NILU: TR 05/2006

NILU: TR 05/2006
REFERENCE: O-103045
DATE: APRIL 2006
ISBN: 82-425-1746-0

ESA Campaign Database (CDB) User Manual

Aasmund Fahre Vik, Terje Krognes, Sam-Erik Walker, Sjur Bjørndalsæter, Christoffer Stoll, Trygve Bårde, Roland Paltiel and Bjørn Gloslie

Preface

The ESA Cal/Val database was developed and implemented at NILU to provide ENVISAT scientist with a common framework and repository for exchange of correlative data, mainly from ground based measurements. The experience from this activity led to a new ESA initiative to develop a more general database, the ESA Campaign Database (CDB). This system is a generalisation and further development of the Cal/Val system used for some ENVISAT calibration and validation campaigns. We have tried to keep the differences to a minimum, to make the transition easy for the user community of the original system. The CDB includes all data and metadata definitions from the previous Cal/Val data centre, but is able to handle data from all ESA campaigns. It is a system for storing and indexing complex data sets from a multitude of sciences, and is no longer a database for correlative data only. Addition of new functionality or redesign of existing components will be an evolutionary process in co-operation with ESA and user representatives. The first step in this process was to accommodate data from 3 pilot campaigns, ESAG02, LARA and DAISEX. Later on, data from several other ESA sponsored campaigns such as SPARC, SEN2FLEX and BACCHUS has been added. The data centre has furthermore been used as a repository for older data sets from the EMAC campaigns of the early nineties.

Contents

ESA Campaign Database (CDB) User Manual

1. Introduction

The current document is meant as an introduction to new users of the ESA Campaign Database (CDB). It gives a brief description to CDB, on how the system is built up, how metadata are described, and what functionality that is provided to the users. It furthermore provides users of CDB data a description on how to use the centre. Finally a detailed description of the CDB metadata definitions is given as an appendix.

The ESA Campaign Database is available at http://nadir.nilu.no/cdb. A user account is required to enter the restricted part of the data centre. The account is personal and will give a user access to data from one or more campaigns. A user registered as data submitter (DS) is allowed to upload data for one or more campaigns and sub-projects, and can browse and download data from other users under the same campaigns. To become a user of CDB you need to sign a data protocol and be registered in the system metadata. Please contact NILU to get this protocol form or download it from the CDB web pages (open area).

In addition to the information and help provided in this document, users of the data centre are encouraged to use the help-desk provided by NILU. The institute has several trained technical and scientific persons ready to assists the users of CDB with any kind of problem. Please feel free to contact the NILU CDB team:

Email: calvalteam@nilu.no
Fax: +47 63 89 80 50
Tlf: +47 63 89 80 00

In case you wish to call us, please ask the switchboard for any of the authors of this document.

2. Description of CDB

2.1. Description of metadata

Metadata are in fact data about data. They provide the information that the data-user needs in order to understand the actual data. For an atmospheric observation, the data can be a series of numbers that does not make sense unless you provide the metadata on what the numbers represent. Typical metadata in this case would be time and location of the measurement, what parameter is measured, what is the uncertainty in the measurement, who did the measurement, what unit is used, etc.

In the ESA project on calibration and validation of ENVISAT (Cal/Val), a comprehensive effort was put down into developing a structure for defining such metadata. The structure was based on previous developments at NILU, mainly through the experiences gained from the EMEP database that has been operative since 1979. The structure for the Cal/Val database specifies all the metadata parameters that are needed for each data file. Table 1 shows the complete list of metadata parameters used in HDF files at the ENVISAT Cal/Val data centre. The structure is very flexible and is designed to store most types of measurements. The first entry in the table is in fact 12 different parameters that are used to identify the owners of the file. The Variable description and Visualisation attributes must be separately declared for each variable, while the other attributes only occurs once in the file. The metadata system is described in more detail in the Appendix A of this document.

Table 1: Metadata parameters used in the CDB database.

Originator Attributes
PI, DO, DS with NAME, AFFILATION, ADDRESS and EMAIL
Dataset Attributes
DATA_DESCRIPTION
DATA_DISCIPLINE
DATA_GROUP
DATA_LOCATION
DATA_SOURCE
DATA_TYPE
DATA_VARIABLES
DATA_START_DATE
DATA_FILE_VERSION
DATA_MODIFICATIONS
DATA_CAVEATS
DATA_RULES_OF_USE
DATA_ACKNOWLEDGEMENT
File Attributes
FILE_NAME
FILE_GENERATION_DATE
FILE_ACCESS
FILE_PROJECT_ID
FILE_ASSOCIATION
FILE_META_VERSION
Variable Description Attributes
VAR_NAME
VAR_DESCRIPTION
VAR_NOTES
VAR_DIMENSION
VAR_SIZE
VAR_DEPEND
VAR_DATA_TYPE
VAR_UNITS
VAR_SI_CONVERSION
VAR_VALID_MIN
VAR_VALID_MAX
VAR_MONOTONE
VAR_AVG_TYPE
VAR_FILL_VALUE
Variable Visualisation Attributes
VIS_LABEL
VIS_FORMAT
VIS_PLOT_TYPE
VIS_SCALE_TYPE
VIS_SCALE_MIN
VIS_SCALE_MAX

In addition to this structure, most metadata parameters are associated with a separate list of legal values that must be used to describe the observation. This makes it easier to store similar or related types of observations in a comparable manner. As an example, a variable containing ozone should be named O3.CONCENTRATION, and not ozone, ozone_concentration, etc. Only the legal values of metadata will be accepted by the database.

CDB builds on the efforts laid down in the Cal/Val project and reuse the same lists of legal parameters. However, new entries to these lists have been provided in order to cope with the different requirements and scope of CDB. This work is continuing as new campaigns are using the data centre. Because of this, a section of continuously updated values of the various metadata parameters are provided on the database web-interface. The lists of legal values found in appendix A of this document may therefore become outdated after some months. It is therefore advices to check the web pages (and the table.dat file that is explained later) first before requesting a change or addition to the metadata parameters.

2.2. Description of database architecture

NILU has designed and implemented a system for organizing ground based measurement data, and for retrieval of the same data by scientists that perform comparisons with measurements from the ENVISAT satellite. The work has been performed in close co-operation with ESA and with representatives of the user community. The system is complex since it entails co-operation between wide spread scientific communities that have separate and different cultures and methods. In the ESA ENVISAT Calibration/Validation effort the measurements of stratosphere physicists, modellers and mathematicians, marine biologists, and space scientists needed to be described within one common frame of reference. As the system evolved through a generalisation process into CDB, an even larger user community had to be incorporated. The system that handles this task is described in the following.

2.2.1 System components

ESA selected the HDF 4.1r3 file format for the file exchange, based on the established use of this format within ESA and some of the user groups. Main software tools have been developed in FORTRAN, IDL, CF (Cold Fusion), SQL and UNIX shell-scripts. The system uses Red Hat Linux, Apache web-server with CF server-side scripting, and a MySQL database.

Through extensive co-operation with the ESA project Officials Rob Koopman, Jolyon Martin and Remo Bianchi, the system design has been extended and adapted according to the CDB user requirements. Luis Alonso and Luis Gomez-Chova have been of particular help with defining the user requirements .

2.2.2 Detailed system description

The system components are here described in a logical order when we follow a data file as it passes from the originator into the storage and forward to an end user. Figure 1 shows a schematic diagram of how the various components are connected.

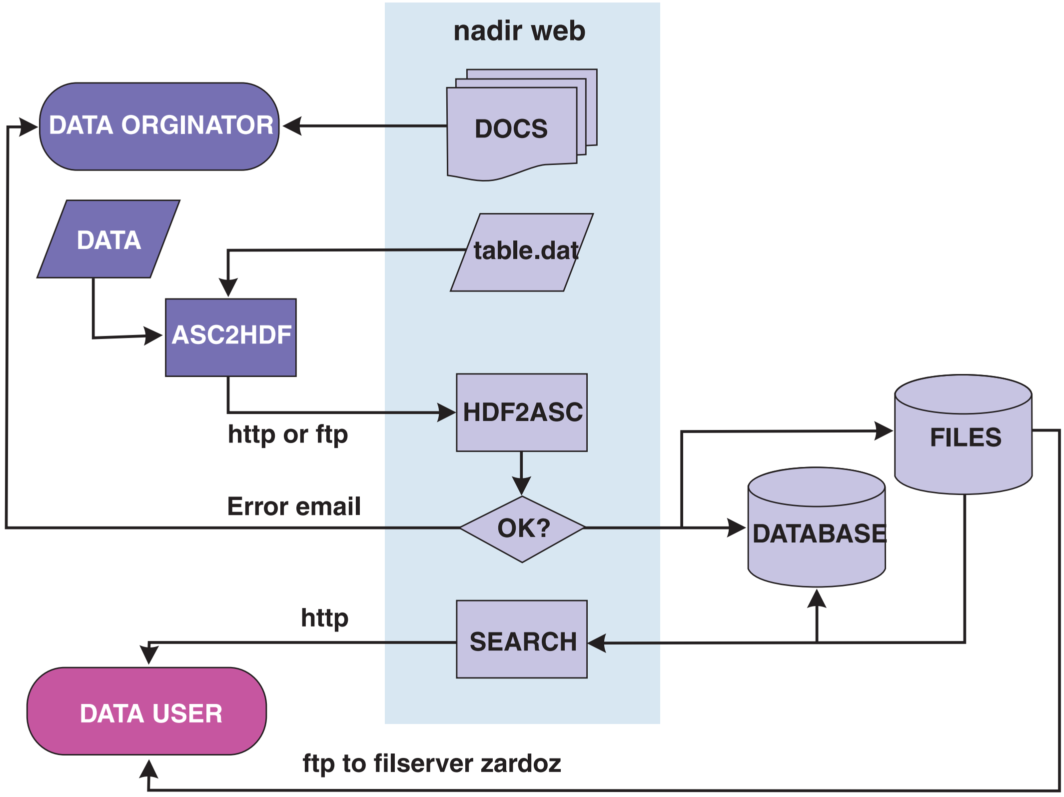

Figure 1: Schematic diagram of data flow from data originator (dark purple modules) to file collection and index database and back to data user (pink module).

The DS (Data Submitter) needs to sign a data protocol and be registered in the system metadata. This allows the user access to the CDB web site, and is also used to give permission to upload data for one or more campaigns and sub-projects (AO's) if requested.

At the CDB website (**http://nadir.nilu.no/cdb/restricted/**) the user will find a database manual (this document), file templates and other documents that help with formatting original data into an HDF file. A software tool named ASC2HDF is available for Windows, Linux, Solaris and HPUX users. This tool accepts data and metadata in two simple text files, and will generate an HDF file after extensively testing the input. A special Excel tool has been developed that allows users to save data directly from the spreadsheet into a correctly formatted HDF file. Both the ASC2HDF program and the Excel tool that work on top of ASC2HDF are described in separate user manuals, also available at the CDB website. One important input file is the so-called table.dat, which contains up-to-date information on all legal values in each metadata field. Whenever metadata are updated at the central site, a new version of table.dat is posted on the web.

When the HDF file has been successfully tested at the local site, it may be uploaded to the CDB site by ftp (**/nadir/esa/incoming**), or through a web upload page (http://nadir.nilu.no/cdb/restricted/index.cfm?fa=upload.main). A set of UNIX shell scripts is started every 5 minutes. These scripts check for new files in the incoming directory, and process each file by launching a FORTRAN program named HDF2ASC, which fetches its business rules and allowed metadata values from the same table.dat file that is published for local use by the data suppliers.

Even files that have been successfully tested by the originator, may be rejected at NILU, mostly due to inconsistencies in the file name (which reflects a subset of the metadata content), or due to duplicate file names or out-of-sequence version numbers. If the data supplier is not accredited for the campaign or sub-project listed in the file, the file will also be rejected. An error report will automatically be emailed to the data supplier and the owner of the logon name that was used, and the file will be moved to a hidden directory.

If all checks out correctly, the received HDF file will be moved to a storage file tree starting at /nadir/esa/data/, and the file name, upload details and central metadata elements are stored in an index database. The system enforces consistent naming of variables and other metadata elements, and consistent spelling of names for people, organisations and sites.

The index database contains the official list of allowed metadata values in the CDB HDF data files, in addition to logs of uploaded/downloaded files, an overview of metadata contents, and the variable list of all accepted HDF data files. All this information is available to dynamic web pages at the web site. The main end-user tools on this site are the pages for data search (http://nadir.nilu.no/cdb/restricted/index.cfm?fa=search.main**)** which allows the user to sort through the data files with advanced criteria selections. Filtering by data supplier, project, location, data source, data type, component and other metadata elements is supported. Data files may also be filtered by a “4-D box algorithm” (any file with data relevant for a given geographical area, altitude/ depth-range and time-frame). Furthermore, files can be filtered by submission date and update status.

All data files that match the search criteria are listed in a new web page, with links to HDF data file download, to comments, and to a variable list. In the variable list page the user may select variables and generate an on-line plot. In the file list the user may also select multiple files for download as a tar-ball. The user may save the search criteria in the index database for convenient re-use at a later time. Metadata content in various files can be browsed on-line and plots can be generated to visualise data content. An example of this is shown in Figure 2.

In addition to the described search interface a graphic interface showing the geographical location of the data may be used for searching. This is a webpage with a Macromedia Flash MX application showing a world map with information on geo-location. The map works in two modes, the Station and the Trajectory mode, and the user may zoom into the map to study details. The station mode displays data where geo-location is constant within a data file and the trajectory mode displays data where geo-location varies over time, but is constant for each time step in the file. The trajectory mode furthermore displays the altitude of the location by a colour scheme. A user may retrieve data by clicking on a dot (in the station mode) or on a trajectory. Within the flash-application a filter-tool with drop-down menus similar to those of the already described text-based search interface is available, and the user may chose one or several parameters. The map in the background will automatically be updated when the user chooses a value for one of the parameters and dots or trajectories not matching the chosen values will vanish. This allows the user to differentiate between all available data files and the user will see the location of all data files before he/she press the “Get file(s)” button at the bottom of the filter. It is also possible to click on any of the remaining dots/trajectories to retrieve data from only that parcel or location. A screen shot of the mapping tool is shown in Figure 3.

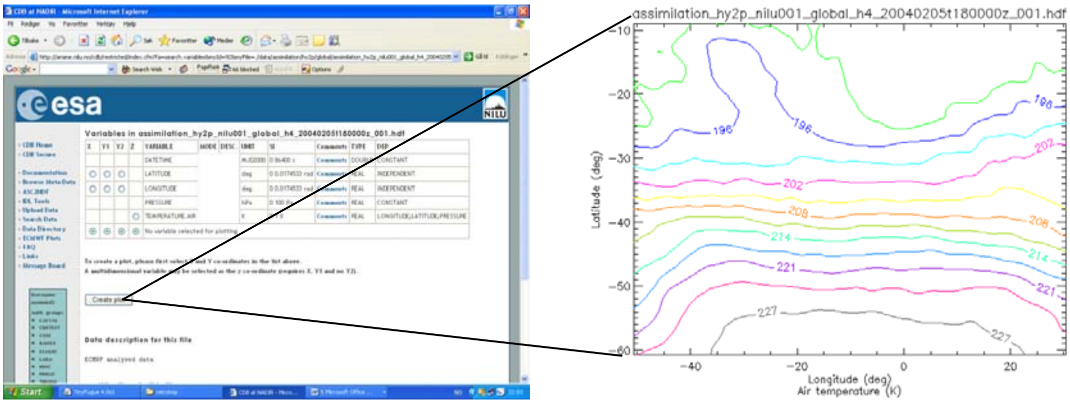

Figure 2: A screenshot from the CDB web pages. Metadata on variables can be browsed on-line, and data may be visualised through built-in plotting routines.

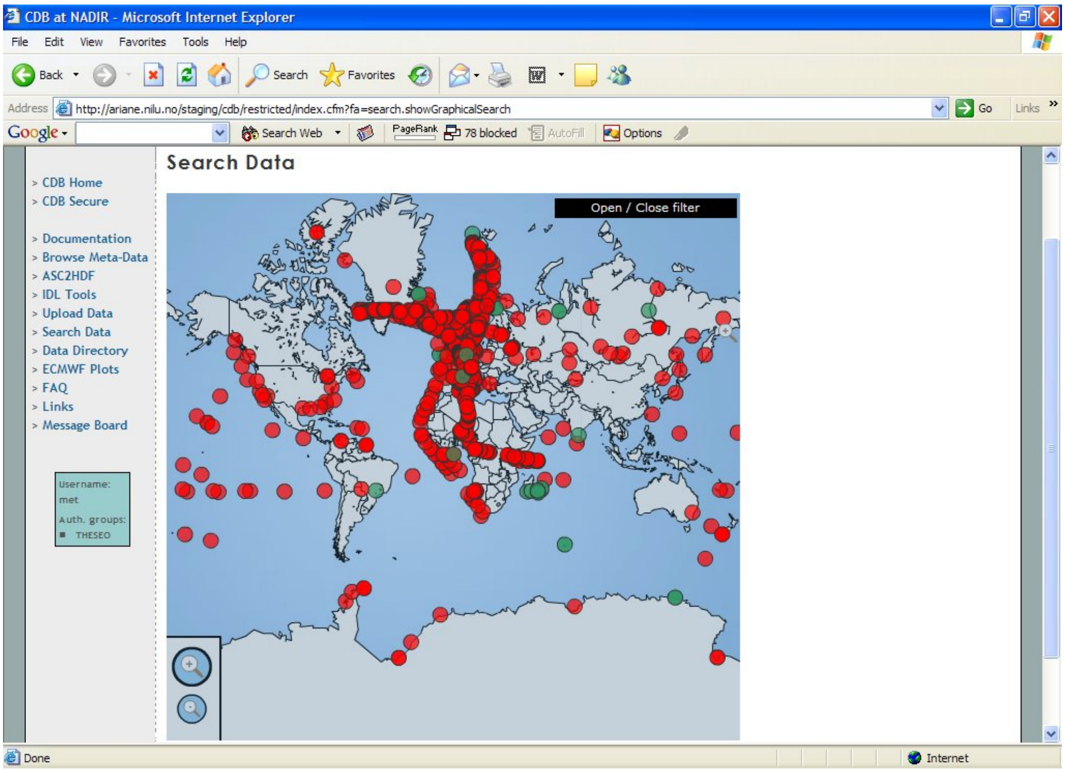

Figure 3: Screenshot of the mapping tool in the station mode. The user sees the location of all files in the database, but will only have access to those marked with green.

Through the construction of the ENVISAT Cal/Val data centre, the possibility of storing data from several different campaigns was considered. The system was therefore designed to archive files from multiple campaigns, and each file is therefore associated with a FILE_ACCESS parameter that declares which campaigns have access to the file. When the web interface decides whether a user should get a match on the file in a search, the FILE_ACCESS variables are used. This functionality was crucial to the generalisation of the data centre into CDB, where several campaigns are included, and cross-linking and shared ownership of data between different projects is likely to occur. The CDB web pages show which campaigns you have been given access to in a box in the lower left corner.

Users that have an IDL license may download IDL scripts for HDF data file formatting (excluding the detailed error checking available in the FORTRAN version) and for plotting of data sets from HDF files (http://nadir.nilu.no/cdb/restricted/index.cfm?fa=secure.idltools).

Users that have signed an additional protocol for access to ECMWF data will find pre-computed T106 ASCII extractions and plots for the last 30 days, as well as plots of isentropic and isobaric forecasts. There is an on-line tool for extracting T106 data into HDF files. There is also an on-line facility for plotting 10-day back trajectories (based on data calculated at DMI by Dr. B. Knudsen). All ECMWF products are available through a web interface (http://nadir.nilu.no/ecmwf). These met- products are mainly adapted to the stratospheric research community, but other data may be made available on users request and agreement with ESA. Please contact NILU to get this protocol.

A new feature of CDB was the implementation of Project Internal Pages (PIP), where users may share campaign specific information through a web-portal. The PIP is available at http://www.nilu.no/pip and contains sections for documents, a link archive, contact information, an image gallery and a discussion board. A user account is also needed to access these pages. Usage of the system is described in section 3.3.6.

2.3. Main differences between the CDB and the Envisat Cal/Val system

The two systems are in fact just one system that currently has separate web interfaces. The Envisat Cal/Val web pages are still maintained and kept operational to avoid unnecessary changes for existing users. Both the Cal/Val and the CDB web portals therefore use the same file processing system, data tree and index database. CDB therefore contains all the Envisat Cal/Val data and the CDB web portal has all the same functionality as the other service, but new additions and modifications have been made. These are generally not visible from the Cal/Val pages.

The main difference between the two web portals and also the main reason for creating a new web interface is that CDB is able to store a multitude of campaign datasets from a variety of scientific disciplines. The Envisat Cal/Val effort is one such campaign. In addition to this, there were needs for developing additional functionality for CDB that was not needed for the Envisat Cal/Val campaign. In addition to the previously described Excel tool for creation of HDF files this includes implementation of quick-looks and storage of image data. This allows upload and archiving of any types of images (i.e. any type of file in any file format) to the data centre. It is also possible to upload thumbnail image or quick-looks and associate this with a large image or HDF file. Both the upload of quick-looks and large image files (or other files) is done through the web-portal at the upload pages. To upload a non-HDF file the users needs to enter all metadata manually in a web form since the system is not able to extract this automatically. Quick-looks may be uploaded and associated to any file (both HDF and non-HDF files) that the specific user has uploaded before. Quick-looks and non-HDF data files are stored in directories separate from the HDF file-tree.

3. Recommendations for optimal use of the database

The previous chapter mainly dealt with what the data centre contains and how it works. The current chapter is an introduction on how to use the database and how to approach the various aspects of data handling. The amount of information you need to get started using the data from the centre is minor, while a data manager of a campaign needs to know the system in detail before he/she can start using it. To avoid overloading users with unnecessary information, we therefore divide our users into three categories: Data users, data submitters and campaign data managers. The following three sections is an introduction to the CDB with specific hints and information for each of these user groups.

3.1. Data users

As mentioned in chapter 1, a personal user-account is required to access the data centre. Such an account can be obtained by filling out the CDB data protocol. Here you may also specify which campaigns you need access for. ESA may, if necessary, approve requests for access to other campaigns than a user participates in. Once you have received your personal user-name and password, you may enter the restricted area on the CDB web pages. The URL of these pages and how they are linked with the data repository is described in chapter 2.

Data are organised in files and values are not stored in the index database. Whenever you want to retrieve a specific dataset, you will have to do this from a data file. The index database is only used to organise these data files, and makes it possible to search for data through the “search data” pages described in Section 2.2.2. The text-based search interface incorporates much of the metadata structure of Table 1 (Section 2.1) and allows the user to select data according to data location, campaign, sub-project (AO), name of persons involved in the measurement, instrument name (data source), data variables, etc. The drop-down menus of the web page are generated on the basis of what is actually stored in the database. As is mentioned on the page, it is not always possible to combine several selection criteria, and some searches may result in no matching files. This is obvious in the graphical search interface where all dots or trajectories disappears when a users selects impossible combinations. Finally, it is possible to limit the text-based search in time and space through selection of a 4-dimensional boundary box. This is not possible in the graphical search interface, but the may click directly on a dot or trajectory to get matching data from that specific station or flight-track.

A successful search will return a list of data files matching the selected criteria, and it is possible to download a single file by simply clicking on the file name. It is also possible to download all files in a zipped tar-ball. In addition to file download, the user may view file content on-line by clicking on the “variables” link to the left of the file-name. This will give the user an overview of which data variables the file contains including possible comments made for each variable. This web page also provides on-line visualisation of the data content, and the user may select values for x and y-axis to get an overview of what the file contains. It is furthermore possible to create plots with two y-variables on one x-axis and 2D, 3D or 4D variables may be visualised by selecting the z-axis. In the latter case, it is necessary to make sure you select the correct x and y-axis in order to get a sensible plot. This could be seen from the right-most column where variable dependencies are displayed.

If you choose to download files, you will need a special tool or program to read the data. As mentioned in Section 2.2, the files are archived in HDF 4.1r3 (Hierarchical Data Format), which is a binary format that allows storage of data in a structure similar to a directory tree. For the CDB, a strict implementation of the HDF format is used and files must be formatted according to the metadata guidelines as described in Section 2.1 and Appendix A. Up-to-date lists of legal values for the various parameters are furthermore available through the web portal. File and variables names are generally self-explanatory, and it is not necessary to know all the details of the metadata definitions to understand the file content. To read the file, it has become common within the scientific community to use the IDL or NOESYS software, both by RSI (http://www.rsinc.com). Another solution is to simply use the basic ncdump program that comes with the HDF libraries and that are available for several platforms through the HDF home page (http://hdf.ncsa.uiuc.edu/release4/obtain.html). This will generate text output that can be viewed directly. A simple search on the internet for an hdf browser will also provide you with a list of freeware programs, but the quality of these are not know to NILU.

As explained in section 2.2.2 CDB also allows upload of non-HDF files through the web portal. The format, data and parameter content and geolocation of these files are stored in the database and available for each file through the “variables” link in the file-list (list of files returned by the file-search pages).

As a data user it is important that you contact either NILU, if you have problems using the data centre, or any of the persons mentioned in the data files (PI, DO or DS) if you have any questions regarding specific data.

3.2. Data submitters

The data submitter (DS) is the person responsible for formatting of HDF files and upload of data to the data centre. This person should preferably be someone with a scientific knowledge of the data and is commonly the same person who performs post processing of measurements or numerical simulations. The DS has to know how to use the data centre, as explained in the previous section, but additional information is needed in order to start formatting of data files.

3.2.1 User account

As for the data user, the DS needs an account at the CDB in order to log in through the password protected web area. Furthermore, he/she needs to be assigned to a specific sub-project under the campaign in order to upload data files successfully. In the ENVISAT Cal/Val activity, these sub-projects were named AOs (AOID158, AOID320, etc.). For other campaigns we are currently also using the organisation acronym name of the DS involved as name of the sub-project. A DS will thereby only be able to upload files that are owned by his/her organisation. Upload rights for data owned by other organisations may be arranged if this is practical, and there is no technical limit to the number or sub-projects one user may have upload right for.

3.2.2 File formatting

Before the DS starts creating HDF files, it is important to contact the campaign data manager (if this is another person) to define what data that are to be archived. A guide for data managers is given in a later section (3.3). The data manager will be responsible for working out a strategy for data archiving together with NILU and should define data templates that tell the DS how to organise the data files. These templates will contain information on what parameters that should be reported in the files, how long should a timeseries be, what temporal and spatial resolution shall be used, etc. As an example, the data from the LaRA campaign (Coincidental airborne laser and Radar altimetry over ice-sheets and sea-ice) contains simultaneous measurements from two airborne instruments. In theory, it is possible to put nearly all the data into one huge HDF file, but this is not optimal for future use of the data. Instead it was decided to split the data into many smaller files with one instrument and a limited series of measurements per file. This makes it easy to search for the measurements of interest and the data are much easier to handle.

The safest way to create an HDF file is by using the ASC2HDF program developed by NILU. This is available for windows and various UNIX-based platforms and can be downloaded from the restricted area of the CDB web portal. The program takes two ASCII files as input, one data and one metadata file, and formats an HDF file according to the guidelines described in Section 2.1 The program comes as a package with precompiled binaries, documentation, example files and Fortran source code. The latter makes it possible to incorporate file formatting directly into existing user side programs. The main advantage with ASC2HDF is that you are always sure that the file is formatted correctly. It is possible to write your own HDF-converter, but it may be difficult to get everything correct and according to the strictly implemented guidelines. NILU also provides a simpler HDF formatting program for users that have a personal IDL license installed. This is called IDL2HDF and does much of the same as ASC2HDF, but it does not have the same extensive tests for data integrity. Both ASC2HDF and IDL2HDF are found on the restricted area of the CDB web portal. Detailed documentation/user manuals are also found on these pages.

3.2.3 Data file content

With the file formatting tool installed at the users computer, there is normally still a barrier to overcome before files can be produced routinely. It is advisable to study the ASC2HDF manual and to look at the example files to get an idea on how to use the program. Once the technical problem is solved, one has to consider exactly how the variables should be expressed in the file, i.e. which variables should be independent and which should be dependant. This is where the importance of a proper template comes in. When in doubt, please contact NILU and we will try to sort out the problem. If you have access to the CDB, you might also want to search for related data sets there to give you a start.

Through the CDB effort, NILU will be responsible for controlling the scientific content of incoming data. Even though the system for uploading of files to the database performs an automatic check on the data, there is always a risk that the uploaded files contain errors beyond simple formatting issues. An example is erroneous use of scales, e.g. using MPa instead of mPa when reporting pressure. Such errors will not be detected automatically, but should be caught by the scientists operating the database. When in doubt about formatting of files, please contact the NILU staff for assistance.

3.2.4 Data file upload

When the HDF file has been successfully tested at the local site by ASC2HDF, it may be uploaded to the CDB site (/nadir/esa/incoming) by ftp, or through a web upload page. A set of UNIX shell scripts is started every 5 minutes. These scripts check for new files in the incoming directory, and process each file by launching a FORTRAN program named HDF2ASC. The program verifies data and metadata according to the predefined rules in the table.dat - the same file that was used by ASC2HDF.

Even files that have been successfully tested by the originator, may be rejected at NILU, mostly due to inconsistencies in the file name (which reflects a subset of the metadata content), or due to duplicate file names or out-of-sequence version numbers. If the data supplier is not accredited for the project listed in the file, the file will also be rejected. An error report will automatically be emailed to the data supplier and the owner of the logon name that was used, and the file will be moved to a hidden directory.

If all checks out correctly, the received HDF file will be moved to a storage file tree starting at /nadir/esa/data, and the file name, upload details and central metadata elements are stored in an index database. The system enforces consistent naming of variables and other metadata elements, and consistent spelling of names for people, organisations and sites.

3.2.5 How to use the Project Internal Pages

Documentation of data is essential to obtain maximum reuse of data after the campaign is finished. For this purpose, we recommend the data submitters to report as much metadata as possible in the actual data files. The structure of Table 1 (Section 2.1) offers plenty of possibilities to report both general file metadata and specific information on each variable. Secondly, it is important also to store an overview and any additional information of the campaign data in the Project Internal Pages (PIP).

The PIP will be set up by the campaign data manager together with NILU, and will typically contain a section for documentation connected to the campaign. The PIP is located at http://www.nilu.no/pip, and a user-name and password is required to access this area. This system is physically located on a different server than the geophysical database, and the two systems do currently not share a password database. Normally, the PIP is accessed with a user-name and password that is common to all member of a specific campaign. The campaign data manager will arrange the "documents" section of the PIP into sub-folder as needed, and all the campaign member will be able to upload any type of files through this system. In addition to documents, the PIP is a natural place to put contact information and links to external key resources on the Internet. A discussion board is available and an image gallery allows users to upload pictures, plots, maps, etc. to the web pages. Please note that the user-side web browser may limit the possibility of displaying special graphic files. Files of the type gif, jpg, tiff, png, etc. are normally displayed properly in most web browsers. Users are encouraged to use the PIP before, during and after the campaign.

3.3. Campaign data managers

The campaign data manager is commonly someone working closely with the scientific campaign coordinator and will thereby have the general overview of all the data collected through the campaign or project. The main task of the data manager is to set down guidelines for reporting of data so that the data submitters know what data that should be archived and how to format their files. The campaign data managers furthermore needs to look beyond the scope of their campaign and try to see how archiving of the campaign data fit in with the objectives of CDB.

3.3.1 Objectives of CDB

The campaign database shall provide an online information system that supports users in managing and exploiting campaign datasets for Earth Observation missions and applications. In a more future perspective the overall aim is to provide a data centre that handles Cal/Val data, satellite data and campaign data in an integrated way. This type of integration will provide an add-on value to all types of measurements, as the data centre becomes a one-stop source to look for data. The centre will in this way increase the dissemination potential for all classes of data. The database is built with a strict quality control of incoming data and options for individual file-formatting is very limited. Using the same principles also for non Cal/Val data, will simplify the use of multi disciplinary data since all files are part of the same uniform data set.

CDB aims to increase the use of geophysical data after a campaign is completed. Measurements are made available for other scientists (only after permission is given from original PI) and data are no longer sent to rest in the drawer of a scientist desk. CDB provides the final archive for the data. Another advantage with using the CDB is the possibility of sharing data within the campaign consortium – both during the campaign and in the analysing phase.

3.3.2 General archiving strategy

The metadata guidelines (Appendix A) define rules for what names and values that can be inserted in a data file. Apart from this, it provides no rules for how the data structure should be defined in a file. As an example, for ozone sondes it is possible to store the ozone concentration (mPa) as a function of altitude, total pressure, time after launch, etc. A template provides the data submitter, i.e. the person responsible for creation and upload of data files, with a guideline for how he or she should define independent and dependent variables and which of these it is necessary to include. An archiving strategy is implemented in order to keep the different templates compatible, so that there is a uniform way of archiving data from different measurements and platforms. This allows for easier comparisons of different observations.

In the ENVISAT Cal/Val activity the data centre was used to store correlative data, later to be used by scientists involved in calibration and validation of some of the satellite instruments. There were clear goals for how the data should be used and a strategy for archiving data was defined based on these. This strategy advised campaign PIs to perform measurements in conjunction with a satellite overpass, and to collect data using a variety of instruments and techniques. This made it possible to assess the various geophysical parameters and to provide an estimate of their uncertainties. It was furthermore important to present the data in a form that was easy to compare against measurements from the satellite instruments. Recommendations for how data should be stored for various instruments were given and these were implemented in a series of templates. These templates are available at the CDB web portal.

3.3.3 Campaign archiving strategy

Before a campaign is performed, the PIs or campaign manager have clear goals for what they want to achieve with the measurements. For the ESA campaigns to be stored at CDB, it is necessary to keep these goals in mind when data are to be converted into HDF files and stored at the data centre. It is furthermore important to keep in mind that CDB is a database for several campaigns, and data from one campaign could be used by other campaigns (please note that data are not automatically shared between campaigns, and that sharing of data only occurs after an agreement with ESA and campaign managers). Such reuse of data may justify upload of more or other types of data. It should also be emphasised that CDB is mainly a geophysical database. All data-files to be converted to HDF and uploaded to the DB must have a specific geo-location and time reference. Auxiliary data may only be uploaded to the PIP and can normally not be shared among different campaigns.

There is also a question regarding the level of data to be archived. The only requirement regarding presence of variables in the CDB data files is that they must contain a time and geo-reference. Data must therefore be of level 1 or higher. A measurement campaign normally includes further processing of data into physical values such as ice thickness, gas concentration, vegetation index, etc. When reporting data to CDB, the data manager must again consider the purpose of the campaign and choose what levels of data that should be archived.

Sometimes, it is beneficial to store all campaign data products, on all available levels. This will allow for future reprocessing of data sets.

3.3.4 Example campaigns

In the first phase of CDB, three additional campaigns (DAISEX, LaRA and ESAG02) were analysed and data were uploaded to the data centre. Together with the ENVISAT Cal/Val activity these three campaigns represent the kind of data that are to be archived in the campaign database. The main focus and archiving issues of these campaigns are discussed in the following. Even though the three campaigns and the ENVISAT Cal/Val campaign represent typical CDB data, it is expected that future campaigns might require additional treatment and that the following archiving guidelines do not apply to all users of CDB.

DAISEX – remote sensing of geo/biophysical parameters

As stated on the DAISEX homepage: "The main scientific objective of the DAISEX campaigns was to demonstrate the feasibility of quantitatively retrieving geo/biophysical variables by accounting for atmospheric effects while at the same time analysing the data for possible additional information on directional anisotropy". The geo/biophysical parameters were measured with a range of different airborne instruments and in-situ air and ground-based measurements (soil-properties, vegetation indices, etc.) were performed in parallel to provide values for validation and atmospheric corrections. In DAISEX, the essential data are those obtained from the airborne measurements, but all data needed to be uploaded to the database. In this way, both the main and correlative data were stored in the same repository. As stated in the previous section, it is possible to store both level 1 data (spectral data) from the airborne sensors in the database, but also higher-level products such as Leaf Area Index (LAI), surface temperatures, etc. In DAISEX, the development of new methods and algorithms for earth observations was the essential goal, and the actual data that was recorded, was of less importance. Still it was considered favourable to store also the raw data in order to make future improvements of algorithms and reprocessing possible. The archiving strategy for DAISEX was therefore to store level 1 and level 2 data from the airborne sensors together with correlative ground-based data. An example on how to format such data in a CDB HDF file is given in Section 3.3.5.

ESAG02 – Airborne gravity measurements and Lidar and laser altimetry over sea-ice

ESAG02 was somewhat similar to DAISEX in the way that the campaign included airborne sensors that were validated against ground-based measurements. Again, it was necessary to store both airborne and correlative of data. The objectives of ESAG02 were twofold: To acquire high-accuracy airborne gravity measurements of the Arctic Ocean and to acquire scanning laser ranging (Lidar) data and profiling laser altimetry over the sea-ice north of Greenland. Both objectives were furthermore in support of the ESA GOCE and CryoSat missions respectively. The objectives of ESAG02 were rather different from DAISEX – the measurements themselves were the essential part of the project, and were not undertaken to demonstrate, develop or validate new methods. For ESAG02 data it was therefore advisable that data were only uploaded to CDB after the crossvalidation between ground-based and airborne measurements had taken place. Data were then structured as gravity and ice elevation (lidar + laser altimetry data) in separate files, both types represented as a function of location and time of the measurement (please note that one file may contain references to several time and locations). Metadata such as altitude/attitude/pitch/roll/etc. of the aircraft were included in the files. Gravity and ice elevation data were expected used by different groups, and it was decided to separate the two data streams into different files.

LaRA – Coincidental airborne laser and Radar altimetry over ice-sheets and sea-ice

The LaRA campaign had a main objective to compare height estimates from coincidental airborne laser and Radar measurements. As for the DAISEX campaign, the actual data, i.e. the measurements of the aircraft cruising altitude, was not the essential results of the project. For LaRA data we therefore proposed to archive the retrieved data from the various instruments as different streams in separate files. Since this was an inter-comparison campaign, the low-level products were probably the most interesting and these were archived to be used for later analyses. The actual altitude and position of the aircraft during the flight (high level product) were probably of less interest, but could also have been reported along with the individual measurements. The Radar level 1b products seemed ideal for storage in the CDB, but the metadata had to be somewhat reformatted and new entries in the list of legal parameters had to be added. This procedure is described in the next section (Section 3.3.5).

3.3.5 How to create a data template

The campaign archiving strategy will provide the data submitters with a clear idea of what they should archive in CDB, but templates should be deduced for each type of data to specify exactly how data should be formatted. The CDB web pages contains references to several templates that were made for the ENVISAT Cal/Val effort, and these templates can be seen as guidelines to how new data should be stored. For the generalisation of the database, the data from the three demonstration campaigns were analysed in details. Below is a resume of this process from the study of LaRA data, showing how raw data in a special campaign format can be reprocessed and stored in CDB HDF files.

The LaRA level 1b Radar data were organised in files with alternating sequences of Data headers and corresponding Waveform data. The first 51 bytes of a sequence were allocated for the header and contained information on time, the aircraft position and attitude, Radar tracking range, tracking shift, receiver attenuation, samples per radar pulse (length) and the Doppler Bin Size. The number of samples per pulse was 512, 256, 128 or 64. As described below, this variation induced an additional difficulty in data formatting. The data array followed the header information and consisted of one real and one imaginary part. The size of the data array was 8 * length (number of pulses) bytes.

In order to avoid storing each waveform in a separate HDF file, it was possible to store the data as a multidimensional array. We proposed to define DATETIME as the main independent variable, and let all other parameters be dependent on this. A problem arose when we used this approach, since we needed another time-variable to represent the waveform (signal as a function of time). This time-variable was on a much smaller scale, and we proposed to solve this with the variable name SIGNAL.SAMPLE. This was an array containing either an index (1,2,3,4…n) or actual time expressed in seconds (or rather a fraction of seconds). We decided to use the index-approach since the time-steps were constant (constant sampling rate). The actual data were then stored in two variables that were dependent on both DATETIME and SIGNAL.SAMPLE. One variable was declared for the real and one for the imaginary part of the data. It was also possible to store these two 2D arrays into one 3D variable, but this solution would complicate the format.

Another problem with using the described approach was that the number of samples per waveform, i.e. the necessary length of the array SIGNAL.SAMPLE varied throughout a measurement series. In some cases there were 512 samples per waveform and only 64 in other cases. We decided to scale the array to the maximum sampling number in order to create a rectangular 2D matrix for the waveform data and to use missing values in cases where less samples had been used. The method is described in the example below.

To visualize the approach, lets say we have 5 waveforms with sampling numbers between 5 and 10. We just use increasing numbers as data to make it easier to trace them. Numbers in parenthesis indicate numbers of samples per wave form.

Wave 1	Datetime = 1	Data = 101, 102, 103, 104, 105, 106	(6)
Wave 2	Datetime = 2	Data = 107, 108, 109, 110, 111, 112, 113, 114	(8)
Wave 3	Datetime = 4	Data = 115, 116, 117, 118, 119	(5)
Wave 4	Datetime = 5	Data = 120, 121, 122, 123, 124, 125, 126, 127, 128, 129	(10)
Wave 5	Datetime = 8	Data = 130, 131, 132, 133, 134, 135, 136	(7)

The DATETIME array will then be: [1,2,4,5,8]
The SIGNAL.SAMPLE array will then be: [1,2,3,4,5,6,7,8,9,10]

The maximum data value is 136 and the VAR_FILL VALUE must then be -99999 (5 nines). See appendix A for rules on how to define the fill values (missing values).

The 2D data array will then be like the table below. **Bold** numbers are for the independent variables that define the axes of the array. *Italic* numbers are for the real data (including missing values).

	1	**2**	**4**	**5**	**8**
1	*101*	*107*	*115*	*120*	*130*
2	*102*	*108*	*116*	*121*	*131*
3	*103*	*109*	*117*	*122*	*132*
4	*104*	*110*	*118*	*123*	*133*
5	*105*	*111*	*119*	*124*	*134*
6	*106*	*112*	*-99999*	*125*	*135*
7	*-99999*	*113*	*-99999*	*126*	*136*
8	*-99999*	*114*	*-99999*	*127*	*-99999*
9	*-99999*	*-99999*	*-99999*	*128*	*-99999*
10	*-99999*	*-99999*	*-99999*	*129*	*-99999*

Such an array can be used to store both the real and the imaginary part of the data (need two variables containing such a 2D array). The drawback with the method is the unnecessary use of missing-values that take up extra space in the HDF file. To avoid this, the data can be stored in a 1D variable with one waveform trailing the previous one. This is similar to how the data were stored in the original LaRA Radar level 1b data (except that they had headers in between the data arrays. This approach would require less storage space, but the files will be more difficult to understand and the data user has to keep track on where one waveform begins and where it ends (like a tape-archive). For the LaRA data we therefore proposed to format the data more like a spreadsheet in a 2D structure like the one just described.

Regarding the actual implementation and storage of the data, a number of new variable names had to be decided on. Below is a list of all the variables that was necessary to include in the HDF file to properly archive the level 1b data. In addition to the ones below, one could also include error estimates of some data in separate variables, and it was naturally necessary to include the global metadata parameters. Variables in bold were necessary to add to the list of legal values for this formatting task.

Variable Name	**Dependence**	**Unit**	**Comment**
DATETIME	INDEPENDENT	MJD2000	Replaced the entry for seconds of the day in the current level 1b data
LATITUDE	DATETIME	Deg	
LONGITUDE	DATETIME	Deg	
VALID.PULSE	DATETIME	DIMENSIONLESS	Flag – 0 or 1
ALTITUDE.INSTRUMENT	DATETIME	Meters	Simply named altitude in current level 1b data
ATTITUDE.PITCH	DATETIME	Deg	The pitch of the aircraft and receiver
ATTITUDE.ROLL	DATETIME	Deg	The roll of the aircraft and receiver
ATTITUDE.YAW	DATETIME	Deg	Replaced the entry named “Heading” in the current level 1b data, and described the flight direction of the aircraft
SIGNAL.RANGE	DATETIME	μs	The length of the tracking range
SIGNAL.DELAY	DATETIME	μs	The length of the tracking shift
SIGNAL.PULSE.LENGTH	DATETIME	μs	Pulse time length. Correlation between pulse length, zero delay and samples/waveform is given in the VAR_NOTES of this variable. This made a SIGNAL.ZERO.DELAY variable obsolete
SIGNAL.ATTENUATION	DATETIME	**dB**	Receiver setting – amplification/ damping. Value in Decibel. SI-conversion difficult.
DOP.BIN.SIZE	CONSTANT or dependent on DATETIME	Meter	Doppler bin size used in the FFT algorithm.

Variable Name	Dependence	Unit	Comment
SIGNAL.SAMPLE	INDEPENDENT	DIMENSIONLESS	We need an independent 1D variable with size equal to the maximum number of samples, as given by SIGNAL.PULSE.LENGTH.
SIGNAL.INTENSITY.REAL	DATETIME, SIGNAL.SAMPLE	DIMENSIONLESS	A 2D array containing one wave-form per datetime. Where the number of pulses is less than the maximum value, the VAR_FILL_VALUE (missing value) should be used
SIGNAL.INTENSITY.IMAGINARY	DATETIME, SIGNAL.SAMPLE	DIMENSIONLESS	A 2D array containing one wave-form per datetime. Where the number of pulses is less than the maximum value, the VAR_FILL_VALUE (missing value) should be used

In addition to the ones above, a series of constant variables describing the D2P Radar Characteristics could be included. This is, however, not normally done. The Radar is, however, not a traditional Radar altimeter and a special instrument name is used instead of ALTIMETER.RADAR. An example could be ALTIMETER.RADAR.D2P to indicate that this is Delay/Doppler Phase-monopulse Radar.

The above example is a brief explanation on how data sets are approached in order to create a data template. In addition to the information given above, the template should contain full metadata descriptions on all variables and a complete definition of all global metadata in the file. The examples found on the CDB web portal should be used as guidelines.

3.3.6 Organising information and documentation

NILU will set up a campaign specific version of the Project Internal Pages (PIP) and will provide the campaign data manager with an administration account. This account will allow you to modify the PIP according to your needs. As a PIP administrator you should especially be aware of the following:

The administrator can create other user accounts for the PIP – both normal users and other administrators of the same campaign. Depending on the needs in the campaign, the administrator can choose to give each individual a separate account or everyone can use the same login and password. Please note that the user names and passwords used in the PIP has no connection to the login accounts used for the geophysical database. Creating a new user is done by following the link “Contact Info” from the main page, and then further by clicking on “New user”. You create a new user by filling in all relevant information in the boxes and by clicking on the “Insert new user” button. White boxes must be filled in while the yellow ones are optional. The “Contact info” page will now be updated with the new user. The administrator can delete users or edit user information by clicking on the respective links to the right of the user names.

The PIP has a module for archiving of documents, and this can be set up with any number of folders and sub folders, much similar to a directory tree structure. A PIP administrator can create new folders and sub folders and can also upload/download files to these. The administrator can also remove directories that are not needed anymore, e.g. any of the three pre-made directories in the documents section. A normal user can only upload/download files to existing folders and will not be able to create or remove directories.

The PIP administrator has the possibility to modify the text in the “Project information” and the “Message” boxes on the PIP main page. This is done by clicking on the corresponding links in the top right corner, and filling out new information in the boxes. Some formatting tools are available to ease creation of e.g. outlined text and hyperlinks. Please contact NILU if you need help with this.

Appendix A

Metadata guidelines for the ESA Campaign Data Base (CDB)

Preface

The ESA Cal/Val database was developed and implemented at NILU to provide ENVISAT scientist with a common framework and repository for exchange of correlative data, mainly from ground based measurements. The experience from this activity led to a new ESA initiative to develop a more general database, the ESA Campaign Database (CDB). This system is a generalisation and further development of the Cal/Val system used for some ENVISAT calibration and validation campaigns. We have tried to keep the differences to a minimum, to make the transition easy for the user community of the original system. The CDB includes all data and metadata definitions from the previous Cal/Val data centre, but is able to handle data from all ESA campaigns. It is a system for storing and indexing complex data sets from a multitude of sciences, and is no longer a database for correlative data only. Addition of new functionality or redesign of existing components will be an evolutionary process in co-operation with ESA and user representatives. The first step in this process was the preparation to accommodate data from 3 pilot campaigns, ESAG02, LARA and DAISEX, and the system is currently used for various measurement campaigns sponsored by ESA.

For maximum compatibility and easy re-use of data, the rules should be common for all campaigns that use the system. Yet, specific project policies are often required. The objective of the CALVAL guideline document was to define specific metadata guidelines for the Validation Campaign of the European Space Agency’s Envisat earth observation mission, in particular for the validation of the AATSR, GOMOS, MERIS, MIPAS and SCIAMACHY sensors. Particular rules were formulated for use by the Envisat Principal Investigators (PIs) Data Originators (DOs) and Data Submitters (DSs). The CDB campaigns may need more general guidelines, but conflict with the CALVAL rules are avoided any impact on existing users..

The current document implements the following highlighting:

The current document is based on the metadata guidelines document developed for the initial Envisat Cal/Val activity, Bojkov et al. (2002). Current metadata definitions and additions are, in agreement with ESA, created by NILU.

Norwegian Institute for Air Research

Terje Krognes
terje.krognes@nilu.no

Sam Erik Walker
sam.erik.walker@nilu.no

Aasmund Fahre Vik
afv@nilu.no

Contents

Metadata guidelines for the ESA Campaign Data Base (CDB) Version April 5th 2006

A1 Introduction

Earth observation satellite campaigns are multidisciplinary, and generally combine selected datasets from satellite instruments with correlative groundbased data. Participants are spread around the globe, and work in different fields of science and in different organisations. This creates a large demand for electronic data exchange, and for indexing and retrieval of many different types of datasets. Common file formats are important tools for efficient indexing and retrieval, although the diversity of the data material is too large for one single file format. Common data definitions (naming conventions and definitions for data and metadata elements) are essential for such complex data exchange. The data definitions constitute a common language, which ensures that the indexing and search terms are subject to one common interpretation by all participants. Furthermore, each data set must be accompanied by metadata that describe the content and context of the data set. These Metadata Guidelines define the meaning of the terms we use. They also define the metadata content that is required in each data set.

The CDB is a generic Campaign Data Base that will hold both selected satellite data sets and data from groundbased measurements and computations. Groundbased in this context covers measurements performed on the ground or inside the atmosphere, with instruments that may be stationary, or may be carried in cars, ships, aircraft, balloons or other vehicles. In many cases the groundbased datasets are created by satellite instruments during tests inside the atmosphere. In some cases groundbased data are created specifically for comparison with satellite data, and are commonly named correlative data. The datasets may be pure measurements, model calculations, or assimilation results (model computations adjusted by assimilation of actual measurements). Depending on the level of finishing, a dataset may have been processed by computer programs that perform anything from simple scaling and calibration to sophisticated outlier removal and assimilation into model computations.

Datasets may be usable for more than one campaign. While some datasets must be protected from viewing by others than campaign or project members, other data sets must be made available for other specified campaigns, or for the entire user community. Extensive mechanisms for user control and data ownership control are included in the system.

These Metadata Guidelines describe a generalised metadata standard based on the Envisat Cal/Val system (Bojkov et al., 2002). The definitions have been carefully chosen to allow new campaign data types to be included, while keeping as much as possible of the original definitions. This is a living document, and modification should be expected to both data definitions, reporting routines and file formats. The CDB consists of a central clearing house for data transfer files, a relational database index, web interfaces for data providers, data users and administrators, Metadata Guidelines and other documentation, software products for creation and quality control of data transfer files, and a group of support personnel at NILU (working under contract for ESA).

A2 Concepts

The multidisciplinary exchange of data in earth observation depends heavily on *good* definitions for data and metadata. Freedom of choice would let different end-users describe similar data sets in very different terms, thus hindering efficient retrieval. To avoid this, we define a small set of data and metadata entities (the structure of our data), and allowed values for each of these entities (the metadata values). The central structural data-definitions are briefly discussed in the following paragraphs.

A2.1 Terminology

metadata Data about data. Parameters that describe, characterize and/or index the data.

parameter A physical or chemical entity that is measured or computed (often pertaining to data), or predefined (often pertaining to metadata).

dataset A set of one or more parameters reported in coincident time and space. In most cases, this refers to a collection of parameters in one single data transfer file, and to the time/space frame covered by this file.
In some cases, however, the time frame of a dataset is larger, and more than one file is needed to define the entire dataset. In some cases the spatial frame or the number of parameters included in the dataset definition may also be larger than what can be accommodated in a single data transfer file. The original definition of dataset above is recommended, but the flexibility of the main data transfer file format is not always sufficient to support a very large or complex dataset in a single file.

variable A data parameter to be reported in a dataset. Characterized by variable name, variable mode, and variable descriptor (see detailed descriptions below).

variable name The primary variable identifier. The name of the physical quantity observed or estimated by the measurement or model calculation

variable mode The mode generally describes how or in what context the variable was measured.

variable descriptor	The descriptor will shift the focus from the normal value of the variable to some other aspect, like its uncertainty, its minimum, a flag, etc.
unit	Ideally, any given combination of a variable name, mode and descriptor should have only one natural, legal unit and scale. The CDB adds the possibility to enforce correct use of units as a part of the campaign policy.
constant	A constant is named as a variable (with name, mode and descriptor, as required). In a global context the constant may actually be a variable entity, but in the context of a given data transfer file (for the range of independent variables covered by that file), the constant can only hold one single value.
independent variable	*Each data file must have at least one independent variable (more than* one if the dependent variable is multidimensional). The dependency is defined in the context of the current data transfer file. In a global context, the variable may not be independent, but it does not depend on the value of any other variable in the current file.
dependent variable	A parameter that is provided as a function of another parameter (for example temperature as a function of time) is called a dependent variable. The parameter on which it depends is an independent variable. The number of independent variables determines the dimensionality of the grid on which the dependent data are provided.
data source	An instrument or a model. Data from the source is normally quality controlled, calibrated and scaled before it is formatted into a data file and submitted to the data centre. Some instruments gather samples that must be analysed in a laboratory before results are reported. The sampler is then considered to be the source. In the CDB we add an option to subdivide the data source name. The subdivision may define several channels as part of an instrument. In assimilation it is often convenient to define the output of each component as a separate “instrument channel”, which can be named by the component name.
data location	The position of the sampled or modelled site (this may be a mobile entity such as a plane or ship).
DO	Data Originator. A defined role for a person that may be referenced in a data file. This role does not automatically give web access or file upload privileges.

DS Data Supplier. A defined role for a person that is registered in metadata with permission to access the CDB web site, and to upload data files for one or more projects or campaigns.

PI Principal Investigator. A defined role for a person that may be referenced in a data file. This role does not automatically give web access or file upload privileges.

A2.2 Data transfer file structure

The main file format is a subset of the HDF 4.1r3 format. The current document limits the user to only use certain features of this format, and to add mandatory metadata information with the variable names and values listed in this document (and the updated on-line versions). In the future, the main file format may be changed to HDF5, which allows more flexibility and more logical formatting of some data types.

When technically feasible (and when required by project policy) the DS will create a data transfer file in the main format for each dataset, and include in this single file both the data and the associated metadata. After checking the file (preferably with the ASC2HDF tool provided from the data centre), the DS will upload the file to the data centre. Sufficient metadata must be available in the header of each file (as specified in Sections 4 and 5). This is required both for proper indexing, and to make the data useful to the end user that retrieves the file. The user will expect to be able to use the data properly without searching for metadata in other sources.

Metadata parameters are divided into Global Attributes (pertaining to an entire dataset contained in one single file), and Variable Attributes (pertaining to one single variable within a dataset). A variable is commonly a chemical component or physical parameter that is reported in a file (the main content of the file). Several variables are normally included in a dataset. The term parameter is in our context normally used for a metadata element (a piece of information about a variable or an entire dataset). The term field is often used for a subdivision of the content of a parameter (for example, a person name parameter consists of both family name and first name). In many cases, a field may be subdivided into sub fields with dot separators.

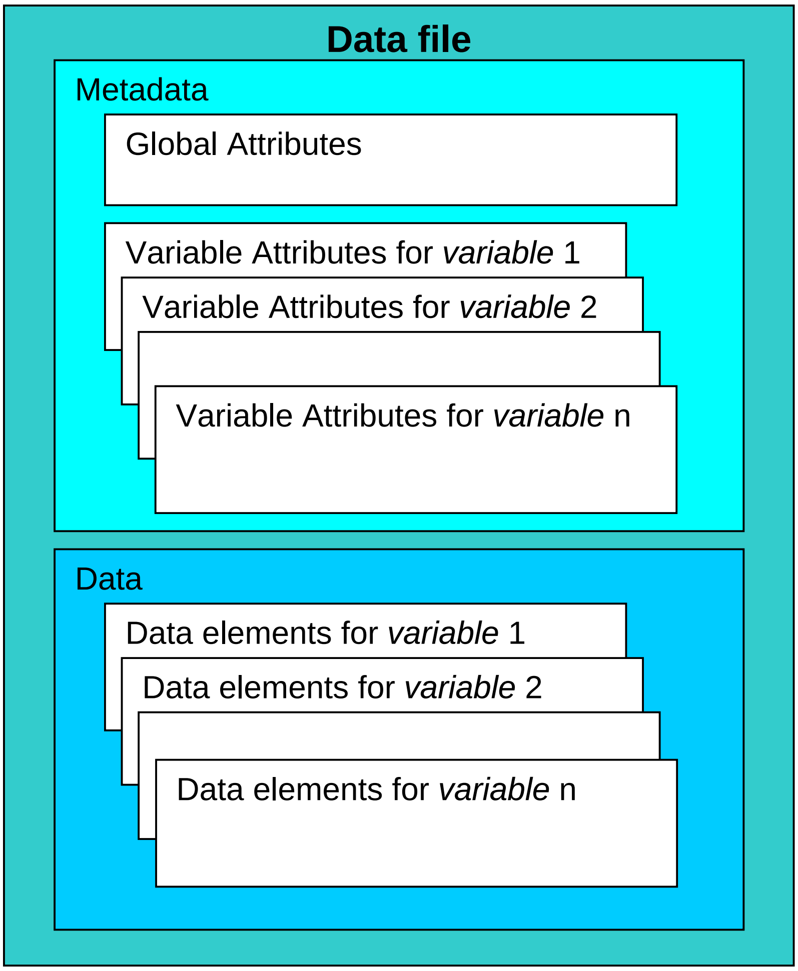

Figure 4: *Simplified view of the file data structure.*

For the purposes described here, a dataset normally consists of all data from one single instrument, auxiliary data (such as related meteorological data), and metadata that describe the data. The main data (measurements or calculations) are often referred to as primary data. The auxiliary data are often referred to as secondary data. One particular class of auxiliary data are time and position information. These variables are often independent variables. The primary data and other secondary data parameters are normally dependent variables.

A2.3 Considerations

In the context of effective data exchange and efficient data management various considerations must be given to the following:

1. The identification (i.e. naming) of the parameter is of great importance. The description (consisting of **variable name** and optional **variable mode** and **variable descriptor**) should allow identification of parameters in various datasets with a similar physical basis. For that reason the **variable name** should contain a basic description in physical terms of the physical quantity estimated and of the geophysical or chemical target that is subject of the measurement, for example TEMPERATURE.AIR. The **variable mode** on the other hand, should emphasise those aspects of the measurement method that prevent simple direct comparison with other estimations: The measurement is

an estimate of the underlying physical quantity, but when comparing estimations obtained with different methods, the differences in variable mode inform the user that differences between the results may actually be due to the estimation method. The third entry, the **variable descriptor**, can be used to construct a related variable that contains additional information (for example: error, uncertainty) on the original variable.

2. The **variable mode** or **variable descriptor** should not be used to distinguish measurement methods that are characterised by the use of specific but potentially different input values of a physical quantity. Typical examples are reference wavelength or pre-defined depths. Instead, these quantities should be provided as independent variables if several values are applicable to the measurements, or otherwise as constants. In practice this means that numeric values will generally not appear in the **variable mode** or **variable descriptor**. The consequence of this consideration is that the data are properly formatted as multidimensional datasets, instead of being presented as one-dimensional slices with independent parameters tucked away in the variable name.

3. A minimal set of time and position variables is mandatory: geolocation must be specified in terms of date, time (in the variable DATETIME), latitude, longitude and altitude or depth. If at all possible this geolocation must describe the effective location of the 'object' that is subject of measurement.

4. Pressure (PRESSURE) or geopotential height (ALTITUDE.GPH) for the measurement or calculation position is acceptable as an alternative if altitude cannot be provided. If this is not available, the geolocation of the instrument and relevant auxiliary parameters must be provided. In this case the geolocation is expressed as LATITUDE.INSTRUMENT, LONGITUDE.INSTRUMENT, ALTITUDE.INSTRUMENT.

5. Data may be reported over several different time scales, such as hourly, daily, monthly or seasonal in length, depending on need. One dataset may be divided into several data files, when this facilitates comparison to satellite data. Since satellite data files typically contain much less than one day of data, correlative data files should generally not contain more than one day of data.

6. There is always a possibility that someone can submit an erroneous dataset that appears to be legal in normal integrity checks. Some types of errors are difficult to detect even with stringent quality control routines. To minimise the workload for data originators and data suppliers, there is a tendency to minimise the amount of mandatory metadata. The system has numerous fields for free text comments and additional information from the data originators. Data originators must use these fields liberally to ensure that users gain sufficient knowledge of the data set and its intended usage.

7. The metadata guidelines may appear complex. However, the guidelines serve to reduce the complexity inherent in the data exchange problem. The majority of typical errors will be detected before the file is indexed and added to a file tree. This constitutes a major improvement in the management efficiency compared to a file tree that is not supported by such an index database. The

resulting metadata index will facilitate both project management and scientific use of the collected data.

A3 Formatting issues

A3.1 Character set

- All metadata entries should be given with characters contained in the US ASCII character set.
- No special national characters are allowed (Å, ñ, ô, ö, etc.).
- Underscore characters "_" are used to separate metadata elements from each other, and cannot be part of a metadata element.
- The period symbol "." is used to separate sub fields from each other inside a metadata element.
- Other special characters ?, #, !, &, %, etc.) should not occur, except in comment text strings.
- Hyphens and apostrophes may occur in names of people, locations or institutions. In other contexts such special characters are not allowed.

A3.2 Capitalisation

- All metadata entries are generally all capitals.
- Variable names and measurement units are defined with specific capitalisation, and the input routines are case sensitive for such elements.
- File names are always set in lower case.
- Names of persons and addresses should be submitted with natural capitalisation.

A3.3 Numeric Type

The currently implemented numerical types are found in Table 3.3. These have been chosen carefully for compatibility in FORTRAN, C, IDL and HDF.

Table 3.3: *Allowed numeric types implemented for the Envisat Cal/Val project.*

Numeric Type	*Comment*
REAL	*HDF: 32 bit floating point numbers (FORTRAN: *4real)*
DOUBLE	*HDF 64-bit floating point numbers (FORTRAN: *8real)*
INTEGER	*HDF: 16-bit signed integers (FORTRAN: *2integer)*
LONG	*HDF: 32-bit signed integers (FORTRAN: *4integer)*
STRING	*character string (Note that the maximum string length is software/tool dependent)*

A3.4 Fill value

Data elements and metadata parameters cannot be left empty. A missing code (also called fill value) is normally used to fill an element when data is not available, but a measurement has been performed.

A3.4.1 Numeric fill values

For numbers, the fill value is negative and consists of nines. In absolute value it must be 2 orders of magnitude larger than the absolute value of the real data. If the **VAR_DATA_TYPE** is of type floating point, then the fractional data of the fill value must be zeroes to the same number of digits as the measurement data.

ATTENTION

Special care must be given to the data format to prevent that the larger fill values exceed the number of positions reserved in the data format.

Example: *General*

Data is of the order `0.1` *the fill value must be:* `-99.0`

Data is of the order `10000` *the fill value must be:* `-9999999`

Example: *Exponentials*

Data is of the order `2.dddE-6` *the fill value is:* `-9.000E-4`

Data is of the order `2.ddE+6` *the fill value is:* `-9.00E+8`

A3.4.2 String fill values

For string variables – the fill value is always “ZZZZZZZZZZ” (10”Z’s”).

Example: *Strings*

The datum is a string *the fill value is:* `ZZZZZZZZZZ`

A3.5 Date formats

There are two date formats used in these guidelines: a numerical format (MJD2000) for data reporting and a string format (ISO 8106) used in the file name construction. The MJD2000 format is used for data records to facilitate calculations and plots.

A3.5.1 MJD2000

The Modified Julian Date (MJD2000) used throughout this document is defined as follows:

```
MJD2000 is 0.000000 on January 1, 2000 at 00:00:00 UTC.
```

A3.5.2 MJD2000 algorithm

The general algorithm to calculate MJD2000 is as follows:

For a given YYYY, MM, DD, hh, mm, ss:

STEP 1: Calculate the Julian date:

```
IF ( MM GT 2 ) THEN
      y = DOUBLE(YYYY)
      m = DOUBLE(MM - 3)
      d = DOUBLE(DD)
ELSE BEGIN
      y = DOUBLE(YYYY - 1)
      m = DOUBLE(MM + 9)
      d = DOUBLE(DD)
ENDELSE

j =   INTEGER( 365.25*( y+4712.0 ) ) +
INTEGER( 30.6*m+0.5 )+ 59.0 + d - 0.5

Check for Julian or Gregorian calendar:
IF ( j LT 2299159.5D0 ) THEN; If Julian calendar.
    jd = j
ELSE            ; If Gregorian calendar.
    gn = 38.0 - INTEGER(3.0*INTEGER( 49.0+y/100.0 )/4.0)
    jd = j + gn
ENDELSE
```

STEP 2: Calculate day fraction

```
df = ( hh*3600.0 + mm*60.0 + ss ) / 86400.0
… for second resolution

or

df = ( hh*3.6E+6 + mm*6.0E+4 + ss*1.0E+3 + ms ) / 8.64E+7
… for milli-second resolution
```

STEP 3: Calculate MJD2000

```
mjd2000 = jd + df – 2451544.5
```

Example: *for* 2002/04/20 at 11:29:23 UTC `mjd2000 = 840.478738`

ATTENTION

Special care must be given to the formatting of MJD2000 by reporting the appropriate number of significant figures to represent the actual time resolution.

A3.5.3 DATETIME (ISO-8106)

The UTC DATETIME representation in ISO-8106 long format is (ISO, 1988):

```
YYYYMMDDThhmmssZ
```

where

YYYY	is the numeric year
MM	is the numeric month
DD	is the numeric day
T	is a delimiter separating time from date
hh	is the numeric hour
mm	is the numeric minute
ss	is the numeric second
Z	is a flag indicating Universal Time (UTC).

ATTENTION

When appropriate, `MM`, `DD`, `hh`, `mm`, `ss` may require a leading zero.

For example `20010101T060501Z`.

A4 Global attributes

To facilitate the understanding of the Global Attributes, three categories have been defined, namely **Originator Attributes** (Section 4.1), **Dataset Attributes** (Section 4.2) and **File Attributes** (Section 4.3). Each metadata parameter in these 3 groups is specified once for each data file. All these attributes (with some very few exceptions) need to be filled in.

Table 4: *Overview of required Global Attributes for the Envisat Cal/Val project. 'X' indicate entries and 'O' indicate optional entries.*

Originator Attributes	Section	Entry	Entry type	Req
PI_NAME	4.1.1	Family name; Given Name	2 semi-colon separated	X
PI_AFFILIATION	4.1.2	Affiliation name, Affiliation Acronym	2 semi-colon separated	X
PI_ADDRESS	4.1.3	Address; Postal code; Country name	3 semi-colon separated	X
PI_EMAIL	4.1.4	E-mail address	Single entry	X
DO_NAME	4.1.5	Family name; Given Name	2 semi-colon separated	X
DO_AFFILIATION	4.1.6	Affiliation name, Affiliation Acronym	2 semi-colon separated	X
DO_ADDRESS	4.1.7	Address; Postal code; Country name	3 semi-colon separated	X
DO_EMAIL	4.1.8	E-mail address	Single entry	X
DS_NAME	4.1.9	Family name; Given Name	2 semi-colon separated	X
DS_AFFILIATION	4.1.10	Affiliation name, Affiliation Acronym	2 semi-colon separated	X
DS_ADDRESS	4.1.11	Address; Postal code; Country name	3 semi-colon separated	X
DS_EMAIL	4.1.12	E-mail address	Single entry	X

Dataset Attributes	Section	Entry	Entry type	Req
DATA_DESCRIPTION	4.2.1	Data description	Single entry	X
DATA_DISCIPLINE	4.2.2	Field; Class; Subclass	3 semi-colon separated	X
DATA_GROUP	4.2.3	Type; Subtype	2 semi-colon separated	X
DATA_LOCATION	4.2.4	Location code name	Single entry	X
DATA_SOURCE	4.2.5	Concatenated:DATA_SOURCE Type + Institute acronym + 3-digit identifier	Concatenated entry	X
DATA_TYPE	4.2.6	Concatenated:Time scale code + Data level code	Single entry	X
DATA_VARIABLES	4.2.7	List of variables in the file	n semi-colon separated	X
DATA_START_DATE	4.2.8	MJD2000	Single entry	X
DATA_FILE_VERSION	4.2.9	3 digit integer	Single entry (ddd)	X
DATA_MODIFICATIONS	4.2.10	Description of the data modifications	Single entry	X
DATA_CAVEATS	4.2.11	Description of the data caveats	Single entry	O
DATA_RULES_OF_USE	4.2.12	Description of the data rules of use	Single entry	O
DATA_ACKNOWLEDGEMENT	4.2.13	Data acknowledgement	Single entry	O
File Attributes	**Section**	**Entry**	**Entry type**	**Req**
FILE_NAME	4.3.1	Concatenated and underscore separated	Concatenated entry	X
FILE_GENERATION_DATE	4.3.2	MJD2000	Single entry	X
FILE_ACCESS	4.3.3	File project association	Semi-colon separated	X
FILE_PROJECT_ID	4.3.4	Custom project identification related to 4.3.3	Single entry	X
FILE_ASSOCIATION	4.3.5	File "other" project association	Semi-colon separated	O
FILE_META_VERSION	4.3.6	Meta data version used	2 semi-colon separated (ddRddd; free format)	X

A4.1 Originator attributes

The Originator Attribute metadata entries describe the ownership of the data found in a given file as well as the guidelines for the use and/or publication of these data.

A4.1.1 PI_NAME

The Global Attribute **PI_NAME** is the data's (or instrument's) Principal Investigator's (PI) Name. The PI has the main scientific and/or institutional responsibility for the given data.

ATTENTION

If there is no instrument PI for the reported data in the file (as is the case for some operational satellite instruments) – then the Data Submitter (DS) must substitute the PI information with the instrument's affiliation coordinates and institute's information.

Type: STRING
Format: Family name; Given names
Entry: The entry consists of two fields separated by a semicolon.
Example: PI_NAME = Vik; Aasmund

A4.1.2 PI_AFFILIATION

The Global Attribute **PI_AFFILIATION** is the Principal Investigator's **official** affiliation name and affiliation acronym.

Type: STRING
Format: Affiliation name; Affiliation acronym
Entry: The entry consists of two fields separated by a semicolon.
Example: PI_AFFILIATION = Norwegian Institute for Air Research; NILU

Table 4.1.2: *Allowed affiliation names and affiliation acronyms of the agencies and institutes participating in the Envisat Cal/Val project.*

AFFILIATION NAME	AFFILIATION ACRONYM
ACRI	ACRI
Airborne Hyperspectral Remote Sensing Systems & Solutions	ITRES
Alfred-Wegener-Institut fuer Polar und Meeresforschung	AWI
Aristotle University of Thessaloniki, Laboratory of Atmospheric Physics	LAP
Australian Institute of Marine Science	AIMS
Belgian Institute for Space Aeronomy	BIRA.IASB
British Antarctic Survey	BAS
Centre National d\'Etudes Spatiales	CNES
Centro de Estudios y Experimentacion de Obras Publicas	CEDEX
Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas	CIEMAT
Chalmers University of Technology	CTH
Commonwealth Scientific and Industrial Research Organisation	CSIRO
Danish Meteorological Institute	DMI
Departement of Geography, University College, London	UCL.GEOG
Department of Meteorology Stockholm University	MISU
Deutscher Wetterdienst	DWD
Deutsches Zentrum fuer Luft- und Raumfahrt	DLR
Dielmo 3D S.L., Valencia	DIELMO
Environmental Research and Services, Florence, Italy	ERS
Estacion Experimenta de Aula Dei-Consejo Superior de Investigaciones Cientificas	EEAD.CSIC
European Centre for Medium-Range Weather Forecasts	ECMWF
European Commission - Joint Research Centre	JRC
European Space Agency	ESA
Finnish Meteorological Institute	FMI
Forschungszentrum Juelich	FZJ
Forschungszentrum Karlsruhe	FZK
Fraunhofer-Institut fuer Atmosphaerische Umweltforschung	IFU
Free University of Berlin	FUB
GKSS Forschungszentrum Geesthacht	GKSS
Hadley Centre	HADCEN
Informus GmbH	INF
Institut fuer Ostseeforschung	IOW
Institut fuer Umweltphysik, Universitaet Bremen	IUP
Institut National de la Recherche Agronomique	INRA
Institute for Environmental Studies - Vrije Universiteit - Amsterdam	IVM

AFFILIATION NAME	AFFILIATION ACRONYM
Institute of Atmospheric Physics - Russian Academy of Sciences	IAP.RAS
Institute of Experimental Meteorology - Russia	IEM
Institute of Meteorology and Water Management	IMWM
Institute of Ocean Sciences	IOS
Instituto de Astrofisica de Andalucia	IAA
Instituto de Desarrollo Regional, Universidad de Castilla-La Mancha	UCLM.IDR
Instituto Nacional de Meteorologia	INM
Instituto Nacional de Tecnica Aerospacial	INTA
Instituto Tecnico Agronomico Provincial (Albacete)	ITAP
International Institute for Geo-Information Science and Earth Observation	ITC
Istituto di Fisica Applicata Carrara	CNR.IFAC
Istituto di Fisica dell Atmosfera del CNR	CNR.ISAC
Istituto di Metodologie per l\'Analisi Ambientale del CNR	CNR.IMAA
Kyrgystan State National University	KSNU
Laboratoire de Meteorologie Dynamique du CNRS	CNRS.LMD
Laboratoire de Physique et Chimie de l\'Environnement du CNRS	CNRS.LPCE
Laboratoire de Physique et Chimie Marines du CNRS	CNRS.LPCM
Laboratoire de Physique Moleculaire et Applications du CNRS	CNRS.LPMA
Laboratoire des Sciences de l\'Image, de l\'Informatique et de la Teledetection (CNRS/ULP)	LSIIT
Laboratoire pour l\'Utilisation du Rayonnement Electromagnetique	LURE
Leibniz Institut fuer Atmosphaerenphysik	IAP
Los alamos national laboratory	LANL
Main Geophysical Observatory - Russia	MGO
Management Unit of the North Sea Mathematical Models	MUMM
Meteorological Service of Canada	MSC
NASA\'s Goddard Space Flight Centre	NASA.GSFC
NASA\'s Jet Propulsion Laboratory	NASA.JPL
NASA\'s Jet Propulsion Laboratory - Table Mountain Facility	NASA.JPL.TMF
NASA\'s Langley Research Centre	NASA.LRC
National Center for Atmospheric Research	NCAR
National Institute for Environmental Studies, Tsukuba, Japan	NIES
National Institute of Public Health and the Environment	RIVM
National Institute of Water and Atmospheric Research	NIWA
National Oceanic and Atmospheric Administration	NOAA
National Physical Laboratory	NPL
National Taras Shevchenko University of Kyiv	KTSU
NOAA National Environmental Satellite Data and Information Service	NOAA.NESDIS
Norwegian Institute for Air Research	NILU
Norwegian Institute for Water Research	NIVA
Observatoire de Bordeaux (INSU/CNRS)	OBORDEAUX
Observatoire de Neuchatel	ON
Plymouth Marine Laboratory	PML
Remote Sensing Application Consultants	RSAC
Royal Meteorological Institute of Belgium	RMI
Royal Netherlands Meteorological Institute	KNMI
Russian Central Aerological Observatory	CAO
Rutherford Appleton Laboratory	RAL
Service Central d\'Exploitation Meteorologique	SCEM
Service d\'Aeronomie du CNRS	CNRS.SA
Smithsonian Astrophysical Observatory	SAO
St.Petersburg State University	SPBSU
Stockholm University	SU
Swedish Environmental Research Institute	IVL
Swedish Institute of Space Physics	IRF
Swiss Federal Institute of Technology - Zurich	ETHZ
Swiss Meteorological Institute	MCH

AFFILIATION NAME	AFFILIATION ACRONYM
United Kingdom Meteorological Office	UKMO
Universidad de Castilla-La Mancha	UCLM
Universita degli Studi di Napoli \"Federico II\"	UNINA
Universite de la Reunion Laboratoire de Physique de l\'Atmosphere	UREUNION.LPA
University of Athens, Department of Physics, Division of Applied Physics	UOA
University of Bern	UBERN
University of Bonn	UBONN
University of Bremen	UBREMEN
University of Cambridge, Department of Chemistry	UCAMB.CHEM
University of Denver	DU
University of Frankfurt	UFRANKFURT
University of Heidelberg	UHEIDELBERG
University of l\'Aquila	UNIVAQ
University of Leicester	ULEICESTER
University of Liege	ULG
University of Lisbon	UL
University of Massachusetts	UMASS
University of Miami	UMIAMI
University of Milano	UNIMI
University of Nagoya	UNAGOYA
University of Oslo	UIO
University of Reading Data Assimilation Research Centre	UREADING.DARC
University of Reims	UREIMS
University of Sao Paulo	UNESP
University of Southampton	USOUTHAMPTON
University of Toronto	UT
University of Valencia	UVAL
University of Valencia, Facultad de Fisica	UVAL.FISICA
University of Valencia, Escuela Tecnica Superior de Ingenieria	UVAL.ETSE
University of Washington	UWAS
University of Wales Aberystwyth	UWA
University of Wollongong	UOW

A4.1.3 PI_ADDRESS

The Global Attribute **PI_ADDRESS** is the Principal Investigator's official mailing address. The country name must be the English entry in ISO 3166-1:1997 (ISO, 1997).

Type: STRING
Format: Address; Postal code; Country name
Entry: Three fields separated by semicolons
Example: PI_ADDRESS = P.O. Box 100; N-2027 Kjeller; Norway

A4.1.4 PI_EMAIL

The Global Attribute **PI_EMAIL** is the Principal Investigator's e-mail address.

Type: STRING
Format: Free format
Entry: Single field
Example: PI_EMAIL = aasmund.vik@nilu.no

A4.1.5 DO_NAME

The Global Attribute **DO_NAME** is the Data Originator's (DO) Name. The DO may or may not be the same person as the PI. It is often important to distinguish the DO from the PI, since the person that has performed the measurements, computed and quality controlled the results, may know details of which the PI is not aware.

Type;	STRING
Format:	Family name; Given names
Entry:	The entry consists of two fields separated by a semicolon.
Example:	DO_NAME = Krognes; Terje

A4.1.6 DO_AFFILIATION

The Global Attribute **DO_AFFILIATION** is the Data Originator's **official** affiliation (the DO_AFFILIATION may differ from the PI_AFFILIATION).

Type:	STRING
Format:	Affiliation name; Affiliation acronym
Entry:	The entry consists of two fields separated by a semicolon.
Example:	DO_AFFILIATION = Norwegian Institute for Air Research; NILU

A4.1.7 DO_ADDRESS

The Global Attribute DO_ADDRESS is the Data Originator's mailing address (the DO_ADDRESS may differ from the PI_ADDRESS). The country name must be the English entry in ISO 3166-1:1997 (ISO, 1997).

Type:	STRING
Format:	Address; Postal code; Country name
Entry:	Three fields separated by semicolons
Example:	DO_ADDRESS = P.O. Box 100; N-2027 Kjeller; Norway

A4.1.8 DO_EMAIL

The Global Attribute **DO_EMAIL** is the Data Originator's e-mail address (the DO_EMAIL may differ from the PI_EMAIL).

Type:	STRING
Format:	Free format
Entry:	Single field
Example:	DO_EMAIL = terje.krognes@nilu.no

A4.1.9 DS_NAME

The Global Attribute **DS_NAME** is the Data Submitter's (DS) Name (the DS may or may not be the same person as the PI or the DO). Sometimes data are processed by and forwarded to the data centre by an additional person or institution. An institution that extracts a subset of the original dataset, may be named a Data Submitter.

ATTENTION

The Data Submitter must be a registered user of the database, either as Principal Investigator or as Co-Investigator. To obtain this status, the DS must submit a signed data protocol to the data centre.

Type;	STRING
Format:	Family name; Given names
Entry:	The entry consists of two fields separated by a semicolon.
Example:	DS_NAME = De Maziere; Martine

A4.1.10 DS_AFFILIATION

The Global Attribute **DS_AFFILIATION** is the Data Submitter's **official** affiliation (he DS_AFFILIATION may differ from the PI_AFFILIATION and DO_AFFILIATION).

Type:	STRING
Format:	Affiliation name; Affiliation acronym
Entry:	The entry consists of two fields separated by a semicolon.
Example:	DS_AFFILIATION = Belgian Institute for Space Aeronomy; BIRA.IASB

A4.1.11 DS_ADDRESS

The Global Attribute DS_ADDRESS is the Data Submitter's mailing address (the DS_ADDRESS may differ from the PI_ADDRESS and DO_ADDRESS). The country name must be the English entry in ISO 3166-1:1997 (ISO, 1997).

Type:	STRING
Format:	Address; Postal code; Country name
Entry:	Three fields separated by semicolons
Example:	DS_ADDRESS = Ringlaan 3; B-1180 Brussels; Belgium

A4.1.12 DS_EMAIL

The Global Attribute **DS_EMAIL** is the Data Submitter's e-mail address (the DO_EMAIL may differ from the PI_EMAIL and the DO_EMAIL).

Type:	STRING
Format:	Free format
Entry:	Single field
Example:	DS_EMAIL = Martine.deMaziere@bira-iasb.oma.be

A4.2 Dataset attributes

The global **Dataset Attributes** provide detailed description of the data contained in the given file. These attributes include the type and identity of the instrument or model, the discipline of the data, a list of the variables included in the file, etc.

A4.2.1 DATA_DESCRIPTION

The Global Attribute **DATA_DESCRIPTION** is a brief sentence describing the data content.

Type:	STRING
Format:	Descriptive text, free format
Entry:	Single field
Example:	DATA_DESCRIPTION= Weekly NILU ozonesonde launch from Orland, Norway.

A4.2.2 DATA_DISCIPLINE

The Global Attribute **DATA_DISCIPLINE** is a character string describing the field of research to which the data in the file belongs. The string refers to the research field and area of the data.

Type:	STRING
Format:	Field; Class; Subclass
Entry:	3 semicolon-separated fields
Example:	DATA_DISCIPLINE = ATMOSPHERIC.CHEMISTRY; INSITU; BALLOON

Table 4.2.2a: *Allowed* ***DATA_DISCIPLINE Field*** *attribute entries. An entry consists of the combination of one of each Field, Class, and Subclass.*

DATA_DISCIPLINE (*Discipline Field*)	***Comment***	***Debate***
ATMOSPHERIC.CHEMISTRY	*Entire atmosphere, chemistry only*	
ATMOSPHERIC.DYNAMICS	*Entire atmosphere, dynamics only*	
ATMOSPHERIC.PHYSICS	*Entire atmosphere, chemistry & dynamics*	
LAND.SURFACE.GEOPHYSICS		
LAND.SURFACE.BIOLOGY	*Covers vegetation and soil characteristics*	*Created for DAISEX demo, December 2003.*
LUNAR.PHYSICS		
OCEANOGRAPHIC.BIOLOGY	*Ocean, biology only*	
OCEANOGRAPHIC.CHEMISTRY	*Ocean, chemistry only*	
OCEANOGRAPHIC.DYNAMICS	*Ocean, dynamics only*	
OCEANOGRAPHIC.PHYSICS	*Ocean, chemistry and dynamics*	
SOLAR.PHYSICS		
STELLAR.PHYSICS		

*Table 4.2.2b: Allowed **DATA_DISCIPLINE Class** attribute entries. An entry consists of the combination of one of each Field, Class, and Subclass.*

DATA_DISCIPLINE (Discipline Class)	Comment
INSITU	
NUMERICAL.SIMULATION	
REMOTE.SENSING	
SAMPLE	

*Table 4.2.2c: Allowed **DATA_DISCIPLINE Subclass** attribute entries. An entry consists of the combination of one of each Field, Class, and Subclass.*

DATA_DISCIPLINE (Discipline Subclass)	Comment
AIRCRAFT	
ASSIMILATION	*data assimilation = combined use of model and experimental data*
BALLOON	
BUOY	
GROUNDBASED	
MODEL	
MOORING	
PLATFORM	*For marine use only*
ROCKET	
SATELLITE	*includes the space shuttle platform*
SHIP	

A4.2.3 DATA_GROUP

The Global Attribute DATA_GROUP is a 2-fields entry, specifying (1) the origin of the data (experimental or model or a combination of both), and (2), the spatial characteristics of the data. The spatial characteristics include the dimensionality of the spatial grid of the dataset for a single data element, in addition to the information whether the ‘footprint’ of the spatial grid varies in space with time, i.e., over the successive data elements.

These concepts are best explained by considering the example of a travelling LIDAR system: At a given point in time, this LIDAR system provides measurements at a single latitude and longitude location but for multiple altitudes. With time, this 1-dimensional spatial grid (fixed latitude and longitude, vector of altitudes), is moving in latitude and longitude. The 2 field entry for this example thus becomes EXPERIMENTAL; PROFILE.MOVING.

NOTE

The dimensionality that is expressed in DATA_GROUP by SCALAR (0D), PROFILE (1D) and FIELD (2D or more) only refers to the spatial dimensionality.

Format: Type; Subtype
Entry: 2 semicolon-separated fields

Example 1: *A timeseries of column measurements from a ground-based instrument will have …*
DATA_GROUP = EXPERIMENTAL; SCALAR.STATIONARY.

Example 2: *A 3D model output on a fixed spatial grid will have …*
DATA_GROUP = MODEL; FIELD.STATIONARY.

Table 4.2.3a: *Allowed* ***DATA_GROUP Type*** *entries. An entry consists of a combination of a Type and a Subtype.*

DATA_GROUP (Group Type)	**Comment**
EXPERIMENTAL	*Measurements*
MIXED	*I.e. assimilation analyses*
MODEL	

Table 4.2.3b: *Allowed* ***DATA_GROUP Subtype*** *entries. An entry consists of a combination of a Type and a Subtype.*

DATA_GROUP (Group Subtype)	**Comment**
SCALAR.MOVING	
SCALAR.STATIONARY	
PROFILE.MOVING	
PROFILE.STATIONARY	
FIELD.MOVING	
FIELD.STATIONARY	

A4.2.4 DATA_LOCATION

The Global Attribute **DATA_LOCATION** is the code of the location, normally based on a fixed location (i.e. a station) or a moving platform name (i.e. a plane, a ship, a buoy, etc.), that the data originates from.

NOTE

Depending on specific campaign policy, the data location for a moving platform (aircraft or ship) may be named after the air strip (where the aircraft is based for the duration of the campaign) or the body of water that the ship is cruising through.

ATTENTION

If the name consists of two or more words, they are separated with periods (.), blanks (space characters) should not occur in the names.

Type: STRING
Format: Refer to Table DATA_LOCATION
Entry: Single field
Example: DATA_LOCATION = ORLAND

Table 4.2.4: *Allowed DATA_LOCATION entry for the Envisat Cal/Val.*

DATA_LOCATION (Location)	Comment	DEBATE	Longitude	Latitude	Elevation
ABERYSTWYTH			-004.1	+052.4	
ADEOS2					
ADRIATIC.SEA					
AIRE.SUR.L.ADOUR	Aire sur l\'Adour				
ALT1	Alert, GPS antenna on Hilton building roof	Specified in ESAG02 documentation.High resolution needed for GPS position.	-62.32675594	82.51143720	56.271
ALT2	Alert, GPS antenna on tripod behid fuel tanks	Specified in ESAG02 documentation.High resolution needed for GPS position.	-62.31546422	82.51110986	42.810
ALOMAR	Alomar, Andøya		+016.0	+069.3	385
ALPILLES					
AMBURLA.SITE1					
ANDENES	Airport, Andøya		+016.2	+069.3	14
ARHANGELSK			+040.5	+068.6	
AROSA			+009.7	+046.8	1840
ARRIVAL.HEIGHTS	Arrival Heights		+166.7	-077.8	190
ATHENS			+023.4	+037.6	
ATLANTIC					
AUSTRALIAN.SEA					
BALTIC.SEA					
BARENTSBURG					
BARRAX	DAISEX study area http://io.uv.es/projects/daisex/	Specified in ESA CDB work statement for DAISEX.			
BAUCE					
BAURU			-049.0	-022.3	300
BELGRANO			-034.6	-077.9	50
BE.130					
BE.230					
BE.MC5					
BERN			+007.5	+047.0	550
BILTHOVEN					

DATA_LOCATION (*Location*)	*Comment*	*DEBATE*	*Longitude*	*Latitude*	*Elevation*
BLANCARES	*DAISEX permanent station http://io.uv.es/projects/daisex/*	*Specified in ESA CDB work statement for DAISEX.*	*-002.1*	*+039.1*	
BLANES					
BRASIL					
BREMEN					
CARIBBEAN					
COLMAR	*http://io.uv.es/projects/daisex/*	*Specified in ESA CDB work statement for DAISEX.*			
COLOMBIA	*Specified for the SAREX campaign*				
COSTA.RICA	*Specified for the SAREX campaign*				
CNP	*Constape Pynt, GPS antenna on roof of Personnel Building*	*Specified in ESAG02 documentation. High resolution needed for GPS position.* *MDB 20040228: Not yet evaluated.* *NILU 20040318: Extended info now in metadata.*	*-22.64819019*	*77.74451119*	*70.770*
DE.BILT	*De Bilt*				
DESERT.ALGERIA.SITE1					
DESERT.ALGERIA.SITE2					
DESERT.ALGERIA.SITE3					
DESERT.ALGERIA.SITE4					
DESERT.ALGERIA.SITE5					
DESERT.ARABIA.SITE1					
DESERT.ARABIA.SITE2					
DESERT.ARABIA.SITE3					
DESERT.EGYPT.SITE1					
DESERT.LIBYA.SITE1					
DESERT.LIBYA.SITE2					
DESERT.LIBYA.SITE3					
DESERT.LIBYA.SITE4					
DESERT.MALI.SITE1					
DESERT.MAURITANIA.SITE1					
DESERT.MAURITANIA.SITE2					
DESERT.NIGER.SITE1					
DESERT.NIGER.SITE2					
DESERT.NIGER.SITE3					
DESERT.SUDAN.SITE1					
DUMONT.D.URVILLE	*Dumont d\'Urville*		*+140.0*	*-066.7*	*20*
DUNHUANG.SITE1					
DYFAMED	*Buoy*				
EGBERT					
EKATERINBURG					
EKRAFANE					
ENGLISH.CHANNEL					
EOS.AQUA	*EOS-AQUA Satellite*				
EOS.AURA	*EOS-AURA Satellite*				
EOS.TERRA	*EOS-TERRA Satellite*				
EP	*Earth Probe satellite*				
ERBS	*Earth Radiation Budget Satellite*				
ERS2	*ESA ERS-2 satellite*				
ESRANGE	*Radar Hill*		*+021.1*	*+067.9*	*485*
EUREKA			*-086.4*	*+080.1*	*610*
FALCON	*DLR Falcon Aircraft*				
FINLAND					
FONTAINEBLEAU	*EMAC site*		*+002.6*	*+048.4*	
FORLI					
FORT.SUMNER	*Fort Sumner*				
FRASCATI	*Specified for the BACCHUS campaign*				
FRENCH.GUIANA	*Specified for the SAREX*				

DATA_LOCATION (*Location*)	*Comment*	*DEBATE*	*Longitude*	*Latitude*	*Elevation*
	campaign				
GAP					
GARDERMOEN					
GARMISCH	*Garmisch-Partenkirchen*				
GEOPHYSICA	*M-55*				
GERMAN.BIGHT					
GILCHING	*EMAC site*		*+011.3*	*+048.2*	
GLOBAL	*Model or satellite global coverage only*				
GOTLAND					
GREENLAND.SITE1					
GSFC	*NASA-GSFC*				
GUYANA	*Specified for the SAREX campaign*				
HALLEY.BAY	*Halley Bay*		*-026.8*	*-075.6*	
HANTY.MANSIYSK					
HARESTUA			*+010.8*	*+060.2*	*580*
HARTHEIM	*http://io.uv.es/projects/daisex/*	*Specified in ESA CDB work statement for DAISEX.*			
HAY.SITE1					
HOBART					
HOHENPEISSENBERG			*+011.0*	*+047.5*	*980*
IGARKA					
INDIAN.OCEAN					
IRKUTSK					
IRSP3	*Indian Satellite IRS-P3*				
ISL.DIKSON					
ISL.HEISA					
ISL.KOTELNIY					
ISSYK.KUL					
IZANA			*-016.5*	*+028.3*	*2367*
JOKIOINEN					
JUNGFRAUJOCH	*International Scientific Station of the Jungfraujoch*		*+008.0*	*+046.6*	*3580*
KARAGANDA					
KARLSRUHE					
KEFLAVIK			*-022.6*	*+064.0*	*38*
KERGUELEN.ISLANDS	*Kerguelen Islands*		*+070.3*	*-049.4*	*10*
KIRUNA			*+020.4*	*+067.8*	*419*
KISLOVODSK			*+042.7*	*+043.7*	
KITT.PEAK			*-111.5*	*+032.0*	*2090*
KRASNOYARSK					
KUS	*Kulusuk airport temporary station, GPS reference*	*ESAG02 Raw Data Report, September 2002.* *MDB 20040228: Request more info, ambiguous name.* *NILU 20040318: Extended info now in metadata.*	*-37.15332542*	*65.57792386*	*72.042*
L.AQUILA	*L\'Aquila*				
LA.REUNION	*Saint-Denis de La Reunion*		*+055.5*	*-020.9*	*10*
LAS.TIESAS.ANCHOR	*Permanent station for for Surface Fluxes and Meteorological Data (http://io.uv.es/projects/daisex/ follow links to "Ground Instruments" and "Permanent Stations")*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040212: Request more info.* *NILU 20040318: Extended info now in metadata.*	*-2.082*	*39.042*	
LAS.TIESAS.LISIMETRO	*Permanent station for for Surface Fluxes and Meteorological Data (http://io.uv.es/projects/daisex/*	*Specified in ESA CDB work statement for DAISEX..*	*-2.090*	*39.058*	

DATA_LOCATION (*Location*)	*Comment*	*DEBATE*	*Longitude*	*Latitude*	*Elevation*
	follow links to "Ground Instruments" and "Permanent Stations")	*MDB 20040212: Request more info* *NILU 20040318: Specified in CDB Specification document from ESA, Appendix A, with different position from LAS.TIESAS.ANCHOR.*			
LAUDER			*+169.7*	*-045.1*	*370*
LEGIONOWO					
LEON					
LOVOSERO					
LULEA	*Radiosonde*		*+022.1*	*+065.6*	
LYR.8.5A	*Svalbard, GPS antenna*	*Specified in ESAG02 documentation. High resolution needed for GPS position.* *MDB 20040228: Ambiguous - what is the full name?* *NILU 20040318: This is the full name used in the ESAG02 Raw data report.*	*15.49307619*	*78.24762997*	*52.516*
LYR.8.5B	*Svalbard, GPS antenna*	*Specified in ESAG02 documentation. High resolution needed for GPS position.* *MDB 20040228: Ambiguous - what is the full name?* *NILU 20040318: This is the full name used in the ESAG02 Raw data report.*	*15.49307719*	*78.24762914*	*52.560*
LYR.9.5	*Svalbard, GPS antenna*	*Specified in ESAG02 documentation. High resolution needed for GPS position.* *MDB 20040228: Ambiguous - what is the full name?* *NILU 20040318: This is the full name used in the ESAG02 Raw data report.*	*15.49307453*	*78.24762906*	*52.550*
MACQUARIE.ISLAND	*Macquarie Island*		*+159.0*	*-054.8*	
MAGADAN					
MALEDIVES					
MARAMBIO					
MARKOVO					
MAUNA.LOA	*Mauna Loa*		*-155.6*	*+019.5*	*3397*
MEDITERRANEAN					
MERIDA					
METEOR.3M	*sattelite*				
METOP1	*sattelite*				
MIR	*Montgolfier InfraRed*				
MONKS.WOOD					
MONTE.CIMONE					
MORETON.BAY	*Moreton Bay*				
MOSCOW			*+037.6*	*+055.8*	
MURMANSK			*+033.1*	*+069.0*	
NAIROBI					

DATA_LOCATION (*Location*)	*Comment*	*DEBATE*	*Longitude*	*Latitude*	*Elevation*
NEUCHATEL			*+007.0*	*+047.0*	*487*
NEUMAYER	*Neumayer Station*		*+008.4*	*-070.6*	
NH	*Northern Hemisphere (model or satellite use only)*				
NH.HIGH.LATITUDE					
NH.LOW.LATITUDE					
NH.MID.LATITUDE					
NIKOLAEVSK					
NOAA14	*Satellite in NOAA TIROS-N program*				
NOAA16	*Satellite in NOAA TIROS-N program*				
NORTH.ATLANTIC					
NORTH.SEA					
NRD1	*Station Nord, GPS antenna on roof of building 7*	*Specified in ESAG02 documentation. High resolution needed for GPS position. MDB 20040228: Ambiguous - what is the full name? NILU 20040318: This is the full name used in the ESAG02 Raw data report.*	*-16.66209092*	*81.60141603*	*70.037*
NRD2	*Station Nord, GPS antenna on roof of building 22 ("Polar2")*	*Specified in ESAG02 documentation. High resolution needed for GPS position. MDB 20040228: Ambiguous - what is the full name? NILU 20040318: This is the full name used in the ESAG02 Raw data report.*	*-16.65691044*	*81.59715722*	*67.514*
NY.ALESUND	*Ny-Ålesund*		*+011.9*	*+078.9*	*15*
O.BORDEAUX	*Observatoire de Bordeaux*		*-000.5*	*+044.8*	*73*
OBERPFAFFENHOFEN					
OBNINSK					
ODIN	*sattelite*				
OHP	*Observatoire de Haute Provence*		*+005.7*	*+043.9*	*679*
OLENEK					
OMSK			*+073.4*	*+054.9*	
OOSTENDE	*EMAC site*		*+002.7*	*+051.3*	
ORLAND	*Ørland*				
OSLO					
PARACOU					
PARAMARIBO					
PAYERN			*+007.0*	*+046.8*	*491*
PECHORA			*+057.1*	*+065.1*	
PENCK	*Ship "Professor Albrecht Penck"*				
PERTH					
PERUGIA					
PETCHORA					
PETROPAVLOVSK					
PLATEAU.DE.BURE	*Plateau de Bure*		*+005.9*	*+044.6*	*2550*
POLARSTERN	*AWI Polarstern research ship*				
POTENZA			*+015.7*	*+040.6*	*820*
PUNTA.ARENAS	*Punta Arenas*				
ROME					
ROTHERA			*-068.1*	*-067.6*	
SALEKHARD			*+066.7*	*+066.5*	*419*
SAMARA					

DATA_LOCATION (*Location*)	*Comment*	*DEBATE*	*Longitude*	*Latitude*	*Elevation*
SAN.PIETRO.CAPOFIUME					
SCO1	*KMS permanent GPS station in Scoresbysund*	*Specified in ESAG02 documentation. High resolution needed for GPS position (missing)* *MDB 20040228: Ambiguous - what is the full name?* *NILU 20040318: This is the full name used in the ESAG02 Raw data report.*			
SCORESBYSUND			*-022.0*	*+070.5*	*10*
SFJ	*Kangerlussuaq, GPS antenna at meteorological hut*	*Specified in ESAG02 documentation. High resolution needed for GPS position.* *MDB 20040228: Ambiguous - what is the full name?* *NILU 20040318: This is the full name used in the ESAG02 Raw data report.*	*-49.29731186*	*67.00601436*	*72.014*
SH	*Southern Hemisphere (model or satellite use only)*				
SH.HIGH.LATITUDE					
SH.LOW.LATITUDE					
SH.MID.LATITUDE					
SIDERADOUGOU					
SODANKYLA	*Sodankÿla*		*+026.7*	*+067.4*	*100*
SONDRESTROMFJORD			*-050.7*	*+067.0*	*180*
SONORASITE1					
SOUTHAMPTON					
SPOT4	*sattelite*				
ST.PETERSBURG					
SUMMIT	*Greenland*		*-038.5*	*+072.6*	*3202*
TABLE.MOUNTAIN	*Table Mountain Facility*		*-117.7*	*+034.4*	*2300*
TARAWA			*+172.9*	*+001.4*	*0*
THANGOO.SITE1					
THESSALONIKI					
THU	*Thule Air Base, GPS antennaon metal rod off Greenland home rule housing building*	*Specified in ESAG02 documentation. High resolution needed for GPS position.* *MDB 20040228: Ambiguous - what is the full name?* *NILU 20040318: This is the full name used in the ESAG02 Raw data report.*	*-68.79667478*	*76.53789506*	*43.884*
THU3	*Thule Air Base, KMS permanent GPS station*	*Specified in ESAG02 documentation. High resolution needed for GPS position.* *MDB 20040228: Ambiguous - what is the full name?* *NILU 20040318: This is the full name used in the ESAG02 Raw data report.*			
THULE			*-068.7*	*+076.5*	*30*

DATA_LOCATION (*Location*)	*Comment*	*DEBATE*	*Longitude*	*Latitude*	*Elevation*
TIKSI					
TINGATINGANA					
TOGO					
TOMELLOSO	*DAISEX study area*	*Specified in ESA CDB work statement for DAISEX.. MDB 20040212: Request more info NILU 20040318: Used in DAISEX data files.*			
TOMELLOSO.ANCHOR	*Permanent station for for Surface Fluxes and Meteorological Data (http://io.uv.es/projects/daisex/ follow links to "Ground Instruments" and "Permanent Stations")*	*Specified in ESA CDB work statement for DAISEX. MDB 20040212: Request more info. NILU 20040318: Specified in CDB Specification document from ESA, Appendix A.*			
TOMSK					
TORONTO			*-079.5*	*+043.8*	*150*
TOWNSVILLE					
TRAPANI					
TROMSO	*EISCAT*		*+019.2*	*+069.6*	
TSUKUBA.NDSC			*140.13*	*36.05*	*+025.0*
TUNISIA	*Specified for the AQUIFER campaign*				
TURA					
UARS	*UARS satellite*				
UCCLE					
VANSCOY			*-106.0*	*+052.0*	
VENEZUELA	*Specified for the SAREX campaign*				
VERNADSKY			*-064.3*	*-065.3*	
VITIM					
VLADIVOSTOK					
VOLGOGRAD					
VORONEZH					
WEILHEIM	*EMAC site*		*+011.1*	*+048.0*	
WMO?????	*TAO Buoy*				
WMO13008	*TAO Buoy*		*-038.0*	*+015.0*	
WMO13009	*TAO Buoy*		*-038.0*	*+008.0*	
WMO13010	*TAO Buoy*		*+000.0*	*+000.0*	
WMO13011	*TAO Buoy*		*-010.0*	*+002.0*	
WMO15001	*TAO Buoy*		*-010.0*	*-010.0*	
WMO15002	*TAO Buoy*		*-010.0*	*+000.0*	
WMO15003	*TAO Buoy*		*-010.0*	*-005.0*	
WMO15004	*TAO Buoy*		*-023.0*	*+000.0*	
WMO15005	*TAO Buoy*		*-010.0*	*-002.0*	
WMO31001	*TAO Buoy*		*-035.0*	*+000.0*	
WMO31002	*TAO Buoy*		*-038.0*	*+004.0*	
WMO32303	*TAO Buoy*		*-095.0*	*+005.0*	
WMO32304	*TAO Buoy*		*-095.0*	*-005.0*	
WMO32305	*TAO Buoy*		*-095.0*	*-008.0*	
WMO32315	*TAO Buoy*		*-110.0*	*+005.0*	
WMO32316	*TAO Buoy*		*-110.0*	*+002.0*	
WMO32317	*TAO Buoy*		*-110.0*	*-002.0*	
WMO32318	*TAO Buoy*		*-110.0*	*-005.0*	
WMO32319	*TAO Buoy*		*-110.0*	*-008.0*	
WMO32320	*TAO Buoy*		*-095.0*	*+002.0*	
WMO32321	*TAO Buoy*		*-095.0*	*+000.0*	
WMO32322	*TAO Buoy*		*-095.0*	*-002.0*	
WMO32323	*TAO Buoy*		*-110.0*	*+000.0*	
WMO41026	*TAO Buoy*		*-038.0*	*+012.0*	
WMO43001	*TAO Buoy*		*-110.0*	*+008.0*	
WMO43301	*TAO Buoy*		*-095.0*	*+008.0*	

DATA_LOCATION (*Location*)	*Comment*	*DEBATE*	*Longitude*	*Latitude*	*Elevation*
WMO46134	*TAO Buoy*				
WMO46146	*TAO Buoy*		*-123.7*	*+049.3*	
WMO51006	*TAO Buoy*		*-140.0*	*+009.0*	
WMO51007	*TAO Buoy*		*-140.0*	*+005.0*	
WMO51008	*TAO Buoy*		*-140.0*	*+002.0*	
WMO51009	*TAO Buoy*		*-140.0*	*-002.0*	
WMO51010	*TAO Buoy*		*-170.0*	*+000.0*	
WMO51011	*TAO Buoy*		*-125.0*	*+000.0*	
WMO51014	*TAO Buoy*		*-140.0*	*-005.0*	
WMO51015	*TAO Buoy*		*-125.0*	*+005.0*	
WMO51016	*TAO Buoy*		*-125.0*	*+002.0*	
WMO51017	*TAO Buoy*		*-125.0*	*-002.0*	
WMO51018	*TAO Buoy*		*-125.0*	*-005.0*	
WMO51019	*TAO Buoy*		*-155.0*	*-005.0*	
WMO51020	*TAO Buoy*		*-155.0*	*+005.0*	
WMO51021	*TAO Buoy*		*-155.0*	*+002.0*	
WMO51022	*TAO Buoy*		*-155.0*	*-002.0*	
WMO51023	*TAO Buoy*		*-155.0*	*+000.0*	
WMO51301	*TAO Buoy*		*-155.0*	*+008.0*	
WMO51302	*TAO Buoy*		*-155.0*	*-008.0*	
WMO51303	*TAO Buoy*		*-170.0*	*+005.0*	
WMO51304	*TAO Buoy*		*-170.0*	*-005.0*	
WMO51305	*TAO Buoy*		*-170.0*	*+002.0*	
WMO51306	*TAO Buoy*		*-170.0*	*-002.0*	
WMO51307	*TAO Buoy*		*-125.0*	*+008.0*	
WMO51308	*TAO Buoy*		*-125.0*	*-008.0*	
WMO51309	*TAO Buoy*		*-170.0*	*+008.0*	
WMO51310	*TAO Buoy*		*-170.0*	*-008.0*	
WMO51311	*TAO Buoy*		*-140.0*	*+000.0*	
WMO52001	*TAO Buoy*		*+165.0*	*+002.0*	
WMO52002	*TAO Buoy*		*+165.0*	*-002.0*	
WMO52003	*TAO Buoy*		*+165.0*	*+005.0*	
WMO52004	*TAO Buoy*		*+165.0*	*-005.0*	
WMO52006	*TAO Buoy*		*+165.0*	*+008.0*	
WMO52007	*TAO Buoy*		*+165.0*	*-008.0*	
WMO52008	*TAO Buoy*		*+156.0*	*+005.0*	
WMO52010	*TAO Buoy*		*+156.0*	*-005.0*	
WMO52011	*TAO Buoy*		*+156.0*	*+002.0*	
WMO52012	*TAO Buoy*		*+156.0*	*-002.0*	
WMO52302	*TAO Buoy*		*+147.0*	*+005.0*	
WMO52307	*TAO Buoy*		*+137.0*	*+002.0*	
WMO52309	*TAO Buoy*		*-180.0*	*+005.0*	
WMO52310	*TAO Buoy*		*-180.0*	*+002.0*	
WMO52311	*TAO Buoy*		*-180.0*	*+000.0*	
WMO52312	*TAO Buoy*		*-180.0*	*-002.0*	
WMO52313	*TAO Buoy*		*-180.0*	*-005.0*	
WMO52315	*TAO Buoy*		*-180.0*	*+008.0*	
WMO52316	*TAO Buoy*		*-180.0*	*-008.0*	
WMO52317	*TAO Buoy*		*+156.0*	*+000.0*	
WMO52318	*TAO Buoy*		*+147.0*	*+000.0*	
WMO52319	*TAO Buoy*		*+156.0*	*+008.0*	
WMO52321	*TAO Buoy*		*+165.0*	*+000.0*	
WMO53001	*TAO Buoy*		*+116.0*	*+018.0*	
WMO53002	*TAO Buoy*		*+114.0*	*+013.0*	
WMO53003	*TAO Buoy*		*+115.0*	*+015.0*	
WOLLONGONG			*+150.9*	*-034.4*	*30*
YAKUTSK			*+129.6*	*+062.0*	
YUKUTSK					
YUZHNO.SAHALINSK					
ZHIGANSK			*+123.4*	*+067.7*	*50*
ZUGSPITZE			*+011.2*	*+047.4*	*2964*
ZVENIGOROD			*+035.8*	*+055.7*	
ZWALM	*EMAC site*		*+003.8*	*+051.0*	

A4.2.5 DATA_SOURCE

The Global Attribute **DATA_SOURCE** consists of three elements. These are the type of instrument or numeric model that created the data (the type may consist of several dot-separated words), the organisation that owns the instrument/model (which may differ from the organisations of the PI, the DO and the DS), and a unique numeric identifier concatenated to the organisation acronym (refer to the Affiliation acronyms in **Table 4.1.2** above).

Each laboratory must assure that no two instruments of the same type have the same identifier, even if they are operated in different locations (a simple number is a sufficient identifier). For example, if NILU acquired a second SAOZ instrument, the entire attribute for NILU's second instrument would become: UVVIS.SAOZ_NILU002

This instrument identification system allows each laboratory to create a worldwide unique identifier for each instrument, without conflict with other laboratories. Any laboratory may operate several instruments of the same type at the same location without identification errors. The instruments may be re-used at different locations, while the instrument history remains traceable. The instruments may be brought to national or international inter-calibration experiments at some common location without naming conflicts. In this particular case, a name is required for each instrument, even if each laboratory has only one. Therefore the naming system must be enforced even for single instruments.

ATTENTION

Instrument names should in general not contain the parameters that it measures. Other metadata entries will ensure that this information is available to the data file users.

RECOMMENDATION

When an instrument is taken out of service, the identifier must not be reused for another instrument.

NOTE

A particular case exists for instruments that are used as "consumables" (for example weather sondes that are often lost after the balloon flight). In such cases a unique identifier may be useless. The identifier 000 is therefore reserved for the NON-UNIQUE case. A laboratory may re-use this particular identifier any number of times.

Type: STRING
Format: Type (from **Table 4.2.5**) and Institute acronym (from **Table 4.1.2**) **concatenated** with a unique 3-digit identifier (for example 001, 007 or 111)
Entry: 2 fields concatenated by an underscore
Example 1: DATA_SOURCE = FTIR_NILU001
Example 2: DATA_SOURCE = UVVIS.SAOZ_NILU002

Table 4.2.5: *Allowed entries for* ***DATA_SOURCE Type*** *in CDB campaigns.*

DATA_SOURCE (*Instrument Type*)	***Comment***	***Debate***
AATSR	*only to be used for non-measurement data like for example the averaging kernels that are derived from theoretical analysis*	*SEN2FLEX adds AATSR data to this database*
AC9		
AHS	*Airborne hyperspectral scanner*	*SEN2FLEX*
AIRFLEX	*Airborne multi-wavelength passive fluorescence detector*	*SEN2FLEX*
AIRMISR		
ALIS		
ALTIMETER.LASER		*Specified in ESAG02 documentation.*
ALTIMETER.RADAR		*Specified in ESAG02 documentation.*
AMON		
AMSR		
AMSU		
ANAIS		
ANEMOMETER	*Wind speed instrument*	*Specified in ESA CDB work statement for DAISEX.*
AOTF	*Acousto-Optic Tunable Filter*	*SEN2FLEX*
APEX		
ASUR		
ATMOINSPECTOR	*New nstrument name CHILD, database uses old name ATMOINSPECTOR*	
ATSR2		
AUTOCHEM	*Chemical data assimilation by UCAMB.CHEM*	
AVHRR		
BAROMETER		*Specified in ESA CDB work statement for DAISEX.*
BB4		
BLACKBODY.EVEREST.1000	*Calibration source for IR thermometer*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: Should not refer to instrument/location in DATA_SOURCE* *NILU 20040318: Already common as extension (after first dot), like in UVVIS.BREWER, PHOTOMETER.PERKINELMER, and as entire name for scientific, "one-off" instruments or models.*
BMP	*Biospherical Multiband Profiler for Subsurface Ed/Lu and R measurements*	
BUCKET.EVAPORIMETRIC		*Specified in ESA CDB work statement for DAISEX.*
BUOY.SST.DRIFTER	*Sea Surface temperature buoy, drifting*	
BUOY.SST.FIXED	*Sea Surface temperature buoy, fixed position*	
BUOY.TAO	*Tropical Atmosphere Ocean Buoy*	
CAESR		
CAMERA		*SEN2FLEX*
CAMERA.VIDEO		*SEN2FLEX*
CASI		

DATA_SOURCE (*Instrument Type*)	*Comment*	*Debate*
CASI3	*Airborne hyperspectral scanner, 400-1050 nm*	*SEN2FLEX*
CCM-200	*Chlorophyll Content Meter*	*SEN2FLEX*
CEILOMETER		
CH4TDL		
CHLOROPHYLL.FLUORESCENCE.PROFILER	*Chlorophyll Fluorescence Profiler*	
CHRIS.PROBA	*Compact high resolution imaging spectrometer on PROBA satellite*	*SEN2FLEX*
COPAS		
CRYOSAMPLER		
CTD		
CYCLOMETER		
CYTOMETERS		
DATALOGGER.CR10		*Specified in ESA CDB work statement for DAISEX. MDB 20040228: Too generic NILU 20040318: The real sources vould be VOLTMETER, GPS, etc. (also generic). The DATALOGGER is often the most tangible source for a group of diverse signals logged during a mission. The source SHOULD be as generic as possible. In addition we have suggested a specific extension to identify one of several logger types used in the same campaign.*
DATALOGGER.CR500		*Specified in ESA CDB work statement for DAISEX. MDB 20040228: Too generic NILU 20040318: The real sources vould be VOLTMETER, GPS, etc. (also generic). The DATALOGGER is often the most tangible source for a group of diverse signals logged during a mission. The source SHOULD be as generic as possible. In addition we have suggested a specific extension to identify one of several logger types used in the same campaign.*
DESCARTES		
DAIS7915	*Digital Airborne Imaging Spectrometer, 79 channels 400nm to 12.3um*	*Specified in ESA CDB work statement for DAISEX. MDB 20040228: Commercial name? What is the physical basis? NILU 20040318: Could change name to SPECTROMETER.DAIS7915. Commercial names are commonly used elsewhere in these metadata.*
ECMWFMODEL.GOMOS		
ECMWFMODEL.MIPAS		
ECMWFMODEL.SCIAMACHY		
ECOC		
ELHYSA		
EMISSIVITY.BOX.1	*Thermal remote sensing unit*	*Specified in ESA CDB work statement for DAISEX. MDB 20040228: BOX.1 not allowed NILU 20040318: Correct comment, could use EMISSIVITY.BOX, and leave the 1 and 2 to the last element in the naming convention in 4.2.5. If NILU has 2 emissivity boxes, they would be named EMISSIVITY.BOX_NILU001 and EMISSIVITY.BOX_NILU002.*

DATA_SOURCE (*Instrument Type*)	*Comment*	*Debate*
EMISSIVITY.BOX.2	*Global change unit*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: BOX.1 not allowed* *NILU 20040318: Correct comment, could use EMISSIVITY.BOX, and leave the 1 and 2 to the last element in the naming convention in 4.2.5. If NILU has 2 emissivity boxes, they would be named EMISSIVITY.BOX_NILU001 and EMISSIVITY.BOX_NILU002.*
ESAR	*Expirimental multi-frequency, dual-polarisation SAR, designed for medium-sized turboprop aircraft.*	
FAR.IR.INTERFEROMETER	*Far Infrared Interferometer*	
FIELSPEC.PRO.FR	*Solar irradiance direct component*	*SEN2FLEX*
FILTRATION		
FIRS2		
FISH	*Airborne alpha-Lyman Hygrometer (balloon)*	
FLUORIMETER		
FOZAN		
FSSP		
FTIR	*Infrared Fourier Transform Spectrometer*	
FTS	*Fourier Transform Spectrometer (UV + IR)*	
GALAI	*Calibration source ULP GALAI 204-P*	*SEN2FLEX*
GASCOD		
GOME	*ESA ERS-2 satellite instrument*	
GOME2		
GOMOS	*only to be used for non-measurement data like for example the averaging kernels that are derived from theoretical analysis*	
GONIOMETER	*Instrument for angle measurements*	*Specified in ESA CDB work statement for DAISEX.*
GPS	*Global Positioning System receiver*	*Specified in ESA CDB work statement for DAISEX.*
GPS.AIR1	*Trimble airborne GPS receiver on forward antenna in ESAG02*	*Specified in ESAG02 documentation..* *MDB 20040228: Not yet evaluated.* *NILU 20040318: Extended info now in metadata. May drop the .AIRx part, and add the internal serial number after the owner isntitution acronym. The comments would then need to be entered in each file, instead of in the central metadata.*
GPS.AIR2	*Ashtec airborne GPS receiver on forward antenna in ESAG02*	*Specified in ESAG02 documentation..* *MDB 20040228: Not yet evaluated.* *NILU 20040318: Extended info now in metadata. May drop the .AIRx part, and add the internal serial number after the owner isntitution acronym. The comments would then need to be entered in each file, instead of in the central metadata.*
GPS.AIR3	*Javad airborne GPS receiver on aft antenna in ESAG02*	*Specified in ESAG02 documentation..* *MDB 20040228: Not yet evaluated.* *NILU 20040318: Extended info now in metadata. May drop the .AIRx part, and add the internal serial number after the owner isntitution acronym. The comments would then need to be entered in each file, instead of in the central metadata.*
GRAVIMETER.LCR	*Airborne gravimeter. Primary LCR data and auxiliary data (platform stabilization, etc) logged on laptop during ESAG02*	*Specified in ESAG02 documentation..* *MDB 20040228: Not yet evaluated.*
HAGAR		

DATA_SOURCE (*Instrument Type*)	*Comment*	*Debate*
HALOE		
HALOX		
HIRDLS		
HPLC		
HUMIDITY.SENSOR		
HY2P	*NILU ECMWF T106 Analysis extraction data on isobaric model levels*	
HY2TH	*NILU ECMWF T106 Analysis extraction data on isentropic model levels*	
HYDROSCAT	*Backscattering measurements*	
HYGROMETER	*Relative humidity*	*Specified in ESA CDB work statement for DAISEX.*
HYMAP	*HyMap Imaging Spectrometer, whisk-broom scanner, 400nm to 2500 nm in 125 bands*	*Specified in ESA CDB work statement for DAISEX.* *NILU 20040328: Could be named SPECTROMETER.HYMAP ?*
IMU	*Inertial Measurement Unit*	*Specified in ESAG02 documentation..* *MDB 20040228: What is it?* *NILU 20040318: Description in Comment field.*
INS	*Inertial Navigating System*	*Specified in ESAG02 documentation..* *MDB 20040228: What is it?* *NILU 20040318: Description in Comment field.*
INS.H764G	*Honeywell H764-G EGI (inertial navigation system)*	*Specified in ESAG02 documentation..* *MDB 20040228: What is it?* *NILU 20040318: Description in Comment field.*
IRRADIANCE.SENSOR		
IRTDL		
ISAMS	*Improved Strat. And Mesos. Sounder aboard UARS*	
ISAR	*Infrared Sea surface temperature Autonomous Radiometer*	
IUE		
LABS		
LAI	*Leaf Area Index measurements*	*Defined under Cal/Val, not used by 20040318.*
LEAF.CHAMBER.ANALYSER	*Portable infrared gas analyzer (model LCA-4, ADC Ltd.)*	*Specified in ESA CDB work statement for DAISEX.*
LCTF		*SEN2FLEX*
LICOR1800UW	*Spectroradiometer for Subsurface Ed and Eu Measurements*	*Defined under Cal/Val, not used by 20040318.*
LICOR.LAI.2000	*Plant Canopy Analyser, non-destructive Leaf Area Index (LAI) measurements*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: Commercial name? What is the physical basis?* *NILU 20040318: Comment added. Could use the LAI entry, and leave the LICOR LAI-2000 name for the free-text comments in the HDF file?*
LIDAR.BACKSCATTER		
LIDAR.DIAL		
LIDAR.OLEX	*Airborne LIDAR (DLR Falcon)*	

DATA_SOURCE (*Instrument Type*)	*Comment*	*Debate*
LIDAR.RIEGL	*Riegl Scanning Lidar*	*Specified in ESAG02 documentation.* *MDB 20040228: Riegel is commercial name? What is the physical basis?* *NILU 20040318: Commercial names already common as extension (after first dot), like in UVVIS.BREWER, PHOTOMETER.PERKINELMER, and as entire name for scientific, "one-off" instruments or models.*
LIDAR.RMR	*Rayleigh-Mie-Raman Lidar*	
LYSIMETER.HERBAL.CROP		*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: Not yet evaluated*
LYSIMETER.LIGNEOUS.CROP		*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: Not yet evaluated*
LYSIMETER.REFERENCE	*Instrument for determining solubility, esp. the amount of water-soluble matter in soil. (http://www.wordreference.com/english/definition.asp?en=lysimeter)* *A lysimeter is essentially a large, stainless steel box or cylinder which is filled with soil, open on the top, and closed at the bottom so all liquid that runs through it can be collected. (http://extoxnet.orst.edu/newsletters/n81_88.htm)*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: Explain* *NILU 20040318: Corrected spelling to English LYSIMETER, added explanatory references*
LPMA	*Balloon-borne experiment operated by LPMA*	
MACSIMS		
MAERI		
MAS		
MERIS	*only to be used for non-measurement data like for example the averaging kernels that are derived from theoretical analysis*	*SEN2FLEX adds MERIS data to this database*
METEOSAT		
MICROWAVE.RADIOMETER		
MIPAS	*only to be used for non-measurement data like for example the averaging kernels that are derived from theoretical analysis*	
MIPAS.B	*MIPAS on balloon*	
MIPAS.STR	*MIPAS on ?*	
MISR		
MIR	*Montgolfier InfraRed*	
MLS		
MODIS	*Moderate ResolutionImaging Spectrometer*	*SEN2FLEX*
MOPITT	*EOS-TERRA Satellite Instrument*	
MOS	*Modular Optoelectronic Scanner (on IRS-P3)*	
MSDOL	*ACRI model*	
MSDOL.ATMOS		
MSDOL.GOMOS		
MSDOL.MIPAS		
MSDOL.SCIAMACHY		
MSDOL.SMR		
MSX		
MVIRI		
OMI	*Ozone satellite instrument*	
OPC		
OPER		
OSIRIS		
OVID		
PARABOLA		
PHOTOMETER		
PHOTOMETER.CIMEL	*Sun-photometer, sky radiance mostly applied to aerosol optical thickness*	
PHOTOMETER.CIMEL.SEAPRISM	*Sun-photometer, radiometric data of sea surface in visible and near infrared; water leaving radiance data*	

DATA_SOURCE (*Instrument Type*)	***Comment***	***Debate***
PHOTOMETER.PERKINELMER		
PHOTOMETER.SUN		
PHOTOMETER.SUN.MICROTOPS.II		*Specified in ESA CDB work statement for DAISEX.*
PHOTOMETER.SUN.REAGAN		*Specified in ESA CDB work statement for DAISEX.*
PLANKTONNET		
PLUVIOMETER	*Precipitation amount*	*Specified in ESA CDB work statement for DAISEX.*
POAM3		
POLDER	*POLarization and Directionality of Earth Reflectances, spectral bands at 443, 500, 550, 590, 670, 700, 720, 800, 864 nm*	*Created for Cal/Val, but not used by 20040318. Needed for DAISEX demo, December 2003*
PROFILOMETER.PIN	*Soil roughness measurements*	*Specified in ESA CDB work statement for DAISEX.*
PSICAM		
PYGIOMETER		
PYRANOMETER		*Created for Cal/Val, but not used by 20040318. Needed for DAISEX demo, December 2003*
RADAR	*Rain radar*	
RADAR.PROFILER	*Windprofiler, MST radar*	
RADIANCE.SENSOR.UPWELLING		
RADIOMETER.BIOSPHERICAL		
RADIOMETER.IR.CIMEL	*CIMEL 312-1, Channel 1: 8 - 13 um, Channel 2: 11.5 - 12.5 um, Channel 3: 10.5 - 11.5 um, Channel 4: 8.2 - 9.2 um. CIMEL 312-2 ASTER slightly different bands*	*Specified in ESA CDB work statement for DAISEX. NILU 20040328: Added .IR in name Use for 312-1 and 312-2, specify type in COMMENTS*
RADIOMETER.IR	*Infrared radiometer (thermometer) typically covering the 8 um to 14 um band.*	*Replaces RADIOMETER.IR.RAYTEK, RADIOMETER.IR.OMEGA, THERMOMETER.IR.EVEREST, THERMOMETER.IR.EVEREST.3400.4ZLC. SPECIFY TYPE IN COMMENTS*
RADIOMETER.SIMBADA		
RADIOMETER.SATLANTIC		
RADIOMETER.TRIOS	*Radiometer UV-A / UV-B / PAR, 280 nm to 720 nm*	*Created and used for Cal/Val. Also specified for DAISEX*
RAMSES	*Hyperspectral Profiler for Subsurface Ed/Lu and R measurements*	*Created for Cal/Val, but not used by 20040318. May be identical to the Trios Radiometer? http://www.trios.de/start.html*
REFLECTOMETER	*For hydric soil content*	*Specified in ESA CDB work statement for DAISEX.*
ROSIS	*Reflective Optics System Imaging Spectrometer, compact airborn, 84 bands in spectral mode, 32 bands in imaging mode from 430nm to 850nm*	*Specified in ESA CDB work statement for DAISEX.*
SABER		
SAFIREA		
SAGE2		
SAGE3		
SALOMON		
SAMPLE.GAS		
SAMPLE.LIQUID		
SATLANTICSENSOR		
SAW		
SBUV2		
SCIAMACHY	*only to be used for non-measurement data like for example the averaging kernels that are derived from theoretical analysis*	
SCINTILLOMETER		
SDLA	*Tunable Diode Laser Spectrometer*	
SEA.ATM.STATE	*placeholder for MAVT aux info*	
SECCHIDISC		
SIMBAD		
SIOUX		

DATA_SOURCE (*Instrument Type*)	*Comment*	*Debate*
SISTER		
SMR		
SMSR	*SeaWiFS Multichannel Surface Reference*	
SOAP		
SODAR	*Windprofiler, sonar principle*	
SOLSPEC		
SOLSTICE2		
SONDE.BACKSCATTER		
SONDE.AIR	*AS-1C-PTH*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: Same as SONED.PTU?* *NILU 20040318: Wright - there seems to be no distincion between the PTH and PTU terms.*
SONDE.O3	*Like Vaisala RS80 ozone*	*Created and used for Cal/Val.* *Also specified for DAISEX*
SONDE.PTU	*PTU sonde, Pressure, Temperature, Humidity (also sometimes referred to as PTH sonde). Carried by balloon, or used as drop sonde. Like Vaisala RS80 radiosonde series.*	*Created and used for Cal/Val.* *Also specified for DAISEX*
SPAD	*Use for SPAD-502, etc (Chlorophyll Content Meter*	
SPECTRALON	*Reference panel. Spectralon is the commercial name for the reflective covering material.*	*Specified in ESA CDB work statement for DAISEX.*
SPECTROMETER		*Specified in ESA Cal/Val (all Cal/Val files could have used SPECTROMETER.IR), also needed for DAISEX demo in CDB*
SPECTROPHOTOMETER		
SPECTRORADIOMETER		
SPECTRORADIOMETER.OL754.PMT	*Optronics Spectroradiometer with PMT monochromator, 200nm - 800 nm*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: What is OL754* *NILU 20040318: Changed from OL754 to OL754.PMT. OL754 is the instrument commercial name, PMT the monochromator type.*
SPECTRORADIOMETER.LICOR1800	*Li-Cor Spectroradiometer with PMT monochromator, 300 nm to 1100 nm range, 6 nm bandwidth.*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: What is OL754* *NILU 20040318: Li-Cor 1800 is the company and instrument commercial name, PMT the monochromator type.*
SPECTRORADIOMETER.GER1500	*Portable spectroradiometer 350 nm - 1050 nm in 512 channels (http://www.ger.com/)*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: What is GER* *NILU 20040318: Corrected from RADIOMETER to SPECTRORADIOMETER. GER is the company name, GER 1500 the instrument name*
SPECTRORADIOMETER.GER2600	*Portable spectroradiometer 350 nm - 2500 nm in 640 channels (http://www.ger.com/)*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: What is GER* *NILU 20040318: Corrected from RADIOMETER to SPECTRORADIOMETER. GER is the company name, GER 2600 the instrument name*

DATA_SOURCE (*Instrument Type*)	*Comment*	*Debate*
SPECTRORADIOMETER.GER3700	*Portable spectroradiometer 350 nm - 2500 nm in 704 channels (http://www.ger.com/)*	*Specified in ESA CDB work statement for DAISEX.* *MDB 20040228: What is GER* *NILU 20040318: GER is the company name, GER 3700 the instrument name*
SPEXTUBE		
SPIRALE		
SPMR	*SeaWiFS Profiling Multichannel Radiometer*	
SSBUV		
SSC		
SSM		
SUSIM		
TDR	*Time Domain Reflectometry, Tektronics Cable tester 1502C or similar*	
TES		
THERMOCOUPLE.TYPEK	*Type K element*	*Requested for the Sparc2004 campaign*
THERMOMETER		*Specified in ESA CDB work statement for DAISEX.*
TM	*Thematic Mapper, Multispectral scanning radiometer on Landsat*	*SEN2FLEX*
TOMS		
TOVS		
TRIOS	*Radiance-Irradiance Spectrometer*	
TRIPLE		
TYCHO		
UNIFIEDMODEL.GOMOS	*UK Met Office Unified Model*	
UNIFIEDMODEL.MIPAS	*UK Met Office Unified Model*	
UNIFIEDMODEL.SCIAMACHY	*UK Met Office Unified Model*	
UVVIS	*UV-visible spectrometer*	
UVVIS.AMAXDOAS	*Airborne DOAS, Cooperation between Universities of Bremen and Heidelberg*	
UVVIS.BREWER		
UVVIS.DOAS		
UVVIS.DOBSON		
UVVIS.FOX		
UVVIS.GUV		
UVVIS.NILUV		
UVVIS.OFFAXIS		
UVVIS.SAOZ		
VEGETATION		

A4.2.6 DATA_TYPE

The Global Attribute **DATA_TYPE** specifies the data time resolution and the data product level. The identifiers are **concatenated into one field**.

Type: STRING, maximum 2 characters
Format: Time Scale Code + Data Level Code
Entry: Single concatenated entry
Example: DATA_TYPE = H2 … *is hourly level 2 data*

Table 4.2.6a: *Time Scale Codes to construct the DATA_TYPE attribute entry. The attribute entry is built by concatenating the Time Scale Code with a Data Level Code.*

DATA_TYPE (Time Scale Code)	Comment
D	*Daily*
H	*Hourly*
M	*Minutes*
S	*Seconds*
O	*Other*

Table 4.2.6b: *Data Level Codes to construct the DATA_TYPE attribute entry. The attribute entry is built by concatenating the Time Scale Code with a Data Level Code.*

DATA_TYPE (Data Level Code)	Comment
0	*Reformatted, time-ordered instrument data*
1	*Geolocated, radiometrically and/or spectrally calibrated instrument data*
2	*Extracted geolocated geophysical data*
3	*Added-value/derived geophysical data, typically gridded data*
4	*Assimilated geophysical data*

A4.2.7 DATA_VARIABLES

The Global Attribute **DATA_VARIABLES** lists the variables, such as the chemical compounds or physical parameters, found in the current data file. This entry contains one field for each variable. Each field consists of the variable name, the variable mode and the variable descriptor (underscore separated). DATETIME, ALTITUDE, LATITUDE and LONGITUDE variables are normally modeless. All other parameters should preferably have a mode, but it is not mandatory. The descriptor is used only when required. Some entries may be subdivided by dots where required (but only in the exact manner stated in the Table 4.2.7 a, b, or c below).

The variable **name** is a basic declaration of the measurable described in the dataset, i.e. the physical property of the measurement subject that is measured or computed by a model. The name includes the chemical or physical identification of the measurement subject. A typical example of a variable name is the concentration of ozone:

```
O3.CONCENTRATION
```

Stringent naming criteria apply to those **independent variables that specify geolocation**. Every datafile must contain a specification of geolocation in four dimensions. In addition to the DATETIME variable, latitude, longitude and a vertical geolocation parameter are mandatory.

- The vertical geolocation should be expressed as ALTITUDE or DEPTH.

- If ALTITUDE is not available, acceptable substitutes are PRESSURE and ALTITUDE.GPH (Geo-Potential Height).

The geolocation provided should specify the location where the measurement variables are sampled (when possible). Only in the event that this information cannot be provided is it acceptable to provide the instrument location with auxiliary information that allows to derive the location of the sampling. In this case the label ".INSTRUMENT" is to be appended to the geolocation parameters. For example:

```
LATITUDE.INSTRUMENT; LONGITUDE.INSTRUMENT;
ALTITUDE.INSTRUMENT.
```

ATTENTION

The mode and the descriptor parts discussed below do not apply to the geolocation variables.

ACCEPTABLE COMBINATIONS OF MANDATORY DATA

1. DATETIME; ALTITUDE; LATITUDE; LONGITUDE

2. DATETIME; ALTITUDE.GPH; LATITUDE; LONGITUDE

3. DATETIME; PRESSURE; LATITUDE; LONGITUDE

4. DATETIME; DEPTH; LATITUDE; LONGITUDE

5. DATETIME; ALTITUDE.INSTRUMENT; LATITUDE.INSTRUMENT; LONGITUDE.INSTRUMENT
(Please provide relevant auxiliary parameters)

6. DATETIME; DEPTH.INSTRUMENT; LATITUDE.INSTRUMENT; LONGITUDE.INSTRUMENT
(Please provide relevant auxiliary parameters)

The **mode** is the context in which the entity is described and should normally be decleared. The mode should contain the information on the measurement method that can lead to differences when comparing with other methods to observe the same quantity. Exceptions are those categories of differences that are already present elsewhere in the metadata, for example the REMOTE.SENSING data are already distinguished from SAMPLE or INSITU in the entry DATA_DISCIPLINE. We may construct several examples compliant with tables 4.2.7a and 4.2.7b where we add typical modes to the ozone variable name:

```
O3.COLUMN_SLANT.SOLAR
O3.COLUMN_VERTICAL.SOLAR
```

Descriptors are needed only when a property is variable over the dataset. As an example, the descriptor DETECTIONLIMIT is used to construct a variable that contains the changing detection limits for a series of measurements. A constant detection limit (or any other static, descriptive information) should be specified in a comment (see sections VAR_DESCRIPTION and VAR_NOTES), and not as a descriptor variable. The descriptor is added only to construct auxiliary variables that describe some particular property of a primary variable (such as the last variable entry H2O_COLUMN_ERROR in the example below). We can create additional examples using the ozone + mode examples above:

```
O3.COLUMN_ SLANT.SOLAR_ UNCERTAINTY.STDEV
O3.COLUMN_ VERTICAL.SOLAR_UNCERTAINTY.STDEV
```

NOTE

The descriptor is not intended to distinguish subsets of a dataset. Such distinctions should be made by providing additional dependent or independent parameters, as outlined in the following examples.

The ozone column obtained by SAOZ measurements are traditionally distinguished in two subsets: measurements at dawn and measurements at dusk. The solar azimuth angle is the parameter is the relevant basis for distinction of these measurements and should be provided together with every measurement of the ozone column.

Irradiance measurements are often performed at specific wavelengths. Wavelength should therefore be an independent parameter if values at more than one wavelength are reported

Water samples are often performed at three depths with optical thickness parameter (DEPTH.SECCHI) 0, 0.5 and 1.0 respectively. Parameters retrieved from these samples and the optical thickness parameter should all be reported as functions of the independent parameter DEPTH.

Variable names, modes, descriptors and units are case sensitive. Please observe the exact capitalisation given in the tables below.

ATTENTION

The combination of a variable name, mode and descriptor must be unique. If the exact combination you need is not yet listed in the table, please contact the authors of this metadata document to declare the combination and assign an appropriate default measurement unit.

Type: STRING
Format: Variable name_Variable mode_Variable descriptor
Entry: Multiple semicolon separated fields (each field constructed according to the format above)
Example: DATA_VARIABLES = DATETIME; LATITUDE; LONGITUDE; ALTITUDE; O3.CONCENTRATION_VERTICAL.SOLAR; H2O.COLUMN_VERTICAL.SOLAR; H2O.COLUMN_VERTICAL.SOLAR_ERROR

Table 4.2.7a: *Allowed DATA_VARIABLES (combinations of Variable Name, Variable Mode and Variable Descriptor).*

DATA_VARIABLES (*Variable Name*)	***Comment***	***Debate***
ABSORPTION.COEFFICIENT		
ACCELLERATION.LINEAR	*Use with modesX, Y or Z*	*Indicated in ESAG02 documentation*
ACCELLERATION.ANGULAR	*Use with modes PITCH, ROLL or YAW*	*Indicated in ESAG02 documentation*
AEROSOL.BACKSCATTER.COEFFICIENT	*Aerosol/cloud backscatter coefficient*	
AEROSOL.BACKSCATTER.RATIO	*Aerosol/cloud Backscatter Ratio*	
AEROSOL.COLOUR.RATIO		
AEROSOL.COLUMN		
AEROSOL.CONCENTRATION	*Aerosol/cloud*	
AEROSOL.DEPOLARIZATION.RATIO	*Aerosol/cloud Depolarization Ratio*	
AEROSOL.EPSILON		
AEROSOL.EXTINCTION.COEFFICIENT	*Aerosol/cloud Extinction Coefficient*	
AEROSOL.EXTINCTION.RATIO	*Aerosol/cloud Extinction Ratio*	
AEROSOL.LIDAR.RATIO	*Aerosol/cloud extinction coefficient over backscatter coefficient*	
AEROSOL.OPTICAL.DEPTH	*Aerosol/cloud Optical Depth*	
AIR.CONCENTRATION	*Air density*	
AIR.MASS.FACTOR		
ALBEDO		
ALTITUDE	*(Modeless)*	
ALTITUDE.GPH	*Geopotential height*	
ALTITUDE.INSTRUMENT	*Altitude of the instrument (Modeless)*	
ALTITUDE.SURFACE	*Altitude of Lake Surface*	
ANGLE	*May be used with modes AZIMUTH and ZENITH, and sometimes with descriptor OFFSET, START, END, MEAN, DELTA or DELTA2CAL.MEAN*	*Indicated in DAISEX documentation. See file /GCUDATA/ANGLES.TXT in archive http://io.uv.es/projects/daisex/database/DB-Daisex99/Temperature_99.zip*
ANGLE.CORSP	*Corresponding angle, corrected for magnetic declination from geographical north. May be used with modes AZIMUTH and ZENITH, and sometimes with descriptor OFFSET, START, END, MEAN, DELTA or DELTA2CAL.MEAN*	*Indicated in DAISEX documentation. See file /GCUDATA/ANGLES.TXT in archive http://io.uv.es/projects/daisex/database/DB-Daisex99/Temperature_99.zip*
ANGLE.ALA	*Average Leave Inclination Angle in degrees*	
ANGLE.LUNAR		
ANGLE.SOLAR		
ANGLE.STELLAR		
ANGLE.VIEW	*View Angle, Line of Sight Angle*	
ATMOSPHERIC.TRANSMISSION		
ATTITUDE.PITCH	*Instrument attitude relative to global or platform coordinate system*	*NILU 20040324: May need to combine with mode to describe if we reference a local platform coordinate system or a global coordinate system. Therefore include PITCH in the name instead of as a mode.*

DATA_VARIABLES (*Variable Name*)	*Comment*	*Debate*
ATTITUDE.ROLL	*Instrument attitude relative to global or platform coordinate system*	*NILU 20040324: May need to combine with mode to describe if we reference a local platform coordinate system or a global coordinate system. Therefore include ROLL in the name instead of as a mode.*
ATTITUDE.YAW	*Instrument attitude relative to global or platform coordinate system*	*NILU 20040324: May need to combine with mode to describe if we reference a local platform coordinate system or a global coordinate system. Therefore include YAW in the name instead of as a mode.*
B.PHASE.FUNCTION		
BACKSCATTERING.COEFFICIENT		
BAROMETRIC.PRESSURE		
BEAM.ATTENUATION.COEFFICIENT		
BEAM.POSITION	*Platform stabilisation data for ESAG0. Combine with mode X, Y, Z, PITCH, ROLL YAW if appropriate.*	*Indicated in DAISEX documentation for LCR gravimetry.*
BIOMASS.DRY		*SEN2FLEX*
BIOMASS.WET		*SEN2FLEX*
BPA	*Bleached particle absorption*	
Br.COLUMN		
Br.CONCENTRATION		
Br2.COLUMN		
Br2.CONCENTRATION		
BrCl.COLUMN		
BrCl.CONCENTRATION		
BrO.COLUMN		
BrO.CONCENTRATION		
BrONO.COLUMN		
BrONO.CONCENTRATION		
BrONO2.COLUMN		
BrONO2.CONCENTRATION		
C2H2.COLUMN		
C2H2.CONCENTRATION	*Acetylene*	
C2H6.COLUMN		
C2H6.CONCENTRATION	*Ethane*	
CFC11.COLUMN		
CFC11.CONCENTRATION	*CFC11 == CFCl3*	
CFC12.COLUMN		
CFC12.CONCENTRATION	*CFC12==CF2Cl2*	
CH2O.COLUMN		
CH2O.CONCENTRATION		
CH3.COLUMN		
CH3.CONCENTRATION		
CH3Br.COLUMN		
CH3Br.CONCENTRATION		
CH4.COLUMN		
CH4.COLUMN.AMF	*air-mass factor*	
CH4.CONCENTRATION	*Methane*	
CH4.CONCENTRATION.AMF	*air mass factor*	
CH4.CONCENTRATION.AVK	*averaging kernel*	
CHL.1.CONCENTRATION		
CHL.1.INDEX	*Algal pigment index valid in Case 1 waters*	
CHL.2.CONCENTRATION		
CHL.2.INDEX	*Algal pigment index valid in Case 2 waters*	
CHL.A.CONCENTRATION	*Chlorophyll*	*Specified in Cal/Val. Also needed for DAISEX.*
CHL.A.INDEX	*Chlorophyll*	*Specified in Cal/Val. Also needed for DAISEX.*
CHL.B.CONCENTRATION	*Chlorophyll*	*Specified in DAISEX data files.*
CHL.B.INDEX	*Chlorophyll*	*Specified in DAISEX data files.*
CHL.TOTAL.CONCENTRATION	*Chlorophyll*	*Specified in DAISEX data files.*
CHL.TOTAL.INDEX	*Chlorophyll*	*Specified in DAISEX data files.*
CHL.FLUORESCENCE	*Chlorophyll-Fluorescence*	
Cl.COLUMN		
Cl.CONCENTRATION	*Chlorine*	
Cl2.COLUMN		
Cl2.CONCENTRATION		

DATA_VARIABLES (*Variable Name*)	*Comment*	*Debate*
Cl2O2.COLUMN		
Cl2O2.CONCENTRATION		
ClO.COLUMN		
ClO.CONCENTRATION	*(Do not confuse the small l with a capital I)*	
ClONO.COLUMN		
ClONO.CONCENTRATION		
ClONO2.COLUMN		
ClONO2.CONCENTRATION		
ClOO.COLUMN		
ClOO.CONCENTRATION		
CLOUD.BOTTOM.HEIGHT	*Cloud Bottom Height*	
CLOUD.BOTTOM.PRESSURE	*Cloud Base Pressure*	
CLOUD.CONDITION	*Text entries only*	
CLOUD.COVER	*Cloud Cover*	
CLOUD.DROPLET.EFFECTIVE.RADIUS	*Cloud droplet effective radius (ref)*	
CLOUD.DROPLET.NUMBER.CONCENTRATION	*Cloud droplet number concentration*	
CLOUD.LAYER.HEIGHT		
CLOUD.LAYER.THICKNESS		
CLOUD.LAYER.TRANSMISSION		
CLOUD.OPTICAL.THICKNESS	*Cloud Optical Thickness*	
CLOUD.TOP.HEIGHT	*Cloud Top Height*	
CLOUD.TOP.PRESSURE	*Cloud Top Pressure*	
CLOUD.TYPE	*WMO codes*	
CN.COLUMN		
CN.CONCENTRATION		
CO.COLUMN		
CO.COLUMN.AMF	*air-mass factor*	
CO.CONCENTRATION	*Carbon monoxide*	
CO.CONCENTRATION.AMF	*air mass factor*	
CO.CONCENTRATION.AVK	*averaging kernel*	
CO2.COLUMN		
CO2.CONCENTRATION	*Carbon dioxide*	
COF2.COLUMN		
COF2.CONCENTRATION		
COLOUR.INDEX	*Colour index f550/f350 after molecular absorption correction*	
COLOUR.RATIO		
CONDUCTIVITY		
CROP.SEED.DENSITY.MASS	*Kg seeds/ha*	*Specified in DAISEX vegetation data files.*
CROP.SEED.DENSITY.PLANTS	*Number of seeds/ha*	*Specified in DAISEX vegetation data files.*
CROP.ROW.SPACING	*Row spacing in meters*	*Specified in DAISEX vegetation data files.*
CROP.ROW.DIRECTION		*SEN2FLEX*
CROP.PLANT.DENSITY		*SEN2FLEX*
CROP.CANOPY.STRUCTURE		*SEN2FLEX*
CROP.PLANT.SPACING	*plant spacing in meters*	*Specified in DAISEX vegetation data files.*
CROP.PLANT.HEIGHT	*Plant height in meters*	*Specified in DAISEX vegetation data files.*
CROP.LEAVES.WEIGHT	*Weight in grammes of leaves (use with mode FRESH or DRY)*	*Specified in DAISEX vegetation data files.*
CROP.LEAVES.MOISTURE	*Water content of leaves in weight %*	*Specified in DAISEX vegetation data files.*
CROP.LEAVES.DRYMATTER	*Dry matter content of leaves in %*	*Specified in DAISEX vegetation data files.*
CROP.STEM.WEIGHT	*Weight in grammes of stem (use with mode FRESH or DRY)*	*Specified in DAISEX vegetation data files.*
CROP.STEM.MOISTURE	*Water content of stem in weight %*	*Specified in DAISEX vegetation data files.*
CROP.STEM.DRYMATTER	*Dry matter content of stem in %*	*Specified in DAISEX vegetation data files.*
CROP.TOTAL.WEIGHT	*Weight in grammes of total plant (use with mode FRESH or DRY)*	*Specified in DAISEX vegetation data files.*
CROP.TOTAL.MOISTURE	*Water content of total plant in weight %*	*Specified in DAISEX vegetation data files.*
CROP.TOTAL.DRYMATTER	*Dry matter content of total plant in %*	*Specified in DAISEX vegetation data files.*
DATETIME	*ENVISAT day in MJD2000, meaning that Jan. 1, 2000 at 00:00:00 hrs = DATETIME 0.000000*	
DAY.MISSION.ELAPSED	*Mission start (e.g., launch) = day 0*	
DAY.OF.YEAR	*Day 1 is January 1st.at 24hrs.*	
DEPTH	*Water depth*	
DEPTH.KD		
DEPTH.SEA.FLOOR	*Depth of the Sea Floor*	

DATA_VARIABLES (*Variable Name*)	*Comment*	*Debate*
DEPTH.SEA.OPT	*OPT depth of samples*	
DEPTH.SECCHI	*Can be dependent or independent. As independent variable it has values 0, 0.5 and 1*	
DISCOLOUR.CODE	*possible values according to MAVT definition*	
DISTANCE		
EMISSIVITY		*Specified in ESA CDB work statement for DAISEX.*
FLAG.ABSOA.CONT		
FLAG.ABSOA.DUST		
FLAG.CASE2.ANOM		
FLAG.CASE2.S		
FLAG.CASE2.Y		
FLAG.TARGET		
FLUORESCENCE		
FLUORESCENCE.687	*687 nm band*	*AIRFLEX*
FLUORESCENCE.760	*760 nm band*	*AIRFLEX*
FLUORESCENCE.RATIO.687.760	*RF/FRF ratio, F687/F760*	*AIRFLEX*
FOAM	*Text entrie only, description of Foam and other Sea Surface Conditions*	
GRAVITY		*Indicated in ESAG02 documentation*
H.COLUMN		
H.CONCENTRATION		
H2.COLUMN		
H2.CONCENTRATION		
H2CO.COLUMN		
H2CO.COLUMN.AMF	*air-mass factor*	
H2CO.CONCENTRATION	*Formaldehyde*	
H2CO.CONCENTRATION.AMF	*air mass factor*	
H2CO.CONCENTRATION.AVK	*averaging kernel*	
H2O.ABOVE.CLOUD	*Water vapour content above clouds*	
H2O.COLUMN		
H2O.COLUMN.AMF	*air-mass factor*	
H2O.CONCENTRATION	*Water Vapour*	
H2O.CONCENTRATION.AMF	*air mass factor*	
H2O.CONCENTRATION.AVK	*averaging kernel*	
H2O.LIQUID.COLUMN		
H2O.LIQUID.CONCENTRATION	*Liquid Water Content*	
H2O.LIQUID.PATH	*Liquid Water Path*	
H2O2.COLUMN		
H2O2.CONCENTRATION		
HBr.COLUMN		
HBr.CONCENTRATION		
HCFC22.COLUMN		
HCFC22.CONCENTRATION		
HCHO.COLUMN		
HCHO.CONCENTRATION		
HCl.COLUMN		
HCl.CONCENTRATION	*(Do not confuse the lower case L with a capital I)*	
HCN.COLUMN		
HCN.CONCENTRATION	*Hydrogen cyanide*	
HCO.COLUMN		
HCO.CONCENTRATION		
HDO.COLUMN		
HDO.CONCENTRATION		
HEADING	*Compass heading*	
HEAVE		
HF.COLUMN		
HF.CONCENTRATION		
HNO3.COLUMN		
HNO3.COLUMN.AMF	*air-mass factor*	
HNO3.CONCENTRATION		
HNO3.CONCENTRATION.AMF	*air mass factor*	
HNO3.CONCENTRATION.AVK	*averaging kernel*	
HNO4.COLUMN.AMF		
HNO4.CONCENTRATION		

DATA_VARIABLES (*Variable Name*)	*Comment*	*Debate*
HO2.COLUMN		
HO2.CONCENTRATION		
HO2NO2.COLUMN		
HO2NO2.CONCENTRATION		
HOBr.COLUMN		
HOBr.CONCENTRATION		
HOCl.COLUMN		
HOCl.CONCENTRATION		
HONO.COLUMN		
HONO.CONCENTRATION		
HUMIDITY		
HUMIDITY.RELATIVE	*Relative humidity*	
ICE.THICKNESS		*Indicated in ESAG02 documentation*
ICE.FREEBOARD.HEIGHT		*Indicated in ESAG02 documentation*
INDEX.PRI		*AIRFLEX*
INDEX.NDVI		*AIRFLEX*
IO.COLUMN		
IO.CONCENTRATION		
IRRADIANCE.DOWNWELLED	*Downwelling irradiance*	
IRRADIANCE.DOWNWELLED.SURFACE		
IRRADIANCE.SURFACE	*Surface irradiance*	
IRRADIANCE.UPWELLED	*Upwelling irradiance*	
LAI	*Leaf Area Index, DIMENSIONLESS*	*Specified in DAISEX vegetation data files.* *MDB20040228: What is this?* *NILU 20040325: See http://io.uv.es/projects/daisex/instr/instrs.htm and http://www.licor.com/env/PDF_Files/LAI2000_150dpi.pdf*
LATITUDE	*(Modeless), Latitude North*	
LATITUDE.EQUIVALENT.PV		
LATITUDE.INSTRUMENT	*(Modeless), Latitude of the Instrument (North)*	
LAYER		
LEVEL		
LONGITUDE	*(Modeless) Longitude East*	
LONGITUDE.INSTRUMENT	*(Modeless) Longitude (East) of the Instrument*	
MeO.COLUMN		
MeO.CONCENTRATION		
MeOCl.COLUMN		
MeOCl.CONCENTRATION		
MeOH.COLUMN		
MeOH.CONCENTRATION		
MeONO2.COLUMN		
MeONO2.CONCENTRATION		
MeOO.COLUMN		
MeOO.CONCENTRATION		
MeOOH.COLUMN		
MeOOH.CONCENTRATION		
N.COLUMN		
N.CONCENTRATION		
N2.COLUMN		
N2.CONCENTRATION		
N2O.COLUMN		
N2O.COLUMN.AMF	*air-mass factor*	
N2O.CONCENTRATION		
N2O.CONCENTRATION.AMF	*air mass factor*	
N2O.CONCENTRATION.AVK	*averaging kernel*	
N2O5.COLUMN		
N2O5.CONCENTRATION	*dinitrogenpentoxide*	
NCO.COLUMN		
NCO.CONCENTRATION		
NH3.COLUMN		
NH3.CONCENTRATION		
NLC.BOTTOM.HEIGHT	*Noctilucent Cloud (NLC)*	
NLC.BOTTOM.PRESSURE		

DATA_VARIABLES (*Variable Name*)	*Comment*	*Debate*
NLC.LAYER.HEIGHT		
NLC.LAYER.THICKNESS		
NLC.LAYER.TRANSMISSION		
NLC.OPTICAL.THICKNESS		
NLC.TOP.HEIGHT		
NLC.TOP.PRESSURE		
NO.COLUMN		
NO.CONCENTRATION		
NO2.COLUMN		
NO2.COLUMN.AMF	*air-mass factor*	
NO2.CONCENTRATION	*nitrogen dioxide*	
NO2.CONCENTRATION.AMF	*air mass factor*	
NO2.CONCENTRATION.AVK	*averaging kernel*	
NO3.COLUMN		
NO3.COLUMN.AMF	*air-mass factor*	
NO3.CONCENTRATION		
NO3.CONCENTRATION.AMF	*air mass factor*	
NO3.CONCENTRATION.AVK	*averaging kernel*	
NOY.COLUMN		
NOY.COLUMN.AMF	*air-mass factor*	
NOY.CONCENTRATION		
NOY.CONCENTRATION.AMF	*air mass factor*	
NOY.CONCENTRATION.AVK	*averaging kernel*	
O.1D.COLUMN		
O.1D.CONCENTRATION		
O.3P.COLUMN		
O.3P.CONCENTRATION		
O2.COLUMN		
O2.COLUMN.AMF	*air-mass factor*	
O2.CONCENTRATION		
O2.CONCENTRATION.AMF	*air mass factor*	
O2.CONCENTRATION.AVK	*averaging kernel*	
O3.COLUMN		
O3.COLUMN.AMF	*air-mass factor*	
O3.CONCENTRATION	*Ozone*	
O3.CONCENTRATION.AMF	*air mass factor*	
O3.CONCENTRATION.AVK	*averaging kernel*	
O4.COLUMN		
O4.CONCENTRATION		
OClO.COLUMN		
OClO.COLUMN.AMF	*air-mass factor*	
OClO.CONCENTRATION	*(Do not confuse the small L with a capital I)*	
OClO.CONCENTRATION.AMF	*air mass factor*	
OClO.CONCENTRATION.AVK	*averaging kernel*	
OCS.COLUMN		
OCS.CONCENTRATION	*Carbonyl sulfide*	
OH.COLUMN		
OH.CONCENTRATION		
OIO.COLUMN		
OIO.CONCENTRATION		
PAN		
PAN.COMPLEX		
PAR	*Photosyntetically available radiation*	
PATH.DIFFERENCE		
PHYTOPLANKTON.PIGMENTS		
PITCH		
PIXEL.NUMBER		*SEN2FLEX*
PLATFORM.ACCELLERATION	*Platform stabilisation data for ESAG0. Combine with mode X, Y or Z if appropriate. In other cases, modes PITCH, ROLL or YAW may also be used*	*Indicated in ESAG02 documentation for LCR gravimeter data.* *MDB 20042028: Explain* *NILU 20040325: We do not know if this is intended to describe movement of the aircraft platform relative to a global coordinate system, or movement of a stabilised instrument platform relative to an aircraft coordinate system. These variable names may potentially also be of use for image geo-referencing?*

DATA_VARIABLES (*Variable Name*)	***Comment***	***Debate***
PMC.BOTTOM.HEIGHT	*Polar Mesospheric Cloud (PMC)*	
PMC.BOTTOM.PRESSURE		
PMC.LAYER.HEIGHT		
PMC.LAYER.THICKNESS		
PMC.LAYER.TRANSMISSION		
PMC.OPTICAL.THICKNESS		
PMC.TOP.HEIGHT		
PMC.TOP.PRESSURE		
POTENTIAL.VORTICITY		
PRESSURE	*Pressure*	
PRESSURE.SURFACE		
PRESSURE.WATER		
PSC.BOTTOM.HEIGHT	*Polar Stratospheric Cloud (PSC)*	
PSC.BOTTOM.PRESSURE		
PSC.LAYER.HEIGHT		
PSC.LAYER.THICKNESS		
PSC.LAYER.TRANSMISSION		
PSC.OPTICAL.THICKNESS		
PSC.TOP.HEIGHT		
PSC.TOP.PRESSURE		
RADAR.BRIGHTNESS	*SAR image intensity*	*For campaign EMAC in CDB*
RADAR.BRIGHTNESS.L	*SAR image intensity, L-band*	*Combine with polarity mode POL.xx*
RADAR.BRIGHTNESS.C	*SAR image intensity, C-band*	*Combine with polarity mode POL.xx*
RADAR.BRIGHTNESS.X	*SAR image intensity, X-band*	*Combine with polarity mode POL.xx*
RADIANCE		
RADIANCE.DOWNWELLED	*Downwelled radiance*	
RADIANCE.DOWNWELLED.SKY		
RADIANCE.SQUARED		
RADIANCE.UPWELLED	*Upwelling radiance*	
RANGE	*distance for e.g. radar, not [min-max]*	*Indicated in ESAG02 documentation*
REFLECTANCE		
REFLECTANCE.687	*687 nm band*	*AIRFLEX*
REFLECTANCE.760	*760 nm band*	*AIRFLEX*
REFLECTANCE.531	*531 nm band*	*AIRFLEX*
REFLECTANCE.570	*570 nm band*	*AIRFLEX*
REFLECTANCE.RHOW		
RELAZ	*Relative Azimuth Transmittance*	
RHOW	*p\'w – water-leaving reflectance*	
ROLL		
SALINITY	*Salinity*	
SEA.STATE		
SF6.COLUMN		
SF6.CONCENTRATION		
SIGNAL		
SIGNAL.NOISE.RATIO	*Signal to noise ratio*	
SIGNIFICANT.WAVE.HEIGHT		
SITE.NAME	*Textual info on sampling sit. Use short name, and include verbose info in comment or description fields.*	*Specified in DAISEX vegetation data files.* *MDB20040228: No* *NILU 20040325: Independent variable name needed to describe name of sampling site*
SITE.ZONE	*Textual info on sampling sit. Use short name, and include verbose info in comment or description fields.*	*Specified in DAISEX vegetation data files.* *MDB20040228: No* *NILU 20040325: Independent variable name needed to indicate named position within sampling site*
SITE.FIELD	*Textual info on sampling site. Use short name, and include verbose info in comment or description fields.*	*Specified in DAISEX vegetation data files.* *MDB20040228: No* *NILU 20040325: Independent variable name needed to indicate named position within sampling site*

DATA_VARIABLES (*Variable Name*)	*Comment*	*Debate*
SITE.CROP	*Textual info on sampling site. Use short name, and include verbose info in comment or description fields.*	*Specified in DAISEX vegetation data files.* *MDB20040228: No* *NILU 20040325: Independent or dependent variable name needed to describe crop at this part of sampling site*
SITE.POINT	*Textual info on sampling site. Use short name, and include verbose info in comment or description fields.*	*Specified in DAISEX vegetation data files.* *MDB20040228: No* *NILU 20040325: Independent variable name needed to indicate named sampling points within sampling site*
SITE.DESCRIPTION	*Short descriptive string that may be used as free-text DEPENDENT or INDEPENDENT variable*	*Specified in DAISEX vegetation data files.* *MDB20040228: No* *NILU 20040325: Dependent variable name needed to describe a sampling site*
SITE.X	*X size of sampling site in meters. Add orientation and other georeferencing info in comment or description fields.*	*Specified in DAISEX vegetation data files.* *MDB20040228: No* *NILU 20040325: Independent variable name needed to indicate position in local coordinate system within sampling site*
SITE.Y	*Y size of sampling site in meters. Add orientation and other georeferencing info in comment or description fields.*	*Specified in DAISEX vegetation data files.* *MDB20040228: No* *NILU 20040325: Independent variable name needed to indicate position in local coordinate system within sampling site*
SKY.CODE	*possible values according to MAVT definition*	
SKY.CONDITION		
SKY.RADIANCE.DISTRIBUTION		
SM	*Suspended matter (marine use)*	
SO2.COLUMN		
SO2.COLUMN.AMF	*air-mass factor*	
SO2.CONCENTRATION		
SO2.CONCENTRATION.AMF	*air mass factor*	
SO2.CONCENTRATION.AVK	*averaging kernel*	
SOIL.DEPTH		*SEN2FLEX*
SOIL.WEIGHT	*Use mode FRESH or DRY*	*Specified in DAISEX vegetation data files.*
SOIL.MOISTURE	*Use mode FRESH or DRY*	*Specified in DAISEX vegetation data files.*
SOIL.ROUGHNESS	*Combine with direction MODE (X or Y) and add orientation and other georeferencing info in comment or description fields.*	*Specified in DAISEX vegetation data files.*
SPECTRAL.ABSORPTION.COEFFICIENT	*Spectral absorption coefficient*	
SPECTRAL.BACKSCATTER.COEFFICIENT	*Spectral backscattering coefficient*	
SPECTRAL.BEAM.ATTENUATION.COEFFICIENT	*Spectral beam attenuation coefficient*	
SPEED	*Velocity (also see VELOCITY.X, VELOCITY.Y and VELOCITY.Z, which can be combined with modes like FREE.INERTIAL or GPS.INTEGRATED.KALMAN.FILTERED)*	
SPM	*Suspended particulate matter (atmospheric use)*	
SPRING.TENSION	*Platform stabilisation data for ESAG02*	*Indicated in ESAG02 documentation*
SURFACE.CODE	*possible values according to MAVT definition*	
SURFACE.CONDITION	*Text entries only*	
SWELL.DIRECTION		
SWELL.HEIGHT		

DATA_VARIABLES (*Variable Name*)	*Comment*	*Debate*
TEMPERATURE	*Temperature*	
TEMPERATURE.AIR		
TEMPERATURE.BRIGHTNESS	*Brightness Temperature. DAISEX will use these with modes SKY or TRANSECT*	*Indicated in DAISEX documentation*
TEMPERATURE.BUCKET	*Bucket Temperature (Ship use)*	
TEMPERATURE.INTERNAL.BOX		
TEMPERATURE.INTERNAL.INSTRUMENT		
TEMPERATURE.LAND.SURFACE		
TEMPERATURE.RADIOMETRIC	*In DAISEX files named Trad(´C)*	*Indicated in DAISEX documentation*
TEMPERATURE.SEA.SUBSURFACE		
TEMPERATURE.SEA.SURFACE		
TEMPERATURE.WATER		
THETA	*Potential Temperature*	
TRANSECT.NAME	*Use in DAISEX GCU-data. Use to define a transect (a straight path to fly over a groundbased sampling site)*	*Indicated in DAISEX documentation* *MDB 20040228: Explain?* *NILU 20040326: Definition of TRANSECT given in the comment field. We need TRANSECT as a variable name to declare a transect by name and initial/final positions. We also need TRANSECT as a mode to indicate that some other variable is measured in a profile along a transect.*
TRANSECT.LATITUDE	*Use to define start and end position of a transect (with modes INITIAL and FINAL)*	
TRANSECT.LONGITUDE	*Use to define start and end position of a transect (with modes INITIAL and FINAL)*	
TROPOSPHERIC.DELAY		
TSM.CONCENTRATION	*Total suspended matter (combine with DRYW,B442)*	
UV.INDEX	*UV Index*	
VEGETATION.INDEX		
VELOCITY	*Can be combined with modes like X, Y or Z*	*Indicated in ESAG02 documentation*
VELOCITY.ANGULAR	*Can be combined with modes like PITCH, ROLL or YAW*	*Indicated in ESAG02 documentation*
VELOCITY.ANGULAR.PITCH	*Can be combined with modes like FREE.INERTIAL or GPS.INTEGR.KALMAN.FILT.INERTIAL*	*Indicated in ESAG02 documentation*
VELOCITY.ANGULAR.ROLL	*Can be combined with modes like FREE.INERTIAL or GPS.INTEGR.KALMAN.FILT.INERTIAL*	*Indicated in ESAG02 documentation*
VELOCITY.ANGULAR.YAW	*Can be combined with modes like FREE.INERTIAL or GPS.INTEGR.KALMAN.FILT.INERTIAL*	*Indicated in ESAG02 documentation*
VELOCITY.X	*Can be combined with modes like FREE.INERTIAL or GPS.INTEGR.KALMAN.FILT.INERTIAL*	*Indicated in ESAG02 documentation*
VELOCITY.Y	*Can be combined with modes like FREE.INERTIAL or GPS.INTEGR.KALMAN.FILT.INERTIAL*	*Indicated in ESAG02 documentation*
VELOCITY.Z	*Can be combined with modes like FREE.INERTIAL or GPS.INTEGR.KALMAN.FILT.INERTIAL*	*Indicated in ESAG02 documentation*
VISIBILITY	*WMO codes*	
VMG		
WAVE.DIRECTION		
WAVE.HEIGHT		
WAVE.PERIOD		
WAVE.TYPE		
WAVELENGTH		
WAVELENGTH.CENTRAL	*Central wavelength of the instrument channels*	*Requested for the Sparc2004 campaign*
WAVELENGTH.FWHM	*Full-Width-Half-Maximum spectral resolution –wavelength- of the instrument channels*	*Requested for the Sparc2004 campaign*

DATA_VARIABLES (*Variable Name*)	***Comment***	***Debate***
WAVENUMBER		
WIND.DIRECTION	*Wind direction*	
WIND.SPEED		
YS	*Yellow substance absorption*	
YSBPA	*Yellow substance and bleached particle absorption*	

Table 4.2.7b: *DATA_VARIABLES Variable mode (not used for DATETIME, ALTITUDE, LATITUDE AND LONGITUDE).*

DATA_VARIABLES (*Variable Mode*)	***Comment***	***Debate***
A442	*optical method for determination of Chl.2.Index*	
ABSORPTION		
ALONG.TRACK		
APRIORI		
ASSIMILATION	*Chemical data assimilation*	
AZIMUTH	*Use with ANGLE*	*Indicated in DAISEX data file ANGLE.TXT, see reference under ANGLE in table 4.2.7a*
BACKSCATTER		
B442	*optical method for determination of TSM*	
BBC??	*Black Body Cavity, where ?? is 00 to 99*	
BOLTZMANN	*method for LIDAR temperature retrieval*	
BULK	*Use with TEMPERATURE to get Bulk Sea Surface temperature (SST)*	
COLLOCATED		
DECLINATION		
DIFFSLANT		
DIFFSLANT.EMISSION		
DIFFSLANT.LIMB		
DIFFSLANT.LUNAR		
DIFFSLANT.SOLAR		
DIFFSLANT.STELLAR		
DIFFSLANT.NADIR		
DIFFSLANT.ZENITH		
DRY	*Used for crop characterisation (opposite of FRESH)*	*Indicated in DAISEX vegetation data files.* *MDB 20040228: Is this a mode?* *NILU 20040326: Yes, we believe so.*
DRYW	*method for determination of TSM*	
ELEVATION		
EMISSION		
FINAL	*Use with DAISEX transect LATITUDE and LONGITUDE*	*Indicated in DAISEX vegetation data files.* *MDB 20040228: Is this a mode?* *NILU 20040326: Yes, we believe so.*
FOV	*Field of View of the instrument*	*Use e.g. ANGEL.VIEW_FOV_START and ANGLE.VIEW_FOV_STOP*
FREE.INERTIAL	*Use with aircraft position, velocity or attitude.*	*Indicated in ESAG02 Raw Data Report for use with INS data*
FRESH	*Used for crop characterisation (opposite of DRY)*	*Indicated in DAISEX vegetation data files.* *MDB 20040228: Is this a mode?* *NILU 20040326: Yes, we believe so.*
GPS.INTEGR.KALMAN.FILT.INERTIAL	*Use with aircraft position, velocity or attitude.*	*Indicated in ESAG02 Raw Data Report for use with INS data*
INITIAL	*Use with DAISEX transect LATITUDE and LONGITUDE*	*Indicated in DAISEX vegetation data files.* *MDB 20040228: Is this a mode?* *NILU 20040326: Yes, we believe so.*
INSITU		
INTERFEROGRAMME	*to be used with PATH.DIFFERENCE*	
HPLC	*method for determination of Chl.2.Index*	
HYDROSTATIC	*method for LIDAR temperature retrieval*	
LIMB		
LINEWIDTH	*method for LIDAR temperature retrieval*	
LUNAR	*with reference to the moon*	
LUNAR.OCCULTATION	*With reference to the moon\'s occultation*	
NADIR		
OFFAXIS	*Off-axis*	

DATA_VARIABLES (*Variable Mode*)	*Comment*	*Debate*
PARALLEL	*Reference to parallel polarisation*	
PERPENDICULAR	*Reference to perpendicular polarisation*	
PITCH	*Attitude angle*	*From ESAG02 documentation* *MDB 20040228: Is this different fromATTITUDE.PITCH?* *NILU 20040324: ATTITUDE.PITCH is a complex name, that can be combined with modes like OFFSET (which may incidentally become a DESCRIPTOR instead), or others. We are currently not certain what is needed by the community. If we need to reference both a global coordinate system and a local one inside an aircraft, more work may be needed to define this properly.* *Use with ACCELLERATION.ANGULAR.*
POL.HH	*Polarity for Radar imaging (usable for others?)*	*Combine with RADAR.BRIGHTNESS.x*
POL.VV	*Polarity for Radar imaging (usable for others?)*	*Combine with RADAR.BRIGHTNESS.x*
POL.HV	*Polarity for Radar imaging (usable for others?)*	*Combine with RADAR.BRIGHTNESS.x*
POL.VH	*Polarity for Radar imaging (usable for others?)*	*Combine with RADAR.BRIGHTNESS.x*
ROLL	*Attitude angle*	*From ESAG02 documentation* *MDB 20040228: Is this different fromATTITUDE.ROLL?* *NILU 20040324: ATTITUDE.ROLL is a complex name, that can be combined with modes like OFFSET (which may incidentally become a DESCRIPTOR instead), or others. We are currently not certain what is needed by the community. If we need to reference both a global coordinate system and a local one inside an aircraft, more work may be needed to define this properly.* *Use with ACCELLERATION.ANGULAR.*
SAMPLE		
SKIN	*Use with TEMPERATURE to get Skin Sea Surface temperature (SST)*	
SKY	*Use with BRIGHTNESS.TEMPERATURE*	*Indicated in DAISEX temperature data files.* *MDB 20040228: Explain context* *NILU 20040326: Brightness temperature is measured with the sonde pointing to the sky, or to the surface. Angle ZENITH or some specific angle from ZENITH is also used. The README.TXT file indicates the mode TRANSECT, we do not know the significance of this.*
SLANT		
SLANT.EMISSION		
SLANT.LIMB		
SLANT.LUNAR		
SLANT.SOLAR		
SLANT.STELLAR		
SLANT.NADIR		
SLANT.ZENITH		
SOLAR	*With reference to the sun*	
SOLAR.OCCULTATION	*With reference to the solar occultation*	
SP	*spectrophotometic method for determination of Chl.2.Index*	
STELLAR	*With reference to a star*	

DATA_VARIABLES (*Variable Mode*)	*Comment*	*Debate*
STELLAR.OCCULTATION	*With reference to a star occultation*	
SURFACE	*Use with BRIGHTNESS.TEMPERATURE*	*Indicated in DAISEX temperature data files.* *NILU 20040326: Brightness temperature is measured with the sonde pointing to the sky, or to the surface. Angle ZENITH or some specific angle from ZENITH is also used. The README.TXT file indicates the mode TRANSECT, we do not know the significance of this.*
TILT		
TOA	*Top Of Atmosphere*	
TRANSECT	*Use with BRIGHTNESS.TEMPERATURE*	*Indicated in DAISEX temperature data files.* *MDB 20040228: Explain context* *NILU 20040326: Brightness temperature is measured with the sonde pointing to the sky, or to the surface. Angle ZENITH or some specific angle from ZENITH is also used. The README.TXT file indicates the mode TRANSECT, we do not know the significance of this.*
U	*velocity component*	
UMKEHR	*Dobson/Brewer specific profiling technique*	
UNPOLARISED		
V	*velocity component*	
VERTICAL		
VERTICAL.EMISSION		
VERTICAL.LIMB	*vertical column retrieved from limb data*	
VERTICAL.LUNAR		
VERTICAL.NADIR		
VERTICAL.SOLAR	*direct-sun observations*	
VERTICAL.SOLAR.FOCUS	*sun-focus observations*	
VERTICAL.STELLAR		
VERTICAL.ZENITH		
W	*velocity component*	
X		
Y		
YAW	*Attitude angle*	*From ESAG02 documentation* *MDB 20040228: Is this different fromATTITUDE.YAW?* *NILU 20040324: ATTITUDE.YAW is a complex name, that can be combined with modes like OFFSET (which may incidentally become a DESCRIPTOR instead), or others. We are currently not certain what is needed by the community. If we need to reference both a global coordinate system and a local one inside an aircraft, more work may be needed to define this properly.* *Use with ACCELLERATION.ANGULAR.*
Z		
ZENITH	*Use with ANGLE*	*Indicated in DAISEX data file ANGLE.TXT, see reference under ANGLE in table 4.2.7a*

Table 4.2.7c: *Variable descriptor (optional).*

DATA_VARIABLES (*Variable Descriptor*)	***Comment***	***Debate***
APPARENT		
ASTRONOMICAL		
BEGIN		
CONTRIBUTION	*relative contribution e.g. of apriori profile to retrieved profile*	
DETECTIONLIMIT		
DIFF.MODEL.OBS	*Difference Model - Observed*	
DIFF.SAT.BUOY	*Difference Satellite - Observed by buoy*	
DIFF.SAT.OBS	*Difference Satellite - Observed by other instrument*	
END		
LIMIT		
MAX	*Maximum value of a set of variables*	
MEAN	*Average*	
MEASUREMENT.SPACING	*space between grid points (note the difference with resolution).*	
MEDIAN	*Median*	
MIN	*Minimum value of a set of variables*	
OFFSET	*Difference between ideal and actual mounting angle when an instrument is mounted on a platform. Use with ATTITUDE.PITCH, ATTITUDE.ROLL and ATTITUDE.YAW for correction constants.* *Other uses (linear offsets) are also conceivable, but currently not specified.*	*Offset angles discussed for the Riegl Lidar in the ESAG02 documentation.*
REGISTRATION.ACCURACY	*use with e.g. ALTITUDE for absolute accuracy of altitude values*	
RESOLUTION	*closest distance between points that can be distinguished.*	
RESOLUTION.ALTITUDE		
RESOLUTION.TIME		
RESOLUTION.X		
RESOLUTION.Y		
SATURATION		
START		
STOP		
UNCERTAINTY.RANDOM	*Random uncertainty*	
UNCERTAINTY.RELATIVE	*Relative uncertainty*	
UNCERTAINTY.RMS	*Root mean square uncertainty*	
UNCERTAINTY.STDEV	*1 sigma (standard deviation) uncertainty*	
UNCERTAINTY.SYSTEMATIC	*Systematic uncertainty == accuracy*	
UNCERTAINTY.TOTAL	*Total uncertainty*	
WHITEREFERENCE	*For use in radiometric measurements of white reference*	

A4.2.8 DATA_START_DATE

The Global Attribute **DATA_START_DATE** specifies the earliest/first measurement date found in the current data file. The date/time format to be used is MJD2000 with fractional days. For resolution in seconds, MJD is to be reported with 6 digits behind the decimal point, for milliseconds 9 decimals should be used.

ATTENTION

An appropriate number of digits after the decimal must be reported to properly represent the desired time resolution

Type:	DOUBLE
Format:	MJD2000 date time specification
Entry:	Single field
Example:	DATA_START_DATE = 800.348678

A4.2.9 DATA_FILE_VERSION

The Global Attribute **DATA_FILE_VERSION** specifies the version of the file submitted to the database.

ATTENTION

DATA_VERSION begins with 001 (leading zeroes), each new version should be incremented by 1. Files submitted out of this sequence will be rejected.

Type:	INTEGER
Format:	DDD with leading zeroes.
Entry:	Single field
Example:	DATA_FILE_VERSION = 003

A4.2.10 DATA_MODIFICATIONS

The Global Attribute **DATA_MODIFICATIONS** describes the data modification history of **DATA_VERSION** found in the data file. Detail of the information is up to the discretion of the data originator.

Type:	STRING
Format:	Free format
Entry:	Single field
Example:	DATA_MODIFICATIONS = Version 002, uses the pump correction table of Komhyr (1986).

A4.2.11 DATA_CAVEATS

The optional Global Attribute **DATA_CAVEATS** refers to potential caveats with the data in the current data file.

Type:	STRING
Format:	Free format
Entry:	Single field
Example:	DATA_CAVEATS = This is near real-time data, final revised data will be available within 3 months.

A4.2.12 DATA_RULES_OF_USE

The optional Global Attribute **DATA_RULES_OF_USE** entry is the PI's (the data owner) guidelines for the data usage.

NOTE

This entry is usually guided through a specific project data protocol.

Type:	STRING
Format:	Free format
Entry:	Single field
Example:	DATA_RULES_OF_USE = Refer to Envisat Cal/Val data protocol, for more information contact nadirteam@nilu.no.

A4.2.13 DATA_ACKNOWLEDGEMENT

The optional Global Attribute **DATA_ACKNOWLEDGEMENT** is the PI's 'desired' acknowledgement of the data when used in publications, presentations, etc.

Type:	STRING
Format:	Free format
Entry:	Single field
Example:	DATA_ACKNOWLEDGEMENT = We thank G. Braathen (WMO) for providing us with the revised ozonesonde data from Orland.

A4.3 File attributes

The global **File Attributes** provide detailed description of the data file. These attributes include the file name and generation date, the names of projects that have access to the file, and the version of the metadata used in the given file.

A4.3.1 FILE_NAME

The Global Attribute **FILE_NAME** is the current data file name. The file should always have the same official name at the NADIR data centre as that used by the DO (to prevent errors when updating files). The name must therefore be generated by the PI, DO or DS according to the following rules:

ATTENTION

The file name is always set in lower case, even if the fields it contains are capitalised.

Type: STRING

Format: **FILE_NAME must be constructed using 6 underscore separated Global Attributes + the correct file extension:**

The **DATA_DISCIPLINE** subclass entry from Table 4.2.2c
The **DATA_SOURCE** entry from Section 4.2.5
The **DATA_LOCATION** entry from Table 4.2.4
The **DATA_TYPE** entry from Section 4.2.6
The **DATA_STARTDATE** entry from Section 4.2.8, **but converted to ISO format.**
The **DATA_VERSION** entry from Section 4.2.9
The **.hdf** file extension (referring in this case to the HDF file format).

Entry: Lower case, underscore separated + ".hdf"

Example: FILE_NAME =groundbased_uvvis.saoz_nilu002_jungfraujoch_h2_19990301t110000z_001.hdf

… illustrating how a NILU instrument could operate at Jungfraujoch without creating identification problems in the metadata or the file naming.

The so-called non-HDF data that is possible to upload through the CDB web portal does not necessarily have the .hdf ending. The user does not have to create the correct file name since the file is automatically renamed by the system upon insertion in the database. The file is renamed according to the metadata entries provided by the user, but the original file extension (e.g. .gif) is not changed.

A4.3.2 FILE_GENERATION_DATE

The Global Attribute **FILE_GENERATION_DATE** is the date of generation of the current file and is to be reported in MJD2000 or iso-format.

Type: DOUBLE
Format: MJD2000 or ISO8106 date/time specification
Entry: Single field
Example: FILE_GENERATION_DATE = 890.857575

A4.3.3 FILE_ACCESS

The Global Attribute **FILE_ACCESS** is a multi-field character string referring to the file project association at the NADIR data centre. FILE_ACCESS is used to define what campaigns in the CDB system that should have access to the file.

Type: STRING
Format: project_1; project_2, project_3, …, project_n
Entry: Multiple fields separated by semicolons
Example: FILE_ACCESS = CALVAL; SEN2FLEX; THESEO

Table 4.3.3: *Allowable project names and equivalent FILE_ACCESS currently active at NADIR data centre.*

FILE_ACCESS (*Group Access Rights*)	*Comment*
AQUIFER	*Multisensor study in Tunisia to improve water management*
ARCHIVE	*Pseudo project with files removed from main data directory*
BACCHUS	*Wine crop surveilance in Frascati*
CALVAL	*ENVISAT Cal/Val Data Centre*
CDBTEST	*Test campaign for CDB system*
COSE	*COSE - Compilation of Atmospheric Observations in Support of Satellite Measurements over Europe*
DAISEX	*CDB Demo Campaign*
EMAC	*CDB Campaign*
ESAG02	*CDB Demo Campaign*
EQUAL	*Currently uses AOID1111 under CALVAL*
LARA	*CDB Demo Campaign*
NDSC	*Network for the Detection of Stratospheric Change*
PUBLIC	*Unrestricted access to the data*
SAREX	*South American Radar Experiment*
SEN2FLEX	*CDB campaign*
SPARC04	*SPARC 2004*
THESEO	

A4.3.4 FILE_PROJECT_ID

The Global Attribute **FILE_PROJECT_ID** is a multi-field string defining the custom projects that have access to the file. The Envisat Cal/Val campaign requires the AOID responsible for providing the file to be given here, other projects should use the affiliation acronym of the data submitter.

For Envisat only one Envisat Cal/Val FILE_PROJECT_ID is allowed.

Type: STRING
Format: id_1; id_2; id_3; …; id_n
Entry: Multiple fields separated by semicolons, but a single entry in the Envisat Cal/Val project
Example: FILE_PROJECT_ID = AOID126

A4.3.5 FILE_ASSOCIATION

The optional Global Attribute **FILE_ASSOCIATION** is a multi-field character string defining the file's other associations such as National Programs, special campaigns, or funding programs.

Type: STRING
Format: project_1; project_2; project_3; …; project_n
Entry: Multiple fields separated by semicolons
Example: FILE_ASSOCIATION = …

A4.3.6 FILE_META_VERSION

The Global Attribute **FILE_META_VERSION** is a single field character string defining the version of the metadata definitions used in the given file and the name of the tool used to generate the file.

Type: STRING
Format: ddRddd; tool name (free format)
Entry: Two fields
Example: FILE_METAVERSION = 03R013; `ASC2HDF ver. 003R007`

A5 Variable attributes

Unlike the global attributes, the variable attributes refer specifically to one single variable. For each variable listed under DATA_VARIABLES in section 4.2.7, there must be one section containing the metadata parameters described under Sections 5.1 and 5.2 below.

Table 5: *Overview of the Variable Attributes.*
'X' indicate entries and 'O' indicate optional entries.

Variable Description Attributes	Section	Entry	Entry type	Req
VAR_NAME	5.1.1	Concatenated, underscore separated	Single entry	X
VAR_DESCRIPTION	5.1.2	Variable description	Single entry	X
VAR_NOTES	5.1.3	Variable notes/warnings	Single entry	O
VAR_DIMENSION	5.1.4	Number of dimensions that the dependent variables depend on	Single entry	X
VAR_SIZE	5.1.5	Number of nodes in each dimension	n semi-colon separated	X
VAR_DEPEND	5.1.6	List of variables that the dimensions depend on	n semi-colon separated	X
VAR_DATA_TYPE	5.1.7	Numeric data type	Single entry	X
VAR_UNITS	5.1.8	Variable units	Single entry	X
VAR_SI_CONVERSION	5.1.9	Conversion factor; SI unit	3 semi-colon separated	X
VAR_VALID_MIN	5.1.10	Valid minimum or detection limit	Single entry	X
VAR_VALID_MAX	5.1.11	Valid maximum or saturation limit	Single entry	X
VAR_MONOTONE	5.1.12	Describes the monotonicity of the variable (3 options)	Single entry	X
VAR_AVG_TYPE	5.1.13	Variable averaging technique used	Single entry	X
VAR_FILL_VALUE	5.1.14	See section description	Single entry	X
Variable Visualisation Attributes	**Section**	**Entry**	**Entry type**	**Req**
VIS_LABEL	5.2.1	Short string to facilitate the identification of the variable	Single entry	X
VIS_FORMAT	5.2.2	FORTRAN like format of the data	Single entry	X
VIS_PLOT_TYPE	5.2.3	Plot type to display the variable	Single entry	X
VIS_SCALE_TYPE	5.2.4	Plot scale type used to display the variable: scale type code; scale order code	2 semi-colon separated	X
VIS_SCALE_MIN	5.2.5	Scale display minimum	Single entry	X
VIS_SCALE_MAX	5.2.6	Scale display maximum	Single entry	X

A5.1 Variable description attributes

A5.1.1 VAR_NAME

The **VAR_NAME** must be identical to one of the entries in section 4.2.7: **DATA_VARIABLES**.

This entry consists of the variable identifier constructed using a variable name, the variable mode and the variable descriptor (not always relevant). See detailed description in section 4.2.7

Type:	STRING
Format:	Refer to section DATA_VARIABLES
Entry:	Up to 3 fields concatenated with an underscore character
Example:	VAR_NAME = O3.COLUMN_VERTICAL.SOLAR

A5.1.2 VAR_DESCRIPTION

The Variable Attribute **VAR_DESCRIPTION** is a verbose description of the variable. This is a free format string that must be provided by the data originator to clearly identify the variable's meaning (preferably inline, or by reference to some easily available document), thus making the data file self-explanatory.

Type:	STRING
Format:	Free format
Entry:	Single field
Example:	VAR_DESCRIPTION = In-situ ozone partial pressure measured by ECC ozonesondes.

A5.1.3 VAR_NOTES

The optional Variable Attribute **VAR_NOTES** is character string containing specific comments about the variable's data elements. Used by the data originator to convey any additional information pertinent to the variable.

Type:	STRING
Format:	Free format
Entry:	Single
Example:	VAR_NOTES = ...

A5.1.4 VAR_DIMENSION

The Variable Attribute **VAR_DIMENSION** is the rank of the variable, defined as the number of independent dimensions required to identify one element of the data variable. If the dimension is given as 3, the VAR_SIZE (see Section 5.1.5) requires 3 elements.

Type:	INTEGER between 1 and 8
Format:	Integer
Entry:	Single
Example:	VAR_DIMENSION = 3

A5.1.5 VAR_SIZE

The Variable Attribute **VAR_SIZE** is a semicolon separated character string containing the specific dimensionalities of the variable. In the following example, the dependent variable is reported for four independent dimensions (time, x, y, z) in a grid of 10*2*3*4 nodes. For a computed field, the VAR_SIZE specifies the number of nodes in the 4D time-space. For a set of measured data and for space coordinates that depend on the time, the VAR_SIZE is the number of data elements in the series. The total number of entries in VAR_SIZE must be equal to VAR_DIMENSION.

Type:	INTEGER(s)
Format:	Integer
Entry:	Semicolon separated, one number per dimension
Example:	VAR_SIZE= 10; 2; 3; 4

A5.1.6 VAR_DEPEND

The Variable Attribute **VAR_DEPEND** is a list of semicolon-separated character strings that describes all independent variables on which the current variable depends. The number of independent variables listed must correspond to VAR_DIMENSION, and the order in which the variables are listed must correspond exactly to the order in which their sizes are given in VAR_SIZE.

ATTENTION

Independent variables must have: ***VAR_DEPEND = INDEPENDENT,***
Constants must have: **VAR_DEPEND = CONSTANT**

Type:	STRING
Format:	Free format
Entry:	Semicolon separated, one name per dimension
Example:	VAR_DEPEND = DATETIME; LONGITUDE; LATITUDE; ALTITUDE

A5.1.7 VAR_DATA_TYPE

The Variable Attribute **VAR_DATA_TYPE** specifies the type of the variable.

Type: STRING
Options: Refer to Table 5.1.7
Entry: Single
Example: VAR_DATA_TYPE = INTEGER

Table 5.1.7: *Variable type options.*

DATA_VARIABLE_TYPE	**Comment**
REAL	*16 bit floating point*
DOUBLE	*32 bit floating point*
INTEGER	*16bit integers*
LONG	*32 bit integers*
STRING	*character string*

A5.1.8 VAR_UNITS

The Variable Attribute **VAR_UNITS** specifies the units in which the data elements are stored in the current data file. The prefix is optional (not needed when reporting in a base unit). While the prefix is concatenated with the unit, multiple units are separated by spaces. Powers of units (signed integer) are concatenated with the unit. No brackets are to be used.

ATTENTION

Units are case sensitive.

The list of accepted units for VAR_SI_CONVERSION has been slightly expanded with respect to SI.

NOTE

Project protocols/templates may restrict this to only one allowed unit and scale for each variable.

Type: STRING
Options: Combination of Tables 5.1.8a and b
Entry: Case sensitive, single field
Example 1: VAR_UNITS = mPa … *for milli Pascal*
Example 2: VAR_UNITS = nm m-2 *for nanometre per square metre*

Table 5.1.8a: *Allowed SI prefix to be used in VAR_UNITS in conjunction with the Units in Table 5.1.8b.*

VAR_UNITS (Base Unit Prefix)	***Comment***
Y	yotta
Z	zetta
E	exa
P	peta
T	tera
G	giga
M	mega
k	kilo
h	hecto
da	deka
d	deci
c	centi
m	milli
u	micro (u is used as a substitute for the greek letter \'mu\')
n	nano
p	pico
f	femto
a	atto
z	zepto
y	yocto

Table 5.1.8b: *Allowed base units to be used in VAR_UNITS.*

VAR_UNITS (Base Unit)	***Comment***	***VAR_SI_CONVERSION***	***Flag***
%	*Percent or Relative Humidity*	0; 0.01; DIMENSIONLESS	
A	*ampere*		base
C	*coulomb*	0;1; s A	base
cd	*candela*		base
d	*day*	0; 86400; s	base
deg	*angular degree*	0; 1.74533E-2; rad	base
degC	*degree Celsius*	273.15 ; 1 ; K	
DIMENSIONLESS	*If dimensionless or no specific unit*	0;1;DIMENSIONLESS	base
DU	*dobson unit*	0; 2.69E20; molec m-2	
g	*gram*		base
Gal	*Galileo, cm s-2*	0, 10-2, m s-2	base
h	*hour*	0; 3600; s	base
Hz	*hertz*	0; 1; s-1	base
J	*joule*	0; 1; m2 kg s-2	base
K	*kelvin*		base
L	*liter*	0; 10-3; m3	base
lm	*lumen*	0; 1; cd sr	base
lx	*lux*	0; 1; cd sr m-2	base
m	*metre*		base
min	*minute*	0; 60; s	base
MJD2000	*Modified Julian Day 2000*	0; 86400; s	base
mol	*mole*		base
molec	*molecule*	0; 1; molec	base
N	*newton*	0; 1; m kg s-2	base
NONE	*Text entries only, otherwise use DIMENSIONLESS*	NONE	
Pa	*pascal*	0; 1; kg m-1 s-2	base
photons		0; 1; photons	base
ppbv	*parts per billion (volume)*	0; 10-9; ppv	
ppmv	*parts per million (volume)*	0; 10-6; ppv	
pptv	*parts per trillion (volume)*	0; 10-12; ppv	
ppv	*parts per volume*	0; 1; ppv	base
psu	*practical salinity unit*	??	base
rad	*radian*	0; 1; DIMENSIONLESS	base
s	*second*		base
sr	*steradian*	0; 1; DIMENSIONLESS	base
V	*volt*	0; 1; m2 kg s-3 A-1	base
W	*watt*	0; 1; m2 kg s-3	base

A5.1.9 VAR_SI_CONVERSION

The Variable Attribute **VAR_SI_CONVERSION** is the conversion factor between the units used for the given data element and the corresponding SI unit. If the measurement unit is identical to the SI unit, the conversion factor is 1 and the constant offset is 0.

In VAR_SI_CONVERION, unit divisions should be factored out to have the shortest possible units string. This means that VAR_UNIT = nm m-2 shall have VAR_SI_CONVERSION = 0; 1.0E-9; m-1 This parameter is intended to facilitate calculations by automated tools, using different data files as input. For plot axis labelling, please refer to the VIS_LABEL metadata variables in section 5.2.1.

ATTENTION

For consistency in the prefixes in VAR_UNITS, kilogram (kg) has been replaced by the gram (g) for consistency with the prefixes in VAR_UNITS.

Type:	STRING
Format:	Offset; Conversion factor; SI unit
Entry:	Single field with 3 semi-colon separated entries
Example:	VAR_SI_CONVERSION = 0; 1.0E-3; Pa *for mPa*

A5.1.10 VAR_VALID_MIN

The Variable Attribute **VAR_VALID_MIN** indicates the valid minimum or detection limit of the data element.

ATTENTION

The number must be specified in the appropriate VAR_UNITS reported in section 5.1.8.

Type:	REAL/DOUBLE/INTEGER/LONG
Format:	Number
Entry:	Single
Example:	VALID_MIN = 10.0

A5.1.11 VAR_VALID_MAX

The Variable Attribute **VAR_VALID_MAX** indicates the valid maximum or saturation limit of the data element.

ATTENTION

The number must be specified in the appropriate VAR_UNITS reported in section 5.1.8.

Type:	REAL/DOUBLE/INTEGER/LONG
Format:	Number
Entry:	Single
Example:	VAR_VALID_MAX = 100

A5.1.12 VAR_MONOTONE

The Variable Attribute **VAR_MONOTONE** indicates if the data element increases or decreases monotonically with respect to DATETIME.

Type:	STRING
Options:	Refer to Table 5.1.12
Entry:	Single
Example:	VAR_MONOTONE = INCREASE

Table 5.1.12: *VAR_MONOTONE categories.*

VAR_MONOTONE	***Comment***
INCREASE	*Increasing time series*
DECREASE	*Decreasing time series*
FALSE	*Neither monotonically increasing or decreasing*

A5.1.13 VAR_AVG_TYPE

The Variable Attribute **VAR_AVG_TYPE** is the averaging 'technique' used in generating the given data element.

Type:	STRING
Format:	Refer to Table 5.1.13
Entry:	Single
Example:	VAR_AVG_TYPE = STANDARD

Table 5.1.13: *VAR_AVG_TYPE Averaging techniques.*

VAR_AVG_TYPE (Applied Averaging Method)	***Comment***
ANGLE.COSINE	Cosine of the average of the arc-cosines of the values
ANGLE.DEGREES	Direction average over 360 deg (i.e., average of 5 and 355 is 0 instead of 180)
ANGLE.HOUR	Direction average over local times (hours) (i.e., average of 2 and 22 is 0 instead of 12)
ANGLE.RADIANS	Direction average over 2 pi
CLEAN	Procedure for computing the mean after eliminating all data above or below a certain standard deviation
DECIBEL	10 times the logarithm of the average of the anti-logarithms of the (values/10)
LOG	Logarithm of the average of the anti-logarithms of the values
NONE	No averaging used
RMS	Square root of the average of the squares of the values
WEIGHTED	
ZONAL.WEIGHTED	
GLOBAL.WEIGHTED	
STANDARD	Simple arithmetic mean

A5.1.14 VAR_FILL_VALUE

The Variable Attribute **VAR_FILL_VALUE** is the number or string inserted if the element is known to be '*erroneous*' or missing. The VAR_FILL_VALUE may be different for each variable in a file, but must be constant for all occurrences within a given variable. For variables with numeric **VAR_DATA_TYPE** the **VAR_FILL_VALUE** is negative and consists of nines**.** In absolute value it must be 2 orders of magnitude larger than the largest absolute value in the real data. If the **VAR_DATA_TYPE** is of type floating point, then the fractional data of the fill value must be zeroes to the same number of digits as the measurement data. For string variables the VAR_FILL_VALUE is "ZZZZZZZZZZ" (i.e.10 times a "Z").

ATTENTION

Consideration must be given to the actual format of the VAR_FILL_VALUE to avoid erroneous formatting in section 5.2.2

Type: REAL/DOUBLE/INTEGER/LONG/STRING
Format: Fixed entry
Entry: Single field
Example1: *for a dataset range* [-82.5428 : 4.2396]…
… *the* `VAR_FILL_VALUE = -9999.0000`
Example2: *for a dataset range* [-1.4E-1 : 2.6E1]…
… *the* `VAR_FILL VALUE = -9.0E3`

A5.2 Variable visualisation attributes

The following metadata entries are defined to facilitate the visualisation of the data content in tables or figures.

A5.2.1 VIS_LABEL

The Variable Attribute **VIS_LABEL** is a short (and concise) character string containing the variable name and unit used to label an axis or a table column.

ATTENTION

The unit must correspond to the appropriate VAR_UNITS reported in section 5.1.8.

Type: STRING
Format: Free format text
Entry: Single field
Example: VIS_LABEL = O3 (ppm)

A5.2.2 VIS_FORMAT

The Variable Attribute **VIS_FORMAT** defines the output format of the data elements to the screen and/or to tables. The values must be chosen to ensure that the specification does not result in truncation of fill values (please refer to VAR_FILL_VALUE in section 5.1.14).

Type: STRING
Format: FORTRAN-*like* format (refer to Table 5.2.2).
Entry: Single field
Example: VIS_FORMAT = F8.3

Table 5.2.2: *Allowed FORTRAN like format types for VIS_FORMAT.*

VIS_FORMAT (Format Type Code)	***Comment***
Ad	Strings (STRING)
Fd.d	Floating point (REAL/DOUBLE)
Ed.d	Exponentials (REAL/DOUBLE/INTEGER/LONG)
Id	Integer (INTEGER/LONG)
Id.d	Integer with leading zeroes (INTEGER/LONG)

A5.2.3 VIS_PLOT_TYPE

The Variable Attribute **VIS_PLOT_TYPE** defines the type of graph to be displayed when plotting the given variable.

Type: STRING
Format: Refer to Table 5.2.3
Entry: Single
Example: VIS_PLOT_TYPE = TIMESERIES

Table 5.2.3: *Available plot types for VIS_PLOT_TYPE.*

VIS_PLOT_TYPE (Plot Type Code)	*Comment*
XY	*2D*
XY.PROFILE	*profile*
XY.TIMESERIES	*timeseries*
XYZ	*3D*
XYZ.COLOUR	
XYZ.CONTOUR	
FALSE	*None*

A5.2.4 VIS_SCALE_TYPE

The Variable Attribute **VIS_SCALE_TYPE** indicates the default scale type when plotting the data element.

Type: STRING
Options: Scale type code; scale order code (refer to Tables 5.2.4a and b)
Entry: 2 semicolon separated fields
Example 1: VIS_SCALE_TYPE = LOG; INCREASE
Example 2: VIS_SCALE_TYPE = FALSE; FALSE *if no suitable scale is available*

Table 5.2.4a: *Available scale type code options for plotting.*

VIS_SCALE_TYPE (Scale Type Code)	*Comment*
LINEAR	*Linear*
LOG	*Logarithm*
FALSE	

Table 5.2.4b: *Available scale order code options for plotting.*

VIS_SCALE_TYPE (Scale Order Code)	*Scale Order*
INCREASE	*Ascending order*
DECREASE	*Descending order*
FALSE	

A5.2.5 VIS_SCALE_MIN

The Variable Attribute **VIS_SCALE_MIN** indicates the default scale minimum when plotting the data element. The number must be specified in the appropriate VAR_UNITS.

Type:	REAL/DOUBLE/INTEGER/LONG
Format:	*Number*
Entry:	Single field
Example:	VIS_SCALE_MIN = 0

A5.2.6 VIS_SCALE_MAX

The Variable Attribute **VIS_SCALE_MAX** indicates the default scale maximum when plotting the data element. The number must be specified in the appropriate VAR_UNITS.

Type:	REAL/DOUBLE/INTEGER/LONG
Format:	*Number*
Entry:	Single field
Example:	VIS_SCALE_MAX = 100

A6 References

Bojkov, B.R., Koopman, R.M. and De Mazière, M., “The Envisat Cal/Val Data Centre”, presented at the *NDSC 2001 Symposium – Celebrating 10 years of atmospheric research.* September 24-27, 2001, Arcachon, France.

Bojkov, B.R., Mazière, M., Koopman, R.M., “Generic Metadata guidelines on atmospheric and oceanographic datasets for the Envisat Calibration and Validation Project”, European Space Agency, ESRIN, Frascati, Italy (2002).

De Mazière, M., “Final Report of the EC-COSE Project (contract ENV4-CT98-0750)”, BIRA-IASB, Brussels, Belgium (2001).

ESA, “Envisat - Caring for the Earth”, European Space Agency, Paris, France (2001a).

ESA, *European Space Agency - Envisat Calibration and Validation Plan home page:* http://envisat.esa.int/support-docs/index.html#calval

ESA, “Second Envisat Rehearsal Campaign Guidelines – Version 1.1”, European Space Agency, Noordwijk, The Netherlands (2001b).

ISO, “Representation of Dates and Ti`mes”, ISO 8106:1988, International Organization for Standardization (ISO), Geneva, Switzerland (1988).

ISO, “Codes for the representation of names of countries and their subdivisions -- Part 1: Country codes”, ISO 3166-1:1997, International Organization for Standardization (ISO), Geneva, Switzerland (1997).

NCSA, National Center for Supercomputing Applications – *HDF 4 home page*: http://hdf.ncsa.uiuc.edu/hdf4.html

Norwegian Institute for Air Research (NILU)
P.O. Box 100, N-2027 Kjeller, Norway

REPORT SERIES SCIENTIFIC REPORT	REPORT NO. TR 05/2006	ISBN 82-425-1746-0 ISSN 0807-7185	
DATE	SIGN.	NO. OF PAGES 100	PRICE NOK 150.-
TITLE ESA Campaign Database (CDB) User Manual		PROJECT LEADER Aasmund Fahre Vik	
		NILU PROJECT NO. O-103045	
AUTHOR(S) Aasmund Fahre Vik, Terje Krognes, Sam-Erik Walker, Sjur Bjørndalsæter, Christoffer Stoll, Trygve Bårde, Roland Paltiel and Bjørn Gloslie		CLASSIFICATION * A	
		CONTRACT REF. ESRIN Contract No. 18525/04/I-OL	
REPORT PREPARED FOR European Space Agency, ESA			

ABSTRACT
The current document gives an introduction to the functionality of the ESA Campaign Data Base (CDB) and furthermore provides detailed information on metadata definitions and standards. The document contains sections for data users, data providers and campaign data managers, and provides hints on how to best use the data centre.

NORWEGIAN TITLE
ESA Campaign Database (CDB) brukermanual

KEYWORDS		
ESA	Campaign	Database

ABSTRACT (in Norwegian)
Dokumentet gir en innføring i funksjonaliteten til CDB databasen som NILU driver for ESA og inneholder dessuten detaljert informasjon om metadata definisjoner og standarder brukt i systemet. Dokumentet har avsnitt for rene brukere av data, leverandører av data og for data administratorer og gir råd om hvordan man bør bruke datasenteret.

** Classification* *A* *Unclassified (can be ordered from NILU)*
B *Restricted distribution*
C *Classified (not to be distributed)*