Customer	**:**	**ESRIN**	**Document Ref**	**:**	**BAQ-TEC-TEN-SER-032**
Contract No	**:**	**4000126749/19/I-NS**	**Issue Date**	**:**	**26 February 2020**
WP No	**:**	**4100**	**Issue**	**:**	**1.0**

Product harmonization procedure description

Abstract : This technical note describes the procedures adopted for the harmonization of products data files of BAQUNIN instrumentation, following the GEOMS guidelines.

Author : Anna Maria Iannarelli
SERCO

Approval : Angelika Dehn
ESA

Distribution :

Serco Italia SpA
Sede Operativa: Via Sciadonna, 24-26 - Frascati (Roma)
Tel: +39 06 98354400 Fax:
www.serco.com

Change History

This document shall be amended by releasing a new edition of the document in its entirety. The Amendment Record Sheet below records the history and issue status of this document.

ISSUE	DATE	REASON
1.0	29 Feb 2021	First version

INTRODUCTION

Document Objective

This document describes the procedure used to harmonize the data obtained by the remote sensing instruments, belonging to BAQUNIN Super Site.

BAQUNIN Super Site hosts several devices belonging to international networks, with their standard output format, and customised or commercial instruments for which a dedicated data analysis software, producing a customised output file, according to the needed of the instrument and scientific scope is developed. The result is a great variety of file layouts, variables names and formats, which introduce some difficulties in the comparison or dissemination of the products.

Document Scope

To standardize the file format, the BAQUNIN team developed an "ad hoc" procedure, described in this document, used to convert the native files/variables format into the format Generic Earth Observation Metadata Standard (GEOMS). The harmonisation of products, following the GEOMS guidelines, must consider the native format files, the content of these files in terms of variables and related information, and the GEOMS available template or guidelines.

Document Structure

Chapter 1 (this one) contains an overview of the Technical note objective and scope.

Chapter 2 introduces a brief description of the GEOMS metadata and structure.

Chapter 3 gives detailed information about the developed procedure, consisting of three main steps.

In particular, in Paragraph 3.1, the reading of native files is described: in this case, the characteristics of files and variables must be considered, as for example the file extension, variables name and units, and so on.

Paragraph 3.2 describes the integration/conversion of attributes and variables according to GEOMS template/guidelines. This second step requires the identification of the appropriate template for each product/instrument, or the creation of a list of attributes and variables to add in the new file to create a GEOMS compliance file.

Paragraph 3.3 describes the creation of the new harmonised file in GEOMS compliance format, containing information about the site and characteristics of the instrument, the retrieval and uncertainties/errors, and variables with standard names and units.

Appendix A gives a panoramic view of the different formats provided by international networks.

REFERENCE DOCUMENTS

The following is a list of documents with a direct bearing on the content of this report. Where referenced in the text, these are identified as RD.n, where 'n' is the number in the list below:

[RD.1] Project Management Plan, (TBD), latest applicable issue

[RD.2] geoms-1.0.pdf

[RD.3] geoms_guidelines_conventions_2.1.pdf

[RD.4] tableattrvalue_04R058.dat (Last version at the date of this document)

ACRONYMS

Acronym	Definition
AAE	Ångström exponent
AERONET	AErosol RObotic NETwork
AOD	Aerosol Optical Depth
APL	Atmospheric Physics Laboratory (at Sapienza)
ASCII	American Standard Code for Information Interchange
BAQUNIN	Boundary-layer Air Quality-analysis Using Network of INstruments
EuBrewNet	European Brewer Network
EuroSkyRad	European Skynet Radiometers network
EVDC	ESA Validation Data Centre
GEOMS	Generic Earth Observation Metadata Standard
JD	Julian Day, decimal number counting days starting from 4713/01/01 BC, 12:00:00 Universal time
MJD	Modified Julian Day, decimal number counting days starting from 1858/11/17, 00:00:00, 00:00:00 Universal Time
MJD2K	Modified Julian Day, decimal number counting days starting from 2000/01/01, 00:00:00 Universal Time, used in GEOMS format
NetCDF	Network Common Data Form
NILU	Norwegian Institute for Air Research
PI	Principal Investigator
TAV	Table of Attributes
WV	Water Vapour

GEOMS METADATA AND STRUCTURE

The Generic Earth Observation Metadata Standard (GEOMS) metadata and data structure requirements can be found at https://evdc.esa.int/documentation/geoms/. On this page, several documents outline the file format ([RD.2],[RD.3]), data structure and metadata developed to improve the portability and accessibility of geophysical datasets, making their contents self-describing.

GEOMS compliance files must contain metadata and data from the instrument measurements, analysis, and auxiliary data. Metadata are divided into global attributes and sets of variable attributes including information about location, instrument characteristics, dataset, and variables. All attributes, both mandatory and optional as their values, are listed in dedicated tables named Table of Attributes (TAV) [RD.4].

In addition to general guidelines describing type, contents and naming of variable and files, also dedicated templates are available for several instrument/products (https://evdc.esa.int/tools/data-formatting-templates/). These templates include all the information and variables of a specific couple of product/instrument required by the scientific community, as for example Aerosol/Lidar (GEOMS-TE-LIDAR-AEROSOL-005.csv), gas/Pandora (GEOMS-TE-PANDORA-DIRECTSUN-GAS-002.csv) or Wind/Sodar (GEOMS-TE-SODAR-001.csv).

BAQUNIN team cooperated with the Norwegian Institute for Air Research (NILU) EVDC team for the development of the last template, GEOMS-TE-SODAR-001.

Moreover, variables names must follow guidelines and refer to dedicated tables, as the variables attributes. Some of the parameters retrieved and used by ground-based instrumentation have not a corresponding name/value listed in the TAV. In this case, a name was created following the GEOMS guidelines. A mismatch with the TAV table can be found also for ground-based remote sensing instruments, as for the pyranometer, for which the source name used is not in TAV table. So, even if all GEOMS rules are followed, the resulting file is not GEOM compliance and it is rejected by the database, which accepts only GEOMS format files. The file containing "ad hoc" variables/names are named *like-GEOMS file*.

Firstly, during the harmonisation procedure, we look for a template dedicated to the specific data file, then to the variable/value listed in the TAV and finally, following the GEOMS guidelines, an appropriate variable name is created. The procedure is summarized in Figure 1.

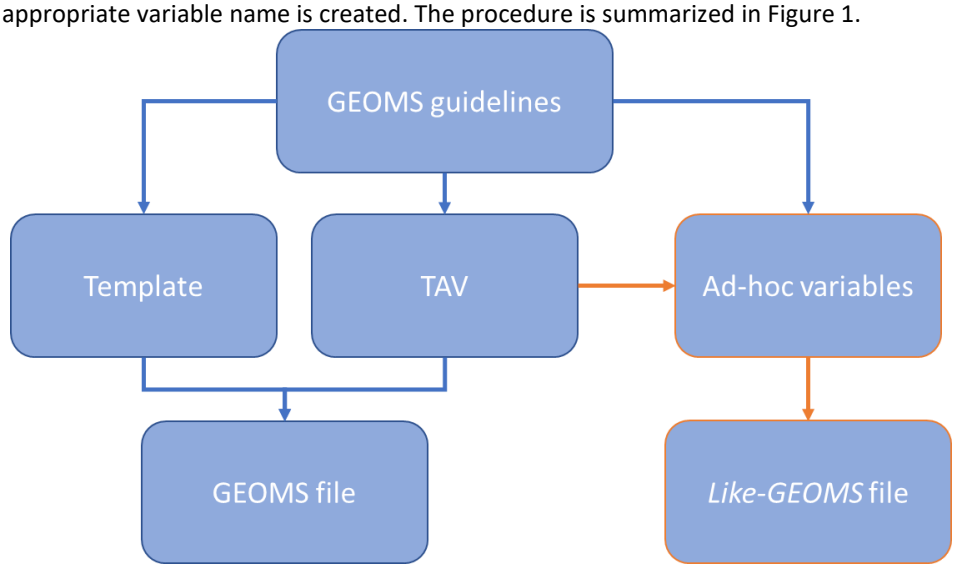

Figure 1: Schematic diagram of the procedure used to create GEOMS and like-GEOMS files.

GENERAL PROCEDURE

The first step of the procedure considers two main aspects:

- information present/absent in the native file to compile the GEOMS metadata structure
- list of variables in the native file and their formats, in terms of units and numerical/string types

In the second step the correct GEOMS template is identified (if it exits) or, in alternative, it is created following the GEOMS guidelines. In both cases, the starting point is the list of the global attributes and variables required in the NetCDF file. For this purpose, three dedicated tables are used: a) general information on the instrument and/or location, b) the list of the Global attributes, c) list of variables with the related attributes. To facilitate the compilation of the various fields, the last two tables report a brief description for each attribute/variable as described in the GEOMS documentation. The format of these tables is shown in Appendix B. These tables must be filled once for each kind of file to convert, and it is an input of the conversion software for each product/file.

The next step consists in the harmonisation of the variables, in terms of format and units. Several routines and functions have been developed to read the native format of a specific variable and convert it into the GEOMS required standard.

Finally, when the metadata are reported in the tables and all the variables are converted, the creation of the file could be completed using other “general” functions, usable for the harmonisation of the product/file of all instruments.

A dedicated software in Python programme language is created for each product/file to perform the convertion. The software includes the specific routine for reading the native file, the conversion routine allowing for the harmonisation of the parameters, and the general routine for reading the input tables (global attributes, list of variables and their attributes) and creating the new file. The scheme given in Figure 2 summarises these steps.

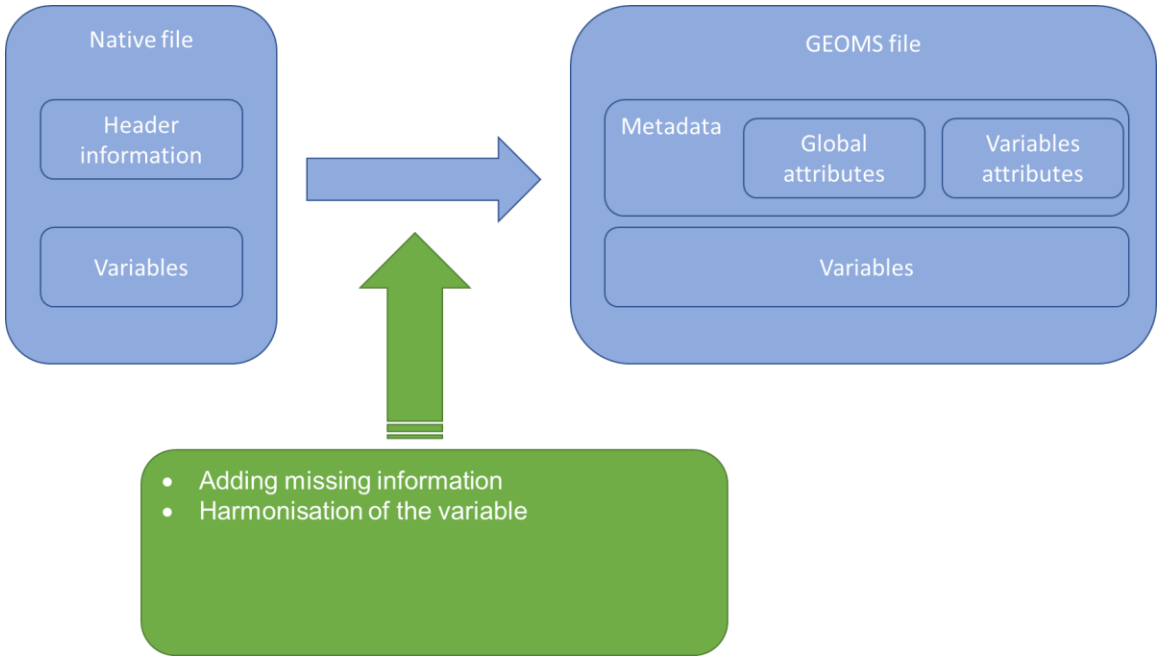

Figure 2: harmonization procedure for the creation of GEOMS files.

Native format files

Data files native can have different formats, for example data downloaded from networks (e.g., AERONET, EuroSkyRad and EuBrewNet) are in ASCII format, while files obtained by internal data analysis (i.e. retrieval performed by BAQUNIN staff) are typically in NetCDF or ASCII format. The files not always contain all the required metadata and sometimes the available information are incomplete. For instance, useful information as processor version or data level, references describing analysis procedure or variables units, location, serial number, and PI name of the instrument, are not present in the native files and must be found on the website or in the references.

Appendix A shows native data file available on AERONET (AOD, AAE and WV file), EuroSkyRad (AOD and AAE) and EUBREWNET (O_3) websites.

Each file contains a different header, reporting information about instrument and/or location, which are given in different format and are listed in different way. As example, in the AERONET files, the information about the instrument serial number and the latitude/longitude of the site, are provided as values of a parameter and repeated for each measurement, instead to be listed in the file's metadata.

Another example of the inhomogeneity of the file is the format: the same parameter can be present in each file with a different format, in one or more variables. The best example is the date/time parameter, which is reported in one or more variables for different native format files: date and time of the measurement are provided dividing time from date, and the type of parameter could be a string or a number. The Table 1 shows the possible date/time format used for the same day.

As said, the native format file could be ASCII, csv, or netCDF and could contain a different number of columns/rows (ASCII or csv format) or variables (netCDF format). For this reason, it is not possible to use the same routine to read all the files. A customised *Script* for each case was developed (in Python Programming Language), and its outputs are a structure containing the header information or metadata, and a matrix with the values of parameter.

Table 1: Example of formats used for the date/time parameter in different products.

VARIABLE TYPE	EXAMPLE	VARIABLES NUMBER	DESCRIPTION
string	20200901T063609Z	1	DATETIME representation in ISO 8601
int-int-int-float	2020 9 1 6.6025	4	Year Month Day Decimal hour
string	01:09:2020 06:36:09	2	day : month: year hours : minute : seconds
float	244.27510416693985	1	Day of the year with decimal
float	2459093.775104167	1	Julian day
float	7549.27510416694	1	Modified Julian day

Harmonisation

The start point is to consider the information given in the native file in order to add missing data. To do this, an Excel file reporting all the useful information is created [*ArchiveRef*_like-GEOMS_format_tables], this file could be used for all files with the same structure, i.e. for all files created by the same instrument/product couple. This file, containing the table showed above, is used as an intermediate step between reading and creating file routines and permits to standardise the conversion procedure.

This Excel file consists of three sheets:

- *Info_instrument* sheet contains all the information concerning data, instrument and location which are not available into the instrument native file and are valid for each data file conversion.
- *GlobalAttributes* sheet consists of two columns with the list the of the global attributes and the description of each of them, and a column with the value of each attribute. If the value of an attribute is reported in the native file, the corresponding value in this table is "*from_software*". It means that the conversion software will be able to assign the correct value read from the native file.
- *Level sheet* reports the level of the data to be converted, considering the option to convert more than one level of the analysis data file. The table contains the variable attributes (columns) and the corresponding value for each variable (rows).

The file is read using the script *Read_excel_info_GEOMS_instrument*. Not all the Global attributes have a fixed value, but it must be retrieved by the content of the file to be converted. For example, the attribute showing start and stop date depends on the single file.

The following step is the identification of the variables/parameters that must be converted, in terms of both format and unit. In this case, a dedicated Python routine has been developed and used according to the needs. As instance, the second date/time format variable shown in the Table

1 uses 4 variables to indicate the instant in which the measurement is carried out that must be merged into a unique variable named DATETIME and with units MJD2K.

Creation of the new file

This is the last step of the conversion procedure. At this stage, Global Attributes, Variable list and related attributes, Variable values are contained in three structures, with a well-defined structure defined according to the other Python routines and are ready to be saved into the new file. In what follows, details about the developed routines are given.

- Create_Name_file_GEOMS

This function is used to create the file name according with the GEOMS guideline. The input of this function is the Global attributes structure because the file name must be composed using the value of some global attributes, as listed below.

[DATA_DISCIPLINE_03]_[DATA_SOURCE]_[AFFILIATION]_[DATA_LOCATION]_

[DATA_START_DATE]_[DATA_STOP_DATE]_[DATA_FILE_VERSION]

- Create_nc

The function creates a netCDF file with extension .nc. The inputs are the complete path of the file, directory, and file name.

- Global_attribute_GEOMS

The function inputs are the file path and the structure containing the global attributes. The routine opens the created file and saves this information with the formatting required by the GEOM standard.

- Var_Dimension

The function creates the dimensions of the variables which must be explicitly defined. A variable has one or more dimensions, and the same dimension could be used for more than one variable, so this routine is used for each needed dimension.

- Def_Variable_GEOMS

The last step is the creation of the variable. The inputs of the function are the file name, the variables list, their attributes (name and value), the dimensions, and the variables value.

A. APPENDIX A

In what follows, a panoramic view of the various formats provided by the international networks is given.

AERONET file content

The file format provided by AERONET website (https://aeronet.gsfc.nasa.gov/)is an archive file (.zip) containing an ASCII file. Table A1 and A2 show the header file information and the list of variables/parameter (grouped by types) content in the file, respectively.

For each parameter listed in Table A2, the values are placed in columns as function of the time, this means that for each time exists a value for all the parameters, also for empty columns. Note that, the number of the wavelengths used by BAQUNIN Cimel is 9, the columns referring to a wavelength are 29, it means that a large number of columns contain a fill value. The total number of the column is 113 but are not present variables containing the uncertainties.

Table A1: AERONET header file information

Description	Example
Data processor version	AERONET Version 3;
Instrument Location	Rome_La_Sapienza
Data version/level	Version 3: AOD Level 1.5
Information about data cloud screening	The following data are cloud cleared and quality controls have been applied but these data may not have final calibration applied. These data may change.
PI's name and contact	Contact: PI=Philippe_Goloub_and_AnnaMaria_Iannarelli_and_Monica_Campanelli; PI Email=philippe.goloub@univ- lille1.fr_ and_annamaria.iannarelli@serco.com_and_m.campanelli@isac.cnr.it
Reference on Points and units	All Points,UNITS can be found at,,, https://aeronet.gsfc.nasa.gov/new_web/units.html

Table A1: AERONET parameters and variables information

	Variables	Lists
1	Date/Time variables	Date(dd:mm:yyyy),Time(hh:mm:ss),Day_of_Year,Day_of_Year(Fraction),
2	AOD [wavelenghts]	AOD_1640nm,AOD_1020nm,AOD_870nm,AOD_865nm,AOD_779nm,AOD_675nm,AOD_667nm,AOD_620nm,AOD_560nm,AOD_555nm,AOD_551nm,AOD_532nm,AOD_531nm,AOD_510nm,AOD_500nm,AOD_490nm,AOD_443nm,AOD_440nm,AOD_412nm,AOD_400nm,AOD_380nm,AOD_340nm,AOD_681nm,AOD_709nm,AOD_Empty,AOD_Empty,AOD_Empty,AOD_Empty,AOD_Empty

	Variables	Lists
3	Ancillary variables related to data analysis	Triplet_Variability_1640,Triplet_Variability_1020,Triplet_Variability_870,Triplet_Variability_865,Triplet_Variability_779,Triplet_Variability_675,Triplet_Variability_667,Triplet_Variability_620,Triplet_Variability_560,Triplet_Variability_555,Triplet_Variability_551,Triplet_Variability_532,Triplet_Variability_531,Triplet_Variability_510,Triplet_Variability_500,Triplet_Variability_490,Triplet_Variability_443,Triplet_Variability_440,Triplet_Variability_412,Triplet_Variability_400,Triplet_Variability_380,Triplet_Variability_340,Triplet_Variability_Precipitable_Water(cm),Triplet_Variability_681,Triplet_Variability_709,Triplet_Variability_AOD_Empty,Triplet_Variability_AOD_Empty,Triplet_Variability_AOD_Empty,Triplet_Variability_AOD_Empty,Triplet_Variability_AOD_Empty
4	Angstrom Exponent	440-870_Angstrom_Exponent,380-500_Angstrom_Exponent,440-675_Angstrom_Exponent,500-870_Angstrom_Exponent,340-440_Angstrom_Exponent,440-675_Angstrom_Exponent[Polar]
5	Data analysis information	Data_Quality_Level Last_Date_Processed
6	Instrument information	AERONET_Instrument_Number Number_of_Wavelengths, Exact_Wavelengths_of_AOD(um)_1640nm,Exact_Wavelengths_of_AOD(um)_1020nm,Exact_Wavelengths_of_AOD(um)_870nm,Exact_Wavelengths_of_AOD(um)_865nm,Exact_Wavelengths_of_AOD(um)_779nm,Exact_Wavelengths_of_AOD(um)_675nm,Exact_Wavelengths_of_AOD(um)_667nm,Exact_Wavelengths_of_AOD(um)_620nm,Exact_Wavelengths_of_AOD(um)_560nm,Exact_Wavelengths_of_AOD(um)_555nm,Exact_Wavelengths_of_AOD(um)_551nm,Exact_Wavelengths_of_AOD(um)_532nm,Exact_Wavelengths_of_AOD(um)_531nm,Exact_Wavelengths_of_AOD(um)_510nm,Exact_Wavelengths_of_AOD(um)_500nm,Exact_Wavelengths_of_AOD(um)_490nm,Exact_Wavelengths_of_AOD(um)_443nm,Exact_Wavelengths_of_AOD(um)_440nm,Exact_Wavelengths_of_AOD(um)_412nm,Exact_Wavelengths_of_AOD(um)_400nm,Exact_Wavelengths_of_AOD(um)_380nm,Exact_Wavelengths_of_AOD(um)_340nm,Exact_Wavelengths_of_PW(um)_935nm,Exact_Wavelengths_of_AOD(um)_681nm,Exact_Wavelengths_of_AOD(um)_709nm,Exact_Wavelengths_of_AOD(um)_Empty,Exact_Wavelengths_of_AOD(um)_Empty,Exact_Wavelengths_of_AOD(um)_Empty,Exact_Wavelengths_of_AOD(um)_Empty,Exact_Wavelengths_of_AOD(um)_Empty
7	Instrument location	AERONET_Site_Name Site_Latitude(Degrees) Site_Longitude(Degrees) Site_Elevation(m)
8	Ancillary parameters	Solar_Zenith_Angle(Degrees) Optical_Air_Mass Sensor_Temperature(Degrees_C)
9	Additional parameter retrieved	Precipitable_Water(cm)
10	Auxiliary data	Ozone(Dobson) NO2(Dobson)

EuroSkyRad file content

The file downloaded by EuroSkyRad website is a double archive file (.tgz/ .tar), containing an ASCII file. The header of the file lists the wavelengths used by the instrument and the parameters (as columns inside the file). No variables describing uncertainties are present. Information regarding site, instrument, units, or format are not available in the file, and must be found in the network website at the link http://www.euroskyrad.net/ in the Metodology or References pages.

For example, the typical header is:

#WL(um)= 0.340 0.400 0.500 0.675 0.870 1.020

#time(utc) (aod x wl), alpha beta year month day

EuBrewNet file content

The file downloadable by EuBrewNet website selecting the site location of the instrument (http://www.eubrewnet.org/eubrewnet/default/index). It is an archive file (.zip), containing an ASCII file. The content of the header file and the list of parameters and variables are shown in Tables A3 and A4, respectively. In this case, some information regarding instrument and location are included in the parameters, while other useful data are available on the website, into the page dedicated to the specific instrument/location.

Table A2: EuBrewNet header file information

Description	Example
Product	ozone_product_1_5
Level	level1.5
Date	September, 2020
Process Date	2020-09-21
DATA - Usage and Guidelines	Notice to users; Recommended guidelines for data use and publication; Using EUBREWNET data; Publishing EUBREWNET data from a 'few' sites
Instrument last configuration / calibration	Config: http://rbcce.aemet.es/eubrewnet/data/get/ConfigbyId?id=xxx Date:2019-07-09, id = 1337;psw prot

Table A3: EuBrewNet parameters and variables information

	Parameter	Description
1	brewerid	Brewer identification number (Brewerid)
2	gmt	UT time of the measure in ISO 8601 format (GMT)
3	configid	Configuration identification number (Configid)
4	n_sum	Index of daily summary (Index)
5	date_index	Continuous date index (0.0 = 0000-00-00T00
6	sza	Solar zenith angle using time function (Degrees)
7	airmass	Calculated airmass using time function (Airmass)
8	temperature	Instrument temperature (C Degrees)

9	filt	Applied neutral density filter (Filt)
10	o3	Calculated Ozone value with algorithm version 2 + SL correction (DU) (DU)
11	std_o3	Standard deviation of the group of measures (DU)
12	so2	Calculated so2 value with Standard algorithm + attenuation filter correction (DU) (DU)
13	std_so2	Standard deviation of the group of measures (DU)
14	latitude	Latitude of the Brewer Location (Degrees)
15	longitude	Longitude of the Brewer Location (Degrees)
16	press	Medium Pressure of the Brewer Location (hP) (Milibars)
17	r6	Standard Lamp R6 ratio (DU) (DU)
18	r6_ref	Standard Lamp R6 reference ratio (DU) (DU)
19	r5	Standard Lamp R5 ratio (DU) (DU)
20	r5_ref	Standard Lamp Reference R5 ratio (DU) (DU)
21	o3_l1	O3 Level 1 (DU) (DU)
22	so2_l1	SO2 Level 1 (DU) (DU)
23	filter_flag	APPLIED FLAGS
24	correction_flag	APPLIED CORRECTIONS
25	configdate	UT Date of the used configuration in ISO 8601 format (GMT)
26	configtype	Type of the used configuration (B header(1), ICF (2), Config (3))
27	process_date	UT process time of the product in ISO 8601 format (process GMT)

B. APPENDIX B

Tables for Global Attributes

In this table, the global attributes required by GEOMS format file are listed. Two columns contain the *Name* and a brief *Description* of the attribute, as reported by GEOMS guideline documents [RD.2]. The *Value* column must be filled in according to the instrument/product file to convert and to the TAV [RD.4].

	Name	Value	Description
1	PI_NAME		The global attribute PI_NAME contains PI name of the instrument.
2	PI_AFFILIATION		The global attribute PI_AFFILIATION contains the Principal Investigator's official affiliation name preferably in English and affiliation acronym.
3	PI_ADDRESS		The global attribute PI_ADDRESS contains the Principal Investigator's official mailing address.
4	PI_EMAIL		The global attribute PI_EMAIL is the Principal Investigator's e-mail address.
5	DO_NAME		The global attribute DO_NAME contains the Data Originator's (DO) Name. The DO is the person that generated, and quality controlled the data. The DO may or may not be the same person as the PI.
6	DO_AFFILIATION		The global attribute DO_AFFILIATION contains the Data Originator's official affiliation and acronym. The DO_AFFILIATION may differ from the PI_AFFILIATION
7	DO_ADDRESS		The global attribute DO_ADDRESS contains the Data Originator's mailing address. The country name must be an official short name entry in English as listed in ISO 3166-1 [14]. The DO_ADDRESS may differ from the PI_ADDRESS.
8	DO_EMAIL		The global attribute DO_EMAIL contains the Data Originator's e-mail address. The DO_EMAIL may differ from the PI_EMAIL.
9	DS_NAME		The global attribute DS_NAME contains the Data Submitter's (DS) name. The Data Submitter is the person that submitted the data to the data centre.
10	DS_AFFILIATION		The global attribute DS_AFFILIATION contains the Data Submitter's official affiliation and acronym. The DS_AFFILIATION may differ from the PI_AFFILIATION and DO_AFFILIATION
11	DS_ADDRESS		The global attribute DS_ADDRESS contains the mailing address of the Data Submitter. The country name must be an official short name entry in English as listed in ISO 3166-1 [14]. The DS_ADDRESS may differ from the PI_ADDRESS and DO_ADDRESS.
12	DS_EMAIL		The global attribute DS_EMAIL contains the Data Submitter's e-mail address. The DO_EMAIL may differ from the PI_EMAIL and the DO_EMAIL.
13	DATA_DESCRIPTION		The optional global attribute DATA_DESCRIPTION contains a brief sentence summarizing the file's data content.
14	DATA_DISCIPLINE		The global attribute DATA_DISCIPLINE describes the field of research to which the data in the file belongs and the data acquisition method. An entry consists of three fields: Discipline field, Acquisition method, and Acquisition platform.
15	DATA_GROUP		The global attribute DATA_GROUP has a two-field entry, specifying the origin of the data (experimental, model, or a combination of both) and the spatial characteristics of the data.

16	DATA_LOCATION		The global attribute DATA_LOCATION contains the identification of the location of the reported geophysical quantities. In general, DATA_LOCATION identifies the fixed location of an instrument, such as a station name; or of a moving platform such as a plane, a ship, a satellite, from which the data originated.
17	DATA_SOURCE		The global attribute DATA_SOURCE consists of two underscore separated fields. The first field describes the type of instrument or numerical model that created the data. The second field is the acronym of the institute/organization that owns the instrument/model.
18	DATA_VARIABLES		The global attribute DATA_VARIABLES lists all data variables reported in the current data file. The list is a succession of fields in the DATA_VARIABLES entry; in other words, the entry consists of one field per variable. Generally, a field consists of the variable name, the variable mode, and the variable descriptor, separated by underscores.
19	DATA_START_DATE		The global attribute DATA_START_DATE specifies the earliest date/time found in the data file. The DATA_START_DATE format is the ISO 8601 data/time described in Sections 3.3.1 and 3.3.2. Fractions of a second shall be rounded down to the nearest second.
20	DATA_STOP_DATE		The global attribute DATA_STOP_DATE specifies the latest date/time found in the data file. The DATA_STOP_DATE format is the ISO 8601 data/time described in Sections 3.3.1 and 3.3.2. Fractions of a second shall be rounded up to the nearest second.
21	DATA_FILE_VERSION		The global attribute DATA_FILE_VERSION specifies the version of the data. It is not associated with a scientific algorithm or a processing algorithm, the attribute entry specifies an arbitrary version of the file, beginning with 001 (with leading zeroes). With each update the data file version shall be incremented by 1.
22	DATA_MODIFICATIONS		The optional global attribute DATA_MODIFICATIONS is intended to describe the data modification history associated with DATA_FILE_VERSION found in the data file. Detail of the information is up to the discretionof the data originator.
23	DATA_CAVEATS		The optional global attribute DATA_CAVEATS refers to potential issues with the data in the current data file and shall inform the user to use this data with caution.
24	DATA_RULES_OF_USE		The optional global attribute DATA_RULES_OF_USE entry is the PI's guidelines for the data usage. This entry is usually guided through a specific project data protocol.
25	DATA_ACKNOWLEDGEMENT		The optional global attribute DATA_ACKNOWLEDGEMENT specifies the PI's "desired" acknowledgment of the data when used in publications, presentations, etc
26	DATA_QUALITY		The global attribute DATA_QUALITY specifies information on quality of the data. This attribute is mandatory for datasets which follow a GEOMS template description, otherwise the use is optional. Currently GEOMS does not regulate the content of DATA_QUALITY, thus the data quality classification is to be defined by the data provider.
27	DATA_TEMPLATE		The global attribute DATA_TEMPLATE specifies information on applicable templates for reported data. This attribute is mandatory for datasets which follow a GEOMS template description, otherwise the use is optional.
28	DATA_PROCESSOR		The optional global attribute DATA_PROCESSOR specifies information on the data processor and the data processor version in a single element field. DATA_PROCESSOR should only be used for concise information on the name and version, while DATA_DESCRIPTION can be used to explain more details on the data processor.
29	FILE_NAME		The global attribute FILE_NAME contains the current data file name. This entry must be identical to the filename in the data archive.
30	FILE_GENERATION_DATE		The global attribute FILE_GENERATION_DATE specifies the date/time of generation of the current file. The FILE_GENERATION_DATE format is the ISO 8601 date/time

31	FILE_ACCESS		The global attribute FILE_ACCESS has a multi-field character string entry referring to the file project association in the data archive. FILE_ACCESS is used to define the file's access rights through data centre interfaces.
32	FILE_META_VERSION		The global attribute FILE_META_VERSION indicates the version of the metadata definitions used in the data file and the tool name used to generate the current HDF or netCDF data file.
33	FILE_DOI		The global attribute FILE_DOI indicates the use of a digital object identifier. The field is currently mandatory, but with empty variable value. A specification of the DOI syntax will be given in a future version of GEOMS.
34	LOCATION_COUNTRY		*Attributes not required*

Table for variables

In this table, the list of variable attributes required by GEOMS is reported. Two columns contain the *Name* and a brief *Description* of the attribute, as reported by GEOMS guideline documents. The *Value* column must be filled in with the information of the variable present in the TAV [RD.4]. For each variable saved into the file, this kind of inputs must be given.

	Name	Value	Description
1	VAR_NAME		Variable name[_Variable mode][_Variable descriptor]
2	VAR_DESCRIPTION		The variable attribute VAR_DESCRIPTION provides a brief description of the variable. This entry should clearly identify the variable's meaning, either explicitly, or by reference to a readily accessible reference document.
3	VAR_NOTES		The optional variable attribute VAR_NOTES conveys any additional information pertinent to the reported variable.
4	VAR_SIZE		The variable attribute VAR_SIZE has a multiple field entry containing the sizes of each dimension of the variable.
5	VAR_DEPEND		The variable attribute VAR_DEPEND lists all independent variables on which the current variable depends. The number of independent variables listed **must be equal to the number of** semicolon-separated fields in VAR_SIZE, and the order in which the variables are listed must correspond to the order in which their sizes are given in VAR_SIZE
6	VAR_DATA_TYPE		The variable attribute VAR_DATA_TYPE specifies the binary representation ("numerical" type) of the data associated with the variable. The used floating-point representations follow the IEEE 754 standard for floating point arithmetic.
7	VAR_UNITS		The variable attribute VAR_UNITS specifies the actual units in which the variable data are reported in the current data file.
8	VAR_SI_CONVERSION		The variable attribute VAR_SI_CONVERSION provides the formula used to convert the data specified in VAR_UNITS to the equivalent data in the corresponding "base units"
9	VAR_VALID_MIN		The variable attribute VAR_VALID_MIN indicates the valid minimum or detection limit of the data variable. The value of VAR_VALID_MIN must be reported in the units specified in VAR_UNITS.
10	VAR_VALID_MAX		The variable attribute VAR_VALID_MAX indicates the valid maximum or saturation limit of the data variable. The value of VAR_VALID_MAX must be reported in the units specified in VAR_UNITS.
11	VAR_FILL_VALUE		The variable attribute VAR_FILL_VALUE is a number, or an empty string value inserted as a substitute data element, if a data element of a variable is reported as a default or missing value. For variables with numeric VAR_DATA_TYPE, the VAR_FILL_VALUE can take any value allowed by the numeric type.

End of Document