

Training session on data formatting and data submission to the EVDC

Ann Mari Fjæraa (amf@nilu.no)

With contribution from EVDC-team



Norsk institutt
for luftforskning

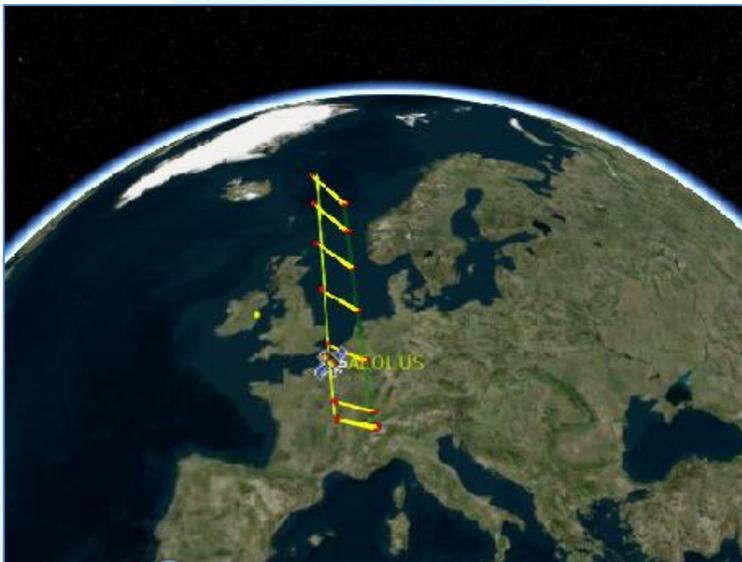
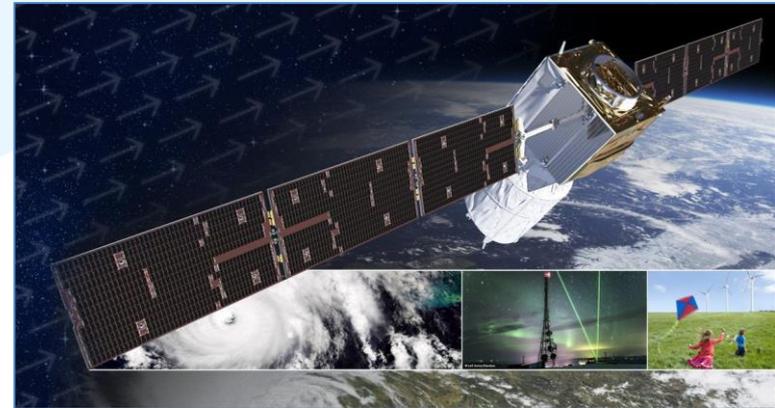
Outline

- About EVDC and GEOMS
- Motivation for a central database and data reporting
- Data formatting and -submission

EVDC: ESA atmospheric Validation Data Centre

EVDC is a central, long-term repository in Europe for archiving and exchange of correlative data for validation of atmospheric composition products from satellite platforms

We focus on ground based and remote sensing in situ data during this course



EVDC - cross cutting activity

Integrating information and measurements from all parts of the project to obtain new information and facilitate easy access to the data

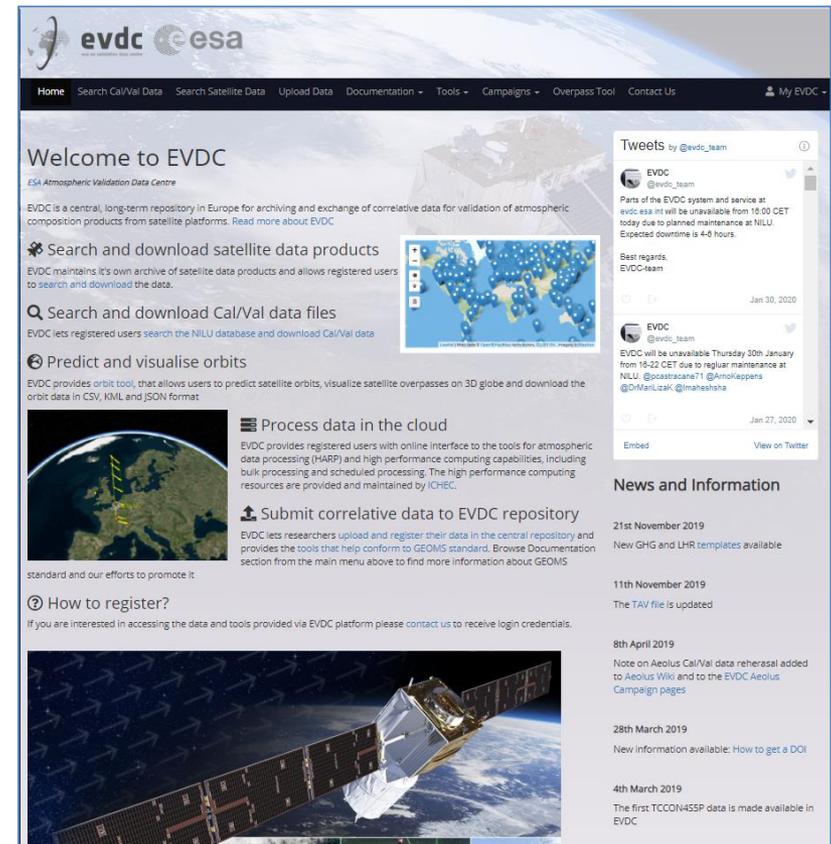
Main services of the data centre

Collect and archive all project data in a long term sustainable system.

Provide free and open access to all data resulting from project.

Complement with data from other relevant networks.

Provide new products and tools for analysis of atmospheric composition.



The screenshot displays the EVDC (ESA Atmospheric Validation Data Centre) website. At the top, there is a navigation bar with links for Home, Search Cal/Val Data, Search Satellite Data, Upload Data, Documentation, Tools, Campaigns, Overpass Tool, and Contact Us. The main content area features a 'Welcome to EVDC' message, followed by several service highlights: 'Search and download satellite data products', 'Search and download Cal/Val data files', 'Predict and visualise orbits', and 'Process data in the cloud'. A 'Submit correlative data to EVDC repository' section is also present. On the right side, there is a 'Tweets by @evdc_team' widget showing recent updates, and a 'News and Information' section with a list of recent news items, including updates on GHG and LHR templates, TAV file updates, and data reherasal additions.

Data reporting, sharing and ownership



NILU Norsk institutt
for luftforskning

Background

Why is data reporting important?

Added value of reporting and sharing data.

Who owns the data in the data base?

Association to projects and the labelling of data e.g:

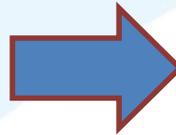
- EVDC data
- NDACC data
- TCCON data

Data policies, Fair and regulated use of data.

Why is data reporting important?

*Lot's of effort behind atmospheric observations,
important that data are used!*

**Atmospheric measurements
are hard work by many people
and considerable amount of
money is invested**



**Reporting makes the data
available for various users, now
and in the future, to contribute to
solve the central environmental
questions within air quality and
climate**

Goal to provide access to data for many users

*Central data base with common standards for import and export facilitate
easy access to data for a wide range of users*

Who owns the data in the data base?

Every dataset created within EVDC (and other programs) is **owned** by the partner/data providers who created this dataset.

**Public data with easy access is the goal,
but it has to be regulated...**

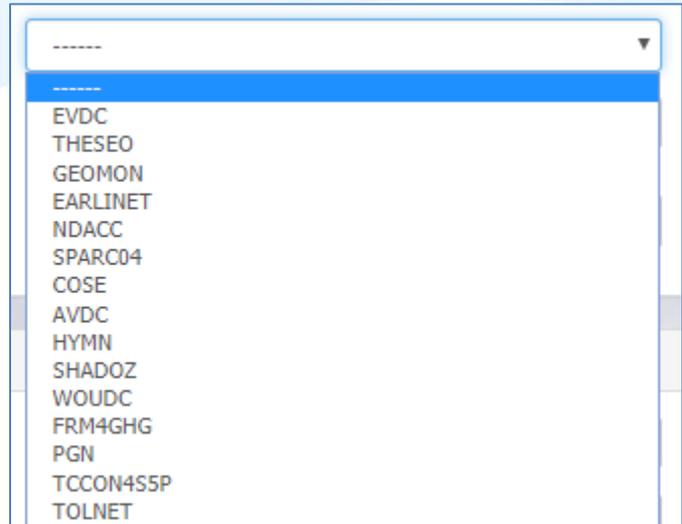
The conditions of use of data is regulated in data policy documents for various programs/networks.

Example: Data associated to NDACC, WOUDC, AVDC

e.g. all measurements performed within the NDACC working groups at NDACC sites will have the NDACC label in EVDC.

WOUDC data have label WOUDC...

Some of the sites overlapping with other programs.



File name link	File size	Submission date	Pi	Campaign
balloon_sonde.o3_awi000_neumayer_20171108t111300z_20171108t132355z_003.h5	153048	2018-02-01 14:29:45.000	Koenig-Langlo, Gert	AVDC, WOUDC
balloon_sonde.o3_awi000_neumayer_20171105t110400z_20171105t125425z_003.h5	135336	2018-02-01 14:29:43.000	Koenig-Langlo, Gert	AVDC, WOUDC
balloon_sonde.o3_awi000_neumayer_20171103t110200z_20171103t131315z_003.h5	153336	2018-02-01 14:29:40.000	Koenig-Langlo, Gert	AVDC, WOUDC
balloon_sonde.o3_awi000_neumayer_20171101t110800z_20171101t125815z_003.h5	135192	2018-02-01 14:29:38.000	Koenig-Langlo, Gert	AVDC, WOUDC
balloon_sonde.o3_awi000_neumayer_20171030t105800z_20171030t124340z_003.h5	131232	2018-02-01 14:29:36.000	Koenig-Langlo, Gert	AVDC, WOUDC
balloon_sonde.o3_awi000_neumayer_20171027t111200z_20171027t130455z_003.h5	137496	2018-02-01 14:29:33.000	Koenig-Langlo, Gert	AVDC, WOUDC
balloon_sonde.o3_awi000_neumayer_20171023t110000z_20171023t125320z_003.h5	137856	2018-02-01 14:29:30.000	Koenig-Langlo, Gert	AVDC, WOUDC
balloon_sonde.o3_awi000_neumayer_20171021t1110200z_20171021t125745z_003.h5	139944	2018-02-01 14:29:28.000	Koenig-Langlo, Gert	AVDC, WOUDC

Data use has to be fair and regulated

Lot of effort behind, **visibility** to the data providers.

Facilitate the involvement of the data providers to ensure proper use of data when necessary.

Reduce misinterpretations (balance between data use, data analysis, depending on use etc).

Make the funding source visible, also important for future funding situation.

Public, open, easy access is the goal...

Data formatting and submission



Norsk institutt
for luftforskning

The EVDC web portal

Home Search Cal/Val Data Search Satellite Data Upload Data Documentation Tools Campaigns Overpass Tool Contact Us My EVDC

Welcome to EVDC

ESA Atmospheric Validation Data Centre

EVDC is a central, long-term repository in Europe for archiving and exchange of correlative data for validation of atmospheric composition products from satellite platforms. [Read more about EVDC.](#)

- Search and download satellite data products**
EVDC maintains its own archive of satellite data products and allows registered users to search and download the data.
- Search and download Cal/Val data files**
EVDC lets registered users [search the NILU database](#) and [download Cal/Val data](#)
- Predict and visualise orbits**
EVDC provides [orbit tool](#), that allows users to predict satellite orbits, visualize satellite overpasses on 3D globe and download the orbit data in CSV, KML and JSON format
- Process data in the cloud**
EVDC provides registered users with online interface to the tools for atmospheric data processing (HARP) and high performance computing capabilities, including bulk processing and scheduled processing. The high performance computing resources are provided and maintained by ICHEC.
- Submit correlative data to EVDC repository**
EVDC lets researchers [upload and register their data in the central repository](#) and provides the [tools that help conform to GEOMS standard](#). Browse Documentation section from the main menu above to find more information about GEOMS standard and our efforts to promote it

[How to register?](#)

Tweets by @evdc_team

EVDC @evdc_team
Parts of the EVDC system and service at [evdc.esa.int](#) will be unavailable from 18:00 CET today due to planned maintenance at NILU. Expected downtime is 4-8 hours.
Best regards, EVDC-team
Jan 30, 2020

EVDC @evdc_team
EVDC will be unavailable Thursday 30th January from 18-22 CET due to regular maintenance at NILU. @pcastracane71 @AmoKeppens @DrMarLizak @maheshasha
Jan 27, 2020

News and Information

21st November 2019
New GHG and LHR [templates](#) available

11th November 2019
The [TAV file](#) is updated

Navigation and login

<https://evdc.esa.int>



The EVDC Cal/Val web interface

Search Cal/Val Data

Search our database of Cal/Val data. Log in to download the data products. Hover over the form fields to see more information about search criteria. Browse Documentation section of EVDC (see the menu above) to find out more about GEOMS standard and metadata.

New Files

0 new files added since your last visit on: 2020-02-11 09:56:00

Search History

Search Date

Data

Location

Data Source Type

Data Discipline Field

Data Discipline Class

Data Originator

Data Supplier

Other

Frameworks

Principal Investigator

AO ID

Time and Location

Date Min

Date Max

Spatial Filter

Off

Point

Bounding Box

Point/Min Longitude (Deg)

Point/Min Latitude (Deg)

Point/Min Altitude (m)

Max Longitude (Deg)

Max Latitude (Deg)

Max Altitude (m)

Stations



Variable

Cal/Val web-interface functions:

- Search datasets by criteria: Location, Data Source Type, Framework, PI..
- Visualise distribution of stations on map.
- Manage access to restricted data.
- List datasets.
- Download data.

Online formatting templates

GEOMS Data Formatting Templates

Data reporting templates are used as an additional guidance to file formatting for specific instrument types. This can be useful to make sure all measurements from similar instruments, e.g. a LIDAR, within a monitoring network report the same mandatory data variables. You may read more about GEOMS templates at the [AVDC site](#).

2011-10-03 GEOMS-TE-BOUY-001.csv

2018-05-28 GEOMS-TE-FTIR-002.csv

2015-12-23 GEOMS-TE-FTIR-ISO-001.csv

2019-02-20 GEOMS-TE-FTIR-TCCON-005.csv

2016-07-21 GEOMS-TE-FTUV-003.csv

2016-07-26 GEOMS-TE-LIDAR-AEROSOL-004.csv

2017-10-28 GEOMS-TE-LIDAR-H2O-005.csv

2017-10-28 GEOMS-TE-LIDAR-O3-005.csv

2017-10-28 GEOMS-TE-LIDAR-TEMPERATURE-005.csv

2014-02-18 GEOMS-TE-MWR-003.csv

2018-02-02 GEOMS-TE-MWR-WIND-001.csv

2018-05-20 GEOMS-TE-PANDORA-DIRECTSUN-GAS-002.csv

2016-12-01 GEOMS-TE-RO-001.csv

2013-07-19 GEOMS-TE-SONDE-002.csv

2018-04-26 GEOMS-TE-UVVIS-DOAS-DIRECTSUN-GAS-007.csv

2018-04-26 GEOMS-TE-UVVIS-DOAS-OFFAXIS-AEROSOL-007.csv

2018-04-25 GEOMS-TE-UVVIS-DOAS-OFFAXIS-GAS-007.csv

2018-04-25 GEOMS-TE-UVVIS-DOAS-ZENITH-GAS-007.csv

2018-11-01 GEOMS-TE-UVVIS-DOBSON-TOTALCOL-001.csv

Data reporting templates are used as an additional guidance to file formatting for specific instrument types.

This can be useful to make sure all measurements from similar instruments, e.g. a LIDAR, within a monitoring network report the same mandatory data variables.

<http://evdc.esa.int/tools/data-formatting-templates/>

Online examples

Examples:

[UWVIS DOAS ZENITH measurements - metadata - data](#)

[AEROSOL LIDAR measurements - metadata_preliminary](#)

[BALLOON SONDE PTU measurements - metadata - data](#)

[OZONE LIDAR measurements - metadata - data](#)

Example files are available as an additional guidance to file formatting for specific instrument types.

```
groundbased_lidar.o3_cnrs.latmos002_haute.provence_20190329t191200z_20190329t233112z_002.meta 15.3 KB
1 ! Output from IDLcr8ASCII application v4.0
2 ! /homevip/annm/projects/EVDC2/avdc_evdc/scripts/groundbased_lidar.o3_cnrs.latmos002_haute.provence_20190329t191200z_20190329t233112z_002.hdf
3 !
4 ! Global Attributes
5 PI_NAME=Godin-Beekmann;Sophie
6 PI_AFFILIATION=Laboratoire Atmospheres, Milieux, Observations Spatiales du CNRS;CNRS.LATMOS
7 PI_ADDRESS=Universite Pierre et Marie Curie, boite 102, 4 Place Jussieu;F-75252 Paris cedex 05;FRANCE
8 PI_EMAIL=sophie.godin-beekmann@latmos.ipsl.fr
9 DO_NAME=Godin-Beekmann;Sophie
10 DO_AFFILIATION=Laboratoire Atmospheres, Milieux, Observations Spatiales du CNRS;CNRS.LATMOS
11 DO_ADDRESS=Universite Pierre et Marie Curie, boite 102, 4 Place Jussieu;F-75252 Paris cedex 05;FRANCE
12 DO_EMAIL=sophie.godin-beekmann@latmos.ipsl.fr
13 DS_NAME=Bodichon;Renaud
14 DS_AFFILIATION=Laboratoire Atmospheres, Milieux, Observations Spatiales du CNRS;CNRS.LATMOS
15 DS_ADDRESS=Universite Pierre et Marie Curie, 4 Place Jussieu;F-75252 Paris cedex 05;FRANCE
16 DS_EMAIL=Renaud.Bodichon@ipsl.fr
17 DATA_DESCRIPTION=Routine lidar stratospheric ozone profile measurements at OHP.
18 DATA_DISCIPLINE=ATMOSPHERIC.CHEMISTRY;REMOTE.SENSING;GROUNDBASED
19 DATA_GROUP=EXPERIMENTAL;PROFILE.STATIONARY
20 DATA_LOCATION=HAUTE.PROVENCE
21 DATA_SOURCE=LIDAR.O3_CNRS.LATMOS002
22 DATA_VARIABLES=LATITUDE.INSTRUMENT;LONGITUDE.INSTRUMENT;ALTITUDE.INSTRUMENT;DATETIME;DATETIME.START;DATETIME.STOP;INTEGRATION.TIME;ALTITUDE;O3.NUM
23 DATA_START_DATE=20190329T191200Z
24 DATA_STOP_DATE=20190329T233112Z
25 DATA_FILE_VERSION=002
26 DATA_MODIFICATIONS=
27 DATA_CAVEATS=
28 DATA_RULES_OF_USE=Follow AVDC data protocol
29 DATA_ACKNOWLEDGEMENT=
30 DATA_QUALITY=Nominal
31 DATA_TEMPLATE=GEOMS-TE-LIDAR-03-005
32 DATA_PROCESSOR=DIAL_V5
33 FILE_NAME=groundbased_lidar.o3_cnrs.latmos002_haute.provence_20190329t191200z_20190329t233112z_002.hdf
34 FILE_GENERATION_DATE=20190411T125540Z
```

The EVDC GEOMS formatting tool

GEOMS tool

Description of the tool

This is an on-line tool to generate GEOMS compliant files. It allows the data provider to upload their data and metadata in a simple single column ASCII format. The tool automatically generates the GEOM file(s) and makes them available for download.

GEOMS File Creation

New Metadata

HDF4 data creation

Description: Upload one or more data files. For multiple files, they must all use the same metadata file input (i.e. measurements from the same instrument, location and data file version). On selecting 'Upload', if a matching metadata file is found it will automatically be chosen for the conversion. Browse and upload a metadata file if no matching file is available or if wanting to use a different file. The conversion program uses the latest [Table Attribute Values \(TAV\)](#) file as an input.

Upload

Data (*.data, *.data.gz)

Ingen fil valgt

Metadata (*.meta)

Ingen fil valgt

The EVDC GEOMS formatting tool

Upload

Data (*.data, *.data.gz) Ingen fil valgt

Metadata (*.meta) Ingen fil valgt

Data file

- groundbased_ftir.ch4_ncar002_mauna.loa.hi_20191101t173913z_20191112t190956z_003.data

Metadata

- groundbased_ftir.ch4_ncar002_mauna.loa.hi_20191101t173913z_20191112t190956z_003.meta

Create HDF4

Data conversion successful

Download

- [groundbased_ftir.ch4_ncar002_mauna.loa.hi_20191101t173913z_20191112t190956z_003.hdf](#)

When file is successfully created →
Upload to EVDC by ssh og web upload
(later slides)

The EVDC GEOMS formatting tool

Data file

- groundbased_ftir_ch4_inoe2000_Poznan190711T073443Z_190711T112009Z.data

Metadata

- groundbased_ftir.ch4_ncar002_mauna.loa.hi_20191101t173913z_20191112t190956z_003.meta

Error: This metadata does not match the data file(s).

Create HDF4

Data file

- groundbased_ftir_ch4_inoe2000_Poznan190711T073443Z_190711T112009Z.data

Metadata

- groundbased_ftir_ch4_inoe2000_Poznan190711T073443Z_190711T112009Z.meta

Create HDF4

Data conversion failed (see Errors in the Log file)

Log file

The EVDC GEOMS formatting tool

```
input data file.  
/tmp/tmpUt41JP/f9f2bcd6-5b74-42c4-8278-698f7fc20b0d/Uploads/gr  
oundbased_ftir_ch4_inoe2000_Poznan190711T073443Z_  
190711T112009Z.data  
INFORMATION: Leading or trailing spaces between DO_ADDRESS  
sub-values removed  
INFORMATION: FILE_META_VERSION updated with corrected TAV  
file version and tool name  
INFORMATION: DATA_TEMPLATE value renamed GEOMS-TE-FTIR-002  
based on DATA_SOURCE value  
INFORMATION: Illegal character(s) in Metadata Attribute  
Entry DATA_QUALITY replaced with valid ISO646-US ASCII  
character(s): €, ", â  
ERROR: PI_AFFILIATION=National Institute of Research and  
Development for Optoelectronics;INOE2000 doesn't match any  
Table Attribute Values under AFFILIATION field  
ERROR: DO_AFFILIATION=National Institute of Research and  
Development for Optoelectronics;INOE2000 doesn't match any  
Table Attribute Values under AFFILIATION field  
ERROR: DS_AFFILIATION=National Institute of Research and  
Development for Optoelectronics;INOE2000 doesn't match any  
Table Attribute Values under AFFILIATION field  
ERROR: DATA_LOCATION=POZNAN doesn't match any Table  
Attribute Values under DATA_LOCATION field  
ERROR: AFFILIATION=FTIR.CH4_INOE doesn't match any Table  
Attribute Values under AFFILIATION field  
ERROR in Check_Metadata/Extract_And_Test procedures: Time:  
Doesn't match any Table Attribute Values under FIELD:  
DATA_VARIABLES_01  
  
IDLcr8HDF stopped - Program Ended on Tue Feb 11 14:30:17 2020
```

Contact
EVDC team
for assistance!

The EVDC GEOMS data upload

Data File Upload Facility

Before uploading a file, first read the [Documents](#) page and the [Browse Meta-data](#) pages. Please check for updated versions of file creations tools and Metadata Guidelines before creating files.

The Data Exchange Protocol can be completed at the following link:
<https://protocol.evdc.nilu.no/>

Upload File

Velg fil groundbased_u...3214z_001.h5

Upload File

100%

<http://evdc.esa.int/upload/>

File Uploaded Successfully, it will be added to the DB in the next 24 hours (as long as it conforms to the GEOMS standards)

