www.project-foci.eu

Co-funded by the European Union

UK Research and Innovation

Horizon Europe

(HORIZON-CL5-2021-D1-01)

Non-CO2 Forcers and their Climate, Weather, Air Quality and Health Impacts

FOCI

Deliverable 1.1

Harmonized dataset of in-situ and column integrated aerosols and gas-phase species

Grant agreement No 101056783

Grant Agreement No.	101056783	
Project acronym	FOCI	
Project full title	Non-CO2 Forcers and their Climate, Weather, Air Quality and Health Impacts	
Call	HORIZON-CL5-2021-D1-01	
Deliverable name	D1.1 Harmonized dataset of in-situ and column integrated aerosols and gas-phase species	
WP contributing to the deliverable	WP1	
Task producing the deliverable	Task 1.1	
Type		Report
		Prototype
		Demonstrator
	X	Other: Data
Dissemination level	X	Public
		Sensitive
		UE/EU-Restricted
Due date of deliverable	Month 12	
Actual submission date	Month 12	
Lead beneficiary	CSIC/FMI	
Author(s)	Marco Pandolfi (CSIC), Dene Bowdalo (BSC), Calum Meikle (BSC), Oriol Jorba (BSC), Jordi Rovira (CSIC), Risto Makkonen (FMI), Tommi Bergman (FMI)	
Other Contributor(s)		
Reviewer(s)	Tuukka Petäjä (UHel)	
Keywords	FOCI, deliverables, harmonized dataset	

ACKNOWLEDGEMENTS
This project has been co-funded by the European Union with funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101056783 and from UKRI under the UK Government's Horizon Europe Guarantee (UKRI Reference Numbers: 10040465, 10053814 and 10050799).

Version	Date	Modified by	Comments
1.0	7 Jun 2023	Marco Pandolfi	First order draft written
2.0	28 Jun 2023	Marco Pandolfi	Revisions made to incorporate reviewer's comments
3.0	3 Jul 2023	Marco Pandolfi, Dene Bowdalo, Oriol Jorba	Final draft ready

	Name	Date
Verification Final Draft by WP leaders	Marco Pandolfi (CSIC), Risto Makkonen (FMI)	29 Aug 2023
Check before upload by project Coordinator	Tomas Halenka (CU)	30 Sep 2023

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This document is the deliverable "D1.1: Harmonized dataset of in-situ and column integrated aerosols and gas-phase species" for the European Union project "FOCI: Non-CO2 Forcers and their Climate, Weather, Air Quality and Health Impacts" (hereinafter also referred to as FOCI, project reference: 101056783).

The observations described in this Deliverable contribute to the FOCI integrated observational and modelling analysis which focuses on the radiative forcing properties of different atmospheric non-CO2 species in the wider context of the warming potential of all key GHGs. The observations described in this Document compose the Milestone M1.1 (Datasets for non-CO2 species) and Deliverable D1.1.

WP1 is already working on the Deliverable D1.2 and on the Milestone M1.2 (Database of advanced aerosol particles optical properties) due by month 24.

This Deliverable reports on the existing datasets of in-situ and column integrated aerosols and gas-phase species that will be used in the framework of the FOCI project. A list of observations included here and links to the available data is briefly reported below, and Chapters 2 to 5 provide the detailed description of the data sets.

a) LIDAR Data:
   - Level 2: https://www.earlinet.org/index.php?id=125
   - Level 3: https://www.earlinet.org/index.php?id=319

   The metadata for each LIDAR stations (Country, station code/name, coordinates, altitude, contact person) and the name, number of measurements, start/end of the aerosol vertical profiles measurements for each available wavelength and total measuring period are available in the FOCI SharePoint.

b) Satellite Data/Products:
   - Regional and global aerosol optical depth (AOD): http://nsdc.fmi.fi/data/data_aod
   - ModIs Dust AeroSol (MIDAS): https://doi.org/10.5281/zenodo.4244106
   - Methane (CH4) observation from GOSAT Project: https://doi.org/10.5285/18ef8247f52a4cb6a14013f8235cc1eb

c) GHOST Database: The GHOST database will be soon available in EBAS and a DOI will be assigned after publication of the dataset (This Deliverable will be updated as soon as the link to GHOST is available).

d) IAGOS database: https://iagos.aeris-data.fr/download/

- Observational Time series : *https://doi.org/10.25326/06*
- Observational Vertical Profiles : *https://doi.org/10.25326/07*
- CO contributions: *https://doi.org/10.25326/3*
- PBL-referenced profiles of O_3 and CO: *https://doi.org/10.25326/4*
- IAGOS CO contribution and interpolated modeled data from ECMWF – L4: https://doi.org/10.25326/3
- IAGOS Climatologies of O_3 and CO - L3: https://doi.org/10.25326/08
- IAGOS PBL referenced profiles of O_3 and CO – L4: https://doi.org/10.25326/4
- IAGOS final quality controlled Observational Data L2 – Time series: https://doi.org/10.25326/06
- IAGOS final quality controlled Observational Data L2 – Vertical profiles: https://doi.org/10.25326/07.

1. INTRODUCTION

This deliverable describes the methodology applied to build a detailed and comprehensive observationally-based dataset on anthropogenic non-CO2 species including climate-relevant gases and aerosol properties that will be used to constrain numerical sensitivity simulations. This Document describes in details the data that WP1 provided to FOCI.

The Deliverable D1.1 "Harmonized dataset of in-situ and column integrated aerosols and gas-phase species" consists of a dataset of in-situ and column integrated aerosols and gas-phase species that WP1 produced in Task 1.1. The dataset was compiled using data from a suite of observational networks/programmes as GAW/WMO, IMPROVE, EMEP, ACTRIS, EARLINET, AERONET, among others.

WP1 is organized in three major tasks:

Task 1.1: Data compilation and harmonization (Lead beneficiary: CSIC)

Task 1.2: Advanced analysis of available data (Lead beneficiary: CSIC)

Task 1.3: Advanced model representation of physiochemical properties of anthropogenic aerosols and their feedbacks (Lead beneficiary: FMI)

One major WP1 outcome (in Task 1.1) is the compilation of existing datasets of in-situ and column integrated aerosols and gas-phase species. The data will be gathered (D.1.1) and made accessible for the advanced analysis of available data (T1.2), advanced model representation (T1.3), analysis and evaluation of model improvements (WP3), performance evaluation of regional climate models (WP4) and regional and urban multiscale climate impact (WP6). For each variable, the measurement requirements (instrumental and operational conditions) will be evaluated according to existing guidelines and protocols, e.g., those from ACTRIS, and data management from ACTRIS Data Centre (ACTRIS DC) complying with the FAIR principles.

Thus, D1.1 consists of data sets, including metadata, of anthropogenic non-CO2 species including climate-relevant gases and aerosol properties.

In this document we describe the key data sources relevant for the FOCI project (ground-based lidar observations, satellite remote sensing products as well as in-situ observation data sets). For each component we summarise the spatiotemporal coverage of the data, relevant data policies, required acknowledgements, data quality control processes etc.

2. LIDAR DATA

This section describes the information from LIDAR (Light Detection and Ranging) instruments available for the FOCI project. LIDAR data for Europe are available from the ACTRIS/EARLINET database in NetCDF format (https://www.earlinet.org/index.php?id=125). The database provides two different levels: Level 2 data

with the vertically resolved aerosol extinction and backscatter profiles at different wavelengths and Level 3 data available since 28 March 2019. Level 3 data are the climatological products obtained from the fully quality controlled (Level 2) data for providing useful aggregated information to the users. The Level 3 data are centrally obtained by the ACTRIS aerosol remote sensing Data Center node of the Italian National Research Council (CNR) in Potenza (Italy). This allows the harmonization and reproducibility of the products. Level 3 products are essential for climatological studies for the different kinds of potential users (related to air quality, meteorological application, climate study, modelling, and so on). In particular, these products provide high-quality information about the vertical distribution and the optical properties of the aerosol over the European continent on a long-term scale, and this could be precious also for the European climate and air quality policymakers. Additionally, these products could also be a term of reference for the stations performing measurements in order to investigate specific processes observed at those stations.

A detailed description of Level 3 data is available in the ACTRIS/EARLINET webpage (https://www.earlinet.org/index.php?eID=tx_securedownloads&p=319&u=0&g=0&t=1674646453&hash=bf2d646ca82a2f25954e883efc15b8e102250f4e&file=fileadmin/user_upload/EARLINET_Database/EARLINET_Level3_v2.0_relNov2022.pdf). For each extinction/backscatter vertical profile, an integrated quantity is calculated: aerosol optical depth (AOD) for extinction profiles, and aerosol integrated backscatter (IB) for backscatter profiles. AOD and IB are the integrals over the altitude of the aerosol extinction and backscatter profiles, respectively. These integrated products represent a proxy for the quantity of aerosol present in the considered portion of the atmospheric column. However, optical depth and integrated backscatter also depend on the type of particles because the extinction and the backscatter efficiencies depend on the size, shape, and refractive index of the particles.

Another integrated quantity is the center of mass, which is a value associated to every backscatter vertical profile. The center of mass of the aerosol content in a certain portion of the atmospheric column is estimated as the backscatter weighted average altitude in the considered altitude range (Mona et al., 2006). This quantity approximates the center of mass of the aerosol layer, and exactly coincides with the true center of mass if both composition and size distribution of the particles are constant with the altitude. This estimate of the center of mass provides a proxy for the altitude where the most relevant part of the aerosol load is located.

Lidar ratio (LR), Angstrom exponent (AE) and particle depolarization cannot be strictly considered integrated quantities, since they are simply averaged in a given portion of the atmospheric column. Anyway, they are part of the integrated product files for sake of simplicity and completeness. The LR, calculated as the ratio between backscatter and extinction, depends on the microphysical properties of aerosol particles as the particle size distribution; the AE is an exponent that expresses the spectral dependence of aerosol extinction and backscatter with the wavelength; the particle depolarization enables distinguishing spherical and non-spherical particles from each other and, furthermore, it can be used to quantify the contributions of different aerosol types to elevated layers.

2.1 Data Policy

EARLINET aerosol profile data are reported in the EARLINET Data base: https://data.earlinet.org/, and are accessible from this repository and from the ACTRIS Data Portal (http://actris.nilu.no). The data policy of these data is harmonized with the ACTRIS data policy.

EARLINET aerosol profile data are freely available to individual users upon acceptance of the data usage policy.

The aerosol profile data are contributions from the ACTRIS and EARLINET consortium. Each aerosol profile site has Principal Investigator(s) (PI), responsible for deployment, maintenance and data collection. The PI has priority use of the data collected at the site. The PI is entitled to be informed of any other use of that site's data. Users of the aerosol profile data must not redistribute these data to third parties, but encourage access and dissemination of EARLINET data directly through ACTRIS Data Portal or from the topical data base at https://data.earlinet.org/ to accept the data policy, and ensure the use of the latest version of the data.

Using ACTRIS and EARLINET aerosol profile data: Please consult with the PI(s) of the data to be used. Publishing ACTRIS and EARLINET aerosol profile data: Please consider authorship for the PI(s). When data from the EARLINET Database are used in a publication, the following acknowledgment should be included:

"The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654109 and previously from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 262254."

"The authors acknowledge EARLINET for providing aerosol LIDAR profiles available at https://data.earlinet.org/".

Cite the following reference: "Pappalardo et al., EARLINET: towards an advanced sustainable European aerosol lidar network, Atmos. Meas. Tech., 7, 2389-2409, doi:10.5194/amt-7-2389-2014, 2014".

We require a reprint of any published papers or reports or a brief description of other uses (e.g., posters, oral presentations, etc.) of data downloaded from the EARLINET database. This will help us to determine the use of EARLINET aerosol profile data, which is helpful in optimizing product development. It also helps us to keep our product-related references current.

2.2 Data Quality Control (QC)

A detailed description of the QA of lidar data is available (https://www.earlinet.org/index.php?eID=tx_securedownloads&p=293&u=0&g=0&t=1674647120&hash=12f85550302481088f896062a8c7ad9e2f5139b0&file=fileadmin/user_upload/EARLINET_QC_v3_0_20210325.pdf). Quality Control procedure on datafile is a useful and needed tool which serves two main groups of users: Data Originator and External users. As concerns EARLINET, the network started its activities in 2000 based on cutting edge lidar instruments for aerosol detection over Europe which agreed on some format of the data

and the database was set up in a collaborative way. Going now in the Research Infrastructure path, the set-up of automatic QC procedure is a stringent request to be met. The needs of the network were different in different periods, and based on that some QC procedures were set up.

A brief description of QC is provided below:

Data quality check procedures for EARLINET/ACTRIS database improved over the time. Here we report the automatic quality check procedures working on the EARLINET database since 25 March 2021.

When a product is submitted to the EARLINET database the following steps occur:

1. technical quality controls (BQC) are executed to ensure the product is compliant mainly from technical point of view with the defined standard.

2. advanced quality controls (AQC) are executed to assess the quality from a physical point of view of the product.

3. The 3 variables “quality_control_level”, “basic_quality_control”, and “advanced_quality_control” are computed by the ARES Data Center according to the results of the previous two steps and added into the product.

4. The product is renamed according to the filename conventions described earlier and definitively stored into the datacenter

There are two types of Quality Control procedures:

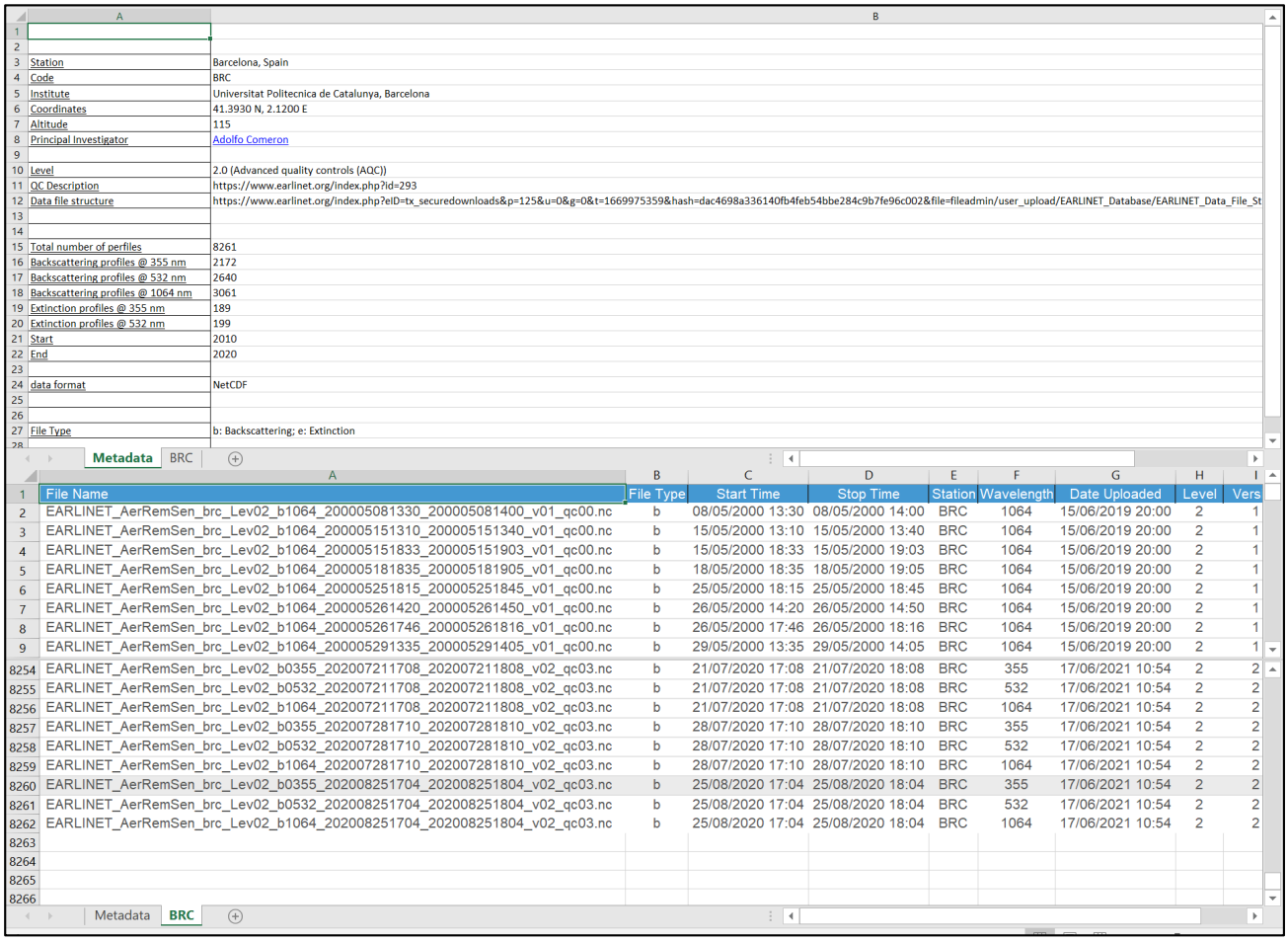
• Basic quality control (BQC): technical quality control on the submitted product. BQC checks that each file contains the mandatory products.

• Advanced quality control (AQC): series of physical checks applied to the submitted product. AQC checks that error on the optical properties is positive for all defined values of the corresponding optical property. A product file does not pass this control if, for defined optical property value, at least for 1 point, the error is negative, zero or not defined.

2.3 Data provided (Deliverable D1.1/Milestone M1.1)

As aforementioned, LIDAR data for Europe are available in NetCDF format in the ACTRIS/EARLINET database. Following the EARLINET Data policy (Chapter 1.1), WP1 cannot directly provide the LIDAR data to FOCI (for example downloading them and uploading them to the FOCI SharePoint). In fact, users of the aerosol profile data must not redistribute these data to third parties. LIDAR data can be directly downloaded from the ACTRIS/EARLINET database if needed. Consequently, WP1 provides for each LIDAR station a metadata reporting the main characteristics of the stations (name, Country, station code, coordinates, altitude,

contact person), name, number, start/end of the measurement of vertical profiles for each available wavelength and total measuring period. This information prepared in WP1 is not available in the EARLINET database and was created for the FOCI project with the main objective of summarizing the information available at each LIDAR station. An example of LIDAR LEVEL 2 metadata is reported in Figure 1.



Station	Barcelona, Spain
Code	BRC
Institute	Universitat Politecnica de Catalunya, Barcelona
Coordinates	41.3930 N, 2.1200 E
Altitude	115
Principal Investigator	Adolfo Comeron
Level	2.0 (Advanced quality controls (AQC))
QC Description	https://www.earlinet.org/index.php?id=293
Data file structure	https://www.earlinet.org/index.php?eID=tx_securedownloads&p=125&u=0&g=0&t=1669975359&hash=dac4698a336140fb4feb54bbe284c9b7fe96c002&file=fileadmin/user_upload/EARLINET_Database/EARLINET_Data_File_St
Total number of perfiles	8261
Backscattering profiles @ 355 nm	2172
Backscattering profiles @ 532 nm	2640
Backscattering profiles @ 1064 nm	3061
Extinction profiles @ 355 nm	189
Extinction profiles @ 532 nm	199
Start	2010
End	2020
data format	NetCDF
File Type	b: Backscattering; e: Extinction

Metadata | BRC

File Name	File Type	Start Time	Stop Time	Station	Wavelength	Date Uploaded	Level	Vers
EARLINET_AerRemSen_brc_Lev02_b1064_200005081330_200005081400_v01_qc00.nc	b	08/05/2000 13:30	08/05/2000 14:00	BRC	1064	15/06/2019 20:00	2	1
EARLINET_AerRemSen_brc_Lev02_b1064_200005151310_200005151340_v01_qc00.nc	b	15/05/2000 13:10	15/05/2000 13:40	BRC	1064	15/06/2019 20:00	2	1
EARLINET_AerRemSen_brc_Lev02_b1064_200005151833_200005151903_v01_qc00.nc	b	15/05/2000 18:33	15/05/2000 19:03	BRC	1064	15/06/2019 20:00	2	1
EARLINET_AerRemSen_brc_Lev02_b1064_200005181835_200005181905_v01_qc00.nc	b	18/05/2000 18:35	18/05/2000 19:05	BRC	1064	15/06/2019 20:00	2	1
EARLINET_AerRemSen_brc_Lev02_b1064_200005251815_200005251845_v01_qc00.nc	b	25/05/2000 18:15	25/05/2000 18:45	BRC	1064	15/06/2019 20:00	2	1
EARLINET_AerRemSen_brc_Lev02_b1064_200005261420_200005261450_v01_qc00.nc	b	26/05/2000 14:20	26/05/2000 14:50	BRC	1064	15/06/2019 20:00	2	1
EARLINET_AerRemSen_brc_Lev02_b1064_200005261746_200005261816_v01_qc00.nc	b	26/05/2000 17:46	26/05/2000 18:16	BRC	1064	15/06/2019 20:00	2	1
EARLINET_AerRemSen_brc_Lev02_b1064_200005291335_200005291405_v01_qc00.nc	b	29/05/2000 13:35	29/05/2000 14:05	BRC	1064	15/06/2019 20:00	2	1
EARLINET_AerRemSen_brc_Lev02_b0355_202007211708_202007211808_v02_qc03.nc	b	21/07/2020 17:08	21/07/2020 18:08	BRC	355	17/06/2021 10:54	2	2
EARLINET_AerRemSen_brc_Lev02_b0532_202007211708_202007211808_v02_qc03.nc	b	21/07/2020 17:08	21/07/2020 18:08	BRC	532	17/06/2021 10:54	2	2
EARLINET_AerRemSen_brc_Lev02_b1064_202007211708_202007211808_v02_qc03.nc	b	21/07/2020 17:08	21/07/2020 18:08	BRC	1064	17/06/2021 10:54	2	2
EARLINET_AerRemSen_brc_Lev02_b0355_202007281710_202007281810_v02_qc03.nc	b	28/07/2020 17:10	28/07/2020 18:10	BRC	355	17/06/2021 10:54	2	2
EARLINET_AerRemSen_brc_Lev02_b0532_202007281710_202007281810_v02_qc03.nc	b	28/07/2020 17:10	28/07/2020 18:10	BRC	532	17/06/2021 10:54	2	2
EARLINET_AerRemSen_brc_Lev02_b1064_202007281710_202007281810_v02_qc03.nc	b	28/07/2020 17:10	28/07/2020 18:10	BRC	1064	17/06/2021 10:54	2	2
EARLINET_AerRemSen_brc_Lev02_b0355_202008251704_202008251804_v02_qc03.nc	b	25/08/2020 17:04	25/08/2020 18:04	BRC	355	17/06/2021 10:54	2	2
EARLINET_AerRemSen_brc_Lev02_b0532_202008251704_202008251804_v02_qc03.nc	b	25/08/2020 17:04	25/08/2020 18:04	BRC	532	17/06/2021 10:54	2	2
EARLINET_AerRemSen_brc_Lev02_b1064_202008251704_202008251804_v02_qc03.nc	b	25/08/2020 17:04	25/08/2020 18:04	BRC	1064	17/06/2021 10:54	2	2

Metadata | BRC

Figure 1. Example of metadata (Deliverables) provided for LEVEL 2 data for each EARLINET LIDAR station

The location of the EARLINET Lidar stations is provided in Figure 2.

Table 1 and Table 2 report the main characteristics of the LIDAR stations where Level 2 data are available and the type and number of available profiles, respectively.

Figure 2. Location of the EARLINET LIDAR stations.

Table 2. Main characteristics of the LIDAR stations where Level 2 data are available

Location	Country	Code	Institute	Coordinates	Altitude [m]
Andøya,	Norway	ARR	ALOMAR/Andøya Space Center http://andoyaspace.no/	69.2780 N, 16.0080 E	380
Antikythera	Greece	AKY	National Observatory of Athens - NOA	35.8600 N, 23.3100 E	193
Athens,	Greece	ATZ	National Technical University of Athens, Physics Department http://www.physics.ntua.gr/~papayannis/	37.9600 N, 23.7800 E	212
Barcelona	Spain	BRC	Universitat Politecnica de Catalunya, Barcelona http://www.tsc.upc.edu/rslab/	41.3930 N, 2.1200 E	115
Belgrade	Serbia	BGD	IPB – Institute of Physics Belgrade http://www.ipb.ac.rs/index.php/en/	44.8557 N, 20.3913 E	89
Belsk	Poland	COG	Institute of Geophysics, Polish Academy of Sciences, Warsaw http://www.igf.edu.pl/	51.8300 N, 20.7800 E	180
Bucharest	Romania	INO	National Institute of R&D for Optoelectronics (INOE) http://www.inoe.ro/en/	44.3480 N, 26.0290 E	93
Cabauw	Netherlands	CBW	KNMI - Royal Netherlands Meteorological Institute http://projects.knmi.nl/earlinet/	51.9700 N, 4.9300 E	0
Clermont-Ferrand	France	PUY	Observatoire de Physique du Globle (OPGC-LaMP) http://www.opgc.fr/SO/mesures/lidarTR.php	45.7610 N, 3.1110 E	420
Cluj-Napoca	Romania	CLJ	Babes-Bolyai University of Cluj Napoca http://www.ubbcluj.ro/en/	46.7682 N, 23.5509 E	405
Dushanbe	Tajikistan	DUS	Leibniz Institute for Tropospheric Research, Leipzig http://www.tropos.de/en/	38.5594 N, 68.8561 E	864
Evora	Portugal	EVO	Institute for Earth Sciences(ICT-Instituto de Ciencias da Terra) http://www.icterra.pt/g1/	38.5678 N, -7.9115 E	293
Garmisch-Partenkirchen	Germany	GAR	Karlsruher Institut für Technologie, KIT, Garmisch-Partenkirchen http://www.imk-ifu.kit.edu/	47.4770 N, 11.0640 E	730
Granada	Spain	GRA	Andalusian Institute for Earth System Research, University of Granada(IISTA-CEAMA) http://www.iista.es/	37.1640 N, -3.6050 E	680
Ispra	Italy	IPR	Joint Research Centre - Institute for Environment and Sustainability, Ispra http://ies.jrc.ec.europa.eu/	45.8167 N, 8.6167 E	209
Kuopio	Finland	KUO	Finnish Meteorological Institute (FMI), Atmospheric Research Centre of Eastern Finland, Kuopio http://en.ilmatieteenlaitos.fi/atmospheric-research-centre-of-eastern-finland	62.7333 N, 27.5500 E	190
Lecce	Italy	SAL	University of Salento, Physics Department, Lecce http://www.matfis.unisalento.it/home_page	40.3330 N, 18.1000 E	30
Leipzig	Germany	LEI	Leibniz Institute for Tropospheric Research, Leipzig http://www.tropos.de/en/	51.3500 N, 12.4330 E	125
Lille	France	LLE	Lille 1 University - Science and Technology http://www.univ-lille1.fr/home/	50.6117 N, 3.1417 E	60
Limassol	Cyprus	CYC	Leibniz Institute for Tropospheric Research, Leipzig AND ERATOSTHENES Centre of Excellence https://www.tropos.de/	34.6771 N, 33.0376 E	11

Limassol	Cyprus	LIM	Cyprus University of Technology (CUT), Limassol http://www.cut.ac.cy/?languageId=2	34.6700 N, 33.0400 E	10
Madrid	Spain	MDR	CIEMAT (Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas), Madrid http://www.ciemat.es/	40.4565 N, -3.7257 E	669
Minsk	Belarus	MAS	B.I. Stepanov Institute of Physiscs, Minsk http://ifan.basnet.by/	53.9170 N, 27.6050 E	200
Naples	Italy	NAP	Dipartimento di Fisica "Ettore Pancini" dell'Università degli studi di Napoli Federico II http://www.fisica.unina.it/	40.8380 N, 14.1830 E	118
Observatory Hohenpeissenberg	Germany	HPB	DWD Meteorological Observatory Hohenpeissenberg http://www.dwd.de/EN/research/observing_atmosphere/composition_atmosphere/hohenpeissenberg/start_mohp_node.html	47.8019 N, 11.0119 E	97
Palaiseau	France	SIR	Centre National de la Recherche Scientifique - Institut Pierre Simon Laplace, Paris http://sirta.ipsl.fr/	48.7130 N, 2.2080 E	156
Payerne	Switzerland	PAY	MeteoSwiss Aerological Station, Payerne, EPFL, Losanne	46.8167 N, 6.9333 E	491
Potenza	Italy	POT	Consiglio Nazionale delle Ricerche - Istituto di Metodologie per l'Analisi Ambientale (CNR-IMAA), Potenza https://www.imaa.cnr.it/	40.6000 N, 15.7200 E	760
Roma-Tor Vergata	Italy	RME	Consiglio Nazionale delle Ricerche - Istituto di Scienze Marine http://www.ismar.cnr.it/	41.8330 N, 12.6500 E	110
Sofia	Bulgaria	SOF	Institute of Electronics, Bulgarian Academy of Sciences, Sofia http://www.ie-bas.org/ie_Eng.htm	42.6500 N, 23.3800 E	550
St. Petersburg	Russia	SPL	Research Park of Saint-Petersburg State University - Observatory of environmental safety http://english.spbu.ru/	59.9427 N, 30.2730 E	35
Thessaloniki	Greece	THE	Aristotle University of Thessaloniki, Thessaloniki http://lap.physics.auth.gr/	40.6300 N, 22.9500 E	50
Warsaw	Poland	WAW	University of Warsaw, Faculty of Physics https://www.igf.fuw.edu.pl/en/instruments/laboratorium-pomiarow-zdalnych-dr-hab-iwona-s-stachlewska-823051-9874/	52.2100 N, 20.9800 E	112

Table 2. Vertical profiles available from EARLINET Level 2 data.

Location	Total Number of profiles	Backscattering profiles @ 355 nm	Backscattering profiles @ 532 nm	Backscattering profiles @ 1064 nm	Extinction profiles @ 355 nm	Extinction profiles @ 532 nm	Measurement period
Andøya,	198	40	94	61	3	0	2005-2021
Antikythera	255	207	48	0	0	0	2020-2022
Athens,	3174	975	889	566	442	302	2000-2021
Barcelona	8261	2172	2640	3061	189	199	2010-2020
Belgrade	36	24	0	0	12	0	2020-2020
Belsk	4134	931	1635	1567	1	0	2000-2021
Bucharest	5626	1429	1436	1606	580	565	2007-2021
Cabauw	1014	199	279	278	188	70	2008-2020
Clermont-Ferrand	1465	1252	213	0	0	0	2012-2020
Cluj-Napoca	118	24	41	28	11	14	2019-2021
Dushanbe	265	51	54	54	53	53	2015-2016
Evora	21335	5109	5964	5410	2224	2628	2009-2021
Garmisch-Partenkirchen	3017	91 (2099*)	654	173	0	0	2000-2019
Granada	7563	2254	2666	2107	230	306	2007-2017
Ispra	5329	547	2439	1286	504	553	2006-2021
Kuopio	1509	253	307	448	270	231	2010-2021
Lecce	2334	663 (366*)	332	342	295 (214**)	122	2000-2020
Leipzig	3935	518	1533	1162	324	398	2000-2020
Lille	391	30	124	39	98	100	2017-2021
Limassol (CYC)	1928	458	454	482	267	267	2021-2021
Limassol (LIM)	128	0	127	0	0	1	2010-2020
Madrid	1706	299	874	429	0	104	2006-2020
Minsk	1465	279	677 (202***)	305	1	1	2000-2020
Naples	2208	377 (859*)	361	27	167 (273**)	144	2000-2017
Observatory Hohenpeissenberg	942	203	237	261	122	119	2016-2021
Palaiseau	4451	2140	1408	586	304	13	2008-2021
Payerne	1019	1019	0	0	0	0	2010-2022
Potenza	4729	1197	1274	1018	796	444	2000-2021
Roma-Tor Vergata	160	1	159	0	0	0	2017-2020
Sofia	2169	1712	226	231	0	0	2002-2021
St. Petersburg	128	3	9	116	0	0	2016-2019
Thessaloniki	2693	752	941	478	361	161	2001-2020
Warsaw	3763	753	666	604	862	878	2013-2021

(*) backscatter profiles at 351 nm; (**) extinction profiles at 351 nm; (***) backscatter profiles at 694 nm.

For Level 2 data WP1 provides for each LIDAR station (33 in total) a metadata reporting the main characteristics of the stations (name, Country, station code, coordinates, altitude, contact person), name, number, start/end of

the measurement of vertical profiles for each wavelength and total measuring period. These Deliverables prepared in WP1 are not available in the EARLINET database and were created for the FOCI project with the main objective of summarizing the information available at each LIDAR station. In fact, as stated in Section 1.1, users of the aerosol profile data must not redistribute these data to third parties, but encourage access and dissemination of EARLINET data directly through ACTRIS Data Portal or from the topical data base at https://data.earlinet.org/ to accept the data policy, and ensure the use of the latest version of the data.

As already mentioned, Level 3 data are organized so that for each station, three types of data are released: profile values, integrated quantities, layer statistics. For each type of data, four different temporal aggregations are provided: seasonal, annual, normal seasonal, normal monthly. These Level 3 data were obtained from the QA/QC Level 2 data for the 33 stations (Table 3).

Table 3. Stations considered for the Level 3 products (Table from Level 3 Data catalogue; https://www.earlinet.org/index.php?id=319).

Station Code	Name	Station Code	Name	Station Code	Name
aby	Aberystwyth	**hpb**	Hohenpeissenberg	**nap**	Naples
arr	Andoya	**ino**	Bucharest	**pot**	Potenza
atz	Athens	**ipr**	Ispra	**puy**	Clermont-Ferr
brc	Barcelona	**kuh**	Kuehlungsborn	**rme**	Roma
cbw	Cabauw	**lei**	Leipzig	**sal**	Lecce
cog	Belsk	**lkp**	Linkoping	**sir**	Palaiseau
dus	Dushanbe	**lle**	Lille	**sof**	Sofia
evo	Evora	**mas**	Minsk	**spl**	St. Petersburg
gar	Garmisch	**mdr**	Madrid	**the**	Thessaloniki
gra	Granada	**mel**	Melpitz	**ucc**	Cork
hbu	Hamburg	**muc**	Maisach	**waw**	Warsaw

Level 3 dataset consists of 4158 files, 126 for each station: 42 contain profile values, 42 contain integrated quantities, 42 contain layer statistics. For each category, 20 files report annual statistics (one file per year, from 2000 to 2019); 20 files report seasonal statistics (one file per year, from 2000 to 2019); 1 file reports normal seasonal statistics; 1 file reports normal monthly statistics.

Level 3 dataset can be downloaded from https://www.earlinet.org/index.php?id=319 ..

3. SATELLITE DATA/PRODUCTS

Satellite instruments provide a vantage point for studying aerosol loading and properties consistently over different regions of the world. This document provides below the satellite data available for FOCI.

3.1 Regional and global aerosol optical depth (AOD) records from major available satellite products.

The Finnish Meteorological Institute (FMI; co-leader of WP1) provided to FOCI monthly Aerosol Optical Depth (AOD) products obtained from different satellite sensors and with different algorithms (gridded monthly AOD merged product for the period 1995–2017).

Fifteen land and ocean regions were defined (Figure 3): Europe (Eur), Boreal (Bor), northern Asia (AsN), eastern Asia (AsE), western Asia (AsW), Australia (Aus), northern Africa (AfN), southern Africa (AfS), South America (SA), eastern North America (NAE), western North America (NAW), Indonesia (Ind), Atlantic Ocean dust outflow (AOd) and Atlantic Ocean biomass burning outflow (AOb). In addition, south-eastern China (ChinaSE), which is part of the AsE region, marked with a blue frame, is considered separately. Land, ocean and global AOD were also considered.

Details about instruments, algorithms and AOD products are available in Sogacheva et al. (2020).

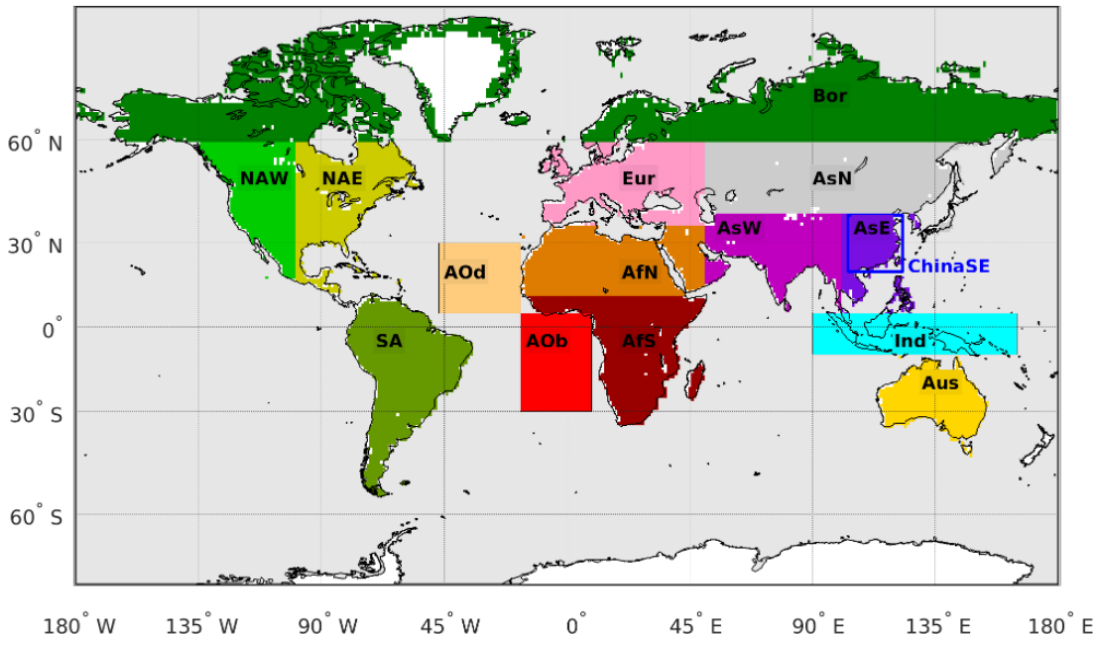


Figure 3. Fifteen land and ocean regions: Europe (Eur), Boreal (Bor), northern Asia (AsN), eastern Asia (AsE), western Asia (AsW), Australia (Aus), northern Africa (AfN), southern Africa (AfS), South America (SA), eastern North America (NAE), western North America (NAW), Indonesia (Ind), Atlantic Ocean dust outflow (AOd) and Atlantic Ocean biomass burning outflow (AOb). In addition, south-eastern China (ChinaSE), which is part of the AsE region, marked with a blue frame, is considered separately. Land, ocean and global AOD were also considered.

3.1.1 Data policy

The monthly Aerosol Optical Depth (AOD) products obtained from different satellite sensors are provided to FOCI by FMI (co-leader of WP1) and are available at http://nsdc.fmi.fi/data/data_aod (Sodankyla National

Satellite Data Centre). When applying the data for commercial, educational or scientific works and publications, please use the following citation: Sogacheva, L., Popp, T., Sayer, A. M., Dubovik, O., Garay, M. J., Heckel, A., Hsu, N. C., Jethva, H., Kahn, R. A., Kolmonen, P., Kosmale, M., de Leeuw, G., Levy, R. C., Litvinov, P., Lyapustin, A., North, P., Torres, O., and Arola, A.: Merging regional and global aerosol optical depth records from major available satellite products, Atmos. Chem. Phys., 20, 2031-2056, doi.org/10.5194/acp-20-2031-2020, 2020. Acknowledge: Dr. L.Sogacheva and AEROSAT team (aerosat.org).

3.1.2 Data Quality Control (QC)

Consistency of regional time records of monthly AOD from 16 different satellite products as TOMS, AVHRR, SeaWiFS, ATSR-2, AATSR, MODIS, MISR, POLDER, VIIRS and EPIC were demonstrated and verified against a ground-based AERONET monthly mean dataset in order to answer the question how well a satellite dataset can reproduce monthly gridded mean AERONET values in a region. To build an AOD product from individual satellite products, two different approaches were introduced and tested. In approach 1, a simple median of the 12 uncorrected and shifted to Terra DT&DR product time records was conducted. In approach 2, the AOD evaluation results (for different aerosol types) against AERONET were used to infer a ranking which was then used to calculate a weighted AOD mean. Two different ranking methods, RM1, simple ranking based on better statistics, and RM2, ranking based on binned statistics, were tested in approach 2. In addition, the order of the processing steps in approach 2 was interchanged (L3 dataset merging or regional merging) to test the stability of the results.

Ten merged L3 AOD monthly products were created and evaluated with AERONET. The evaluation shows that the quality of the merged products (except for one created with the approach 1 for shifted AOD) is as good as that of the most highly ranked individual AOD products in each region.

Details about AOD products QC are available in Sogacheva et al. (2020).

3.1.3 Data provided (Deliverable D1.1/Milestone M1.1)

WP1 provides gridded monthly AOD merged product at 550 nm for the period 1995–2017 (Sogacheva et al., 2020). The monthly Aerosol Optical Depth (AOD) products obtained from different satellite sensors are provided to FOCI by FMI (co-leader of WP1) and are available at http://nsdc.fmi.fi/data/data_aod .

3.2 ModIs Dust AeroSol (MIDAS): A global fine resolution dust optical depth dataset

Monitoring and describing the spatiotemporal variability of dust aerosols is crucial to understand their multiple effects, related feedbacks and impacts within the Earth system. The MIDAS (ModIs Dust AeroSol) dataset has been recently made available (Gkikas et al., 2021; 2022).

From Gkikas et al. (2021):

"MIDAS provides columnar daily dust optical depth (DOD) at 550 nm at global scale and fine spatial resolution (0.1° x 0.1°) over a 15-year period (2003-2017). This new dataset combines quality filtered satellite aerosol optical depth (AOD) retrievals from MODIS-Aqua at swath level (Collection 6.1, Level 2), along with DOD-to-AOD ratios provided by MERRA-2 reanalysis to derive DOD on the MODIS native grid. The uncertainties of MODIS AOD and MERRA-2 dust fraction with respect to AERONET and LIVAS, respectively, are taken into account for the estimation of the total DOD uncertainty…… MIDAS expands, complements and upgrades existing observational capabilities of dust aerosols and it is suitable for dust climatological studies, model evaluation and data assimilation."

Figure 4 reports the defined regional subdomains presented in Gkikas et al. (2021).

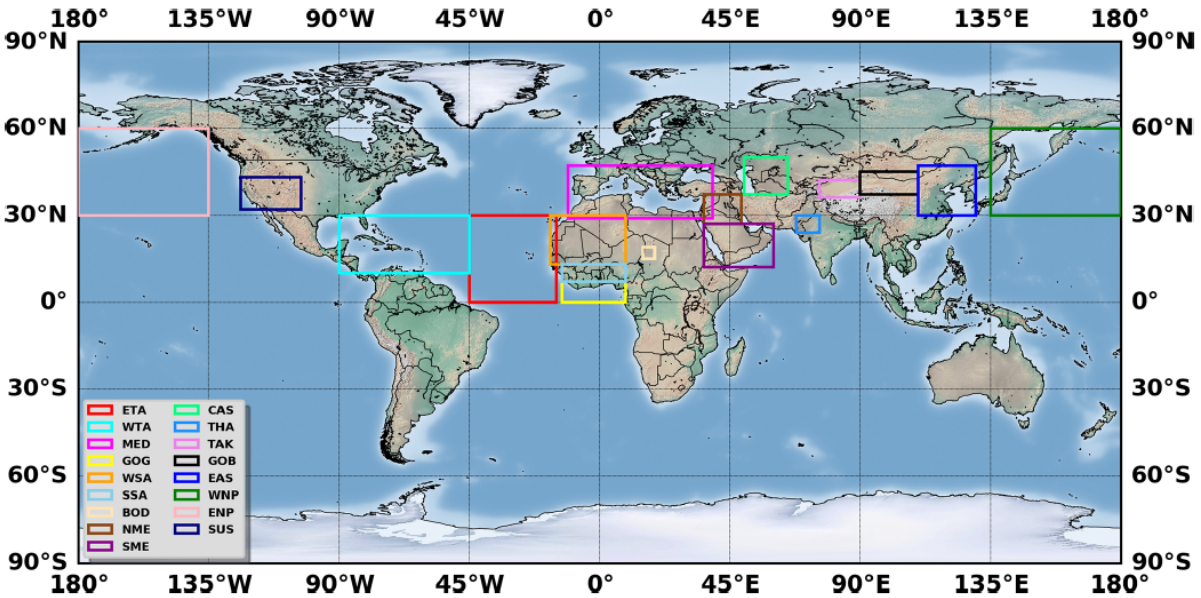


Figure 4. Regional domains of the east tropical Atlantic (ETA), west tropical Atlantic (WTA), Mediterranean (MED), Gulf of Guinea (GOG), western Sahara (WSA), sub-Sahel (SSA), Bodélé Depression (BOD), northern Middle East (NME), southern Middle East (SME), central Asia (CAS), Thar Desert (THA), Taklamakan Desert (TAK), Gobi Desert (GOB), East Asia (EAS), western North Pacific (WNP), eastern North Pacific (ENP), and southwestern United States (SUS).

3.2.1 Data policy

The MIDAS data set is available at https://doi.org/10.5281/zenodo.4244106 (Gkikas et al., 2020; Antonis Gkikas, Emmanouil Proestakis, Vassilis Amiridis, Stelios Kazadzis, Enza Di Tomaso, Alexandra Tsekeri, Eleni Marinou, Nikos Hatzianastassiou, & Carlos Pérez García-Pando. (2020). ModIs Dust AeroSol (MIDAS): A global fine resolution dust optical depth dataset [Data set]. Zenodo. https://doi.org/10.5281/zenodo.4244106).

License is available at https://creativecommons.org/licenses/by/4.0/legalcode .

References: Gkikas et al. (2021, 2022)

3.2.2 Data Quality Control (QC)

The description of the datasets (MODIS, MERRA-2, LIVAS, AERONET) used and of the uncertainties of the products are available in Gkikas et al. (2021). The uncertainties of the MODIS AOD and MERRA-2 dust fraction, with respect to the AEronet RObotic NETwork (AERONET) and LIdar climatology of vertical Aerosol Structure for space-based lidar simulation (LIVAS), respectively, are considered for the estimation of the total DOD uncertainty.

3.2.3 Data provided (Deliverable D1.1/Milestone M1.1)

AOD and DOD for the period 2003 – 2017 are available in NetCDF4 format at https://zenodo.org/record/4244106 .

3.3 Methane (CH_4) observation from GOSAT Project

The Greenhouse Gases Observing Satellite "IBUKI" (GOSAT; https://www.gosat.nies.go.jp/en/) is the world's first spacecraft to measure the concentrations of carbon dioxide and methane, the two major greenhouse gases, from space. The spacecraft was launched successfully on January 23, 2009, and has been operating properly since then. GOSAT observational data provide the global distribution of carbon dioxide (CO_2) and methane (CH_4) since 2009. The GOSAT Project is a joint effort of the Ministry of the Environment (MOE), the National Institute for Environmental Studies (NIES), and the Japan Aerospace Exploration Agency (JAXA).

A recent publication acts as a reference for everyone making use of the GOSAT CH_4 data (Parker et al., 2020).

GOSAT observes infrared light reflected and emitted from the earth's surface and the atmosphere. Column abundances of CH_4 are calculated from the observational data (Figure 5). The column abundance of a gas species is expressed as the number of the gas molecules in a column above a unit surface area. The observation instrument onboard the satellite is the Thermal And Near-infrared Sensor for carbon Observation (TANSO).

TANSO is composed of two subunits: the Fourier Transform Spectrometer (FTS) and the Cloud and Aerosol Imager (CAI).

Absorption spectra are obtained from the FTS observational data. The CAI data provide information on clouds and aerosols. These data are used together to calculate column abundances of CH_4 over observation points where interferences of clouds and aerosols are small. Sources and sinks as well as three-dimensional distributions of CH_4 are estimated using a global atmospheric transport model.

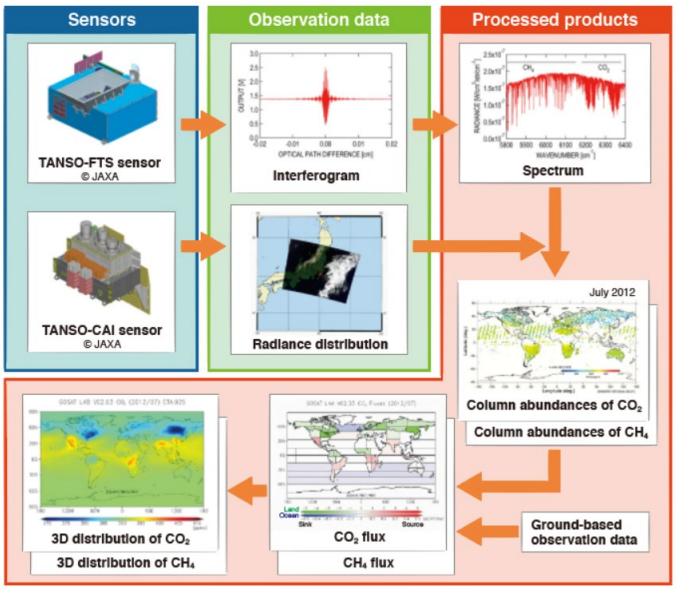


Figure 5. Process flow of data obtained with FTS and CAI.

3.3.1 Data policy

The University of Leicester GOSAT Proxy v9.0 CH_4 data are available from the Centre for Environmental Data Analysis data repository at https://doi.org/10.5285/18ef8247f52a4cb6a14013f8235cc1eb (Parker, R.; Boesch, H. (2020): University of Leicester GOSAT Proxy XCH4 v9.0. Centre for Environmental Data Analysis, 07 May 2020. doi:10.5285/18ef8247f52a4cb6a14013f8235cc1eb). Since the launch of the GOSAT satellite in 2009, these data have been produced by the UK National Centre for Earth Observation (NCEO) as part of the ESA Greenhouse Gas Climate Change Initiative (GHG-CCI) and Copernicus Climate Change Services (C3S) projects.

Use of these data is covered by the following license: http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/.

3.3.2 Data Quality Control (QC)

Here we consider the FTS SWIR Level 2 data products that store the column abundances of CH_4 retrieved from the radiance spectra in the band 1 through 3 of FTS.

In order for the GOSAT data products to be utilized meaningfully in the science community, uncertainties associated with the data products need to be clarified through validation. High-precision data obtained independently by ground-based instruments and aircrafts are used to validate the data products. Improvements in the data processing algorithms are made based on the validation results. For validating the Level 2 column abundances of CH_4, data from ground-based high-resolution Fourier transform spectrometers and in-situ observation instruments installed on aircrafts were used. Properties of clouds and aerosols calculated in the routine data processing are checked against the data obtained with remote sensing instruments such as ground-based sky radiometers and lidars. The estimated source/sink strengths and three-dimensional distributions of CH4 are evaluated by comparing with other model outputs (Figure 10; https://www.gosat.nies.go.jp/en/about_6_validation.html). Figure 6 shows the sounding locations of the ground-based high-resolution Fourier transform spectrometers in the Total Carbon Column Observing Network and other independent sites (https://tccon-wiki.caltech.edu/) that were used for Level 2 data validation.

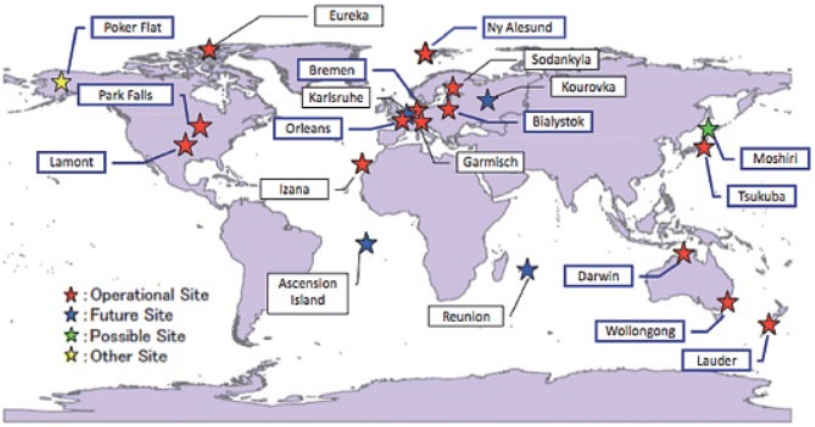


Figure 6. Sounding locations of the ground-based high-resolution Fourier transform spectrometers in the Total Carbon Column Observing Network and those operating independently. (https://tccon-wiki.caltech.edu/).

3.3.3 Data provided (Deliverable D1.1/Milestone M1.1)

Daily averages concentrations in ppb of columnar CH_4 from 2009 to 2021 are provided as Deliverable. Data in NetCDF format are available from the Centre for Environmental Data Analysis data repository at https://doi.org/10.5285/18ef8247f52a4cb6a14013f8235cc1eb (Parker and Boesch, 2020).

Figure 7 below shows an example of CH_4 daily averages concentrations available from the above link.

	X	Y	xch4_quality_flag	solar_zenith_angle	xch4_uncertainty	gain	exposure_id	raw_xco2	retr_flag	raw_xco2_error	xch4	raw_xch4	sensor_zenith_angle	model_xco2	raw_xch4_error	model_xco2_median_diff	model_xco2_range	time
2	-48.97261429	78.67469788	1	85.725922	29.813797	0	2.00904E+21	383.24603	0	2.9244735	1732.3041	1718.0615	15.385242	388.44455	26.503389	0.1756897	0.22537231	23/04/2009 0:18
3	-44.42263031	79.8483429	1	85.391418	36.502617	0	2.00904E+21	383.07901	0	3.4901099	1795.63	1779.4832	1.506498	388.50586	32.338009	0.13314819	0.18560791	23/04/2009 0:18
4	-55.26330948	79.36251831	1	84.40554	33.294758	0	2.00904E+21	379.91281	0	3.469058	1749.5457	1718.8431	14.782787	388.70187	28.699114	0.20159912	0.25640869	23/04/2009 0:18
5	-58.73358917	78.08162689	1	84.698593	32.034496	0	2.00904E+21	372.71655	0	3.3023274	1753.4884	1690.9518	28.138823	388.4982	27.015797	0.1423645	0.24353027	23/04/2009 0:18
6	-65.07168579	78.54172516	1	83.385468	28.684673	0	2.00904E+21	370.41849	0	3.1270335	1748.6561	1673.5718	28.988403	389.04294	23.538557	0.18780518	0.23703003	23/04/2009 0:18
7	-79.03792572	79.05157471	1	80.750244	34.188583	0	2.00904E+21	378.82785	0	4.1053524	1747.8115	1709.3165	28.983404	389.36771	27.835337	0.17373657	0.21081543	23/04/2009 0:19
8	-93.46453857	78.90093231	1	78.116234	23.602215	0	2.00904E+21	380.72943	0	3.0385592	1735.1372	1704.2253	28.980642	389.65973	18.772303	0.035644531	0.035888672	23/04/2009 0:20
9	-94.61740112	80.35162354		77.818481	21.067669	0	2.00904E+21	378.29117	0	2.5930293	1733.9636	1693.5182	15.366262	389.34995	16.989061	0.13250732	0.20181274	23/04/2009 0:20
10	-100.424408	78.59364319	1	76.797615	22.751282	0	2.00904E+21	380.37534	0	2.8374751	1773.058	1739.4102	28.137106	389.71515	18.160381	0.19677734	0.28500366	23/04/2009 0:20
11	-144.1237488	68.7157135	1	62.404613	17.875755	0	2.00904E+21	276.25027	0	1.6923599	1815.7072	1295.2665	28.953659	389.18091	9.9823942	0.33157349	0.34655762	23/04/2009 0:24
12	-155.2505493	70.75848389	1	61.372078	10.375836	0	2.00904E+21	370.59378	0	1.0987031	1774.8751	1694.7856	14.824647	390.08829	8.5390387	0.13711548	0.14761353	23/04/2009 0:25
13	-160.6537018	69.87758636	1	59.657299	13.60371	0	2.00904E+21	304.62485	0	1.2947335	1812.6198	1423.0311	28.692165	389.96304	8.8021507	0.18225098	0.31210327	23/04/2009 0:25
14	-149.3790588	68.29935455	1	60.780346	15.178244	0	2.00904E+21	318.69806	0	1.6278652	1800.8484	1481.8221	14.77599	389.26056	9.934557	0.24343872	0.41201782	23/04/2009 0:25
15	-164.4365845	64.62765503	1	54.418396	9.5324211	0	2.00904E+21	351.6051	0	0.88380915	1781.0134	1613.2708	28.680828	390.13889	7.6231302	0.51132202	0.55737305	23/04/2009 0:26
16	175.4265594	-39.13415909	0	52.430473	10.877333	0	2.00904E+21	388.34821	0	1.3462406	1717.8423	1747.4373	28.227404	383.78497	9.2592258	0.31997681	0.5397644	23/04/2009 0:55
17	172.8486328	-41.47351074	0	54.404499	12.171397	0	2.00904E+21	382.89014	0	1.488712	1732.1364	1737.075	14.859683	383.79904	10.167282	0.28363037	0.49664307	23/04/2009 0:55
18	172.4539185	-42.8299408	0	55.70974	10.600492	0	2.00904E+21	385.83258	0	1.2791978	1704.3365	1722.165	15.507851	383.86856	9.0627222	0.24459839	0.43188477	23/04/2009 0:56
19	169.6122284	-45.14064407	0	57.790466	10.192952	0	2.00904E+21	381.69644	0	1.1658926	1716.7496	1716.0137	1.5460099	383.87579	8.736887	0.22686768	0.40133667	23/04/2009 0:56
20	145.9868469	-68.55345917	1	82.051865	17.802204	0	2.00904E+21	372.17044	0	1.8926991	1663.4177	1625.1827	28.729485	383.00153	15.303815	0.29812622	0.38034058	23/04/2009 1:03
21	149.5920105	-69.23145294	1	82.353485	22.611233	0	2.00904E+21	377.20822	0	2.4927468	1651.2289	1635.3508	15.34947	382.96085	19.61351	0.29931641	0.37582397	23/04/2009 1:03
22	147.6439056	-70.41713715	1	83.659424	25.050858	0	2.00904E+21	366.42621	0	2.4287457	1705.1188	1640.4902	14.730777	382.88602	21.509134	0.30975342	0.40026855	23/04/2009 1:04
23	143.8716278	-69.6811676	1	83.340721	20.239874	0	2.00904E+21	368.74796	0	2.0644774	1620.1849	1568.9709	28.130857	382.91437	17.521524	0.29910278	0.3843689	23/04/2009 1:04
24	141.5581665	-70.82910919	1	84.686462	27.007963	0	2.00904E+21	364.73407	0	2.5443392	1677.476	1606.799	28.727476	382.83438	23.315681	0.29821777	0.38360596	23/04/2009 1:04
25	145.4381104	-71.59545898	1	84.988396	28.04212	0	2.00904E+21	366.75781	0	2.7308118	1593.1674	1534.991	15.355988	382.82321	24.481651	0.30490112	0.38995361	23/04/2009 1:04
26	142.995636	-72.72954559	1	86.289856	33.473942	0	2.00904E+21	356.70511	0	2.8545039	1648.0967	1544.3903	14.731583	382.75104	28.830275	0.2875061	0.37374878	23/04/2009 1:05
27	138.9476929	-71.8993454	1	85.974586	31.274002	0	2.00904E+21	354.78464	0	2.7590802	1642.1584	1530.3833	28.131456	382.798	26.604595	0.28161621	0.35745239	23/04/2009 1:05
28	-68.02561951	77.87898254	1	87.02372	75.733482	0	2.00904E+21	343.40784	0	6.9368815	1713.9202	1521.0293	14.783914	389.00342	59.776089	0.07232666	0.10705566	23/04/2009 1:56
29	-79.8073349	79.36251831	1	84.38356	48.517792	0	2.00904E+21	372.77951	0	5.6170855	1670.5444	1609.3046	14.782307	389.06424	39.956593	0.17855835	0.2517395	23/04/2009 1:56
30	-83.26928711	78.08113861	1	84.678459	40.510941	0	2.00904E+21	383.72729	0	4.5860605	1775.0586	1757.7726	28.13835	389.47916	34.176048	0.24853516	0.4029541	23/04/2009 1:56
31	-94.48888397	80.29831696	1	81.748001	31.836735	0	2.00904E+21	360.94211	0	3.6015418	1747.6301	1628.7537	14.782086	389.2941	24.82439	0.10406494	0.19992065	23/04/2009 1:57
32	-172.3424683	66.34370422	1	59.786854	10.407323	0	2.00904E+21	343.17337	0	1.0085349	1778.5935	1573.9382	28.946182	389.77118	7.9639478	0.28720093	0.53317261	23/04/2009 2:03
33	173.6721497	68.59130859	1	58.347397	13.733718	0	2.00904E+21	332.97122	0	1.5050931	1786.493	1532.9019	28.20812	390.02377	9.531888	0.22146606	0.25088501	23/04/2009 2:04
34	179.7170868	66.42529297	1	57.796982	10.914761	0	2.00904E+21	341.27338	0	1.1392529	1763.5055	1551.7502	1.3113174	389.83731	8.08741	0.20706177	0.28549194	23/04/2009 2:04
35	178.4985352	65.14830017	1	56.494328	9.3863945	0	2.00904E+21	356.05511	0	0.93079209	1764.1476	1618.3448	1.2911136	390.1272	7.4996405	0.23477173	0.31802368	23/04/2009 2:04
36	175.2181549	65.61235809	1	56.12027	9.6383419	0	2.00904E+21	359.46228	0	0.94748259	1780.5804	1648.2959	14.832393	390.28732	7.7930269	0.34677124	0.51235962	23/04/2009 2:04
37	174.2436676	64.30577087	1	54.804291	11.616174	0	2.00904E+21	324.41019	0	1.1375849	1769.5643	1477.9991	15.338361	390.39569	8.201932	0.37319946	0.45373535	23/04/2009 2:04
38	177.3826599	63.86213303	1	55.191906	9.9042063	0	2.00904E+21	339.58231	0	0.92441922	1785.7224	1561.0511	1.2843468	390.42734	7.5434914	0.12399292	0.22332764	23/04/2009 2:04

Figure 7. Example of GOSAT output showing CH_4 concentrations in ppb (column "xch4") for 23 of April 2009.

4. OBSERVATIONS PROVIDED TO FOCI INCLUDED IN THE GHOST DATABASE

There have been numerous model evaluation studies which utilise data from one or more surface measurement networks. However, there is typically little to no detail given about the methodology used in combining data/metadata from across different networks, or regarding the Quality Assurance (QA) or station classifications employed to subset the data. Therefore, evaluation efforts from different groups, are often incomparable and non-reproducible. In response to this, GHOST was established: **Globally Harmonised Observations in Space and Time.**

GHOST harmonises global network surface measurements of components across 13 distinct measurement component matrices. In total, the dataset comprises of 6,193,648,409 processed measurements, for 220 components, across 32 networks, from January 1970 through to January 2022. This represents one of the biggest collection of harmonised atmospheric surface measurements ever undertaken. The full list of components per matrix is shown in Figure 8, and an illustration of the spatial coverage of some of these components is shown in Figure 9.

The central concept of the work in GHOST is an effort to standardise the data/metadata from the major public reporting networks which provide atmospheric surface measurements. Unlike other major synthesis efforts, no data is screened out. Rather, each processed measurement is additionally associated with flags which pertain to a plethora of standardised data flags/quality assurance flags, providing users a way to subset data in a flexible and reproducible manner. In this way also, any subsets of observations used in model evaluation efforts can be traced directly back to a documented project, and cross-group evaluation efforts can be directly compared.

Data will made available to the community via a repository. A DOI will be assigned soon to the repository after publication of the Dataset. The data is available in 2 forms. The first is separated out per network, per component. The second form is merged data across networks per component. Both forms are provided in a NetCDF per YYYYMM.

Matrix	GHOST Component Name
gas	sconco3, sconcno, sconcno2, sconcso2, sconcco, sconcch4, sconcc2h4, sconcc2h6, sconcc3h6, sconcc3h8, sconcisop, sconcc6h6, sconcc7h8, sconcc10h16, sconcnmvoc, sconcvoc, sconnmhc, sconchc, sconcnh3, sconchno3, sconcpan, sconchcho, sconchcl, sconchf, sconch2s
pm	sconcal, sconcas, sconcbc, sconcebc, sconcc, sconcca, sconccd, sconccl, sconccobalt, sconccr, sconccu, sconcec, sconcfe, sconchg, sconck, sconcmg, sconcmn, sconcmsa, sconcna, sconcnh4, sconcnh4no3, sconcni, sconcno3, sconcoc, sconcpb, sconcse, sconcso4, sconcso4nss, sconcso4ss, sconcv, sconczn
pm10	pm10, pm10al, pm10as, pm10bc, pm10ebc, pm10c, pm10ca, pm10cd, pm10cl, pm10cobalt, pm10cr, pm10cu, pm10ec, pm10fe, pm10hg, pm10k, pm10mg, pm10mn, pm10msa, pm10na, pm10nh4, pm10nh4no3, pm10ni, pm10no3, pm10oc, pm10pb, pm10se, pm10so4, pm10so4nss, pm10so4ss, pm10v, pm10zn
pm2.5	pm2p5, pm2p5al, pm2p5a, pm2p5bc, pm2p5ebc, pm2p5c, pm2p5ca, pm2p5cd, pm2p5cl, pm2p5cobalt, pm2p5cr, pm2p5cu, pm2p5ec, pm2p5fe, pm2p5hg, pm2p5k, pm2p5mg, pm2p5mn, pm2p5msa, pm2p5na, pm2p5nh4, pm2p5nh4no3, pm2p5ni, pm2p5no3, pm2p5oc, pm2p5pb, pm2p5se, pm2p5so4, pm2p5so4nss, pm2p5so4ss, pm2p5v, pm2p5zn
pm1	pm1, pm1al, pm1as, pm1bc, pm1ebc, pm1c, pm1ca, pm1cd, pm1cl, pm1cobalt, pm1cr, pm1cu, pm1ec, pm1fe, pm1hg, pm1k, pm1mg, pm1mn, pm1msa, pm1na, pm1nh4, pm1nh4no3, pm1ni, pm1no3, pm1oc, pm1pb, pm1se, pm1so4, pm1so4nss, pm1so4ss, pm1v, pm1zn
aod	od500aero, od500aerocoarse, od500aerofine, fm500frac, od380aero, od440aero, od550aero, od675aero, od870aero, od1020aero, ae440-870aero
extaod	extod440aero, extod440aerocoarse, extod440aerofine, extod675aero, extod675aerocoarse, extod675aerofine, extod870aero, extod870aerocoarse, extod870aerofine, extod1020aero, extod1020aerocoarse, extod1020aerofine, extae440-870aero
absaod	absod440aero, absod675aero, absod870aero, absod1020aero, absae440-870aero
ssa	sca440aero, sca675aero, sca870aero, sca1020aero
asy	asy440aero, asy440aerocoarse, asy440aerofine, asy675aero, asy675aerocoarse, asy675aerofine, asy870aero, asy870aerocoarse, asy870aerofine, asy1020aero, asy1020aerocoarse, asy1020aerofine, sphaero
rin	rinreal440, rinreal675, rinreal870, rinreal1020, rinimag440, rinimag675, rinimag870, rinimag1020
vconc	vconcaero, vconcaerofine, vconcaerocoarse
size	vconcaerobin1, vconcaerobin2, vconcaerobin3, vconcaerobin4, vconcaerobin5, vconcaerobin6, vconcaerobin7, vconcaerobin8, vconcaerobin9, vconcaerobin10, vconcaerobin11, vconcaerobin12, vconcaerobin13, vconcaerobin14, vconcaerobin15, vconcaerobin16, vconcaerobin17, vconcaerobin18, vconcaerobin19, vconcaerobin20, vconcaerobin21, vconcaerobin22

Figure 8. GHOST processed atmospheric components per data matrix.

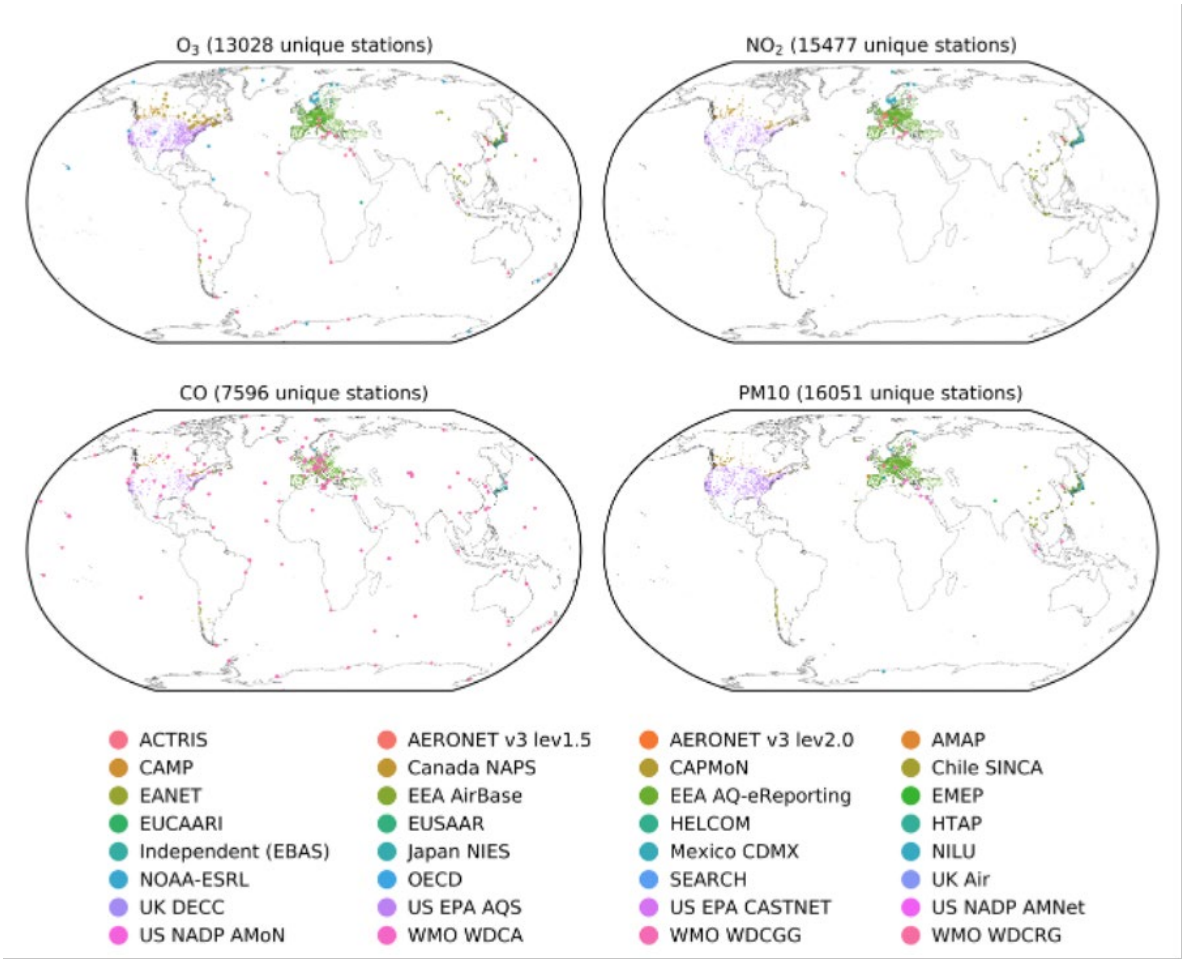


Figure 9. Spatial distribution of the available stations per component across the time array (1970-2022), for 4 key components (O_3, NO_2, CO, PM10).

4.1 Data provided (Deliverable D1.1/Milestone M1.1)

Data will be made available to the community via a repository. GHOST database will be soon available in the EBAS database and a DOI will be assigned after publication of the dataset.

5. IAGOS DATABASE

IAGOS (In-service Aircraft for a Global Observing System; https://www.iagos.org/iagos-data/) combines the expertise of European scientific institutions and weather services with the global infrastructure of civil aviation to provide essential data on climate change and air quality. It builds on the scientific and technological experience gained within the research projects MOZAIC (Measurement of Ozone and Water Vapour on Airbus in-service Aircraft), which was funded by the EC between 1993 and 2004 under FP 4 and FP 5, and CARIBIC (Civil Aircraft for the Regular Investigation of the Atmosphere Based on an Instrument Container).

IAGOS delivers a time and spatially resolved multi-component dataset on Essential Climate Variables (ECV) and Air Pollutants. The data provide information on the distribution and long-term changes in the troposphere and lower stratosphere and regular vertical profiles over major cities.

Table 4 reports the IAGOS observed variables that can be downloaded from https://iagos.aeris-data.fr/download/. For all variables reported in Table 4 the error and the flags are also available.

Table 4. IAGOS observed variables.

IAGOS variable	Long name	CF standard name	Unit	Processing levels	Available since	Source
O3	Ozone mixing ratio	mole_fraction_of_ozone_in_air	ppb	L0, L1, L2, NRT	Aug 1994	P1[(a)], PM[(b)], PC[(c)]
CO	Carbon monoxide mixing ratio	mole_fraction_of_carbon_monoxide_in_air	ppb	L0, L1, L2, NRT	Dec 2001	P1, PM, PC
NO	Nitrogen monoxide mixing ratio	mole_fraction_of_nitrogen_monoxide_in_air	ppb	L0, L1, L2	Aug 1994	P1
NO_x	NO_x expressed as nitrogen mixing ratio	mole_fraction_of_nox_expressed_as_nitrogen_in_air	ppb	L0, L1, L2	Aug 1994	P1
NOy	Total reactive nitrogen mixing ratio	mole_fraction_of_total_reactive_nitro gen_in_air	ppb	L0, L1, L2	Aug 1994	P1
cloud	Cloud particles total cocnetration	Number_concentration_of_cloud_liquid_water_particles_in_air	cm^{-3}	L0, L1, L2	Jul 2011	P1
CH4	Methane mixing ratio	mole_fraction_of_methane_in_air	ppb	L0, L1, L2		P2[(d)], PC
PNSD	Particle size distribution	number_size_distribution_of_dried_ aerosol_particles_at_STP_in_air	1e6 m^{-3}	L0, L1, L2	May 2005	P1
PNC	Particle number concentration	Particle number concentration between 140 and 1050 nm diameter	1e6 m^{-3}	L0, L1, L2	May 2005	P1
PSC	Particle surface concentration	Particle number concentration between 140 and 1050 nm diameter	1e-6 m^{-3}	L0, L1, L2	May 2005	P1
PVC	Particle volume concentration	Particle number concentration between 140 and 1050 nm diameter	1e-12	L0, L1, L2	May 2005	P1

(a) P1: Package 1; (b) PM: Package MOZAIC; (c) PC: Package CARIBIC; (d) P2: Package 2

IAGOS observational data are submitted to several automatic and manual processes during its life cycle. Each step corresponds to a different processing level.

Figure 10 and Table 5 describe the data processing levels workflow (source: https://iagos.aeris-data.fr/documents/).

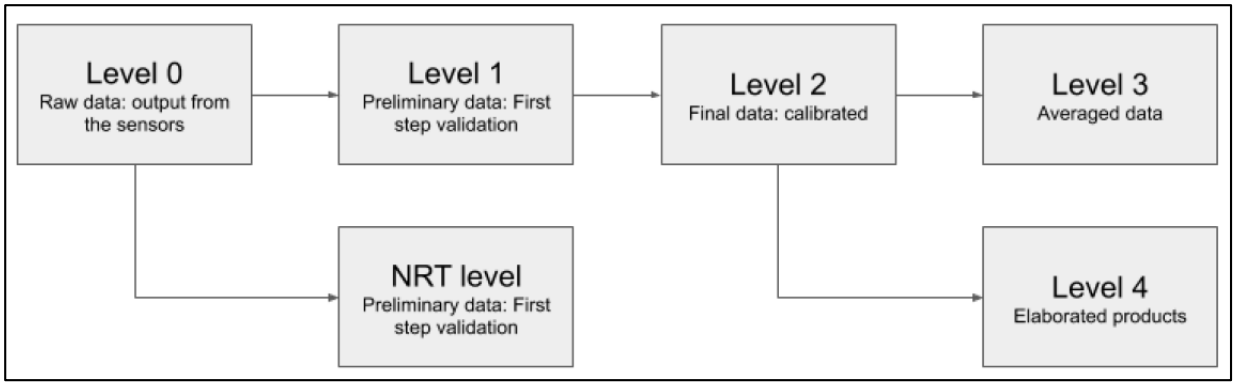


Figure 10. Description of the data processing levels workflows.

Table 5. Definitions of the IAGOS data processing levels.

Level	Description	Type of product
Level 0A	Raw sensor output in sensor or physical units	In situ Observation
Level 0B	Raw sensor output in sensor or physical units with minimum level of quality control	In situ Observation
NRT	Quality assured data with minimum level of automatic quality control provided within three days after the acquisition	In situ Observation
Level 1	Quality assured data with minimum level of quality control by the instrument PI or automated	In situ Observation
Level 2	Approved, calibrated and fully quality controlled data product or geophysical variable	In situ Observation
Level 3	Averaged global gridded products	Gridded climatologies (long-term) of O3, CO and H2O
Level 4	Elaborated data products derived by post-processing of Level 2 data and data from other sources like ECMWF (re-)analyses, FLEXPART lagrangian model	● Footprints ● CO contributions ● ECMWF variables ● Representation error ● PBL profiles

The datasets managed by the Data Centre are of two sorts:

● final citable datasets: Level 2 to Level 4

● internal datasets: Level 0 to Level 1

All the datasets are documented by rich metadata including the following information:

● spatial and temporal extent

● contacts information

● keywords (name of the variables, etc.)

● instruments and platforms identification

● provenance information (original dataset, history of processing, etc.)

Only the final data sets are assigned with a DOI. So far, the following DOI have been assigned:

- Observational Time series : https://doi.org/10.25326/06

- Observational Vertical Profiles : https://doi.org/10.25326/07

- CO contributions: https://doi.org/10.25326/3

- PBL-referenced profiles of O3 and CO: https://doi.org/10.25326/4

A DOI is also assigned to the IAGOS Data Portal: https://doi.org/10.25326/20

For the internal datasets non-persistent homemade identifiers are generated. A new system based on ePIC handles is currently under implementation.

Derived and elaborated products

Data product is composed of variables of two sort: Carbon Monoxide contribution and interpolated modeled data from ECMWF.

Carbon Monoxide contribution: quantification of IAGOS CO origin calculated with SOFT-IO v1.0

Source attribution using FLEXPART and carbon monoxide emission inventories (Sauvage et al., 2017) is a tool based on the FLEXPART particle dispersion model (Stohl et al.al., 2005) coupled on emission inventories provided in the scientific community (e.g. GVAS v1.2, GFED4, MACCity, EDGAR). It has been developed for the atmospheric community to quantify source/receptor links for atmospheric trace gases considered as passive, such as for in-situ measurements of carbon monoxide (CO). SOFT-IO has been first applied and evaluated over the IAGOS data-sets (MOZAIC, CARIBIC, IAGOS).

SOFT-IO simulates the global contributions of anthropogenic and biomass burning emissions from the ECCAD emission inventory database (http://eccad.aeris-data.fr/) for all measured CO mixing ratio.

This product will help in the interpretation of the subtantial IAGOS data-set. In particular it helps quantifying the CO geographical origin and emission sources that drive the observed CO distributions in the troposphere and in the lower stratosphere.

For each CO measurements (every 0.5° in latitude or longitude at cruising altitude, every 10hPa during ascent or descent of the flight), CO origin is automatically calculated and separated by type of source (biomass burning and anthropogenic emissions) and by geographical regions (14 over the globe, see Fig.1). Ancillary data of CO contribution is automatically calculated and now available in the IAGOS data base, from December 2001 to present day. IAGOS also provides systematic ancillary parameters based on ECMWF data and on FLEXPART back-trajectories along each flight measurements. These ancillary parameters are available from 1994 to present day. The following ECMWF data are interpolated along flight trajectories using 6 hours ECMWF analysis and 3 hours ECMWF forecasts:

- Air potential temperature
- Geopotential height at 500 hPa
- Orography
- Pressure at 4 different PV values (1.5, 2, 3 and 4)
- Surface Pressure

- Vertical wind velocity
- Boundary layer height

Below the list of derived and elaborated IAGOS products that can be downloaded from https://iagos.aeris-data.fr/download/:

1) IAGOS Carbon Monoxide contribution and interpolated modeled data from ECMWF – L4
   Information/download: https://doi.org/10.25326/3
2) IAGOS Climatologies of Ozone and Carbon Monoxide - L3
   Information/download: https://doi.org/10.25326/08
3) IAGOS Planetary boundary layer referenced profiles of Ozone and Carbon Monoxide – L4
   Information/download: https://doi.org/10.25326/4
4) IAGOS final quality controlled Observational Data L2 – Time series
   Information/download: https://doi.org/10.25326/06
5) IAGOS final quality controlled Observational Data L2 – Vertical profiles
   Information/download: https://doi.org/10.25326/07

5.1 Data policy

Use of the IAGOS data is free for non-commercial users. Access to the data base is granted after registration and upon acceptance of the data protocol.

IAGOS data is licensed under the Creative Commons Attribution 4.0 International licence (CC BY 4.0).

The summary of (and not a substitute for) the licence can be found at https://creativecommons.org/licenses/by/4.0/legalcode

Using the IAGOS data, users are free to:

- Share — copy and redistribute the material in any medium or format
- Adapt — remix, transform, and build upon the material for any purpose, even commercially

Under the following terms:

- Attribution — You must give appropriate credit, provide a link to the licence, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the licence permits.

The licensor cannot revoke these freedoms as long as you follow the licence terms.

5.1.1 How to cite IAGOS data

Use of the data requires proper reference and citation of the IAGOS data, using the exact citation (including the provided DOI) as provided at the moment of upload from IAGOS, if applicable.

For all of the IAGOS data IAGOS will provide with the citation to use through the landing page of the data object; from this page users will always be able to download the data object again. The Digital Object Identifier (DOI) of the data object will always resolve to its landing page.

5.1.2 Aknowledgements

By downloading the IAGOS data product users agree to the licencing conditions that apply to the data. Under this license derived products and redistribution are allowed, but users are required to always inform of the original source of the data used, refer them to the license text and the original source at IAGOS for possible updates or uploads.

IAGOS ask users to inform the data providers, traceable through the metadata connected to the provided DOI, when the data is used for publication(s), and to offer them the possibility to comment and/or offer them co-authorship or acknowledgement in the publication when this is justified by the added value of the data for your results.

In accordance with the IAGOS data policy, users of IAGOS data products are required to:

1. include the following acknowledgements in publications: "MOZAIC/CARIBIC/IAGOS data were created with support from the European Commission, national agencies in Germany (BMBF), France (MESR), and the UK (NERC), and the IAGOS member institutions (http://www.iagos.org/partners). The participating airlines (Lufthansa, Air France, Austrian, China Airlines, Hawaiian Airlines, Iberia, Cathay Pacific, Air Namibia, Sabena) supported IAGOS by carrying the measurement equipment free of charge since 1994. The data are available at http://www.iagos.fr thanks to additional support from AERIS."
2. offer co-authorship to the IAGOS Principal Investigators if the IAGOS data play a significant role in the publication
3. identify themselves and provide contact information (valid email address)
4. provide a short description of the intended research

5.2 Data Quality Control (QC)

The IAGOS data products are provided "as is", without warranty of any kind, express or implied, including but not limited to the warranties of merchantability, fitness for a particular purpose, title and non-infringement. In no event shall the copyright holders or anyone distributing the IAGOS data products be liable for any damages or other liability, whether in contract, tort or otherwise, arising from, out of or in connection with the IAGOS data products. Every effort is made to produce the most accurate and precise measurements possible. However, IAGOS reserves the right to make corrections to the data based on recalibration standards or for other reasons

deemed scientifically justified. IAGOS is not responsible for results and conclusions based on the use of these data.

For each observation measured by IAGOS-CORE and IAGOS-CARIBIC instruments, uncertainty and quality flags are provided. The uncertainty is provided for all the measurements made by IAGOS sensors. Table 6 reports the list of flags associated to the data.

Table 6. List of flags associated to the data

Value	State of validity	Definition
0	Good	Passed documented required QC tests
2	Limited (questionable/suspect)	Failed non-critical documented metric or subjective test(s)
3	Erroneous/bad	Failed critical documented QC test(s) or as assigned by the data provider
4	Not validated	Used for data when no QC test performed or the information on quality is not available
7	Missing value	Used as placeholder when data are missing

5.3 Data provided (Deliverable D1.1/Milestone M1.1)

IAGOS data described above can be selected and downloaded from https://iagos.aeris-data.fr/download/. Chapter 4.0 above provides DOI and links to different data or products.

6. SUMMARY AND OUTLOOK

The observations described in this Deliverable (D1.1: Harmonized dataset of in-situ and column integrated aerosols and gas-phase species) contribute to the FOCI integrated observational and modelling analysis which focuses on the radiative forcing properties of different atmospheric non-CO2 species in the wider context of the warming potential of all key GHGs. The observations described in this Document compose the Milestone M1.1 (Datasets for non-CO2 species) and Deliverable D1.1.

WP1 is already working on the Deliverable D1.2 and on the Milestone M1.2 (Database of advanced aerosol particles optical properties) due by month 24.

7. REFERENCES

Gkikas, A., Proestakis, E., Amiridis, V., Kazadzis, S., Di Tomaso, E., Tsekeri, A., Marinou, E., Hatzianastassiou, N., and Pérez García-Pando, C.: ModIs Dust AeroSol (MIDAS): a global fine-resolution dust optical depth data set, Atmos. Meas. Tech., 14, 309–334, https://doi.org/10.5194/amt-14-309-2021, 2021.

Gkikas, A., Proestakis, E., Amiridis, V., Kazadzis, S., Di Tomaso, E., Marinou, E., Hatzianastassiou, N., Kok, J. F., and García-Pando, C. P.: Quantification of the dust optical depth across spatiotemporal scales with the MIDAS global dataset (2003–2017), Atmos. Chem. Phys., 22, 3553–3578, https://doi.org/10.5194/acp-22-3553-2022, 2022.

Mona, L., Amodeo, A., Pandolfi, M., Pappalardo, G.: Saharan dust intrusions in the Mediterranean area: Three years of Raman lidar measurements, JGR: Atmosphere, https://doi.org/10.1029/2005JD006569, 2006.

Parker, R. J., Webb, A., Boesch, H., Somkuti, P., Barrio Guillo, R., Di Noia, A., Kalaitzi, N., Anand, J. S., Bergamaschi, P., Chevallier, F., Palmer, P. I., Feng, L., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Petri, C., Pollard, D. F., Roehl, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: A decade of GOSAT Proxy satellite CH4 observations, Earth Syst. Sci. Data, 12, 3383–3412, https://doi.org/10.5194/essd-12-3383-2020, 2020.

Parker, R. and Boesch, H.: University of Leicester GOSAT Proxy XCH4 v9.0, Centre for Environmental Data Analysis, https://doi.org/10.5285/18ef8247f52a4cb6a14013f8235cc1eb, 2020.

Sauvage, B., Fontaine, A., Eckhardt, S., Auby, A., Boulanger, D., Petetin, H., Paugam, R., Athier, G., Cousin, J.-M., Darras, S., Nédélec, P., Stohl, A., Turquety, S., Cammas, J.-P., and Thouret, V.: Source attribution using FLEXPART and carbon monoxide emission inventories: SOFT-IO version 1.0, Atmos. Chem. Phys., 17, 15271–15292, https://doi.org/10.5194/acp-17-15271-2017, 2017.

Sogacheva, L., Popp, T., Sayer, A. M., Dubovik, O., Garay, M. J., Heckel, A., Hsu, N. C., Jethva, H., Kahn, R. A., Kolmonen, P., Kosmale, M., de Leeuw, G., Levy, R. C., Litvinov, P., Lyapustin, A., North, P., Torres, O., and Arola, A.: Merging regional and global aerosol optical depth records from major available satellite products, Atmos. Chem. Phys., 20, 2031–2056, https://doi.org/10.5194/acp-20-2031-2020, 2020.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys., 5, 2461–2474, https://doi.org/10.5194/acp-5-2461-2005, 2005.