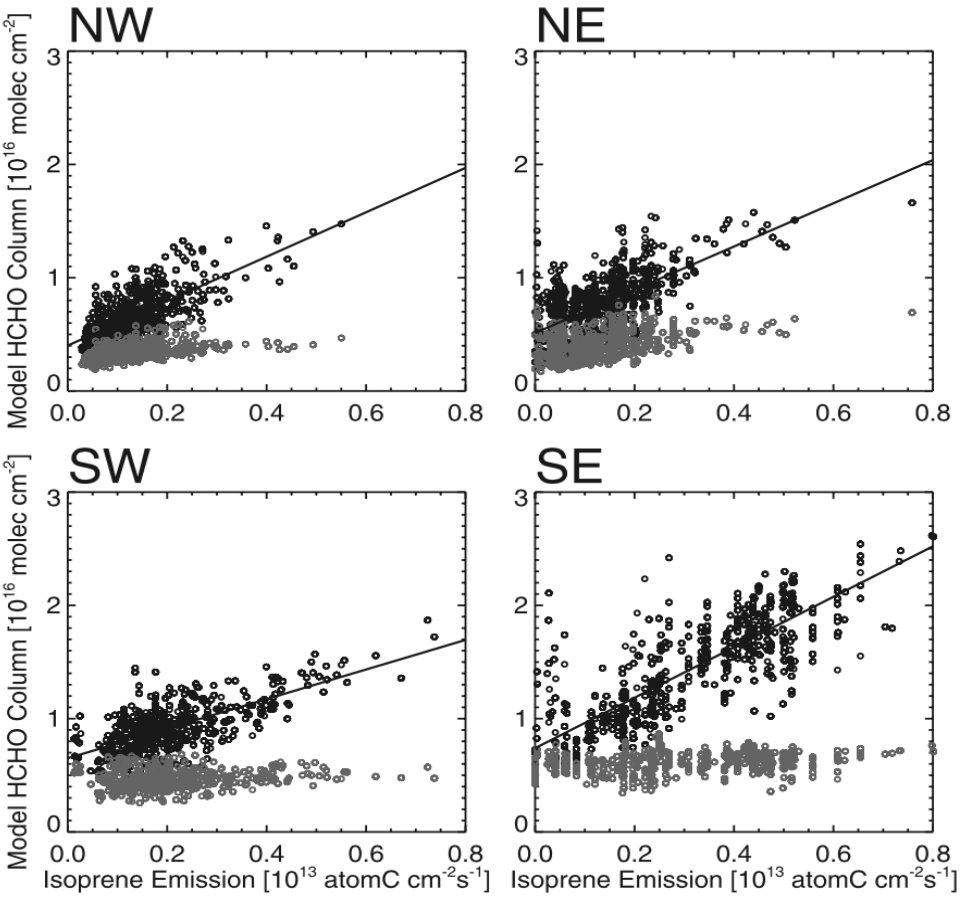


Figure 3-26 Relationship between isoprene emissions and HCHO concentrations over different regions of the US [RD.99].

3.5.6. PARTICULATE MATTER (PM)

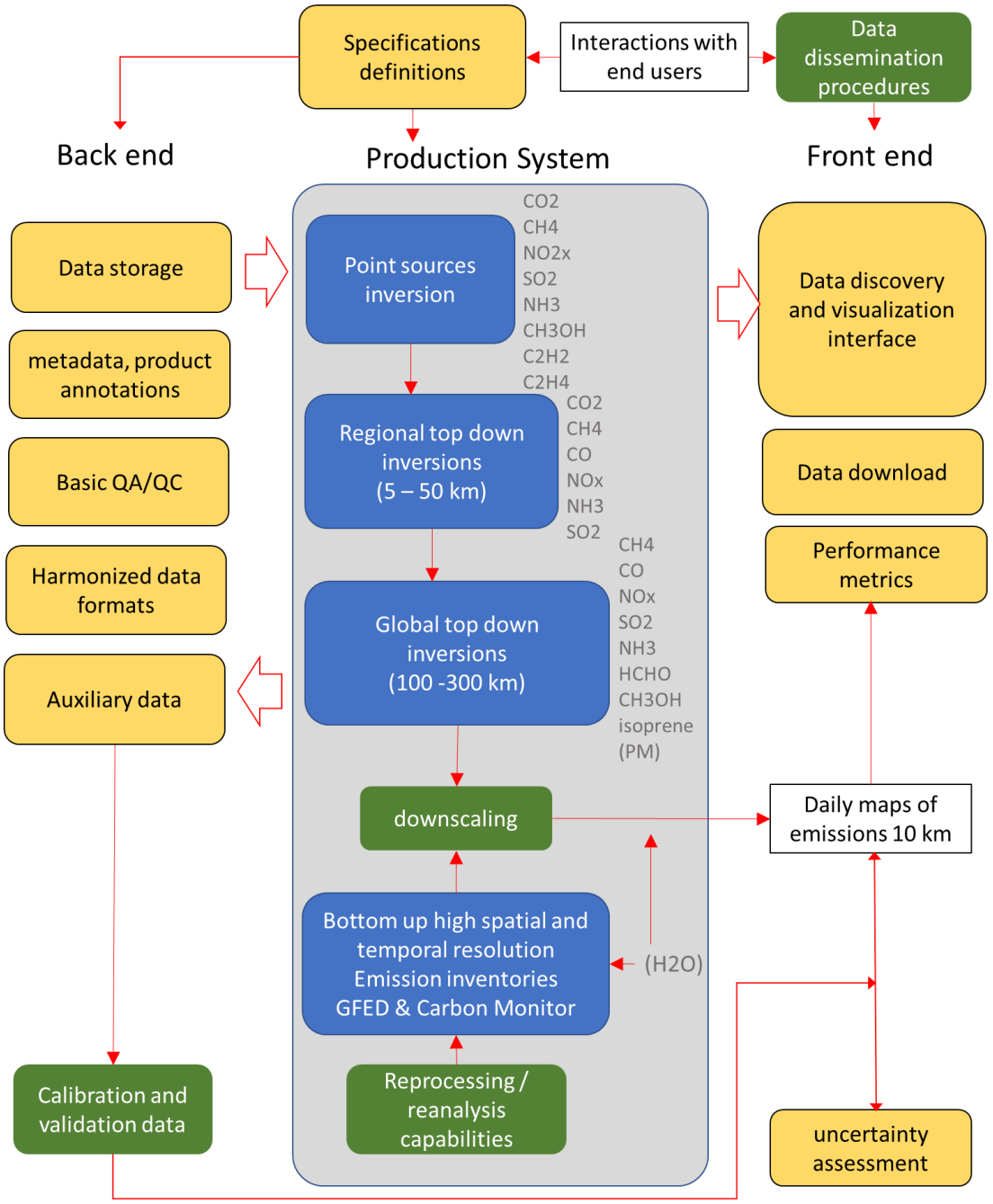
World Emission does not directly optimize emissions of particulate matter (PM) with an inversion model to deliver gridded regional or global scale emission maps, given the uncertainty of intertwined primary and secondary emissions, e.g., from the conversion of NH_3, biogenic precursors and SO_2 into aerosols, complex aerosol mixtures, non-linear heterogeneous chemistry and deposition processes that determine both PM10 and PM2.5. Rather, it takes advantage of optimized global PM-precursor SO_2 and NH_3 emissions to simulate the PM10 and PM2.5 distribution using the LMDZ-INCA model at 1° by 1° globally. Aerosols and gases are treated in the same code to ensure coherence between gas phase chemistry and aerosol dynamics as well as possible interactions between gases and aerosol particles. INCA accounts for the following four basic properties of the ambient aerosol matter: size, chemical composition, hygroscopicity and mixing state of the particles. A superposition of lognormal modes is used, each mode is described by two variables: a mass median diameter and a geometric standard deviation. This multimodal approach allows the representation of the coexistence of externally and internally mixed particles with a limited number of tracers. Sulphate and organic carbon are treated as external mixtures. Black carbon is considered either as externally mixed either or as internally mixed. The treatment of the formation of secondary organic aerosols was done in order to optimize the number of tracers treated in the full chemistry scheme. It considers two natural precursors (isoprene and a-pinene which plays the role of a terpene surrogate). Each one produces 2 gas-phase and 2 aerosol-phase products. The scheme to produce volatile organic carbon the precursors of these species is described in [RD.83] and [RD.78]. Submicron aerosols are transported into two distinct modes, one that is soluble (rather hygroscopic) and a second one that is insoluble. The carbonaceous aerosol fraction is composed of various products of incomplete combustion of fossil fuels and biomass. Direct particle emissions contain both soot and organic matter. INCA keeps track of both soluble and insoluble BC and POM. It assumes that primary, insoluble carbonaceous particles become soluble with time. This ageing process transfers smaller size insoluble accumulation mode particles into the larger size soluble accumulation mode, decreasing slightly the mode diameter of the latter. The shape of the dust size distribution at emission is constrained by fitting the data set recently used by [RD.101]. PM10 and PM2.5 are estimated considering the specificity of the aerosol size distribution of each aerosol component. The model simulations with initial and optimized precursor emissions (NH_3, SO_2, CHOCHO, isoprene and HCO) is compared with AOD data from TROPOMI [RD.102], ESA and MODIS products, with primary PM emissions being scaled for test regions in order to match modelled with observed AOD, as a prospective study. In addition, World Emission will attempt to estimate anthropogenic emissions of PM10 and PM2.5 using NOx as a proxy and compare the results with local and regional air quality stations ([RD.103], [RD.104]) and the ground-based remote sensing Aerosols Robotic NETwork (AERONET) ([RD.105], [RD.106]).

3.5.7. WATER VAPOR (H2O)

World Emission does not address water vapor emissions and sinks from the atmospheric water cycle, but it uses water vapor column densities from OMI [RD.107] and preliminary TROPOMI products [RD.108] to compute dew point temperatures, without the need of LMDZ-INCA here. The Carbon Monitor inventory exploits these dew point temperatures to estimate residential cooling energy requirements and pertaining CO_2 emissions from the power sector during heatwaves in high humidity environments, such as coastal regions, monsoon regions and the wet tropics. Cooling-induced emissions depend on both air temperature and humidity [RD.109], with compound extremes of hot temperatures and elevated relative humidity increasing deadly risks for populations [RD.110]. Water vapor in the boundary layer is also an important driver of plant stress [RD.111], reducing gross primary productivity from stomatal closure and enhancing emissions of BVOCs during drought episodes. World Emission uses water vapor retrieved from TROPOMI transformed into water pressure deficit (VPD) to investigate the correlations between HCO, CH_3OH column concentration and isoprene concentration / retrieved emissions in order to identify regions where water stress on plant photosynthesis and BVOC emissions is controlled by VPD rather than by other drivers such as radiation and soil moisture deficits. Preliminary analyses of the drivers of plant stress in Tropical South America [RD.112] indicates an extensive role of VPD over large regions of the Amazon Forest.

3.6. PRODUCT GENERATION WORKFLOW

The World Emission platform is based on a single top-down emission inversion system that combines all atmospheric composition data with bottom-up inventories to deliver a wall-to-wall quantification for all the species of interest, at very high spatial and temporal resolution, over the whole globe, and with rapid availability of data to users would be impractical, given the different lifetimes of reactive gases and long-lived greenhouse gases, and the resolution of atmospheric dispersion models that must be adapted to each problem. For instance, a plume of NOx, SO2 or NH3 from an industrial site must be simulated at kilometric scale to constrain local emissions, whereas diffuse large scale CH4 and CO2 emissions require 3D global transport-chemistry models. Therefore, we propose a suite of sub-systems suitable to cover anthropogenic and natural emissions of CO2, CH4, CO, SO2, NOx, NH3, as well as of CH3OH, and isoprene, going from point sources to the globe. shows the integration of the products generation within the overall analysis flow of World Emission system.



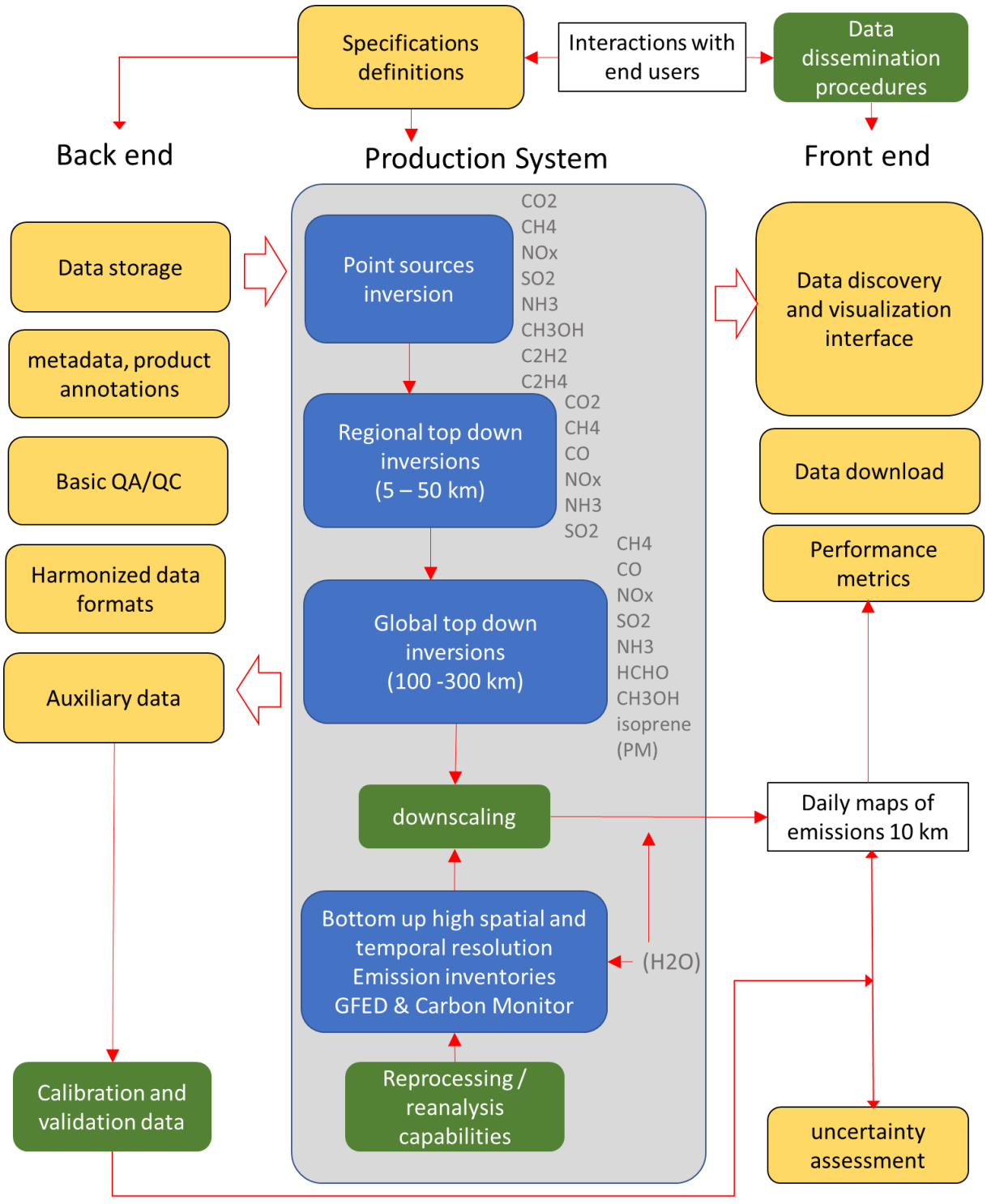


Figure 3-27: Overall system flow

- A sub-system to assess emissions from **localized point sources** like power plants and industrial sites that produce localized atmospheric plumes of emitted species. These plumes detected by spaceborne imagers (TROPOMI) or crossed by sounders (OCO-2) will be analysed systematically to retrieve all possible point sources emissions over the entire globe.
- A sub-system to diagnose gridded weekly emission fields for the **selected regions** with industrial activities and high population densities, or intense wildfires and deforestation. In those domains, emissions produce regional enhancements of atmospheric concentration of each

species at scales of 1 to ≈ 500 km. Those enhancements integrate the contributions of emissions from point sources, area sources such as megacities and conurbations, forest and savannas fires, together with the signal of more diffuse regional emissions. Here, satellite data will be inverted into emissions using 3D mesoscale regional models of atmospheric transport and chemistry.

- A subsystem for **global emissions**, exploiting synoptic and hemispheric-scale atmospheric concentration gradients at scales of 100 to 1000 km. Those large-scale signals will be exploited by a common top-down inversion system to constrain emissions with a 3D global chemistry transport model integrated at a spatial resolution of 100 km, combined with the new GFED and Carbon Monitor near real time bottom-up global inventories for anthropogenic combustion emissions and wildfires.

All the species are analysed as an integrated whole being part of an almost unique processing chain with a kernel that is common to all the species. Such configuration is pursued to ease vertical scalability (i.e., towards adding more species).

Note that local and regional estimates can be considered as intermediate products still reaching the overall goal of global coverage.

The detailed analysis methods for different species with particular living time and emission characteristics are attached below in the following figures.

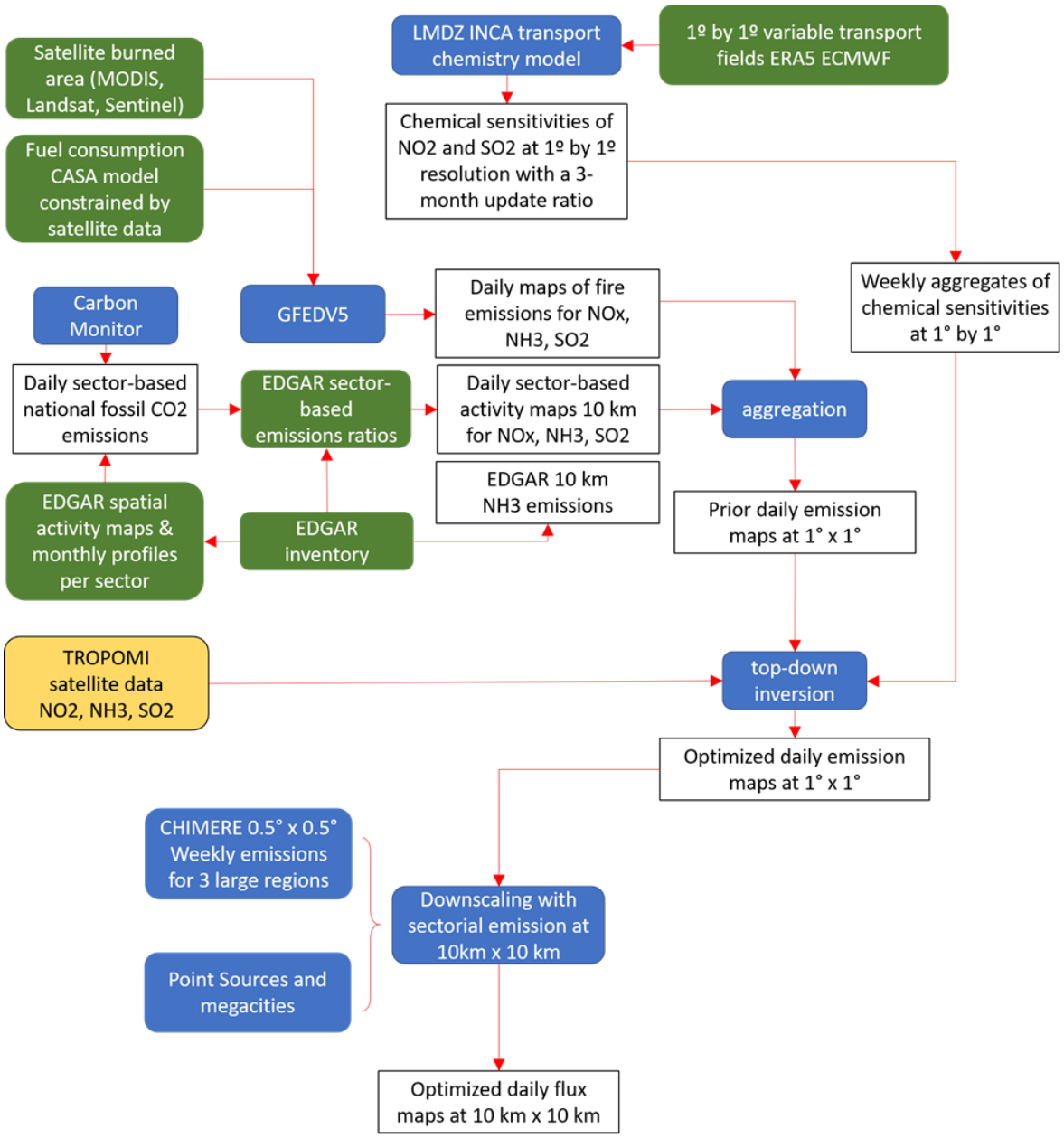


Figure 3-28: Detailed analysis flow for NOx, NH3, SO2.

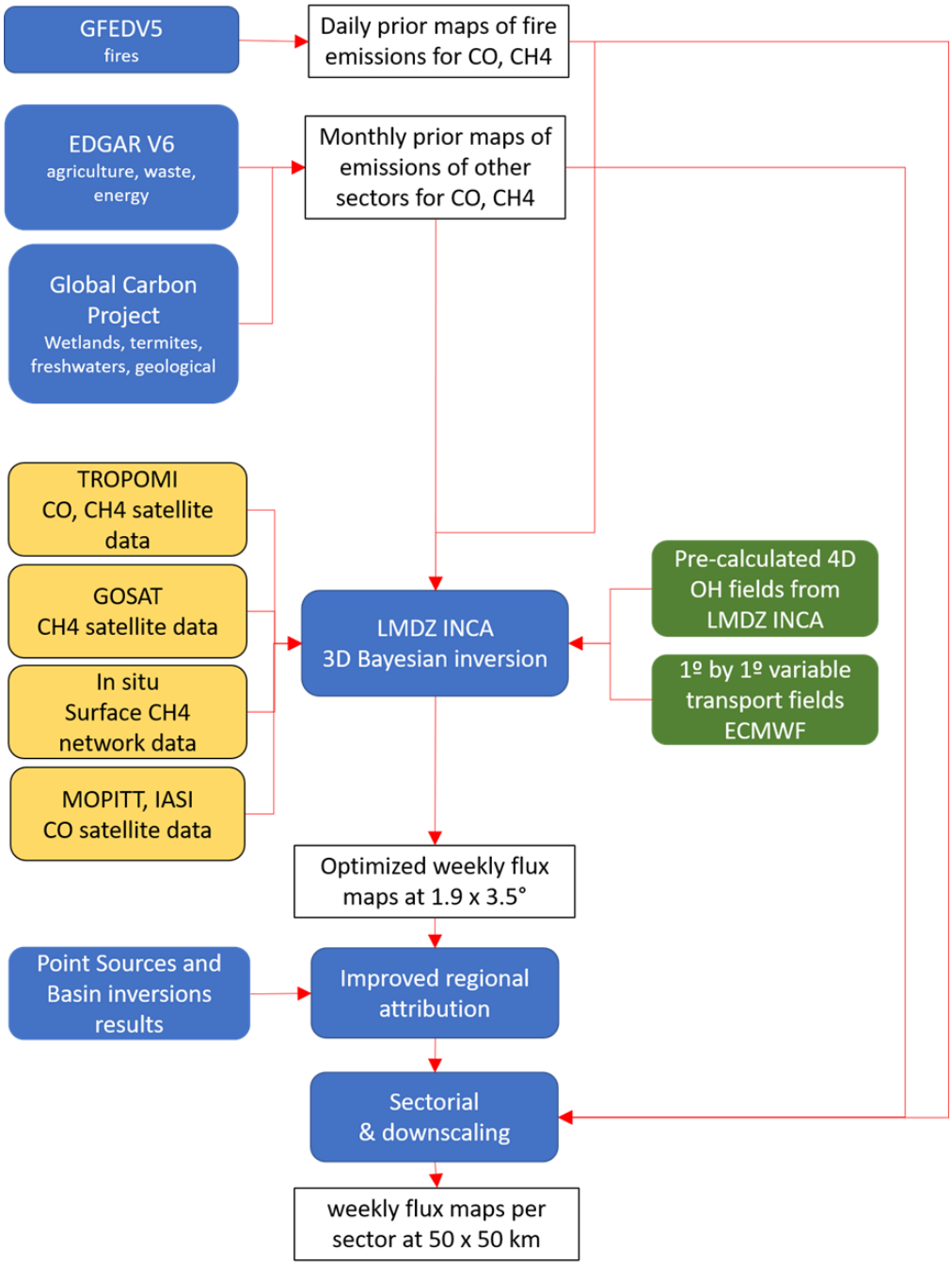


Figure 3-29: Detailed analysis flow for CH4, CO.

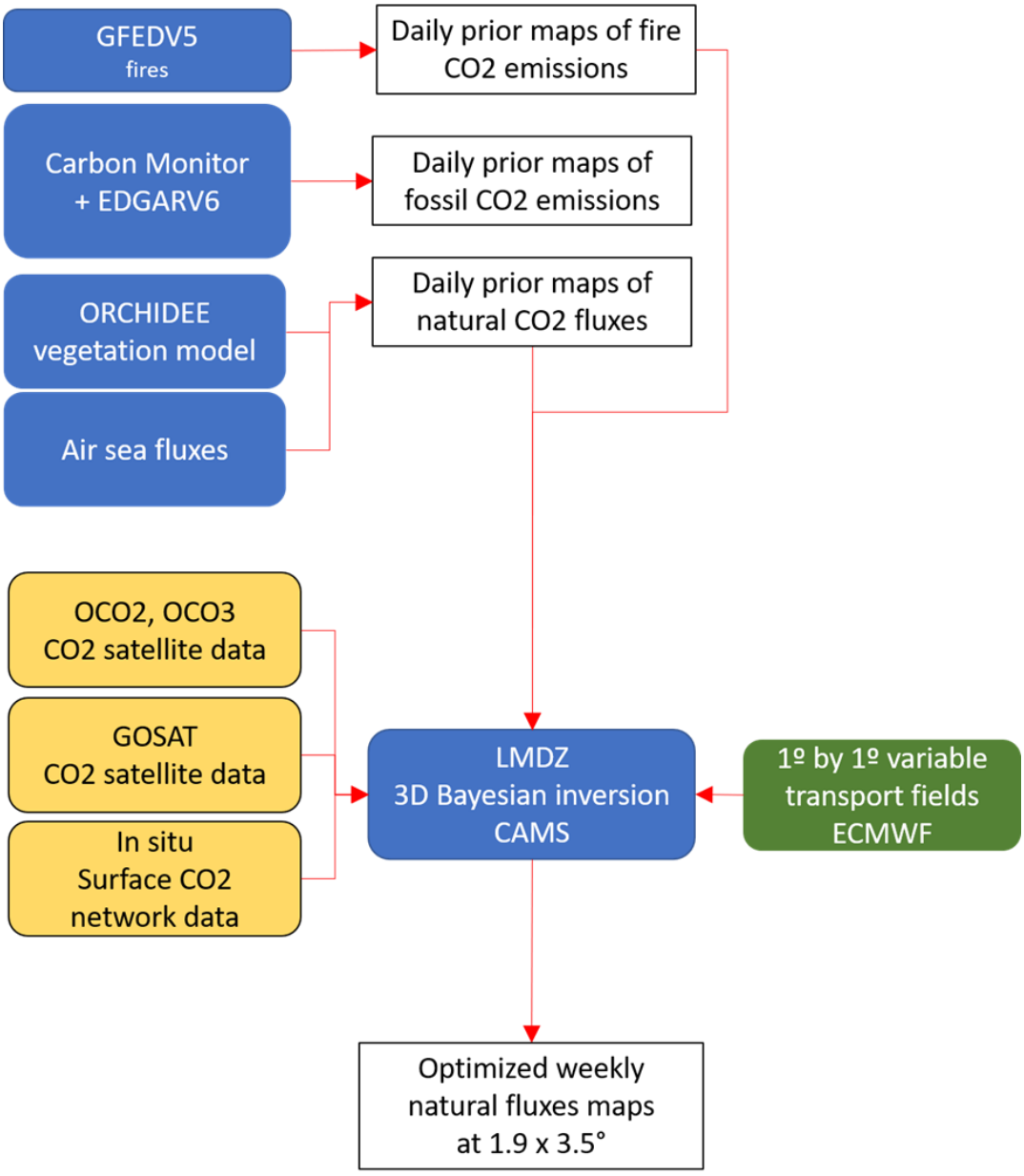


Figure 3-30: Detailed analysis flow for CO2.

END OF DOCUMENT