



KNMI HDF5 Data Format Specification, v3.5

Hans Roozkrans and Iwan Holleman

Internal report, November 2003

1 Contents

| | | |
|------|---|----|
| 1 | Contents..... | 2 |
| 2 | Change notes..... | 4 |
| 2.1 | Change notes for version 3.5 | 4 |
| 2.2 | Change notes for version 3.4..... | 4 |
| 3 | KNMI HDF 5 implementation | 6 |
| 3.1 | What is HDF ?..... | 6 |
| 3.2 | Versions | 6 |
| 3.3 | Tag naming..... | 6 |
| 3.4 | Purpose of KNMI HDF 5 Data Format | 6 |
| 4 | Overview | 7 |
| 4.1 | Mandatory groups | 7 |
| 4.2 | Repeated groups | 7 |
| 5 | Description of KNMI HDF 5 Data Format | 9 |
| 5.1 | Compression of datasets..... | 9 |
| 5.2 | Data types | 9 |
| 5.3 | Mandatory and optional fields..... | 9 |
| 5.4 | Overview group | 10 |
| 5.5 | Geographic group..... | 11 |
| 5.6 | Image group | 13 |
| 5.7 | Profile group..... | 17 |
| 5.8 | Discharge group | 17 |
| 5.9 | Visualisation group | 18 |
| 5.10 | Satellite group..... | 18 |
| 5.11 | Radar group | 20 |
| 5.12 | Station group | 20 |
| 5.13 | Classification group | 21 |
| 6 | Data types | 22 |
| 6.1 | Table..... | 22 |
| 7 | Conventions..... | 23 |
| 7.1 | Timestamps | 23 |
| 7.2 | HDF 5 file naming convention..... | 23 |

7.3 Product_group_name convention.....24
7.4 Product_name convention.....28

2 Change notes

2.1 Change notes for version 3.5

The following changes are introduced with respect to KNMI HDF5 **version 3.4**:

The main change in this version is the extension of the KNMI HDF5 Image Format Specification beyond image data format only. The datamodel is extended to be able to store wind profile data, lightning localization data, and polar scan data from radar. The title of document has therefore been changed to "KNMI HDF5 Data Format Specification". In addition, a few minor changes have been applied to the datamodel which will not affect the current systems.

- The repeated **image group is no longer mandatory**, it will off course always be present for HDF5 files containing image data.
- Two new repeated groups have been added: **profile group and discharge group**, intended for storage of wind profile and lightning localizations, respectively.
- The name of the **lightning group is changed to station group**, and the lightning subgroup in the image group has been removed. The naming of the tags in the station group has been changed and two tags for storage of availability information have been added.
- The **grid, point, and vector groups** have been deleted as they are not used.

- A "**image_datetime_valid**" tag has been added in the image group(s) for specifying the date and time that a forecast product is valid.
- The tags for the new repeated groups have been added in the overview group: **number_profile_groups** and **number_discharge_groups**. In addition, the **number_lightning_groups** tag has been renamed to **number_station_groups**. Finally, **number_grid_groups**, **number_point_groups**, and **number_vector_groups** have been removed from the overview group.
- A tag "**geo_par_pixel**" describing the geophysical parameters of the image coordinates has been added to the geographical group.
- The "**DISPLAY_ORIGIN**" subtags to the HDF5 images in the overview, image, classification groups has been changed from mandatory to only mandatory for images with an unusual orientation.
- A few tags describing the characteristics of a polar radar scan have been added to the **radar subgroup** of the image group.
- A **profile group** with contents has been added which is intended for storage of profile data from weather radar, sodar, and profiler. It can be extended for profiles from radiosonde and models.
- A **discharge group** with contents has been added which is intended for storage of timeseries of localizations from lightning detection systems.
- Two tags, one for storage of limited availability

information and one for adjustment information, have been added to the **radar group**.

2.2 Change notes for version 3.4

The following changes are introduced with respect to KNMI HDF5 **version 3.2**:

- **The naming convention of repeated groups.** The change is that potential repeated groups **always** get a number at the end of the groupname. In contrary to the definition in version 3.2, now the groupname ends with 1 in case there is only one group included in the file (e.g. image1 instead of image). This is only valid for groups that are allowed to be repeated (image, visualisation, satellite, radar, lightning, classification, grid, point, and vector).
- **The implementation of calibration tables.** If the values in a calibration table are linear calibrated then not all pixel values need to be included in the table. The minimum and maximum values of linear range(s) may be included only and then the values in between can be calculated by interpolation. This convention may save a lot of table space (e.g. in case of 16-bit pixel values).
An example:

Old table:

| Integer value | Calibrated value |
|---------------|------------------|
| 0.0 | -50.0 |
| 1.0 | -49.5 |
| 2.0 | -49.0 |
| ... | ... |

| | |
|-------|------|
| 251.0 | 75.5 |
| 252.0 | 76.0 |
| 253.0 | 76.5 |
| 254.0 | 77.0 |

New table:

| Integer value | Calibrated value |
|---------------|------------------|
| 0.0 | -50.0 |
| 255.0 | 77.5 |

- **Definition of order of geo_product_corners** (in Geographic group). The first corner is always the southwest corner . Then northwest, northeast and southeast. In earlier versions it was defined as lowerleft, upperleft, etc.
- **HDF5 file naming convention.** The timestamp convention was not defined in earlier versions (the convention that is specified in chapter 7 is not applicable in file names). In chapter 8 the timestamp convention as used within KNMI is described.

- **Tables in HDF5 can only contain elements of the same data type.** This means that annotation/classification tables may contain only strings. So, also the pixel values must be written as strings.
- **The names of two attributes in the Overview group have been changed:**
Old name (version 3.2):
New name (version 3.3):
 KNMIhdf5_version_number
 hdf5tag_version_number
 KNMI_hdf_url
 hdf5tag_url

The following changes are introduced with respect to KNMI HDF5 **version 3.3**:

- **In the Overview group:** “number_xxx_groups” tags are only mandatory in case the value is larger than 0. In previous version these groups were mandatory in any case.

3 KNMI HDF5 implementation

In this document the HDF5 Data format as implemented within KNMI is described. Some more specific background information is given.

3.1 What is HDF ?

HDF (Hierarchical Data Format) is a multi-object file format for sharing scientific data in a distributed environment. HDF is developed by the National Center for Supercomputing Applications (NCSA) in the USA. An important feature of HDF files is that the files are self-describing. The aim of HDF is to make files, containing e.g. image data, self-contained. By grouping related pieces of information together and forcing a certain structure of the file, it will be easy for other users to make use of the data in the files. Self-contained also means that all meta-data associated with the dataset and/or needed for further processing of the dataset is included in the file. Another attractive feature of HDF is the "free" availability of software libraries to handle HDF files. Libraries are developed by NCSA for different operating systems (Unix, Linux, Windows). More general information on HDF can be found on: <http://hdf.ncsa.uiuc.edu>.

3.2 Versions

The implementation of the file format described in this document is based on HDF version 5.0 (HDF5). KNMI has based its definition of HDF5 on the MEOS-HDF format of Kongsberg Spacetec. MEOS (Multi-mission Earth Observation System) is Kongsberg Spacetec's line of systems for earth observation satellites ground stations. Spacetec is

the builder of KNMI's Omnivoor/Beelden image database system.

For users of the KNMI HDF5 files the following HDF5 utilities are most useful: "h5view" and "h5dump". These utilities (and many others) can be downloaded from: <http://hdf.ncsa.uiuc.edu/hdf5>.

3.3 Tag naming

All information placed within an HDF5 file can be stored and retrieved by a named tag. Tag names should be unambiguous and self-explanatory. In general, tags should be lower case with the underscore character used as the word spacer. Use of abbreviations is common when naming tags. However, KNMI and Spacetec suggest that abbreviations only are used when it can be used unambiguously and when it's quite obvious what it stands for. E.g. "clong" could be used to represent "centre longitude", but we suggest that "center_longitude" is used instead.

3.4 Purpose of KNMI HDF5 Data Format

The purpose of the KNMI HDF5 Data Format is to store the (image) data in a platform independent and self-contained way. The HDF5 Data Format as described in this document is implemented in the image data infrastructure of KNMI (BIK) for all EUMETSAT and NOAA satellite data, weather radar data, and lightning image data. This document is distributed to users of KNMI HDF5 data files to enable them to process the data.

4 Overview

The layout of the HDF5 files is similar to a directory structure in a file system. A folder, or divider, similar to a directory name, is used as a placeholder for information that is logically grouped together. The presence or absence of fields gives information on what can be done with the file. A minimum required set of tags and groups must be present in each KNMI HDF5 file.

| Groups | Subgroups |
|----------------|----------------|
| overview (M) | |
| geographic (M) | |
| | map_projection |
| image (R) | |
| | calibration |
| | statistics |
| | satellite |
| | radar |
| profile (R) | |
| discharge (R) | |

| | |
|--------------------|-------------|
| visualisation (R) | |
| satellite (R) | |
| | sensor |
| radar (R) | |
| station (R) | |
| classification (R) | |
| | calibration |

M These groups are **mandatory**

R These groups can be **repeated**

4.1 Mandatory groups

The following groups are the minimal required content of the KNMI HDF5 files.

Overview group

This group provides an overview of the dataset. It contains references, identifiers, a quicklook and fields describing the content of the file.

Geographic group

All information about the geographic reference of the dataset is placed in this group. This includes map projection and parameters used in this operation.

4.2 Repeated groups

All other groups are optional. This implies these groups will be only included in case it is relevant or applicable to the type of data. These optional groups are called repeated groups, where a repetition of zero indicates the absence of a group. The **Overview** group contains fields stating the number of repeated groups present in the HDF5 file. The names of the groups will be numbered as follows: **groupnamen** where **groupname** is the name of the group, and **n** is the group number. Numbering starts at 1 and runs until the number of groups as stated in the overview group.

Image group

This group forms the core of a HDF5 image file. It contains most relevant image information. The image data (pixel values) itself and metadata about the image are placed in this group. There can be more than one image group in one HDF5 file. In the KNMI HDF5 implementation the following rules for the storage of multiple image-layers are defined:

- One HDF5 file can only contain multiple images (2D arrays) in case these images all have the same **geographical** properties (spatial resolution, area coverage and map projection). This implies that the images, stored in **one** HDF5 file, **all** have an **equal** number of pixel columns and rows.
- One Image group can only contain multiple images (2D arrays) in case these images all have the same **product_name** (so, the subgroups, calibration, etc. applies to all images in this image group). This only makes sense for animations (timed sequences of images).

For the moment KNMI does not have plans to

store time series of images in one HDF5 file and this technical note does not describe the HDF5 implementation for the storage of time series.

The KNMI implementation implies that in case more than one image-layer is stored in one HDF5 file (e.g. all AVHRR channel images), each image-layer is stored in a separate (repeated) Image group. The storage of image-layers in the image_data tag is implemented following the NCSA standards as described in the document "HDF5 Image and Palette Specification" (see: <http://hdf.ncsa.uiuc.edu/HDF5/doc/ADGuide/ImageSpec.html>).

Each Image group contains a Calibration subgroup that contains all relevant metadata in relation to the calibration of the pixelvalues of the image(s), stored in the group. The transformation of (binary) pixel values to real geophysical values is described here.

Profile group

This group is intended for storage of (wind) profile data from weather radars, profilers, sodars, and models.

Discharge group

This group is intended for storage of localization data from a lightning detection system.

Visualisation group

In this group a colour palette is stored (following the NCSA specs; see: <http://hdf.ncsa.uiuc.edu/HDF5/doc/ADGuide/ImageSpec.html>).

There can be more Visualisation groups in one HDF5 file.

Satellite group

In this group, all relevant information regarding the satellite is placed. This includes the position of the satellite at the time of data capturing. Orbit prediction data are included here. Also descriptions of the on-board instruments/sensors can be placed here (e.g. spectral bands related to channels, etc.). If data of more than one satellite is included in the HDF5 file then more Satellite groups will be included in the file.

Radar group

Idem as satellite group for (weather) radar systems.

Station group

Idem as satellite group for general observation systems, e.g., lightning detection stations.

Classification group

If automatic classification has been done, results from this process can be placed here. E.g. land/sea/cloud masks. There can be more Classification groups in one HDF5 file. This group is currently used for storage of geographical overlays for satellite images.

5 Description of KNMI HDF5 Data Format

In the HDF5 structure, besides groups, two other data types can be recognised: datasets and attributes. Datasets are meant for storage of the real image data and metadata entities larger than 16Kb. Attributes belong to a group or a dataset and are meant for storage of small entities of metadata. In the following tag description it is defined for each tag whether it is a dataset or an attribute (*in the D/A column*).

5.1 Compression of datasets

HDF5 offers the possibility to perform a lossless compression on all datasets. The compression is performed using the "zlib" library. The zlib library is a free, general-purpose lossless data-compression library for use on virtually any computer hardware and operating system (for information see: <http://www.gzip.org/zlib/>). The compression level is specified by an integer between 1 and 9 (inclusive) where 1 results in the fastest compression while 9 results in the best compression ratio. The default value is 6 being an optimum between speed and compression ratio. Upon extraction of datasets from a HDF5 file, the decompression is handled automatically by the HDF5 library. Compression should be performed to all datasets in the KNMI HDF5 format.

5.2 Data types

The Type column defines the data type of the field. This is one of Int, Long, Float, Double, String, Table

and Image for integer, long integer, floating point number, floating point number with double precision, text, tables and images, respectively. The physical size occupied on storage by the given field is not a concern for this document. The underlying software hides this fact from the user. When reading a float value from the HDF file, the retrieved value will be in the native float format on the target computer, regardless of where the file was originally created.

5.3 Mandatory and optional fields

The issue of mandatory and optional fields in the file is not straightforward. A general HDF browser will be able to read data from the file no matter which fields are required (mandatory) and which are optional. This will also be the case for browsers (viewers) which understand the structure of KNMI HDF5 files. For example, if an image is map projected, the map projection group must be present and contains a set of fields, which are required in that context. If the image is not map projected, the map projection group does not have to be present.

Explanation of M/O column:

M= Mandatory field. This flag is used to indicate whether this field is either required for further processing using HDF5 libraries or required by the Omnivoor/Beelden database and processing system. All fields, having an M attached, should be included in the KNMI HDF5 tag. If not, the field is not applicable to the type of image data set stored in the file or the group (in which the field is included) is optional and not present in the HDF5 file.

O= Optional field. This flag indicates fields that are included in the tag for the purpose of the user, to provide background information or to help the user.

5.4 Overview group

This group provides an overview of the dataset. It contains references, identifiers, a representative quicklook of the image data and fields describing the content of the file.

| Tag Name | D/A | Type | O/M | Description | Convention/Example |
|------------------------|-----|--------|-----|--|---|
| product_group_name | A | String | M | Name of the productgroup; used as dataset identifier in the OMBE system | See chapter 7; e.g. "METEOSAT_7_MVIRI_ECMWF" |
| products_missing | A | String | M | Product_name of products included in the productgroup but missing in the HDF file. In case of more than one missing product, names are comma separated. | See chapter 7; e.g. METEOSAT_7_MVIRI_VIS_ECMWF, METEOSAT_7_MVIRI_VIS_ECMWF |
| product_datetime_start | A | String | M | Start date and time stamp of product(s) in HDF5 file | See chapter 7; e.g. "01-JAN-2002;12:34:40.120" |
| product_datetime_end | A | String | M | End date and time stamp of product(s) in HDF5 file (maybe the same as start date) ; Start and end date together define the period for which all products (observations and/or forecasts) included in the HDF5 file are valid; e.g. in case of a time serie the start date/time defines the date/time of the first image in the time serie and the end date/time defines the date/time of the last image. | See chapter 7; e.g. "01-JAN-2002;12:49:34.880" |
| abbttitle | A | String | O | Dataset identification: abbreviated title for the dataset | |
| product_group_title | A | String | O | Dataset identification: the explicit name of the geographic dataset, to sufficiently identify it by the users | |
| product_group_doc | A | String | O | Documentation reference | |
| hdftag_version_number | A | String | M | Version number of KNMI HDF-tag | ="3.5" |
| hdf5_url | A | String | O | Internet reference (for the documentation) of HDF5 | ="http://hdf.ncsa.uiuc.edu/hdf5" |
| hdftag_url | A | String | O | Internet reference of KNMI HDF5 tag | ="http://www.knmi.nl/onderzk/imag efomat/bik_web/HDF5_info/HDF5 _intro.htm" |
| dataset_summary | A | String | O | dataset overview: a brief textual description of the geographic dataset which summarises the content of the geographic dataset | |
| dataset_org_descry | A | String | O | dataset overview: description of the organisation responsible for commissioning production (together with a statement of status if appropriate) | |
| dataset_raster_type | A | String | O | dataset overview: type of raster data (spatial data, aerial data, semantic data or grid data, etc.) | |
| dataset_raster_descry | A | String | O | dataset overview: description of raster data, depending on the type of raster data | |

| | | | | | | |
|------------------------------|------------------|--------|----------------|--|---|--|
| dataset_sample | D | Image | M ¹ | Sample of image_data as stored in the Image group. The sample is meant to show an overview of the dataset. This can be a RGB color composite of image layers stored in the HDF file. | The NCSA specs are implemented here; see: http://hdf.ncsa.uiuc.edu/HDF5/doc/ADGuide/ImageSpec.html | |
| CLASS | CLASS | A | String | M | Identifier to interpret dataset as image | = "IMAGE" |
| | IMAGE_SUBCLASS | A | String | O | Indicator of type of palette that should be used | = "IMAGE_TRUECOLOR" or "IMAGE_INDEXED" |
| | IMAGE_COLORMODEL | A | String | O | Indicator of color model of palette that should be used | = "RGB" |
| | IMAGE_VERSION | A | String | M | Version number of image specification | = "1.2" |
| | DISPLAY_ORIGIN | A | String | M ³ | Indicator at which corner pixel (0,0) should be viewed | = "UL" or "LL" or "UR" or "LR" |
| | PALETTE | A | REF_OBJ | O | Object reference pointer to a color palette in "Visualisation" group | |
| | INTERLACE_MODE | A | String | O | Indicator of storage layout of image data | = "INTERLACE_PIXEL" or "INTERLACE_PLANE" |
| dataset_sample_descr | A | String | O | Description of dataset_sample, channels used, ... | e.g. "R: channel_1,G: channel_2,B: channel_4" | |
| dataset_meta_language | A | String | O | metadata language: language used in the textual statements | e.g. "CEN" | |
| number_image_groups | A | Int | M ² | Number of image groups present in file | 1 or 2 etc. | |
| number_profile_groups | A | Int | M ² | Number of profile groups present in file | | |
| number_discharge_groups | A | Int | M ² | Number of discharge groups present in file | | |
| number_visualisation_groups | A | Int | M ² | Number of visualisation groups present in file | | |
| number_satellite_groups | A | Int | M ² | Number of satellite groups present in file | | |
| number_radar_groups | A | Int | M ² | Number of radar groups present in file | | |
| number_station_groups | A | Int | M ² | Number of station groups present in file | | |
| number_classification_groups | A | Int | M ² | Number of classification groups | | |

¹ : Mandatory, only in case the size of the images, stored in the HDF5 file, makes the inclusion of a quicklook useful.

² : Mandatory, only in case the relevant group is represented in the file (so, in case the value of the tag is zero than the tag is not mandatory).

³ : Mandatory, only in case the (0,0) pixel should **not** be displayed in the upper-left corner.

5.5 Geographic group

All information about the geographic reference of the dataset is placed in this group. This includes map projection and parameters used in this operation. By definition all image data included in one HDF5 dataset have the same georeferences. So, the metadata included in this group applies to all images, point, grid, vector and classification data that are stored in the HDF5 file (only **one** Geographic group is included in an HDF5 file).

| Name | D/A | Type | O/M | Description | Comment |
|--------------------|-----|------|-----|---|----------|
| geo_number_columns | A | Int | M | Number of pixels per line in images (x dimension) | e.g. 256 |
| geo_number_rows | A | Int | M | Number of pixel lines in images (y dimension) | e.g. 256 |

| | | | | | |
|-------------------------|---|----------------|----------------|--|---|
| geo_pixel_size_x | A | Float | M | Horizontal pixel size in projection plane (x-axis positive towards east) | e.g. 2.5 |
| geo_pixel_size_y | A | Float | M | Vertical pixel size in projection plane (y-axis positive towards north) | e.g. -2.5 |
| geo_par_pixel | A | String | M ³ | Geophysical parameters of image coordinates (horizontal,vertical) | e.g. "X,Y" |
| geo_dim_pixel | A | String | M | Dimensions of image pixel size (horizontal,vertical) | e.g. "KM,KM" |
| geo_column_offset | A | Float | M | Geographic column offset of pixel (0,0) from origin of projection plane | e.g. 200 |
| geo_row_offset | A | Float | M | Geographic row offset of pixel (0,0) from origin of projection plane | e.g. 100 |
| geo_pixel_def | A | String | M | Definition of lat/lon position inside the image pixels | e.g. "CENTRE" or "LU" etc. |
| geo_product_center | A | Table of Float | M ¹ | Latitude and longitude at the centre point of the image | Tab(Lon, Lat) |
| geo_product_corners | A | Table of Float | M ¹ | Latitude and longitude of each of the four product corners (starting with southwest corner and then clockwise) | Tab(Lon1, Lat1, Lon2, Lat2, Lon3, Lat3, Lon4, Lat4) |
| geo_ref_tiepoints | D | Table of Float | M ² | List of points mapping pixel co-ordinates (x,y) to (lon,lat). The purpose of this table is to enable a geo navigation in the image by users. | Tab(X1,Y1,LON1,LAT1,X2,Y2,LON2,LAT2,) |
| geo_navigation_accuracy | A | Int | O | Estimated navigation error (number or pixels) | e.g. 1 |

¹ : One of these tags is mandatory

² : Mandatory, only in case of geographical image data without a valid projection

³ : Mandatory, only in case of non-geographical image data

Map projection subgroup

| Name | D/A | Type | O/M | Description | Convention/Example |
|----------------------------|-----|--------|----------------|--|--|
| projection_indication | A | String | M | Projection indication | "Y" or "N" |
| projection_name | A | String | M | Name of the projection according to strict naming convention | One of: "STEREOGRAPHIC", "MERCATOR", "SATELLITE_VIEW" |
| projection_descr | A | String | O | Description of the projection (formulas) or reference to URL | " http://www..... " |
| projection_proj4_params | A | String | M ¹ | List of space separated tag=value pairs to be used with the PROJ4 projection library. Each tag is preceded with a +. The URL of the PROJ4 document and library is referenced in "projection_descr" (in this case http://www.remotesensing.org/proj/) | e.g. "+proj=stere +a=6378.4 +b=6356.9 +lat_0=90.0 +lon_0=0.0 +lat_ts=60.0" |
| projection_semi_major_axis | A | Float | O ² | Semi major axis (earth radius at Equator) in "geo_dim_pixel" units | 6378.4 |
| projection_semi_minor_axis | A | Float | O ² | Semi minor axis (earth radius at pole) in "geo_dim_pixel" units | 6356.9 |
| projection_fplat | A | Float | O ² | Fundamental point in decimal latitude degrees (-90.0 - 90.0) | e.g. 90.0 |
| projection_fplon | A | Float | O ² | Fundamental point in decimal longitude degrees (-180.0 - 360.0) | e.g. 0.0 |
| projection_lat_true_scale | A | Float | O ² | Latitude of true scale in decimal degrees (-90.0 – 90.0) | e.g. 60.0 |
| projection_def_v1 | A | Float | O ² | deflection of vertical 1 | |
| projection_def_v2 | A | Float | O ² | deflection of vertical 2 | |

| | | | | | |
|---------------------------|---|--------|-----|---|----------------|
| projection_def_v3 | A | Float | ○ 2 | deflection of vertical 3 | |
| projection_std_meridian_1 | A | Float | ○ 2 | standard meridian 1 | |
| projection_std_meridian_2 | A | Float | ○ 2 | standard meridian 2 | |
| projection_std_meridian_3 | A | Float | ○ 2 | standard meridian 3 | |
| projection_std_par_1 | A | Float | ○ 2 | standard parallel 1 | |
| projection_std_par_2 | A | Float | ○ 2 | standard parallel 2 | |
| projection_std_par_3 | A | Float | ○ 2 | standard parallel 3 | |
| projection_scale_factor | A | Float | ○ 2 | Scale factor at the central meridian | |
| projection_zone | A | String | ○ 2 | UTM zone number | |
| projection_height | A | Float | ○ 2 | Height of satellite platform in "geo_dim_pixel" units | e.g. 35785.831 |

¹ and ²: The **projection_proj4_params** tag is only mandatory in case the image data is projected. The tag contains all projection information needed to perform a projection using the PROJ4 library. If the **projection_proj4_params** tag is not available, then the other tags containing projection parameters must be filled.

5.6 Image group

All data related to the image data and the displaying of such must be placed in this group. There can be more than one Image group included in the HDF5 file.

| Name | D/A | Type | O/M | Description | Convention/Example |
|-----------------------|-----|------------------|----------------|---|--|
| image_product_name | A | String | M | Name of product stored in the Image group | See chapter 9; e.g. "METEOSAT_7_MVIRI_VIS_ECMWF" |
| image_source_ref | A | Table of REF_OBJ | O | Reference to "source" metadata as stored in Satellite, Radar or Lightning groups (see paragraphs 5.8-5.10). In case of "fused" products this tag may contain more than one reference to different groups. | |
| image_data | D | Image | M | Image, stored in 2-dimensional array of "Nrows x Ncolumns" | The NCSA specs are implemented here; see: http://hdf.ncsa.uiuc.edu/HDF5/doc/ADGuide/ImageSpec.html |
| CLASS | A | String | M | Identifier to interpret dataset as image | = "IMAGE" |
| IMAGE_SUBCLASS | A | String | O | Indicator of type of palette that should be used | = "IMAGE_TRUECOLOR" or "IMAGE_INDEXED" or "IMAGE_BITMAP" |
| IMAGE_COLORMODEL | A | String | O | Indicator of color model of palette that should be used | = "RGB" |
| IMAGE_WHITE_IS_ZERO | A | Int | O | Used in case of "IMAGE_BITMAP" to define bit | 0 = false, 1 = true |
| IMAGE_VERSION | A | String | M | Version number of image specification | = "1.2" |
| DISPLAY_ORIGIN | A | String | M ⁴ | Indicator of at which corner pixel (0,0) should be viewed | = "UL" or "LL" or "UR" or "LR" |
| PALETTE | A | REF_OBJ | O | Object reference pointer to colour palette in "Visualisation" group | |
| image_size | A | Long | M | The total size of the image data in bytes | e.g. 1123078 |
| image_bytes_per_pixel | A | Int | M | Number of bytes used to represent a pixel value | e.g. =1, for one-byte-per-pixel data |

| | | | | | | |
|------------------------|---------------------|-----------------|----------------|--|---|--|
| image_geo_parameter | A | String | M | Geophysical parameter or quantity with unit represented in image | e.g. = "TEMPERATURE_[K]" | |
| image_preview | D | Image | M ¹ | Thumbnail of image_data (for quicklook purposes). E.g. including every 10th pixel column and row of original image. | The NCSA specs are implemented here; see: http://hdf.ncsa.uiuc.edu/HDF5/doc/ADGuide/ImageSpec.html | |
| | CLASS | A | String | M | Identifier to interpret dataset as image | = "IMAGE" |
| | IMAGE_SUBCLASS | A | String | O | Indicator of type of palette that should be used | = "IMAGE_TRUECOLOR" or "IMAGE_INDEXED" or "IMAGE_BITMAP" |
| | IMAGE_COLORMODEL | A | String | O | Indicator of color model of palette that should be used | = "RGB" |
| | IMAGE_WHITE_IS_ZERO | A | Int | O | Used in case of "IMAGE_BITMAP" to define bit | 0 = false, 1 = true |
| | IMAGE_VERSION | A | String | M | Version number of image specification | = "1.2" |
| | DISPLAY_ORIGIN | A | String | M ⁴ | Indicator of at which corner pixel (0,0) should be viewed | = "UL" or "LL" or "UR" or "LR" |
| | PALETTE | A | REF_OBJ | O | Object reference pointer to colour palette in "Visualisation" group | |
| image_start_obs | A | String | M ² | Date and time when image observation is started | See chapter 7; e.g. "01-JAN-2002;12:34:40.120" | |
| image_end_obs | A | String | M ² | Date and time when image observation is ended | See chapter 7; e.g. "01-JAN-2002;12:49:34.880" | |
| image_datetime_valid | A | String | M ⁵ | Date and time for which the forecast product being stored in this image group is valid. | See chapter 7; e.g. "01-JAN-2002;12:49:34.880" | |
| image_number_image_obs | A | Int | M ³ | Number of image observations included in composites. | e.g. 10 | |
| image_obs_timestamp | A | Table of String | M ³ | Timestamp for each image included in the composite. UTC time stamps of each image, included in composite, are listed. Timestamps are comma separated | For time stamps convention see chapter 7. E.g. Tab(01-JAN-2002;12:34:40.120,02-JAN-2002;12:24:32.130,...) | |

¹ : Mandatory, only in case the size of the image, stored in the Image group, makes the inclusion of a preview useful (e.g. > 256 x 256 pixels in an original image).

² : Mandatory, only in case the observation times of the image deviate from the observation times included in the Overview group

³ : Mandatory, only in case of a composite image.

⁴ : Mandatory, only in case the (0,0) pixel should **not** be displayed in the upper-left corner.

⁵ : Mandatory, only in case of a forecast image

Calibration subgroup

| Name | D/A | Type | O/M | Description | Convention/Example |
|-----------------------|-----|--------|----------------|---|--|
| calibration_flag | A | String | M | Is image data calibrated? Is there a fixed relation between the pixel bytes and a geophysical parameter? | "Y" = data is calibrated "N" = data is uncalibrated (raw) |
| calibration_level | A | String | O | Level of calibration? E.g. the NASA definition: 0 = raw, 1 = sensor calibrated, 2 = atmospheric corrected, etc. | e.g. "NASA level 0" |
| calibration_reference | A | String | O | Reference to calibration data used (e.g. NOAA tables); e.g. NOAA URL | " http://www..... " |
| calibration_formulas | A | String | M ¹ | Text representation of formulas used for conversion of pixel values (PV) to geophysical parameter or quantity (GEO). The layout is fixed. | e.g. "GEO=0.933*PV+1.444" |

| | | | | | |
|-------------------------------|---|-----------------|----------------|---|---------------------------------------|
| calibration_table | D | Table of float | M ¹ | Table containing mapping between pixel values and calibrated geophysical parameter or quantity | e.g. see chapter 6, table 1.1 and 1.2 |
| calibration_missing_data | A | Int | M | Pixel value representing missing data (e.g. bad or missed lines) | e.g. 255 |
| calibration_out_of_image | A | Int | M | Pixel value representing "out of image" or "out of range" | e.g. 255 |
| calibration_annotation_tables | D | Table of string | O | Table containing mappings from pixel values to some textual values (for the purpose of annotation). | e.g. see chapter 6, table 1.3 |

¹ Either a formula or table is needed in this subgroup. In case of a simple formula (e.g. linear cases) a formula is preferred, otherwise a table.

Statistics subgroup

| Name | D/A | Type | O/M | Description | Convention/Example |
|------------------|-----|---------------|-----|---|-----------------------------|
| stat_min_value | A | Float | M | Minimum geophysical pixel value present in image (one value for each image) | e.g. 260.5 |
| stat_max_value | A | Float | M | Maximum geophysical pixel value present in image (one value for each image) | e.g. 315.6 |
| stat_min_value_5 | A | Float | O | Minimum geophysical pixel value after 5% of the pixels are clipped from the low end of the histogram (one value for each image) | e.g. 261.5 |
| stat_max_value_5 | A | Float | O | Maximum geophysical pixel value after 5% of the pixels are clipped from the high end of the histogram (one value for each image) | e.g. 299.7 |
| stat_histogram | A | Table of Long | O | Histogram for image in the image group | e.g. Tab(0,0,25,50,300,...) |
| stat_bin_count | A | Integer | O | Number of bins used in the histogram | e.g. 15 |
| stat_bin_size | A | Integer | O | Size of each bin (in counts) | e.g. 17 |
| stat_std_dev | A | Float | O | Standard deviation of pixel values in image | e.g. 80.22 |
| stat_mean | A | Float | O | Mean of pixel values in image | e.g. 155.5 |

Satellite subgroup

All information related to the quality and to the processing of the image/VAP is placed in this group. This includes level of processing, e.g. level1 (pan corrected and orthorectified), level2 (map projected). Other information that can be placed here is a reference to the algorithms used by the processor. Reference to external processing procedures, e.g. NOAA. There can be more than one Satellite group included in the HDF5 file.

| Name | D/A | Type | O/M | Description | Convention/Example |
|------------------------|-----|--------|-----|--|----------------------------------|
| satellite_product | A | String | O | Description of satellite product contained in image | e.g. "AVHRR_CHANNEL_1" |
| image_acquisition_time | A | String | O | UTC when the input data for this image dataset was received on site (e.g. at KNMI) | For UTC convention see chapter 7 |
| image_generation_time | A | String | O | UTC when this image dataset was produced | For UTC convention see chapter 7 |
| image_processing_level | A | String | O | Level of processing | Any convention can be used here |

| | | | | | |
|-------------------------------------|---|--------------|---|---|--|
| image_processing_software | A | String | O | Processing software used: reference to algorithms used | Can be a URL (where a document is located) |
| image_processing_hist | A | String | O | Processing history: listing of processing steps | e.g. "CALIBRATED, NAVIGATED, PROJECTED" |
| navigation_processing_hist | A | String | O | Navigation processing history: description of navigation steps | e.g. "LANDMARK_CORRECTION" |
| image_accuracy_indication | A | String | O | Accuracy indication of geophysical pixel value: e.g. for SST image (static information) | e.g. "0.5 C" |
| image_quality_indication | A | String | O | Indication if quality metadata available | "Y" = quality fields included "N" = quality fields not included |
| image_quality_ingest | A | String | O | Ingest quality parameters | |
| image_quality_ber | A | Double | O | Bit Error Rate | |
| image_nbr_missing_lines | A | Int | O | Number of detected missing lines | |
| image_missing_line_numbers | A | Table of Int | O | The line numbers of the missing lines | Tab(1, 103, ...) |
| missing_lines_correction_procedures | A | String | O | Procedure used to correct for missing or bad lines. | e.g. "DUPLICATE_LAST" or "IGNORE" or "BLEND" or "NONE" |
| lineage_id | A | Long | O | image quality: unique id for the process history of the data set | |
| lineage_method | A | String | O | image quality: method used for creating or processing data set | |
| lineage_organisation | A | String | O | image quality: organisation applying the method on the dataset | |
| lineage_start_date | A | String | O | image quality: startdate of process | |
| lineage_end_date | A | String | O | image quality: enddate of process | |
| lineage_purpose | A | String | O | image quality: purpose | |
| lineage_confidence | A | String | O | image quality, metaquality: confidence | |
| lineage_reliability | A | String | O | image quality, metaquality: reliability | |
| lineage_methodology | A | String | O | image quality, metaquality: methodology | |
| lineage_abstraction | A | String | O | image quality, metaquality: abstraction effect | |

Radarsubgroup

| Name | D/A | Type | O/M | Description | Convention/Example |
|-------------------------|-----|--------|-----|--|----------------------|
| radar_product | A | String | O | Description of radar product contained in image | e.g. "PCAPPI" |
| radar_product_parameter | A | String | O | Parameter specifying radar product, e.g. height of CAPPI | e.g. "0.8KM" |
| radar_method | A | String | O | Description of method used to derive radar product | e.g. "Interpolation" |
| radar_quality | A | String | O | Indication of quality | e.g. "The_best" |
| radar_elevation | A | Float | O | Elevation of radar beam in degrees | e.g. 0.3 |

| | | | | | |
|----------------|---|-------|---|---|-----------|
| radar_rotation | A | Float | O | Azimuthal speed of radar antenna in degrees/sec. | e.g. 18 |
| radar_prf_low | A | Float | O | Low PRF used during acquisition of radar data in Hz. | e.g. 750 |
| radar_prf_high | A | Float | O | High PRF used during acquisition of radar data in Hz. | e.g. 1000 |

5.7 Profile group

This group is intended for storage of vertical profiles from e.g. weather radars, sodars, or profilers, and it can be extended from radiosonde observations of model profiles. The values of a variable at the different levels/heights in a profile are stored in a datasets. Different datasets are proposed for several variables and additional datasets can be included. The only mandatory dataset is the one described the geometric height of the vertical levels in the profile.

| Name | D/A | Type | O/M | Description | Convention/Example |
|-----------------------|-----|----------------|-----|---|-----------------------------|
| profile_number_levels | A | Int | M | Number of vertical levels in profile, i.e., length of profile datasets | e.g. 15 |
| profile_missing_data | A | Float | M | Float value indicating "missing data" in profile datasets | e.g. -9999.0 |
| profile_height | D | Table of Float | M | Dataset with geometric height of vertical levels in profile in meter | e.g. Tab(100,300, ...) |
| profile_u_wind | D | Table of Float | O | Dataset with east-west component of wind in m/s | e.g. Tab(-5.0,13.0,...) |
| profile_v_wind | D | Table of Float | O | Dataset with north-south component of wind in m/s | e.g., Tab(20.0, 15.4,...) |
| profile_w_wind | D | Table of Float | O | Dataset with vertical component of wind in m/s | e.g. Tab(0.5, -1.0,...) |
| profile_radial_stddev | D | Table of Float | O | Dataset with standard deviation of radial velocity as deduced from the wind model fit used in profile extraction in m/s | e.g. Tab(1.0,1.5, ...) |
| profile_reflectivity | D | Table of Float | O | Dataset with reflectivity factor at levels in profile in dBZ | e.g. Tab(-31.5, -20.5, ...) |
| profile_number | D | Table of Int | O | Dataset with number of points used to compile values at each level | e.g. Tab(1000,400, ...) |

5.8 Discharge group

This group is intended for storage of the timeseries of the localizations observed by a lightning detection system. The timeseries of each parameter detected by the system is stored as an array in a dataset. The data and time of each discharge event (localization) is given as a time offset in seconds against a reference data and time which is stored in this group as well. The time offset is stored in a dataset of doubles to allow for storage of a high accuracy time offset (microsec.) over a daily period. A number of parameters describing the characteristics of a discharge have been defined, but addition of other datasets is possible.

| Name | D/A | Type | O/M | Description | Convention/Example |
|-------------------|-----|------|-----|--|--------------------|
| number_discharges | A | Int | M | Number of discharges in timeseries, i.e., length of discharge datasets | e.g. 100 |

| | | | | | |
|--------------------|---|-----------------|---|--|---|
| reference_datetime | A | String | M | Date and time stamp against which discharges are referenced | See chapter 7; e.g. "01-JAN-2002;12:34:40.120" |
| time_offset | D | Table of Double | M | Dataset with time offsets of discharges with respect to reference date and time in sec., double allows for microsec accuracy | e.g., Tab(0.001,0.0011, ...) |
| longitude | D | Table of Float | M | Dataset with geographical longitudes of discharges in decimal degrees | e.g., Tab(4.78,5.17,...) |
| latitude | D | Table of Float | M | Dataset with geographical latitudes of discharges in decimal degrees | e.g. Tab(51.1,52.9, ...) |
| event_type | D | Table of Char | O | Dataset with types of observed discharges: "0" single-point, "1" start of CC, "2" CC discharge, "3" end of CC, "4" CG stroke, "5" CG return stroke | e.g. Tab(1,2,2,3,...) |
| position_error | D | Table of Float | O | Dataset with position errors of CG stroke localizations as deduced by detection system in meter | e.g. Tab(500,1500,..) |
| rise_time | D | Table of Float | O | Dataset with rise times of induced current for detected CG strokes in sec. | e.g. Tab(0.000010,0.000020, ...) |
| decay_time | D | Table of Float | O | Dataset with decay times of induced current for detected CG strokes in sec. | e.g. Tab(0.000125,0.0005,...) |
| current | D | Table of Float | O | Dataset with estimated currents of CG strokes in ampere | e.g., Tab(100000,250000,) |

5.9 Visualisation group

In this group a color_palette is stored. Documentation on how the use of color_palettes is implemented in the HDF5 tag is described in chapter 8. There can be more than one Visualisation group stored in one HDF5 file.

| Name | D/A | Type | O/M | Description | Convention/Example |
|----------------|-----|--------|-----|---|---|
| color_palette | D | Table | M | Color palette in RGB format: two-dimensional array of "Ncolors x 3" | The NCSA specs are implemented here; see: http://hdf.ncsa.uiuc.edu/HDF5/doc/ADGuide/ImageSpec.html |
| CLASS | A | String | M | Identifier to interpret dataset as color palette | = "PALETTE" |
| PAL_VERSION | A | String | M | Version number of palette specification | = "1.2" |
| PAL_TYPE | A | String | M | Definition of type of palette, usually a directly indexed array | = "STANDARD8" |
| PAL_COLORMODEL | A | String | M | Definition of the color model used in palette, usually RGB | = "RGB" |

5.10 Satellite group

In this group, mostly static information regarding the satellite is placed. This includes the position of the satellite at the time of data capturing. Also descriptions of the on-board instruments can be placed here (e.g. spectral bands related to channels). Moreover, in case of polar orbiting satellites actual orbit information is stored.

| Name | D/A | Type | O/M | Description | Convention/Example |
|-------------------------------|-----|----------------|----------------|--|---|
| satellite_name | A | String | M | Name of satellite (e.g. NOAA) | e.g. "NOAA" or "METEOSAT", etc. |
| satellite_id | A | String | M | ID of satellite (e.g. 16) | e.g. 16 or 7 |
| satellite_description | A | String | O | Textual description of satellite or reference to URL | e.g. http://www.eumetsat.de/ |
| satellite_agency | A | String | O | Owner/agency of satellite (e.g. RADARSAT International) | e.g. "EUMETSAT" |
| satellite_platform_type | A | String | O | Polar orbiting, geostationary, ground based | e.g. "GEOSTATIONARY" |
| satellite_platform_position | A | String | M | Position relative to ground of the satellite (or instrument in case of weather radar) | e.g. "AT CROSSING POINT OF EQUATOR AND MERIDIAN" |
| satellite_launch_date | A | String | O | Start Date of the source identifier (can be launch date of satellite or first date of first operational data delivery). | e.g. "17 MARCH 1996" |
| satellite_acquisition_station | A | String | O | Groundstation where the data is gathered from the satellite | e.g. "KNMI in De Bilt" |
| satellite_asc_desc_flag | A | String | M ¹ | Satellite pass is ascending or descending | "A" = ascending; "D" = descending; |
| satellite_subtrack | A | Table of Float | M ¹ | Satellite sub track (at nadir) ; Two pairs of Lat/Lon are included: the position of the satellite at start_obs and at end_obs. | Tab(Lat(start), Lon(start), Lat(end), Lon(end)) |
| satellite_orbit_number | A | Int | O ¹ | Orbit number for pass | e.g. 12388 |
| satellite_orbit_prop_ind | A | String | O ¹ | Indicator which orbit proposition data is present in group | "TBUS" = TBUS data is present "TLE" = Two line elements is present "SV" = State vector is present "NONE" = no orbit prop data is present |
| satellite_orbit_prop | A | String | O ¹ | State vector, or TBUS message or two line elements used in the navigation processing. For info see: http://noaasis.noaa.gov/NOAASIS/ml/navigation.html | The TBUS, state vector or two-line elements in the form of a text string |
| satellite_orbit_prop_age | A | String | O ¹ | Age of state vector (or TBUS or two-line elements). Date of the TBUS message, state vector or two-line elements. | 20-JAN-2002 |

¹ These fields are included only in case of polar orbiting satellites.

Sensor subgroup

| Name | D/A | Type | O/M | Description | Convention/Example |
|------------------------|-----|--------|-----|--|--|
| sensor_name | A | String | M | Name of instrument/sensor (of which data is included in dataset) | e.g. "MVIRI" or "AVHRR" |
| sensor_type | A | String | O | Type of sensor | e.g. "IMAGER" or "SOUNDER" |
| sensor_descr | A | String | O | Short description sensor or instrument | e.g. "ADVANCED VERY HIGH RESOLUTION RADIOMETER" |
| sensor_descr_long | A | String | O | Long description sensor or instrument or reference to URL | e.g. http://www.eumetsat.de/ |
| sensor_number_channels | A | Int | O | Number of channels | e.g. 5 |

| | | | | | |
|------------------------|---|--------|---|--|-------------------------------|
| sensor_desc_channels | A | String | O | Description of channels(waveband, signal/noise ratio) | See chapter 7 |
| sensor_radiometric_res | A | String | O | Radiometric resolution of sensor (Number of bits) | e.g. "10 BITS" |
| sensor_spatial_res | A | String | O | Spatial resolution (NADIR in km) | e.g. "1.1 KM" |
| sensor_temporal_res | A | String | O | Temporal resolution or repetition frequency | e.g. "9 DAYS" or "30 MINUTES" |

5.11 Radar group

| Name | D/A | Type | O/M | Description | Convention/Example |
|-------------------|-----|----------------|----------------|---|------------------------------|
| radar_id | A | String | M ¹ | ID of radar station or WMO station number | e.g. "NL50" |
| radar_name | A | String | M ¹ | Name of radar station | e.g. "DE_BILT" |
| radar_location | A | Table of float | M | Longitude and latitude of radar station in decimal degrees | e.g. Tab(5.17, 52.10) |
| radar_height | A | Float | O | Height above mean-sea level of radar antenna feed in meters | e.g. 48.0 |
| radar_num_contrib | A | Int | O | Number of scans that this radar has contributed to image data | e.g. 96 |
| radar_adjustment | A | String | O | String describing adjustment applied to radar reflectivity observations | e.g. "F=0.0+0.01*range" |
| radar_system | A | String | O | Description of radar hardware and manufacturer | e.g. "METEOR_360AC" |
| radar_software | A | String | O | Description of radar processing software | e.g. "RAINBOW" |
| radar_wavelength | A | Float | O | Wavelength of radar in cm | e.g. 5.2 |
| radar_beamwidth | A | Float | O | 3dB width of radar beam in degree | e.g. 1.0 |
| radar_angles | A | Table of float | O | Elevations of radar antenna used in degree | e.g. Tab(0.3, 1.1, 2.0, 3.0) |

¹ One of these tags is required.

5.12 Station group

| Name | D/A | Type | O/M | Description | Convention/Example |
|----------------------|-----|----------------|----------------|--|-----------------------|
| station_id | A | String | M ¹ | ID of detection station or WMO station number | e.g. "NL06" |
| station_name | A | String | M ¹ | Name of detection station | e.g. "HOOGVEEN" |
| station_location | A | Table of float | M | Longitude and latitude of detection station in decimal degrees | e.g. Tab(5.17, 52.10) |
| station_height | A | Float | O | Height above mean-sea level of detection station in meters | e.g. 48.0 |
| station_availability | D | Table of Char | O | Array with availability reports in 1-minute intervals. For each 1-minute interval the number of seconds with a properly functioning station is reported. | e.g. Tab(60,60,...) |

| | | | | | |
|-----------------------------|---|--------|---|---|---------|
| station_number_availability | A | Int | O | Number of 1-minute availability reports in "station_availability" array | e.g. 5 |
| station_system | A | String | O | Description of detection hardware and manufacturer | "SAFIR" |

¹ One of these tags is required.

5.13 Classification group

This group is meant to store "classified" image data. E.g. a land/sea/cloud mask that can be used for presentation purposes or to enable the user to do calculations only for certain surfaces (e.g. SST's only for cloudfree sea pixels). The classification group contains one image, where the pixel values are encoded to represent different features in the image.

| Name | /A | Type | /M | Description | Convention/Example |
|----------------------|----|-----------------|----------------|---|---|
| classification_flag | A | String | M | YES means that the image has been classified, NO otherwise | "Y" or "N" |
| classification_type | A | String | M | Type of classification | e.g. "LANDSEACLOUDMASK" |
| classification_image | D | Image | M | An image mask with encoded values (X/Y array) | The NCSA specs are implemented here; see: http://hdf.ncsa.uiuc.edu/HDF5/doc/ADGuide/ImageSpec.html |
| CLASS | A | String | M | Identifier to interpret dataset as image | = "IMAGE" |
| IMAGE_SUBCLASS | A | String | O | Indicator of type of palette that should be used | = "IMAGE_TRUECOLOR" or "IMAGE_INDEXED" or "IMAGE_BITMAP" |
| IMAGE_COLORMODEL | A | String | O | Indicator of color model of palette that should be used | = "RGB" |
| IMAGE_WHITE_IS_ZERO | A | Int | O | Used in case of "IMAGE_BITMAP" to define bit color | 0 = false, 1 = true |
| IMAGE_VERSION | A | String | M | Version number of image specification | = "1.2" |
| DISPLAY_ORIGIN | A | String | M ¹ | Indicator of at which corner pixel (0,0) should be viewed | = "UL" or "LL" or "UR" or "LR" |
| PALETTE | A | REF_OBJ | O | Object reference pointer to colour palette in "Visualisation" group | |
| classification_table | D | Table of string | M | A table with information on how to interpret the encoded values in the image mask | e.g. see chapter 6, table 1.3 |

¹ : Mandatory, only in case the (0,0) pixel should **not** be displayed in the upper-left corner.

Processing subgroup

| Name | D/A | Type | /M | Description | Convention/Example |
|----------------------|-----|--------|----|---|---|
| processing_software | A | String | O | Reference to the processing software. E.g. a URL | " http://www... " |
| processing_algorithm | A | String | O | Reference to algorithm or method used. E.g. a URL | " http://www... " |

6 Data types

In this section some of the data types found in KNMI HDF5 Data Format Specification are described.

| Data types | Description |
|------------|---|
| Int | Integer value |
| Long | Long integer value |
| Float | Float value |
| Double | Double value |
| String | Character-array |
| Image | Two-dimensional array of 1,2, 4 byte values |
| Table | One/two dimensional array of data values |

6.1 Table

A table is a one or two dimensional array of data elements.

Calibration/annotation/classification tables are two dimensional arrays. These tables contain two columns. The first column contains the pixel values of a product. The second column contains the related calibrated value or a text representation. In HDF5 all elements of a table need to be of the same data

type. So, if text is included in the second column then the elements in the first column (the pixel values) need to be also string typed.

Not all pixel values need to be represented in the table. E.g. in case of linear calibrations it is allowed to exclude one or pixel values in the first column (the idea behind this is that tables become a lot shorter). The excluded pixel values can then be calculated by the user by interpolation between the neighbouring pixel values that are included in the table.

Table 1.1 Table showing calibration of pixel values in channel 4 to °C

| Integer value | Calibrated value |
|---------------|------------------|
| 0.0 | -50.0 |
| 1.0 | -49.5 |
| 2.0 | -49.0 |
| ... | ... |
| 251.0 | 75.5 |
| 252.0 | 76.0 |
| 253.0 | 76.5 |
| 254.0 | 77.0 |
| 255.0 | 77.5 |

Table 1.2 Same example as table 1.1, but now only the first and last pixel value are included.

The values in between can be calculated by interpolation.

| Integer value | Calibrated value |
|---------------|------------------|
| 0.0 | -50.0 |
| 255.0 | 77.5 |

Table 1.3 Table showing annotation of values in a classification image. Table showing annotation of values in a classification image.

| Data value | Classification |
|------------|----------------|
| "0" | "Unclassified" |
| "1" | "Sea" |
| "2" | "Land" |
| "3" | "Clouds" |
| "4" | "Snow" |

7 Conventions

7.1 Timestamps

When timestamps are represented as text they shall have the following convention: **DD-MON-YYYY;HH:MM:SS.sss**.

| | |
|-------------|---|
| DD | two digit number representing the day of the month, e.g 05. |
| MON | three character normal english abbreviated month name, e.g. JAN, FEB, MAR, and so on. |
| YYYY | four digit year number, e.g. 2000. |
| HH | two digit number representing the hour of the day, e.g. 08, |
| MM | two digit number representing the minute of the hour, e.g 58. |
| SS | two digit number representing the seconds of the minute, e.g 23. |
| sss | three digit number representing the milliseconds of the second, e.g 549 |

Full example: 05-JAN-2000;08:58:23.549

7.2 HDF5 file naming convention

Within the image data infrastructure of KNMI (BIK), files are generated, stored and dispatched. Within BIK there is a convention for the names of the generated HDF5 files:

`<product_group_name>[subsystem][processing_param]<timestamp>.H5`

The fields within `<..>` are mandatory while the one within `[..]` are optional. Note that there can be more than one `processing_param`, they will then be separated by `'_'`.

Explanation of fields:

- `product_group_name` = Product group name as defined within OMNIVOOR-beelden (see chapter 9)
- `subsystem` =
 - `automatic_production` -> AP,
 - `interactive_production` (via Web) -> USERNAME
 - `rt server` -> RT,
- `processing param` (when file is generated in production shell of Omnivoor/Beelden)=
 - `IMAGE_WARP`, `SUBSET`,
 - `COMPOSITION`,
 - `PIXEL_CONVERT`
 - `LAYER_EXTRACT`
- `timestamp` = time when file was generated in BIK. Convention: YYYYDDMMHHMM (year-month-day-hour-minute)

7.3 Product_group_name convention

Product_group_names and product_names are to be generated by the frontend systems and then to be included in the HDF5 files. The convention for the names is defined in the Omnivoor/Beelden database project and the frontends need to confirm to this convention. The product_group_name is primarily used to link the dynamic metadata in the HDF5-file to the Omnivoor metadata database. Each productgroup contains a unique set of one or more products as defined by the administrator/ manager of BIK. The product_group_name will be a unique static identifier to classify an image-productgroup and will be included in the "overview" group of the HDF5 tag.

The following **general** convention rules are implemented in BIK:

- The product_group_name is uniquely linked to a **defined set of products**.
- The product_group_name is a string that has **variable length** but a **maximum length of 50** characters.
- The product_group_name always contains **three** "underscores".
- In case it is impossible to fill in something, NA (not applicable) has to be filled in.
- The product_group_name is always written in **CAPITALS**.

The naming convention and examples for three data types (satellite, radar and lightning) is described.

I. Satellite Data

The product_group_name is divided in four parts, describing:

1. the **source type** (e.g. NOAA, METEOSAT, MSG)
2. the mission **ID** of satellite (e.g. 16, 7, 1)
3. type of **sensor(s)** (e.g. AVHRR, MVIRI, SEVIRI)
4. a **miscellaneous** field for geography or other information (e.g. FULL PASS, GLOBE, ECMWF)

Examples:

| Sourcetype | _ID | _sensor | _miscellaneous | product_group_name |
|------------|-----|---------|----------------|------------------------|
| METEOSAT | 7 | MVIRI | GLOBE | METEOSAT_7_MVIRI_GLOBE |
| MSG | 1 | SEVIRI | ECMWF | MSG_1_SEVIRI_ECMWF |
| MSG | 1 | SEVIRI | HRVEUROPE | MSG_1_SEVIRI_HRVEUROPE |
| NOAA | 16 | AVHRR | FULL PASS | NOAA_16_AVHRR_FULLPASS |

II. Radar Data

The product_group_name is divided in four parts, describing:

- The **Source type**. For radar data it is always equal to: **RAD**
- The **Radar ID**. The radar ID consists of four characters: **AAII**. The first two letters indicate the country from which the radar data is originating, and the last two numbers indicate the specific radar or composite content. Country abbreviations and number are taken from OPERA (see below).
- The **Product type**. The type of radar product, e.g. pseudoCAPPI, echotops, is indicated by a three character string in agreement with current CRIS definition (see below)

- A **Miscellaneous field**. The miscellaneous field is used for instance to indicate accumulation time.

Table Radar I. *The AA subfield of the ID is used to define from which country the radar product is originating. Abbreviations are taken from OPERA.*

| AA | Country |
|----|----------------|
| NL | Netherlands |
| DL | Germany |
| UK | United Kingdom |
| BX | Belgium |
| FR | France |
| DN | Denmark |

Table Radar II. *The II subfield of the ID is used to define which radar site or composite image is contained by datafile. Definition is in agreement with OPERA standards.*

| II | Content |
|-------|---|
| 00 | Not used. |
| 01-19 | Not used for radar data (normally used for 'global distribution') |
| 20-39 | Used to identify national and regional composites. |
| 40-89 | Used to identify radar site data. |
| 90-99 | Reserved but frequently used for test bulletins |

Table Radar III. *The Product type field (PPP) is used to specify the type of radar product in accordance with definition in current CRIS/Rainbow.*

| PPP | Product |
|-----|---|
| PPZ | Plan-Position Indicator of reflectivity Z, or similarly for velocity V and spectral width W |
| PCP | PseudoCAPPI of reflectivity |

| | |
|-----|---|
| CLT | Cluttermap corresponding to pseudoCAPPI |
| ETH | Echotop product (general) |
| ETW | Echotop for Hail Warning (45dBZ) |
| HAW | Hail Warning Product |
| VP2 | Wind Profile |
| RAU | Rainfall Accumulation (uncorrected) |
| RAC | Rainfall Accumulation (corrected) |
| CAL | Calibration Data |
| BIT | BITE messages |
| LOG | LOG messages |

Examples of radar product_group_names:

| Source | _ID | _Product | _miscellaneous | product_group_name |
|--------|------|----------|----------------|--------------------|
| RAD | NL50 | ETH | NA | RAD_NL50_ETH_NA |
| RAD | NL21 | PCP | NA | RAD_NL21_PCP_NA |
| RAD | NL51 | RAU | 24H | RAD_NL51_RAU_24H |
| RAD | NL50 | CAL | DBZ | RAD_NL50_CAL_NA |
| RAD | FR21 | PCP | NA | RAD_FR21_PCP_NA |
| RAD | EU20 | PCP | FIRST | RAD_EU20_PCP_FIRST |

III. Lightning Data

The product_group_name is divided in four parts, describing:

- The **Source type**. For lightning data it is always equal to: **LGT**

- The **Lightning ID**. The Lightning ID consists of four characters: **AAII**. The first two letters indicate the country from which the lightning data is originating, and the last two numbers indicate the specific content. Lightning data are considered to be a composite, thus it will be in the range 21..40. Country abbreviations and number are taken from OPERA (see below table L.I and L.II).
- The **Product type**. The type of lightning product, e.g. Localization files, Density maps or Hazardous weather forecast, is indicated by a three character string, meaning is stated below in table L.III.
- A **Miscellaneous field**. The miscellaneous field is used for instance to indicate accumulation time.

Table Lightning I. *The AA subfield of the ID is used to define from which country the lightning product is originating. Abbreviations are taken from OPERA. only countries of interest for the Dutch lightning detection system are named here.*

| AA | Country |
|----|-------------|
| NL | Netherlands |
| DL | Germany |
| BX | Belgium |
| FR | France |

Table Lightning II. *The II subfield of the ID is used to define which lightning composite image is contained by datafile. Definition is in agreement with OPERA standards.*

| II | Content |
|-------|--|
| 20-39 | Used to identify national and regional composites. |
| 90-99 | Reserved but frequently used for test bulletins |

Table Lightning III. *The Product type field (PPP) is used to specify the type of lightning product in accordance with definitions in the current SAFIR system,*

| | |
|-----|---------------------------------|
| PPP | Product |
| LAP | Lightning Accumulated Positions |
| LOG | LOG messages |

Examples of lightning product_group_names:

| Source | _ID | _Product | _miscellaneous | product_group_name |
|--------|-----|----------|----------------|--------------------|
|--------|-----|----------|----------------|--------------------|

| | | | | |
|-----|------|-----|-----|------------------|
| LGT | NL21 | LAP | 15M | LGT_NL21_LAP_15M |
| LGT | NL21 | LAP | 24H | LGT_NL21_LAP_24H |

7.4 Product_name convention

The product_name will be a unique static identifier to classify an image-product and will be included in the "Image" group of the HDF tag .

The following **general** convention rules are implemented in BIK:

- The product_name is a string that has **variable length** but a maximum length of **50** characters.
- The Product_name always contains **four** "underscores".
- In case it is impossible to fill in something, NA (not applicable) has to be filled in.
- The product_name is always written in **capitals**.

I. Satellite Data

The product_name is divided in five parts, describing:

1. the **source type** (e.g. NOAA, METEOSAT, MSG)
2. the mission **ID** of satellite (e.g. 16, 7, 1)
3. type of **sensor(s)** (e.g. AVHRR, MVIRI, SEVIRI)
4. description of **product type** (e.g. VIS, IR, CH1, CH5)
5. a **miscellaneous** field for geography or other information (e.g. GLOBE, EUROPE, FULL PASS)

Examples:

| Source type | _ID | _sensor | _type | _miscellaneous | Product_name |
|-------------|-----|---------|--------|----------------|----------------------------|
| METEOSAT | 7 | MVIRI | VIS | GLOBE | METEOSAT_7_MVIRI_VIS_GLOBE |
| MSG | 1 | SEVIRI | VIS0.6 | ECMWF | MSG_1_SEVIRI_VIS0.6_ECMWF |
| MSG | 1 | SEVIRI | HRV | EUROPE | MSG_1_SEVIRI_HRV_EUROPE |
| NOAA | 16 | AVHRR | CH1 | FULLPASS | NOAA_16_AVHRR_VIS_FULLPASS |

II. Radar Data

The product_name is divided in five parts, describing:

1. The **Source type**. For radar data it is always equal to: **RAD**
2. The **Radar ID**. The radar ID consists of four characters: **AAII**. The first two letters indicate the country from which the radar data is originating, and the last two numbers indicate the specific radar or composite content. Country abbreviations and number are taken from OPERA (see Tables R.I and R.II).
3. The **Product type**. The type of radar product, e.g. pseudoCAPPI, echotops, is indicated by a three character string in agreement with current CRIS definition (see Table R.III)
4. The **Product parameter**. Parameter describing the elevation of a PPI, the height of a pseudoCAPPI, threshold of an echotop product, etc...

5. A **Miscellaneous field**. The miscellaneous field is used for instance to indicate accumulation time. Four parts of the **product_name** are equal to the **product_group_name**, and only the product parameter is different. The convention for the product parameter is given below.

Table Radar IV. *The Product Parameter field (PARM) is used to detail the product information. It always consists of one letter and a float (AF.F). The letter indicates the type of parameter and the float its value.*

| PPP | PARM | Description |
|-----|------|--|
| PPZ | E0.3 | PPI taken at elevation of 0.3 degrees |
| PCP | H0.8 | CAPPI taken at altitude of 0.8 km |
| ETH | Z7.0 | Echotops using reflectivity threshold of 7.0 dBZ |

Examples:

| Typesource | _ID | _Product | _Param | _miscellaneous | Product_name |
|------------|------|----------|--------|----------------|----------------------|
| RAD | NL50 | PPZ | E0.5 | NA | RAD_NL50_PPZ_E0.5_NA |
| RAD | NL21 | ETH | Z7.0 | NA | RAD_NL21_ETH_Z7.0_NA |
| RAD | BX21 | PCP | H1.0 | NA | RAD_BX21_PCP_H1.0_NA |

III. Lightning Data

The product_name is divided in five parts, describing:

- The **Source type**. For lightning data it is always equal to: **LGT**
- The **lightningID**. The lightning ID consists of four characters: **AAll**. The first two letters indicate the country from which the lightning data is originating, and the last two numbers indicate the specific radar or composite content. Country abbreviations and number are taken from OPERA (see Tables L.I and L.II).
- The **Product type**. The type of lightning product, (see Table L.III)
- The **Product parameter**. Parameter describing the type of strokes included in the product (see Table L.IV).
- A **Miscellaneous field**. The miscellaneous field is used for instance to indicate accumulation time.

Four parts of the **product_name** are equal to the **product_group_name**, and only the product parameter is different. The convention for the product parameter is given below in the examples

Table Lightning IV. *The Product Parameter field (_Param) is used to describe the stroke types included in the product.*

| _Param | Description |
|--------|--|
| ALL | All kind of strokes are included |
| CC | Only cloud-cloud strokes are included |
| CG | Only cloud-ground strokes are included |

| | |
|----|----------------|
| NA | Not applicable |
|----|----------------|

Examples:

| Typesource | _ID | _Product | _Param | _miscellaneous | Product_name |
|------------|------|----------|--------|----------------|----------------------|
| LGT | NL21 | LAP | ALL | 05M | LGT_NL21_LAP_ALL_05M |
| LGT | NL21 | LAP | CG | 24H | LGT_NL21_LAP_CG_24H |
