

Need for metadata in NORS - adoption of GEOMS

Martine De Mazière
Belgian Institute for Space Aeronomy
martine@oma.be



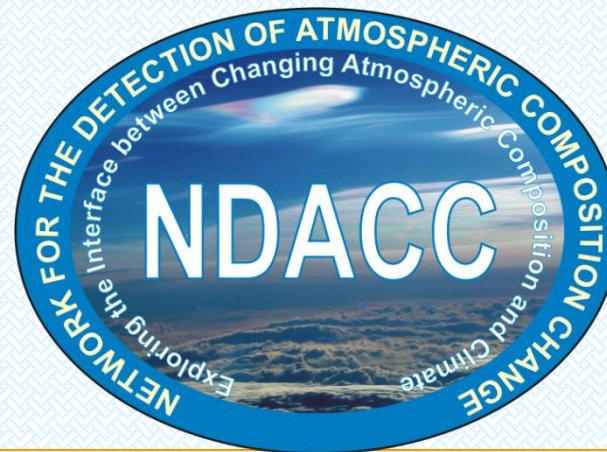
What is NORS ?

- EU FP7 project: Demonstration Network Of ground-based Remote Sensing Observations in support of the GMES Atmospheric Service

- General Objective:

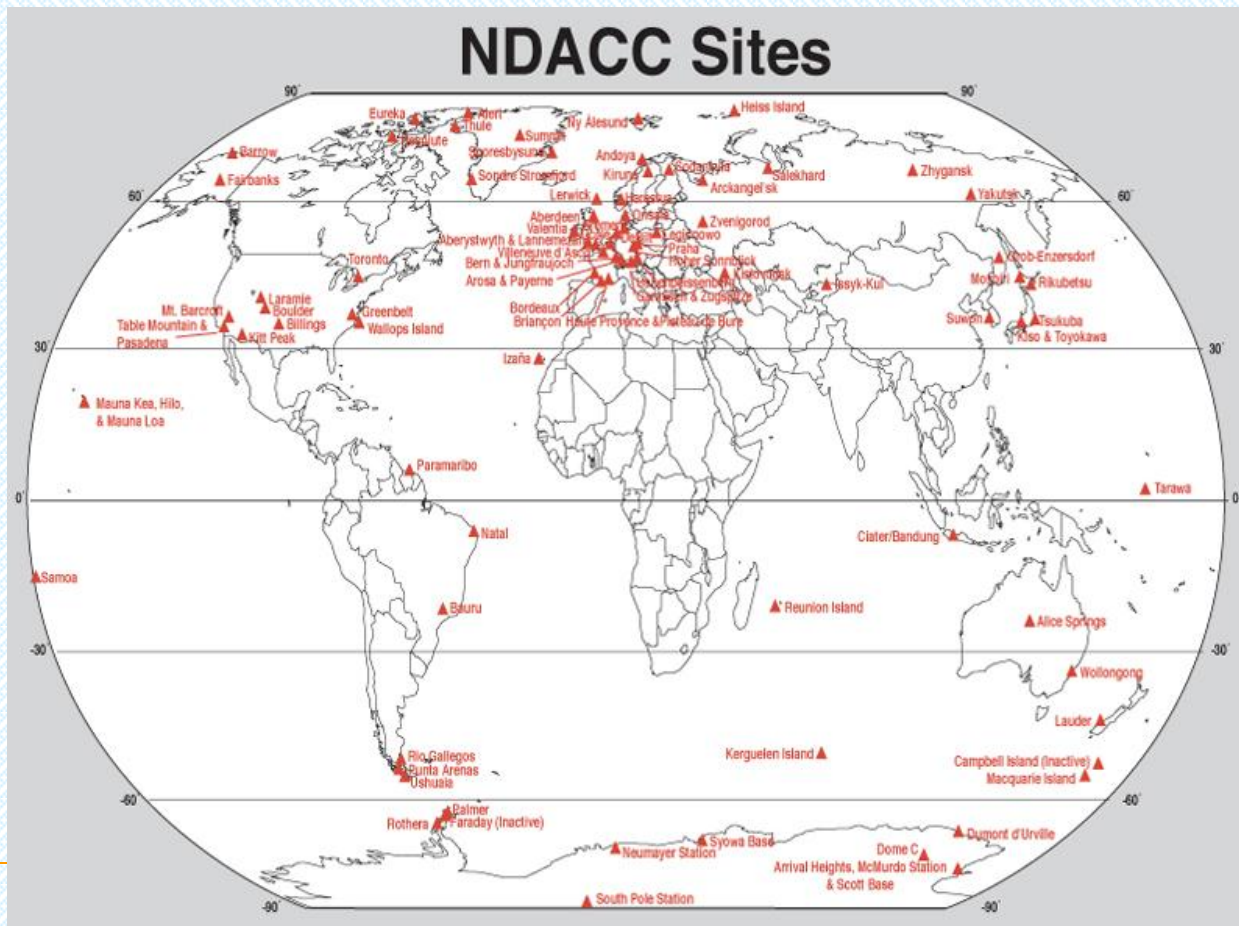
Demonstrate the suitability of NDACC data for supporting the quality assessment of the GMES Atmospheric Service (MACC-II) products.

NDACC: Network for the Detection of Atmospheric Composition Change



What is NDACC ?

NDACC is A global network of more than 70 high-quality atmospheric monitoring stations equipped with ground-based remote sensing instruments and O₃ sondes, that started coordinated operations in 1991.



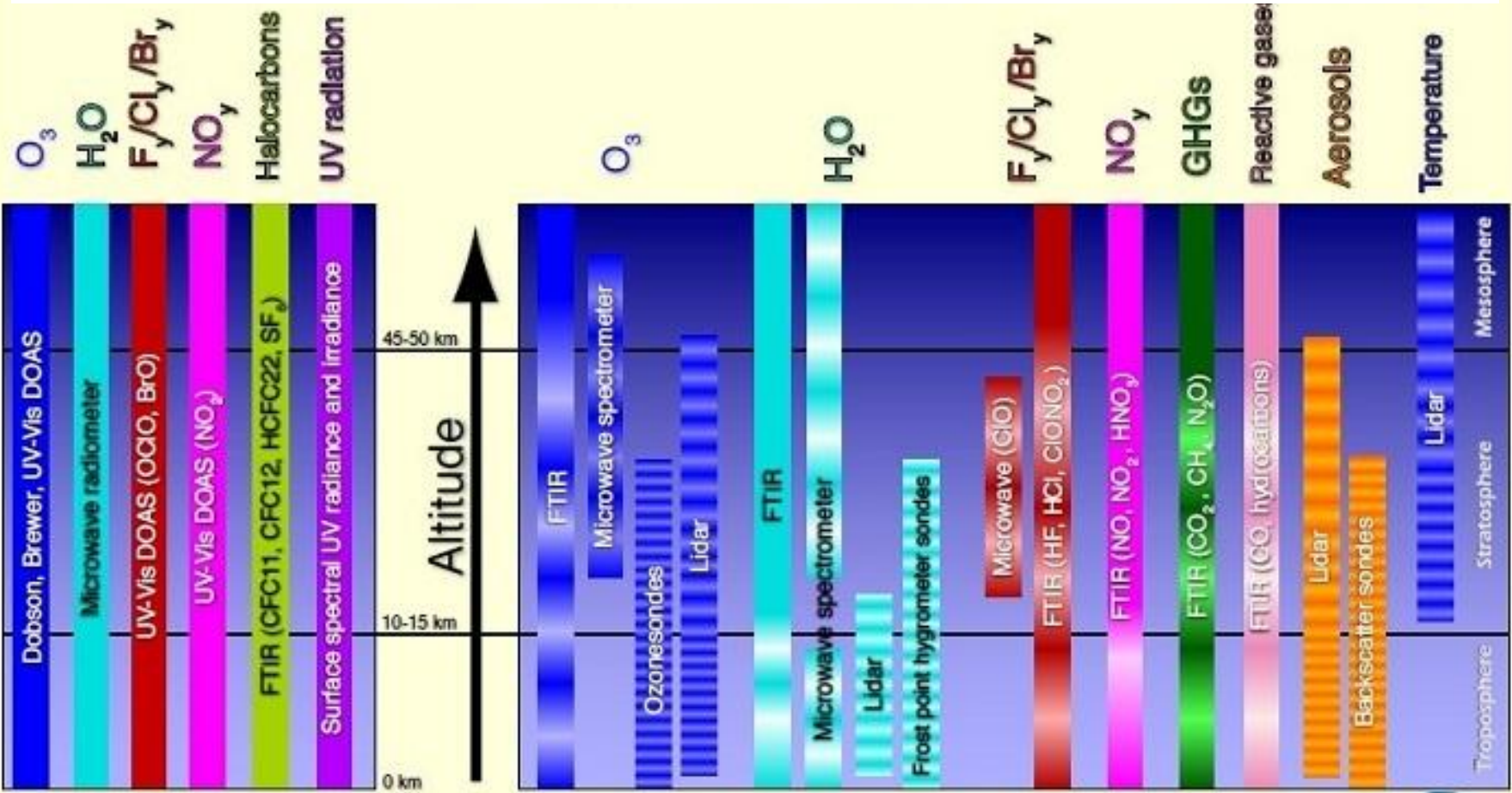
NDACC objectives

- *To study the temporal and spatial variability of atmospheric composition and structure*
- *To provide early detection and subsequent long-term monitoring of changes in the chemical and physical state of the stratosphere and upper troposphere, thereby providing the means to discern and understand the causes of such changes*
- *To establish the links between changes in stratospheric O₃, UV radiation at the ground, tropospheric chemistry, and climate*

- *To provide independent validation, calibration, and complementary data for space-based sensors of the atmosphere*
- *To support process-study field campaigns occurring at various latitudes and seasons.*
- *To provide verified data for testing and improving multidimensional chemistry and transport models of the stratosphere and troposphere, thus enabling reliable forecasting of the atmosphere's evolution.*

NDACC observational capabilities

Lidar, Microwave radiometer, FTIR, (MAX)DOAS, Dobson/Brewer, UV spectrometer, O₃ sondes



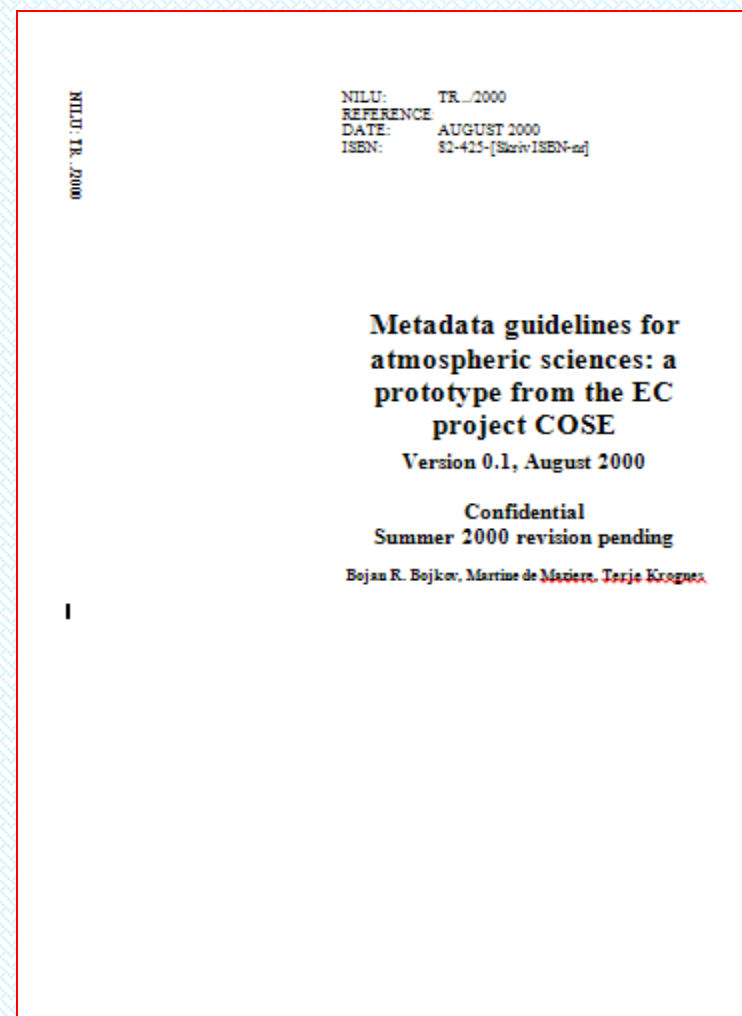
Ripples indicate approximate vertical resolution.

Total Column

Vertical Profiles

NDACC for (satellite) validation

- Need for harmonised metadata and formatting guidelines came up in the **EU FP4 project COSE** (Compilation of atmospheric Observations in support of Satellite measurements over Europe; **1998-2000**) and was strongly supported by the NDACC community.
- Under COSE, we developed **common formatting and metadata guidelines** that could deal with the different types of data

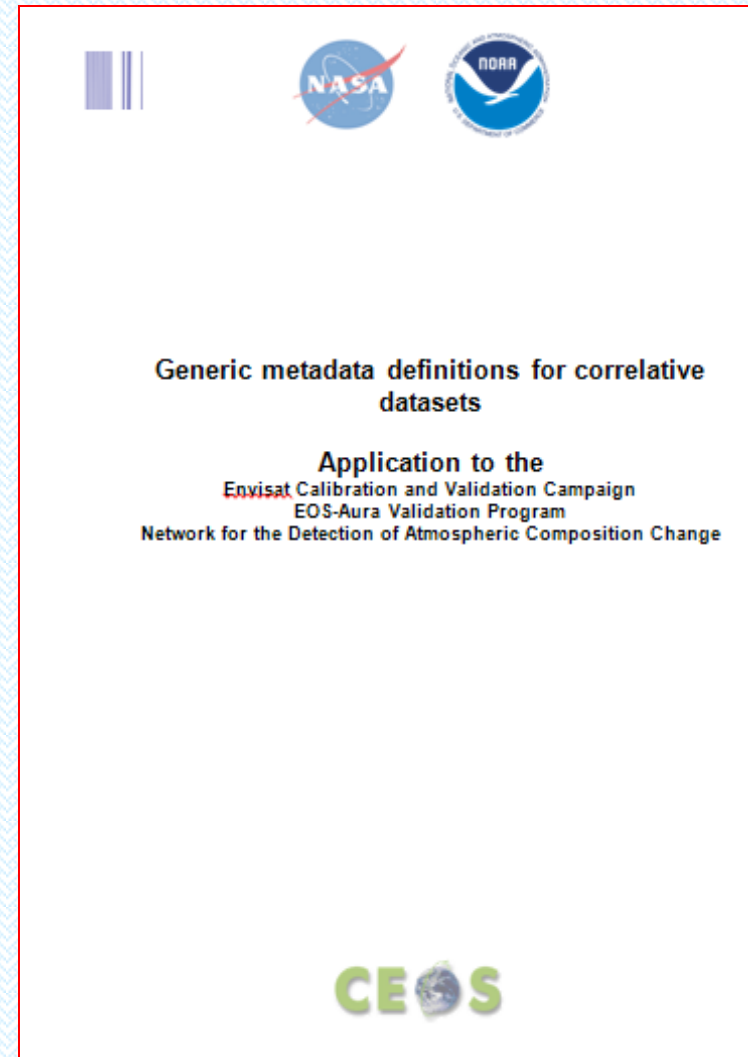


Implementation of metadata and formatting guidelines

- March 1, 2002: Launch of Envisat
- 2002: Start of the Envisat Calibration and Validation Campaign
- A dedicated validation server was set up at NILU under ESA contract, the so-called Envisat Validation Database or (later) **EVDC or Envisat Validation Data Center** (<http://nadir.nilu.no/calval/index.php>)
- The **COSE guidelines and HDF format** were adopted by ESA (and NDACC) as the standard for the validation data in the EVDC

Evolution of metadata guidelines

- NASA establishes the **Aura Validation Data Center (AVDC)** and adopts the same guidelines
- Adopted also by CEOS



Evolution of metadata guidelines

By 2007:

- it is realised that the formatting and metadata guidelines in EVDC and AVDC have evolved slightly differently
- the experience gained shows some deficiencies with the guidelines

⇒ A common effort between ESA, NASA and the NDACC community* led to the GEOMS guidelines
(harmonised and revised wrt previous versions)

* in the frame of an ESA project called 'Generic Environment for Calibration/Validation Analysis' (GECA)

to GEOMS

- Generic Earth Observation Metadata Standard (March 2011)

<http://avdc.gsfc.nasa.gov/GEOMS>



The Generic Earth Observation Metadata Standard (GEOMS)

Version 1.0
March 21, 2011

Authors

Christian Retscher, UMBC, NASA/GSFC
Martine De Mazière, BIRA-IASB
Yasjka Meijer, ESA/ESRIN
Aasmund F. Vik, NILU
Ian Boyd, NIWA
Sander Niemeijer, S&JT
Robert M. Koopman, ESA/ESRIN
Bojan Bojkov, ESA/ESRIN

Contributors

Terje Krognæs, NILU
Rolf Von Kuhlmann, ESA/ESRIN
Thorsten Fehr, ESA/ESRIN
Geoff Buswell, Logica UK Limited

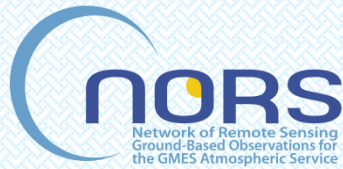
Concept of GEOMS

- GEOMS defines the **metadata and data structure requirements** developed to facilitate the use of geophysical datasets by improving their **portability and accessibility**, and by making their contents **self-describing**.
- It has been designed to **support all measurements from Earth observation instruments (atmosphere, ocean, ...)**.
- In recent years, the HDF format has become the *de-facto* satellite data exchange format for the ESA and the NASA Earth observation missions. Next to HDF, another hierarchical data format, netCDF, is extensively used in Earth observation
 - ⇒ **The GEOMS metadata guidelines are implemented using the HDF4, HDF5 or netCDF file formats, but are not limited to these formats.**

Implementation of GEOMS

- GEOMS is adopted as the standard for submission/archiving of data in EVDC, AVDC and NDACC
- NORS – as it is dealing with NDACC data – has also adopted the GEOMS standard for the metadata with the HDF file format.
- NORS data will be available from the public NDACC database

- An important outcome of NORS will be the so-called **NORS Validation Server**:
 - A Web-based engine that will automatically and interactively produce validation reports of the MACC-II products, based on comparisons with the NORS data.
- The metadata are needed in this NORS Validation Server to get all the necessary information about the data like quality flags, information about vertical resolution, uncertainties, etc.
- QA/QC procedures upon submission of the datafiles to the database will verify compliance of the NORS data with the GEOMS standard; as such, there is a guarantee that all required variables and metadata attributes are available in the datafiles

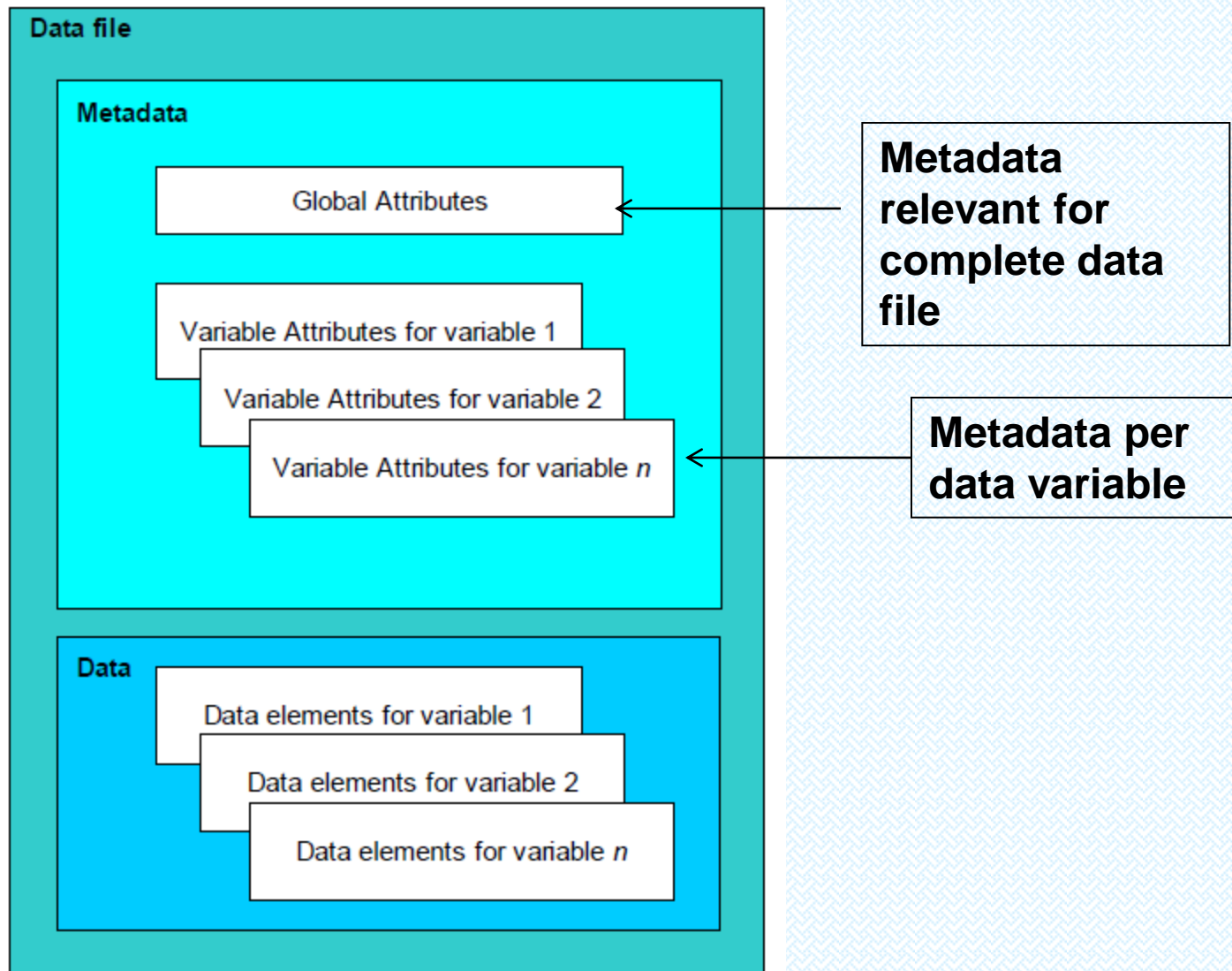


Brief presentation of GEOMS

- Generic guidelines
- Table of Attribute Values
- Templates



GEOMS data structure



Attributes

- **Global attributes**
 - Global Originator Attributes
 - Global Dataset Attributes
 - Global File Attributes
- **Variable Description Attributes**

Attributes

- Attributes have fixed names

e.g., DATA_QUALITY
VAR_UNITS

- Allowed entries for attributes are free strings or must be taken from a list of pre-defined entries from a **Table of Attribute Values (TAV)** which is kept as a living table on the AVDC Website

e.g.,

DATA_QUALITY = NRT as defined in Henne et al, 2000
VAR_UNITS = ppmv

Variables

- Variable names have a fixed syntax:

Primary parameter _ mode _ descriptor

e.g.,

O3.COLUMN_ABSORPTION.SOLAR_AVK

O3.COLUMN_ABSORPTION.SOLAR_FLAG

This naming convention is different from the netCDF-CF conventions but one could establish a one-to-one relationship

Templates

- Every data file **must** contain a specification of geolocation in **three or four dimensions**
- The geolocation variables must comply with **stringent naming criteria.**
- **The other variables to be reported in a datafile are defined in templates, agreed per instrumental technique, with the data providers and users**

e.g., FTIR template ; LIDAR template

- Templates are maintained on the AVDC Website; they have version numbers.
- The datafile contains an attribute that identifies the template with which the datafile is compliant.

Summary - 1

- The GEOMS Standard has been designed to cover all Earth Observation data (not only from instruments but also models)
 - For data exchange and archiving
 - Self-describing
- It is implemented and used since a decade (with some evolution)
- It has been adopted by the satellite validation communities (EVDC and AVDC) and NDACC, and therefore also by NORS

Summary - 2

- At present, it seems to come up to the needs of the data providers and users.
- In particular:
The fact that the NORS Validation Server finds all required information in the NORS datafiles to do an automatic or interactive validation job, demonstrates the usefulness and suitability (fit-for-purpose) of the GEOMS standard.



aeronomie.be



Global attributes: data originator

- PI_NAME
- PI_AFFILIATION
- PI_ADDRESS
- PI_EMAIL
- DO_NAME
- DO_AFFILIATION
- DO_ADDRESS
- DO_EMAIL
- DS_NAME
- DS_AFFILIATION x
- DS_ADDRESS DS_EMAIL

Global attributes: dataset attributes

- DATA_DESCRIPTION
- DATA_DISCIPLINE
- DATA_GROUP
- DATA_LOCATION
- DATA_SOURCE
- DATA_VARIABLES
- DATA_START_DATE
- DATA_STOP_DATE
- DATA_FILE_VERSION
- DATA_MODIFICATIONS
- DATA_CAVEATS
- DATA_RULES_OF_USE
- DATA_ACKNOWLEDGEMENT
- DATA_QUALITY
- DATA_TEMPLATE
- DATA_PROCESSOR

Global attributes: file attributes

- FILE_NAME
- FILE_GENERATION_DATE
- FILE_ACCESS
- FILE_PROJECT_ID
- FILE_ASSOCIATION
- FILE_META_VERSION
- FILE_DOI

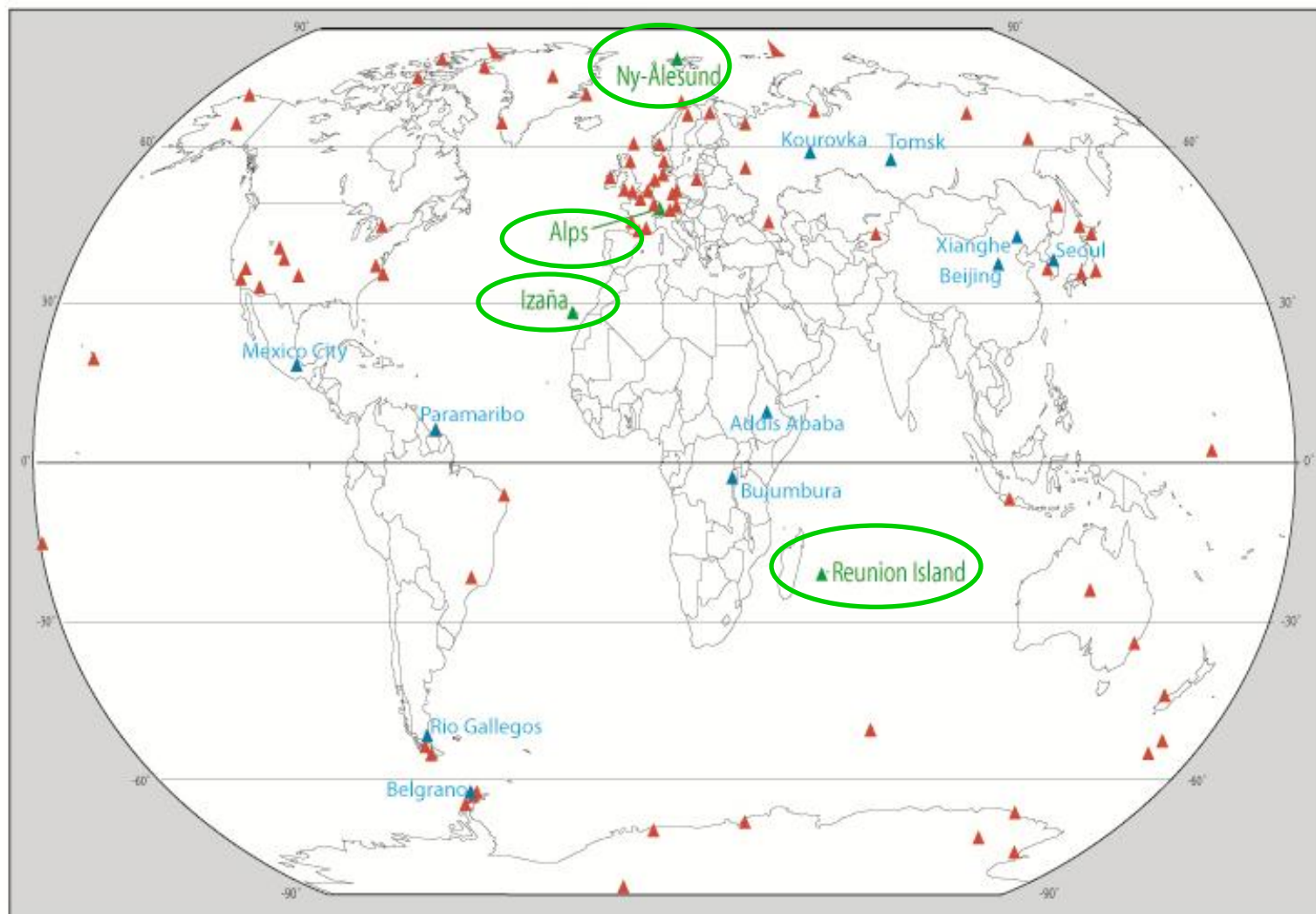
Example of FILE_NAME:

groundbased_ftir.ch4_bira.iasb002_la.reunion_
20120423t073156z_20120424t110952z_001.hdf

Variable attributes

- VAR_NAME
- VAR_DESCRIPTION
- VAR_NOTES
- VAR_SIZE
- VAR_DEPEND
- VAR_DATA_TYPE
- VAR_UNITS
- VAR_SI_CONVERSION
- VAR_VALID_MIN
- VAR_VALID_MAX
- VAR_FILL_VALUE

NORS map



- ▲ Operational NDACC stations
- ▲ NDACC stations selected as pilot stations in NORS
- ▲ Stations to be developed in NORS to potentially become NDACC stations