FIDUCEO Multi-sensor Match up System (MMS) Manual

Tom Block, Sabine Embacher

Brockmann Consult GmbH

8/22/2019

FIDUCEO has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no. 638822

1 Content

2 Introduction 7
2.1 Scope 7
2.2 Version Control 7
2.3 Software Changelog 7
2.3.1 Updates from version 1.4.9 to 1.5.0 7
2.3.2 Updates from version 1.4.8 to 1.4.9 7
2.3.3 Updates from version 1.4.7 to 1.4.8 7
2.3.4 Updates from version 1.4.6 to 1.4.7 8
2.3.5 Updates from version 1.4.5 to 1.4.6 8
2.3.6 Updates from version 1.4.4 to 1.4.5 8
2.3.7 Updates from version 1.4.3 to 1.4.4 8
2.3.8 Updates from version 1.4.2 to 1.4.3 8
2.3.9 Updates from version 1.4.1 to 1.4.2 8
2.3.10 Updates from version 1.4.0 to 1.4.1 8
2.3.11 Updates from version 1.3.9 to 1.4.0 8
2.3.12 Updates from version 1.3.8 to 1.3.9 9
2.3.13 Updates from version 1.3.7 to 1.3.8 9
2.3.14 Updates from version 1.3.6 to 1.3.7 9
2.3.15 Updates from version 1.3.5 to 1.3.6 9
2.3.16 Updates from version 1.3.4 to 1.3.5 9
2.3.17 Updates from version 1.3.3 to 1.3.4 9
2.3.18 Updates from version 1.3.2 to 1.3.3 9
2.3.19 Updates from version 1.3.1 to 1.3.2 9
2.3.20 Updates from version 1.3.0 to 1.3.1 9
2.3.21 Updates from version 1.2.9 to 1.3.0 9
2.3.22 Updates from version 1.2.8 to 1.2.9 9
2.3.23 Updates from version 1.2.7 to 1.2.8 10
2.3.24 Updates from version 1.2.6 to 1.2.7 10
2.3.25 Updates from version 1.2.5 to 1.2.6 10
2.3.26 Updates from version 1.2.4 to 1.2.5 10
2.3.27 Updates from version 1.2.3 to 1.2.4 10
2.3.28 Updates from version 1.2.2 to 1.2.3 10

2.3.29 Updates from version 1.2.1 to 1.2.2 10
2.3.30 Updates from version 1.2.0 to 1.2.1 10
2.3.31 Updates from version 1.1.3 to 1.2.0 11
2.3.32 Updates from version 1.1.2 to 1.1.3 11
2.3.33 Updates from version 1.1.1 to 1.1.2 11
2.3.34 Updates from version 1.1.0 to 1.1.1 11
2.3.35 Updates from version 1.0.5 to 1.1.0 11
2.3.36 Updates from version 1.0.4 to 1.0.5 11
2.3.37 Updates from version 1.0.3 to 1.0.4 11
2.3.38 Updates from version 1.0.2 to 1.0.3 11
2.3.39 Updates from version 1.0.1 to 1.0.2 11
2.3.40 Updates from version 1.0.0 to 1.0.1 12
2.4 Applicable and Reference Documents....12
2.5 Glossary....12
3 Software Overview....14
3.1 Ingestion Tool....14
3.2 Matchup Tool....15
3.3 Post Processing Tool 15
3.4 Supported Satellite Sensors....16
3.4.1 AIRS-Level-1B 16
3.4.2 AMSR-2....23
3.4.3 AMSR-E....24
3.4.4 AMSU-B 25
3.4.5 (A)ATSR....25
3.4.6 AVHRR 26
3.4.7 AVHRR FCDR....28
3.4.8 AVHRR FRAC....29
3.4.9 CALIOP CLay 30
3.4.10 CALIOP VFM 31
3.4.11 HIRS 31
3.4.12 HIRS FCDR....32
3.4.13 IASI 32
3.4.14 MHS....34

3.4.15 MODIS MxD06 35
3.4.16 SLSTR 41
3.4.17 SSMT2 42
3.5 Supported Insitu Data 44
3.5.1 GRUAN reference files 44
3.5.2 OceanRain 44
3.5.3 SST-CCI 44
4 Configuration 46
4.1 Configuration Files 46
4.1.1 System Configuration 46
4.1.2 Database Configuration 47
4.1.3 MMD Writer Configuration 48
4.1.3.1 Variables Configuration 48
4.1.3.1.1 Renaming a sensor 49
4.1.3.1.2 Define a separator per sensor namespace 49
4.1.3.1.3 Exclude variables 49
4.1.3.1.4 Rename variables 50
4.1.3.1.5 Rename attributes of variables 50
4.1.4 UseCase Configuration 50
4.1.4.1 Conditions 53
4.1.4.1.1 Border Distance Condition 54
4.1.4.1.2 Spherical Distance Condition 55
4.1.4.1.3 Time Delta Condition 55
4.1.4.1.4 Overlap Remove Condition 56
4.1.4.1.5 UniqueSamples Condition 57
4.1.4.2 Screenings 57
4.1.4.2.1 Angular Screening 57
4.1.4.2.2 Angular Cosine Proportion Screening 58
4.1.4.2.3 HIRS LZA Angular Screening 59
4.1.4.2.4 Buehler Cloud Screening 59
4.1.4.2.5 Pixel Value Screening 60
4.1.4.2.6 Window Value Screening 62
4.1.5 Post Processing Configuration 63
4.1.5.1 Post Processing Plugins 63

4.1.5.1.1 Spherical Distance Plugin 64
4.1.5.1.2 SST Insitu Timeseries Plugin 64
4.1.5.1.3 NWP plugin 66
4.1.5.1.4 Cloud Flagging for HIRS 70
4.1.5.1.5 ElevationToSolzen 72
4.1.5.1.6 AddAmsrAngles 73
4.1.5.1.7 AddAmsr2ScanQualityData 73
4.1.5.1.8 AddIasiSpectrum 74
4.1.5.1.9 AddDistanceToLand 74
4.1.5.1.10 CALIOP Feature Classification Flags Midpoints Extract 75
4.1.5.1.11 CALIOP add L2 CLay variables 76
4.1.5.1.12 AIRS add channel variables 76
4.1.5.1.13 AddGruanSource 78
4.1.5.1.14 AddAvhrrCorrCoeffs 78
5 MMD file format 79
5.1 Global Attributes 79
5.1.1 sensor-name 79
5.2 Extra MMD Variables 79
5.2.1 <sensor-name>_x, <sensor-name>_y 79
5.2.2 <sensor-name>_file_name 79
5.2.3 <sensor-name>_processing_version 79
5.2.4 <sensor-name>_acquisition_time 79
5.3 CF Confomity 79
5.3.1 CF conform usage of “units” attribute 79
6 Installation and Operation on CEMS 80
6.1 Workspace Structure 80
6.2 Building the MMS on CEMS 81
6.3 Operating the MMS on CEMS 81
6.3.1 Setting up the environment 81
6.3.2 CEMS Specific Configuration 81
6.3.2.1 System Configuration 82
6.3.2.2 Database Configuration 82
6.3.3 Operating the MMS 82

6.3.4 Workflow Files 83

2 Introduction

This document is the user manual for the Fiduceo Multisensor Matchup System software (MMS). It covers all tools, configuration files, supported data formats and plugin configurations used by the software. In addition, a description of the installation and operation procedures on the CEDA JASMIN cluster is given.

2.1 Scope

This document is thought as a living document, updates will be integrated whenever applicable or necessary. The contents of this version of the manual refers to the MMS software version 1.5.0.

2.2 Version Control

Version	Reason	Reviewer	Date of Issue
1.0	Initial version		
1.1	New minor version		
1.2	New minor version, added post-processing		2017-01-17
1.3	New minor version, cumulative updates		2017-10-13
1.4	Updates to V 1.4		2019-04-04
1.4	New minor version, cumulative updates		2019-06-20
1.5	Updates to v 1.5		2019-08-22

2.3 Software Changelog

This chapter lists the changes, updates and bugfixes applied to the MMS software.

2.3.1 Updates from version 1.4.9 to 1.5.0

- corrected bug in SLSTR reader - read over product borders
- added data caching for SLSTR reader

2.3.2 Updates from version 1.4.8 to 1.4.9

- added support for compressed SLSTR L1B data
- corrected SeedPointMatchupStrategy coordinate rounding
- corrected SLSTR L1B time coding calculations
- updated AVHRR FCDR naming regular expression
- corrected post-processing for AVHRR FCDR correlation coefficients

2.3.3 Updates from version 1.4.7 to 1.4.8

- implemented support for segmented geo-location data for AVHRR GAC

- added support for AVHRR GAC v2.10.2 data

2.3.4 Updates from version 1.4.6 to 1.4.7

- Added inverse cosine scaling for Sobol random sequences
- Added post-processing for AVHRR FCDR correlation coefficients
- Added support for SLSTR L1B data
- Added support for AVHRR_GAC v2.10.1 data
- Changed MongoDb socket timeout to 2 minutes

2.3.5 Updates from version 1.4.5 to 1.4.6

- Added option to switch Sobol sequences between equal-area and equal-angle sampling.

2.3.6 Updates from version 1.4.4 to 1.4.5

- Updated AVHRR FCDR reader to follow new filename regular expression

2.3.7 Updates from version 1.4.3 to 1.4.4

- Updated AVHRR FCDR reader to support format changes

2.3.8 Updates from version 1.4.2 to 1.4.3

- Added support for compressed AVHRR FRAC data products
- Corrected NetCDF lib fill value initialisation during write operations

2.3.9 Updates from version 1.4.1 to 1.4.2

- Added reader for AVHRR FIDUCEO FCDR products
- Added reader for AVHRR HIRS FCDR products
- Added support for GRUAN data reference format
- Corrected bug in distance-to-land: NPE when two or more instances are allocated
- Added support for AVHRR FRAC data products
- updated scripts to new JASMIN file system
- added optional paging and offset to database drivers
- updated to SNAP 6.0.8
- updated 3dr party libraries
- added DB maintenance tool

2.3.10 Updates from version 1.4.0 to 1.4.1

- migrated to asynchronous PMonitor
- AIRS reader implemented
- AIRS post processing to add channel data implemented
- Bugfix BowTiePixellocator

2.3.11 Updates from version 1.3.9 to 1.4.0

- NWP Post Processing improvements
- Post processing to add AIRS channel data to matchup implemented
- AIRS Level-1B IR Radiances reader implemented
- Bug - null pointer exception if MODIS products with bow tied geocoding has gaps - fixed

2.3.12 Updates from version 1.3.8 to 1.3.9

- Improved performance of IASI geo-coding
- Speed up in polar-orbiting matchup strategy
- Corrected handling of multiple intersecting areas in one orbit file

2.3.13 Updates from version 1.3.7 to 1.3.8

- Added post-processing that copies the AMSR2 Scan_Data_Quality variable to the MMD
- Renamed PostProcessing Plugin AddAmsreSolarAngles, expanded to be applicable also for AMSR-2

2.3.14 Updates from version 1.3.6 to 1.3.7

- Update to use SNAP version 6.0.0
- Corrected bug in PostGRES database driver
- Maintenance of core modules

2.3.15 Updates from version 1.3.5 to 1.3.6

- Implemented support for AMSR 2 data
- Added functionality to extract all overflights over a single location
- Added global temp-file handling with automated cleanup
- Improved geocoding for MxD06 sensor data, improved accuracy

2.3.16 Updates from version 1.3.4 to 1.3.5

- Improve check if product is already ingested to speed up database ingestion
- Improve geocoding for MxD06 data, bow tie case, to handle with data gaps

2.3.17 Updates from version 1.3.3 to 1.3.4

- Bugfix. Wrong day of year (DOY) handling at PathContext.

2.3.18 Updates from version 1.3.2 to 1.3.3

- Bugfix. Fix read for AcquisitionInfo of CALIOP_xxx_Reader.
  Caliop reader must be able to create product bounding geometry for very small products.

2.3.19 Updates from version 1.3.1 to 1.3.2

- Urgent workaround to resolve NetCDF v4.6.10 bug

2.3.20 Updates from version 1.3.0 to 1.3.1

- Caliop L2 CLay Post Processing to add CLay data to an MMD file which already contain Caliop VFM Data.

2.3.21 Updates from version 1.2.9 to 1.3.0

- Added condition plugin to ensure unique samples per MMD file
- Implemented support for AVHRR GAC data in version 1.5

2.3.22 Updates from version 1.2.8 to 1.2.9

- Added satellite sensor reader for CALIOP L2 Clay products
- Implemented support for MxD06 Modis cloud products
- Fixed bug in NWP post processing plugin configuration parsing

2.3.23 Updates from version 1.2.7 to 1.2.8

- Implemented support for AVHRR GAC data in version 1.3 and 1.4
- Updated SST-In-situ reader to support v04.0 data

2.3.24 Updates from version 1.2.6 to 1.2.7

- Implemented support for OceanRain insitu SST data

2.3.25 Updates from version 1.2.5 to 1.2.6

- MMD – Writer ensures CF conform usage of variable attribute "units"
- Initialize ReaderCache from system or writer configuration
- Corrected SeedPointStrategy to use numPointsPerDay
- Post Processing for CALIOP Feature_Classification_Flags

2.3.26 Updates from version 1.2.4 to 1.2.5

- Optimized resource usage for distance-to-land calculations
- Updated NWP post processing to accept scaled integer geolocation data
- Added post processing that calculates the distance to land
- Added support for matchup with multiple secondary sensors
- Added satellite sensor reader for CALIOP L2 VFM products
- Border distance condition is ready to support multiple secondary sensors
- Time Delta condition is ready to support multiple secondary sensors
- Overlap remove condition is ready to support multiple secondary sensors
- Window value screening is ready to support multiple secondary sensors

2.3.27 Updates from version 1.2.3 to 1.2.4

- Fixed bug in IASI reader that caused crashes when reading files significantly larger than 2 GB
- Updated third party libraries: NetCDF, MongoDB, H2, MySQL and SNAP

2.3.28 Updates from version 1.2.2 to 1.2.3

- Added "Sun_Glint_Angle" variable to AMSR-E reader
- Extended IASI reader to support v4 and v5 MDR data
- Implemented post processing to add IASI spectrum to MMD
- Added IASI reader for EUMETSAT format
- Fixed bug in AngularScreening that removed too many pixels in some cases.

2.3.29 Updates from version 1.2.1 to 1.2.2

- Extended Archive class to understand DAY_OF_YEAR elements
- Added post processing plugin to convert elevation to zenith angles
- Improved PostProcessingTool – now creates target directory if not existing

2.3.30 Updates from version 1.2.0 to 1.2.1

- Updated BorderDistance condition to allow distinguishing between reference and secondary sensor
- Added RFI glint variables to AMSR-E reader
- Added post processing plugin to add ERA-interim NWP data to matchup datasets

2.3.31 Updates from version 1.1.3 to 1.2.0

- Implemented post processing engine
- Added spherical distance post-processing plugin
- Added AMSRE solar angles post-processing plugin
- Added SST in-situ time series extraction post-processing plugin
- Added WindowValue screening plugin

2.3.32 Updates from version 1.1.2 to 1.1.3

- Added reader cache size parameter to MMD writer configuration
- Updated MMDWriter to store CF conforming default fill value attributes

2.3.33 Updates from version 1.1.1 to 1.1.2

- Fixed bug in AVHRR reader that causes crashes when extracting 1x1 pixel windows

2.3.34 Updates from version 1.1.0 to 1.1.1

- Implemented full support for PostGIS databases

2.3.35 Updates from version 1.0.5 to 1.1.0

- Added condition plugin for overlap removal
- Added support for SST-CCI insitu data

2.3.36 Updates from version 1.0.4 to 1.0.5

- Migrated system configuration from properties to XML format
- Implemented configurable archiving rules

2.3.37 Updates from version 1.0.3 to 1.0.4

- Added MMD Writer configuration file
- Added MMD variable renaming engine

2.3.38 Updates from version 1.0.2 to 1.0.3

- Added reader for SSM/T-2 data
- Added full support for Apache H2 database
- Renamed parameter for AngularScreening plugin
- Added mmd writer configuration
- Fixed bug in HIRS reader plugin that prevented HIRS NOAA 18 data to be handled

2.3.39 Updates from version 1.0.1 to 1.0.2

- Added HIRS “lza” angular screening
- Added HIRS L1C reader
- Updated regular expression for AMSU-B/MHS reader
- Added reader for ATSR1, ATSR2 and AATSR L1B data in ENVISAT format
- Added optional calculation of matchup center distance variable
- Added PixelValue screening plugin
- Corrected fill value handling to follow CF conventions
- Added reader form AMSR-E L2A data in HDF format

- Added support for multiple processing version of sensor data

2.3.40 Updates from version 1.0.0 to 1.0.1

- Implemented cloud screening algorithm for AMSU-B, MHS and SSMIS according to University of Hamburg
- Bugfixes:
    - Fixed issue where input files were not closed correctly
    - Fixed performance bottleneck when adding secondary observation pixels

2.4 Applicable and Reference Documents

The following documents are applicable (AD) or references (RD) in this manual

RD 1	FIDUCEO_MMS_IP_v_1_1	Fiduceo Multisensor Matchup System Implementation Plan, version 1.1
RD 2	Buehler2007	S. A. Buehler, M. Kuvatov, T. R. Sreerekha, V. O. John, B. Rydberg, P. Eriksson, and J. Notholt: A cloud filtering method for microwave upper tropospheric humidity measurements, Atmos. Chem Phys. Discuss., 7, 7509–7534, 2007
RD 3	Hans2016	TN: Computing the Tb thresholds for cloud detection for the different viewing angles, Imke Hans, University of Hamburg
RD 4	SST_CCI-TN-UOL-003	ESA CCI Phase 2 (SST) - MMD Content Specification
AD 1	EUM/OPS-EPS/MAN/04/0032	IASI Level 1: Product Guide, issue v4C, 3 December 2014

2.5 Glossary

CDO	Climate Data Operators
CEDA	Centre for Environmental Data Analysis
CEMS	Climate and Environmental Monitoring from Space
JDBC	Java DataBase Connectivity
JDK	Java Development Kit
LZA	Local Zenith Angle
MMD	Multisensor Matchup Dataset
MMS	Multisensor Matchup System
NetCDF	Network Common Data Format
NWP	Numerical Weather Prediction
TN	Technical Note
VZA	Viewing Zenith Angle
XML	eXtensible Markup Language

3 Software Overview

The MMS is a software system that is composed of several interacting components. This chapter gives a high-level overview of the purpose and interactions of the components.

To operate the MMS tools, some external components need to be installed and accessible to the tools. This is mainly a unified archive of input data (see 4.1.1) and a database server engine (currently supported MongoDB, Apache H2 and PostGIS) with a user account accessible to the tools.

The calculation of matchup datasets is a three-step process:

1. Ingest satellite metadata into the database. This has to be run once for every input dataset that should be used for matchup analysis. This task is executed by the IngestionTool (chapter 3.1)
2. Find matchups and generate MMD files. This is done by the MatchupTool (chapter 3.2).
3. (optionally) Run a post-processing job on the generated MMDs. This is done by the PostProcessingTool (chapter 3.3).

3.1 Ingestion Tool

This tool extracts satellite metadata from the input files, pre-processes the metadata and stores the records to the database. The tool is invoked from the command line using the script *ingestion-tool.sh/bat* which is located in the *bin* directory of the MMS installation. The command line parameters can be plotted using the –h/--help option:

```
ingestion-tool version 1.5.0

usage: ingestion-tool <options>
Valid options are:
   -c,--config <arg>           Defines the configuration directory. Defaults to './config'.
   -end,--end-time <Date>      Define the ending time of products to inject.
   -h,--help                   Prints the tool usage.
   -s,--sensor <arg>           Defines the sensor to be ingested.
   -start,--start-time <Date>  Define the starting time of products to inject.
   -v,--version <arg>          Define the sensor data processing version.
```

--config: define the path to the configuration files directory

--sensor: define the sensor key to be ingested. Sensor keys are defined by the product reader modules available, see chapter 3.4.

--version: defines the sensor data processing version to be ingested. Must match one of the versions stored in the archive, see 4.1.1.

--start-date: defines the start date for the ingestion (format: “yyyy-DDD”)

--end-date: defines the end date for the ingestion (format: “yyyy-DDD”)

All satellite data files matching the type and version with an acquisition time start that falls into the interval defined by *start-date* and *end-date* will be processed.

The ingestion tool uses the following configuration files:

- system-config.xml (chapter 4.1.1)
- database.properties (chapter 4.1.2)

3.2 Matchup Tool

The matchup tool calculates matchups and generates MMD files. The tool is invoked from the command line using the script *matchup-tool.sh/bat* which is located in the *bin* directory of the MMS installation. The command line parameters can be plotted using the –h/--help option:

```
matchup-tool version 1.5.0

usage: matchup-tool <options>
Valid options are:
   -c,--config <arg>          Defines the configuration directory. Defaults to './config'.
   -end,--end-time <arg>      Defines the processing end-date, format 'yyyy-DDD'
   -h,--help                  Prints the tool usage.
   -start,--start-time <arg>  Defines the processing start-date, format 'yyyy-DDD'
   -u,--usecase <arg>         Defines the path to the use-case configuration file. Path is
relative to the
                              configuration directory.
```

--config: define the path to the configuration files directory

--start-time: defines the start date for the matchup processing (format: “yyyy-DDD”)

--end-time: defines the end date for the matchup processing (format: “yyyy-DDD”)

--usecase: defines the usecase configuration file to use, see chapter 4.1.4

The matchup tool uses the following configuration files:

- system-config.xml (chapter 4.1.1)
- database.properties (chapter 4.1.2)
- mmd-writer-config.xml (chapter 4.1.3)

3.3 Post Processing Tool

The post processing tool allows to add/remove/modify variables in a set of MMD files already generated. It is a command line tool using the following parameters (the same printout can be achieved by running the tool using the –h parameter). The tool is started using shells scripts located in the bin directory of the installation named *post-processing-tool.sh/bat*.

```
post-processing-tool version 1.5.0

usage: post-processing-tool <options>
Valid options are:
   -c,--config <arg>          Defines the configuration directory. Defaults to './config'.
   -end,--end-date <arg>      Defines the processing end-date, format 'yyyy-DDD'.
                              DDD = day of year.
```

```
  -h,--help                    Prints the tool usage.
  -i,--input-dir <arg>         Defines the path to the input mmd files directory.
  -j,--job-config <arg>        Defines the path to post processing job configuration file.
                               Path is relative to the configuration directory.
  -start,--start-date <arg>    Defines the processing start-date, format 'yyyy-DDD'.
                               DDD = day of year.
```

--config: define the path to the configuration files directory

--job-config: define the job-configuration file to use; the format of this file is described in chapter 4.1.5.

--input-dir: defines the input data directory. All files inside this directory which match the file naming pattern "*mmd\\d{1,2}.*_.*_.*_\\d{4}-\\d{3}_\\d{4}-\\d{3}.nc*" are collected as input files for post processing. The directory search runs recursively through all subdirectories.

--start-date: defines the start date for the post-processing (format: “yyyy-DDD”)

--end-date: defines the end date for the post processing (format: “yyyy-DDD”)

All MMD files with a start date within the period defined by start-date and end-date will be taken into account for processing. If any severe error occurs while computing a post processing, the processing will report the error and continue processing with the next input file, if possible.

The post processing tool uses the following configuration files:

- system-config.xml (chapter 4.1.1)
- post processing configuration (chapter 4.1.5)

3.4 Supported Satellite Sensors

This chapter lists the input satellite data source types available for matchup processing. In subsequent chapters, the sensors, the data format and the available variables are described in more detail.

3.4.1 AIRS-Level-1B

AIRS Level-1B Version v5.0.0.0

IR Calibrated Radiance Products (AIRIBRAD)

Sensor Key	Source
airs-aq	AIRS Level-1B data from Aqua platform

Variable Name	Description
Latitude	Footprint boresight geodetic Latitude in degrees North (-90.0 ... 90.0)
Longitude	Footprint boresight geodetic Longitude in degrees East (-180.0 ... 180.0)
Time	Footprint "shutter" TAI Time: floating-point elapsed seconds since Jan 1, 1993
scanang	Scanning angle of AIRS instrument with respect to the AIRS Instrument for

	this footprint (-180.0 ... 180.0, negative at start of scan, 0 at nadir)
satheight	Satellite altitude at nadirTAI in km above reference ellipsoid (e.g. 725.2)
satroll	Satellite attitude roll angle at nadirTAI (-180.0 ... 180.0 angle about the +x (roll) ORB axis, +x axis is positively oriented in the direction of orbital flight completing an orthogonal triad with y and z.)
satpitch	Satellite attitude pitch angle at nadirTAI (-180.0 ... 180.0 angle about +y (pitch) ORB axis. +y axis is oriented normal to the orbit plane with the positive sense opposite to that of the orbit's angular momentum vector H.)
satyaw	Satellite attitude yaw angle at nadirTAI (-180.0 ... 180.0 angle about +z (yaw) axis. +z axis is positively oriented Earthward parallel to the satellite radius vector R from the spacecraft center of mass to the center of the Earth.)
satgeoqa	Satellite Geolocation QA flags: Bit 0: (LSB, value 1) bad input value; Bit 1: (value 2) PGS_TD_TAItoUTC() gave PGSTD_E_NO_LEAP_SECS; Bit 2: (value 4) PGS_TD_TAItoUTC() gave PGS_E_TOOLKIT; Bit 3: (value 8) PGS_EPH_EphemAttit() gave PGSEPH_W_BAD_EPHEM_VALUE; Bit 4: (value 16) PGS_EPH_EphemAttit() gave PGSEPH_E_BAD_EPHEM_FILE_HDR; Bit 5: (value 32) PGS_EPH_EphemAttit() gave PGSEPH_E_NO_SC_EPHEM_FILE; Bit 6: (value 64) PGS_EPH_EphemAttit() gave PGSEPH_E_NO_DATA_REQUESTED; Bit 7: (value 128) PGS_EPH_EphemAttit() gave PGSTD_E_SC_TAG_UNKNOWN; Bit 8: (value 256) PGS_EPH_EphemAttit() gave PGSEPH_E_BAD_ARRAY_SIZE; Bit 9: (value 512) PGS_EPH_EphemAttit() gave PGSTD_E_TIME_FMT_ERROR; Bit 10: (value 1024) PGS_EPH_EphemAttit() gave PGSTD_E_TIME_VALUE_ERROR; Bit 11: (value 2048) PGS_EPH_EphemAttit() gave PGSTD_E_NO_LEAP_SECS; Bit 12: (value 4096) PGS_EPH_EphemAttit() gave PGS_E_TOOLKIT; Bit 13: (value 8192) PGS_CSC_ECItoECR() gave PGSCSC_W_BAD_TRANSFORM_VALUE; Bit 14: (value 16384) PGS_CSC_ECItoECR() gave PGSCSC_E_BAD_ARRAY_SIZE; Bit 15: (value 32768) PGS_CSC_ECItoECR() gave PGSTD_E_NO_LEAP_SECS; Bit 16: (value 65536) PGS_CSC_ECItoECR() gave PGSTD_E_TIME_FMT_ERROR; Bit 17: (value 131072) PGS_CSC_ECItoECR() gave PGSTD_E_TIME_VALUE_ERROR; Bit 18: unused (set to zero); Bit 19: (value 524288) PGS_CSC_ECItoECR() gave PGSTD_E_NO_UT1_VALUE; Bit 20:(value 1048576) PGS_CSC_ECItoECR() gave PGS_E_TOOLKIT;

	Bit 21: (value 2097152) PGS_CSC_ECRtoGEO() gave PGSCSC_W_TOO_MANY_ITERS; Bit 22: (value 4194304) PGS_CSC_ECRtoGEO() gave PGSCSC_W_INVALID_ALTITUDE; Bit 23: (value 8388608) PGS_CSC_ECRtoGEO() gave PGSCSC_W_SPHERE_BODY; Bit 24: (value 16777216) PGS_CSC_ECRtoGEO() gave PGSCSC_W_LARGE_FLATTENING; Bit 25: (value 33554432) PGS_CSC_ECRtoGEO() gave PGSCSC_W_DEFAULT_EARTH_MODEL; Bit 26: (value 67108864) PGS_CSC_ECRtoGEO() gave PGSCSC_E_BAD_EARTH_MODEL; Bit 27: (value 134217728) PGS_CSC_ECRtoGEO() gave PGS_E_TOOLKIT; Bit 28-31: not used
glintgeoqa	Glint Geolocation QA flags: Bit 0: (LSB, value 1) bad input value; Bit 1: (value 2) glint location in Earth's shadow (Normal for night FOVs); Bit 2: (value 4) glint calculation not converging; Bit 3: (value 8) glint location sun vs. satellite zenith mismatch; Bit 4: (value 16) glint location sun vs. satellite azimuth mismatch; Bit 5: (value 32) bad glint location; Bit 6: (value 64) PGS_CSC_ZenithAzimuth() gave any 'W' class return code; Bit 7: (value 128) PGS_CSC_ZenithAzimuth() gave any 'E' class return code; Bit 8: (value 256) PGS_CBP_Earth_CB_Vector() gave any 'W' class return code; Bit 9: (value 512) PGS_CBP_Earth_CB_Vector() gave any 'E' class return code; Bit 10: (value 1024) PGS_CSC_ECItoECR() gave any 'W' class return code except PGSCSC_W_PREDICTED_UT1 (for Glint); Bit 11: (value 2048) PGS_CSC_ECItoECR() gave any 'E' class return code (for Glint); Bit 12: (value 4096) PGS_CSC_ECRtoGEO() gave any 'W' class return code (for Glint); Bit 13: (value 8192) PGS_CSC_ECRtoGEO() gave any 'E' class return code (for Glint); Bit 14: (value 16384) PGS_CSC_ECItoECR() gave any 'W' class return code except PGSCSC_W_PREDICTED_UT1 ; Bit 15: (value 32768) PGS_CSC_ECItoECR() gave any 'E' class return code
moongeoqa	Moon Geolocation QA flags: Bit 0: (LSB, value 1) bad input value; Bit 1: (value 2) PGS_TD_TAItoUTC() gave PGSTD_E_NO_LEAP_SECS; Bit 2: (value 4) PGS_TD_TAItoUTC() gave PGS_E_TOOLKIT; Bit 3: (value 8) PGS_CBP_Sat_CB_Vector() gave PGSCSC_W_BELOW_SURFACE; Bit 4: (value 16) PGS_CBP_Sat_CB_Vector() gave PGSCBP_W_BAD_CB_VECTOR; Bit 5: (value 32) PGS_CBP_Sat_CB_Vector() gave PGSCBP_E_BAD_ARRAY_SIZE; Bit 6: (value 64) PGS_CBP_Sat_CB_Vector() gave

	PGSCBP_E_INVALID_CB_ID; Bit 7: (value 128) PGS_CBP_Sat_CB_Vector() gave PGSMEM_E_NO_MEMORY; Bit 8: (value 256) PGS_CBP_Sat_CB_Vector() gave PGSCBP_E_UNABLE_TO_OPEN_FILE; Bit 9: (value 512) PGS_CBP_Sat_CB_Vector() gave PGSTD_E_BAD_INITIAL_TIME; Bit 10: (value 1024) PGS_CBP_Sat_CB_Vector() gave PGSCBP_E_TIME_OUT_OF_RANGE; Bit 11: (value 2048) PGS_CBP_Sat_CB_Vector() gave PGSTD_E_SC_TAG_UNKNOWN; Bit 12: (value 4096) PGS_CBP_Sat_CB_Vector() gave PGSEPH_E_BAD_EPHEM_FILE_HDR; Bit 13: (value 8192) PGS_CBP_Sat_CB_Vector() gave PGSEPH_E_NO_SC_EPHEM_FILE; Bit 14: (value 16384) PGS_CBP_Sat_CB_Vector() gave PGS_E_TOOLKIT; Bit 15: not used
ftptgeoqa	Footprint Geolocation QA flags: Bit 0: (LSB, value 1) bad input value; Bit 1: (value 2) PGS_TD_TAItoUTC() gave PGSTD_E_NO_LEAP_SECS; Bit 2: (value 4) PGS_TD_TAItoUTC() gave PGS_E_TOOLKIT; Bit 3: (value 8) PGS_CSC_GetFOV_Pixel() gave PGSCSC_W_MISS_EARTH; Bit 4: (value 16) PGS_CSC_GetFOV_Pixel() gave PGSTD_E_SC_TAG_UNKNOWN; Bit 5: (value 32) PGS_CSC_GetFOV_Pixel() gave PGSCSC_W_ZERO_PIXEL_VECTOR; Bit 6: (value 64) PGS_CSC_GetFOV_Pixel() gave PGSCSC_W_BAD_EPH_FOR_PIXEL; Bit 7: (value 128) PGS_CSC_GetFOV_Pixel() gave PGSCSC_W_INSTRUMENT_OFF_BOARD; Bit 8: (value 256) PGS_CSC_GetFOV_Pixel() gave PGSCSC_W_BAD_ACCURACY_FLAG; Bit 9: (value 512) PGS_CSC_GetFOV_Pixel() gave PGSCSC_E_BAD_ARRAY_SIZE; Bit 10: (value 1024) PGS_CSC_GetFOV_Pixel() gave PGSCSC_W_DEFAULT_EARTH_MODEL; Bit 11: (value 2048) PGS_CSC_GetFOV_Pixel() gave PGSCSC_W_DATA_FILE_MISSING; Bit 12: (value 4096) PGS_CSC_GetFOV_Pixel() gave PGSCSC_E_NEG_OR_ZERO_RAD; Bit 13: (value 8192) PGS_CSC_GetFOV_Pixel() gave PGSMEM_E_NO_MEMORY; Bit 14: (value 16384) PGS_CSC_GetFOV_Pixel() gave PGSTD_E_NO_LEAP_SECS; Bit 15: (value 32768) PGS_CSC_GetFOV_Pixel() gave PGSTD_E_TIME_FMT_ERROR; Bit 16: (value 65536) PGS_CSC_GetFOV_Pixel() gave PGSTD_E_TIME_VALUE_ERROR;

	Bit 17: (value 131072) PGS_CSC_GetFOV_Pixel() gave PGSCSC_W_PREDICTED_UT1; Bit 18: (value 262144) PGS_CSC_GetFOV_Pixel() gave PGSTD_E_NO_UT1_VALUE; Bit 19: (value 524288) PGS_CSC_GetFOV_Pixel() gave PGS_E_TOOLKIT; Bit 20: (value 1048576) PGS_CSC_GetFOV_Pixel() gave PGSEPH_E_BAD_EPHEM_FILE_HDR; Bit 21: (value 2097152) PGS_CSC_GetFOV_Pixel() gave PGSEPH_E_NO_SC_EPHEM_FILE; Bit 22-31: not used
zengeoqa	Satellite zenith Geolocation QA flags: Bit 0: (LSB, value 1) (Spacecraft) bad input value; Bit 1: (value 2) PGS_CSC_ZenithAzimuth(S/C) gave PGSCSC_W_BELOW_HORIZON; Bit 2: (value 4) PGS_CSC_ZenithAzimuth(S/C) gave PGSCSC_W_UNDEFINED_AZIMUTH; Bit 3: (value 8) PGS_CSC_ZenithAzimuth(S/C) gave PGSCSC_W_NO_REFRACTION; Bit 4: (value 16) PGS_CSC_ZenithAzimuth(S/C) gave PGSCSC_E_INVALID_VECTAG; Bit 5: (value 32) PGS_CSC_ZenithAzimuth(S/C) gave PGSCSC_E_LOOK_PT_ALTIT_RANGE; Bit 6: (value 64) PGS_CSC_ZenithAzimuth(S/C) gave PGSCSC_E_ZERO_INPUT_VECTOR; Bit 7: (value 128) PGS_CSC_ZenithAzimuth(S/C) gave PGS_E_TOOLKIT; Bit 8: (value 256) (Sun) bad input value; Bit 9: (value 512) (suppressed) PGS_CSC_ZenithAzimuth(Sun) gave PGSCSC_W_BELOW_HORIZON (This is not an error condition - the sun is below the horizon at night); Bit 10: (value 1024) PGS_CSC_ZenithAzimuth(Sun) gave PGSCSC_W_UNDEFINED_AZIMUTH; Bit 11: (value 2048) PGS_CSC_ZenithAzimuth(Sun) gave PGSCSC_W_NO_REFRACTION; Bit 12: (value 4096) PGS_CSC_ZenithAzimuth(Sun) gave PGSCSC_E_INVALID_VECTAG; Bit 13: (value 8192) PGS_CSC_ZenithAzimuth(Sun) gave PGSCSC_E_LOOK_PT_ALTIT_RANGE; Bit 14: (value 16384) PGS_CSC_ZenithAzimuth(Sun) gave PGSCSC_E_ZERO_INPUT_VECTOR; Bit 15: (value 32768) PGS_CSC_ZenithAzimuth(Sun) gave PGS_E_TOOLKIT
demgeoqa	Digital Elevation Model (DEM) Geolocation QA flags: Bit 0: (LSB, value 1) bad input value; Bit 1: (value 2) Could not allocate memory; Bit 2: (value 4) Too close to North or South pole. Excluded. (This is not an error condition - a different model is used); Bit 3: (value 8) Layer resolution incompatibility. Excluded;

	Bit 4: (value 16) Any DEM Routine (elev) gave PGSDEM_E_IMPROPER_TAG; Bit 5: (value 32) Any DEM Routine (elev) gave PGSDEM_E_CANNOT_ACCESS_DATA; Bit 6: (value 64) Any DEM Routine (land/water) gave PGSDEM_E_IMPROPER_TAG; Bit 7: (value 128) Any DEM Routine (land/water) gave PGSDEM_E_CANNOT_ACCESS_DATA; Bit 8: (value 256) Reserved for future layers; Bit 9: (value 512) Reserved for future layers; Bit 10: (value 1024) PGS_DEM_GetRegion(elev) gave PGSDEM_M_FILLVALUE_INCLUDED; Bit 11: (value 2048) PGS_DEM_GetRegion(land/water) gave PGSDEM_M_FILLVALUE_INCLUDED; Bit 12: (value 4096) Reserved for future layers; Bit 13: (value 8192) PGS_DEM_GetRegion(all) gave PGSDEM_M_MULTIPLE_RESOLUTIONS; Bit 14: (value 16384) PGS_CSC_GetFOV_Pixel() gave any 'W' class return code except PGSCSC_W_PREDICTED_UT1; Bit 15: (value 32768) PGS_CSC_GetFOV_Pixel() gave any 'E' class return code
nadirTAI	TAI time at which instrument is nominally looking directly down. (between footprints 15 & 16 for AMSU or between footprints 45 & 46 for AIRS/Vis & HSB) (floating-point elapsed seconds since start of 1993)
sat_lat	Satellite geodetic latitude in degrees North (-90.0 ... 90.0)
sat_lon	Satellite geodetic longitude in degrees East (-180.0 ... 180.0)
scan_node_type	'A' for ascending, 'D' for descending, 'E' when an error is encountered in trying to determine a value.
satzen	Spacecraft zenith angle (0.0 ... 180.0) degrees from zenith (measured relative to the geodetic vertical on the reference (WGS84) spheroid and including corrections outlined in EOS SDP toolkit for normal accuracy.)
satazi	Spacecraft azimuth angle (-180.0 ... 180.0) degrees E of N GEO)
glintlat	Solar glint geodetic latitude in degrees North at nadirTAI (-90.0 ... 90.0)
glintlon	Solar glint geodetic longitude in degrees East at nadirTAI (-180.0 ... 180.0)
sun_glint_distance	Distance (km) from footprint center to location of the sun glint (-9999 for unknown, 30000 for no glint visible because spacecraft is in Earth's shadow)
topog	Mean topography in meters above reference ellipsoid
topog_err	Error estimate for topog
landFrac	Fraction of spot that is land (0.0 ... 1.0)
landFrac_err	Error estimate for landFrac
state	Data state: 0:Process, 1:Special, 2:Erroneous, 3:Missing
CalScanSummary	Bit field. Bitwise OR of CalFlag over the all channels with ExcludedChans < 3. Zero means all "good" channels were well calibrated for this scanline Bit 7 (MSB): scene over/underflow; Bit 6: (value 64) anomaly in offset calculation; Bit 5: (value 32) anomaly in gain calculation; Bit 4: (value 16) pop detected;

	Bit 3: (value 8) DCR Occurred; Bit 2: (value 4) Moon in View; Bit 1: (value 2) telemetry out of limit condition; Bit 0: (LSB, value 1) cold scene noise
spaceview_selection	Indicates which footprints were included for this scan. Each bit is high when the corresponding space view is used in the spaceview offset calculation. (See L1B Processing Requirements, section 6.2); LSB is first space view.
Rdiff_swindow	Radiance difference in the 2560 cm**-1 window region used to warn of possible errors caused by scene non-uniformity and misalignment of the beams: radiance(Rdiff_swindow_M1a_chan) - radiance(Rdiff_swindow_M2a_chan). (radiance units)
Rdiff_lwindow	Radiance difference in the longwave window(850 cm**-1) used to warn of possible errors caused by scene non-uniformity and misalignment of the point beams: radiance(Rdiff_lwindow_M8_chan) - radiance(Rdiff_lwindow_M9_chan). (radiance units)
SceneInhomogeneous	Threshold test for scene inhomogeneity, using band-overlap detectors (bit fields).; Bit 7 (MSB, value 128): scene is inhomogeneous, as determined by the Rdiff_swindow threshold. For v5.0 the test is abs(Rdiff_swindow) > 5 * sqrt(NeN(Rdiff_swindow_M1a_chan)^2 + NeN(Rdiff_swindow_M2a_chan)); Bit 6 (value 64): scene is inhomogeneous, as determined by the Rdiff_lwindow threshold. For v5.0 the test is abs(Rdiff_lwindow) > 5 * sqrt(NeN(Rdiff_lwindow_M8_chan)^2 + NeN(Rdiff_lwindow_M9_chan)); Bits 5-0: unused (reserved)
dust_flag	Flag telling whether dust was detected in this scene; 1: Dust detected; 0: Dust not detected; -1: Dust test not valid because of land; -2: Dust test not valid because of high latitude; -3: Dust test not valid because of suspected cloud; -4: Dust test not valid because of bad input data
dust_score	Dust score. Each bit results from a different test comparing radiances. Higher scores indicate more certainty of dust present. Dust probable when score is over 380. Not valid when dust_flag is negative.
spectral_clear_indicator	Flag telling whether scene was flagged as clear by a spectral filter. Only ocean filter is validated; 2: Ocean test applied and scene identified as clear; 1: Ocean test applied and scene not identified as clear; 0: Calculation could not be completed. Possibly some inputs were missing or FOV is on coast or on the edge of a scan or granule; -1: Unvalidated land test applied and scene not identified as clear; -2: Unvalidated land test applied and scene identified as clear
BT_diff_S02	Brightness temperature difference Tb(1361.44 cm-1) - Tb(1433.06 cm-1) used as an indicator of SO2 release from volcanoes. Values under -6 K have likely volcanic SO2. (Kelvins)

OpMode	Instrument Operations Mode. See AIRS Command Handbook, section 6.4 for a definition of each bit. Bits 0 (LSB)-2 cal phase; bits 3-6 Cal Func; Bit 7 quicklook (expedited) flag; bits 8-11 submode Bits 12-14 Mode (0=standby, 1=ready, 2=operate, 3=checkout, 4=decontaminate, 5=off, 6=survival); bit 16 transition flag
EDCBOARD	EDC A/B Powered on Indicator:; 0: Both sides off; 1: Side A; 2: Side B; 3: Invalid; 65534: No value downlinked

3.4.2 AMSR-2

AMSR-2 L1R brightness temperatures in HDF5 format as supplied by JAXA. This reader only handles data in the low resolution, the 89 GHz A and B data are ignored.

Sensor Key	Source
amsr2-gcw1	AMSR-2 data from GCOM-W1 platform

Variable Name	Description
Area_Mean_Height	Altitude data
Brightness_Temperature_(resXX,6.9GHz,Y)	Observed brightness temperature at 6.9 GHz, resampled to resolution XX, polarisation Y
Brightness_Temperature_(resXX,7.3GHz,Y)	Observed brightness temperature at 7.3 GHz, resampled to resolution XX, polarisation Y
Brightness_Temperature_(resXX,10.7GHz,Y)	Observed brightness temperature at 10.7 GHz, resampled to resolution XX, polarisation Y
Brightness_Temperature_(resXX,18.7GHz,Y)	Observed brightness temperature at 18.7 GHz, resampled to resolution XX, polarisation Y
Brightness_Temperature_(resXX,23.8GHz,Y)	Observed brightness temperature at 23.8 GHz, resampled to resolution XX, polarisation Y
Brightness_Temperature_(resXX,36.5GHz,Y)	Observed brightness temperature at 36.5 GHz, resampled to resolution XX, polarisation Y
Brightness_Temperature_(resXX,89.0GHz,Y)	Observed brightness temperature at 89.0 GHz, resampled to resolution XX, polarisation Y
Earth_Incidence	Incidence angle of observation in degrees
Earth_Azimuth	Azimuth angle of observation in degrees
Sun_Elevation	Sun elevation angle in degrees
Sun_Azimuth	Sun azimuth angle in degrees
Land_Ocean_Flag_6	Land coverage flag for 6.9 GHz channel
Land_Ocean_Flag_10	Land coverage flag for 10.7 GHz channel
Land_Ocean_Flag_23	Land coverage flag for 23.8 GHz channel
Land_Ocean_Flag_36	Land coverage flag for 36.5 GHz channel

Latitude_of_Observation_Point_for_89A	Latitude in degrees, subsampled to match the low resolution raster size
Longitude_of_Observation_Point_for_89A	Longitude in degrees, subsampled to match the low resolution raster size
Pixel_Data_Quality_6_to_36	Pixel data quality flags, rearranged to SHORT data type by merging the two BYTE data arrays
Position_in_Orbit	The satellite position on the orbit
Scan_Time	Acquisition time per scanline in TAI seconds since 1993

3.4.3 AMSR-E

AMSR-E L2A brightness temperatures in HDF4 format. This reader only handles the "Low_Res_Swath" data.

Sensor Key	Source
amsre-aq	AMSR-E data from Aqua platform

Variable Name	Description
Time	Acquisition time per scanline in TAI seconds since 1993
Latitude	Latitude of pixel in decimal degrees
Longitude	Longitude of pixel in decimal degrees
Earth_Incidence	Incidence angle of observation in degrees
Earth_Azimuth	Azimuth angle of observation in degrees
Sun_Elevation	Sun elevation angle in degrees
Sun_Azimuth	Sun azimuth angle in degrees
Land_Ocean_Flag_6	Land coverage flag for 6 GHz channel
X_XV_Res_1_TB	Brightness temperatures vertical polarisation in low resolution for channels X_X (6_9, 10_7, 18_7, 23_8, 36_5, 89_0)
X_XH_Res_1_TB	Brightness temperatures horizontal polarisation in low resolution for channels X_X (6_9, 10_7, 18_7, 23_8, 36_5, 89_0)
Scan_Quality_Flag	Quality flags per scan
Channel_Quality_Flag_XXV	Quality flags per vertical polarisation channel XX (6, 10, 18, 23, 36, 89)
Channel_Quality_Flag_XXH	Quality flags per horizontal polarisation channel XX (6, 10, 18, 23, 36, 89)
Res1_Surf	Land coverage percentage for resolution one
Geostationary_Reflection_Latitude	RFI sources glint angles latitude
Geostationary_Reflection_Longitude	RFI sources glint angles longitude
Sun_Glint_Angle	Angular separation between the Reflected Satellite View Vector and the Sun Vector. When Sun_Glint_Angle is zero, the instrument views the center of the specular (mirror-like) sun reflection.

3.4.4 AMSU-B

AMSU-B L1c data from NOAA CLASS storage, converted to HDF using “atovin” and “convert_to_hdf5”.

Sensor Key	Source
amsub-n15	AMSU-B NOAA15
amsub-n16	AMSU-B NOAA16
amsub-n17	AMSU-B NOAA17

Variable Name	Description
btemps_ch1	Brightness temperature in channel 1 (2.9684 m^{-1})
btemps_ch2	Brightness temperature in channel 2 (5.0032 m^{-1})
btemps_ch3	Brightness temperature in channel 3 (6.1146 m^{-1})
btemps_ch4	Brightness temperature in channel 4 (6.1146 m^{-1})
btemps_ch5	Brightness temperature in channel 5 (6.1146 m^{-1})
chanqual_ch1	Channel quality for channel 1
chanqual_ch2	Channel quality for channel 2
chanqual_ch3	Channel quality for channel 3
chanqual_ch4	Channel quality for channel 4
chanqual_ch5	Channel quality for channel 5
instrtemp	Instrument temperature
qualind	Quality indicator
scanqual	Scan line quality
scnlin	Scan line number
scnlindy	Day of scan line acquisition
scnlintime	Time of scan line acquisition
scnlinyr	Year of scan line acquisition
Latitude	Latitude of pixel
Longitude	Longitude of pixel
Time	Acquisition time in TAI93 seconds
Satellite_azimuth_angle	The satellite azimuth angle
Satellite_zenith_angle	The satellite zenith angle
Solar_azimuth_angle	The solar azimuth angle
Solar_zenith_angle	The solar zenith angle

All per-scan line variables are extended to cover the full pixel grid.

3.4.5 (A)ATSR

ATSR1, ATSR2 and AATSR data in ENVISAT native format (*.N1). The MMS supports data in processing version v2.1 and v3 format.

Sensor Key	Source
atsr-e1	ATSR1 on ERS1
atsr-e2	ATSR2 on ERS2
aatsr-en	AATSR on ENVISAT

Variable Name	Description
btemp_nadir_1200	Brightness temperature, nadir view (11500-12500 nm)
btemp_nadir_1100	Brightness temperature, nadir view (10400-11300 nm)
btemp_nadir_0370	Brightness temperature, nadir view (3505-3895 nm)
reflec_nadir_1600	Reflectance, nadir view (1580-1640 nm)
reflec_nadir_0870	Reflectance, nadir view (855-875 nm)
reflec_nadir_0670	Reflectance, nadir view (649-669 nm)
reflec_nadir_0550	Reflectance, nadir view (545-565 nm)
confid_flags_nadir	Confidence flags, nadir view
cloud_flags_nadir	Cloud flags, nadir view
btemp_fward_1200	Brightness temperature, forward view (11500-12500 nm)
btemp_fward_1100	Brightness temperature, forward view (10400-11300 nm)
btemp_fward_0370	Brightness temperature, forward view (3505-3895 nm)
reflec_fward_1600	Reflectance, forward view (1580-1640 nm)
reflec_fward_0870	Reflectance, forward view (855-875 nm)
reflec_fward_0670	Reflectance, forward view (649-669 nm)
reflec_fward_0550	Reflectance, forward view (545-565 nm)
confid_flags_fward	Confidence flags, forward view
cloud_flags_fward	Cloud flags, forward view
lat_corr_nadir	Latitude corrections, nadir view
lon_corr_nadir	Longitude corrections, nadir view
sun_elev_nadir	Solar elevation, nadir view
view_elev_nadir	Satellite elevation, nadir view
sun_azimuth_nadir	Solar azimuth, nadir view
view_azimuth_nadir	Satellite azimuth, nadir view
lat_corr_fward	Latitude corrections, forward view
lon_corr_fward	Longitude corrections, forward view
sun_elev_fward	Solar elevation, forward view
view_elev_fward	Satellite elevation, forward view
sun_azimuth_fward	Solar azimuth, forward view
view_azimuth_fward	Satellite azimuth, forward view
latitude	Latitudes
longitude	Longitudes
altitude	Topographic altitude

All tie-point raster variables are interpolated to the full raster resolution.

3.4.6 AVHRR

AVHRR L1C GAC data in NetCDF format, converted using routines developed at University of Reading. The MMS supports data of processing versions 1.2, 1.3, 1.4-cspp, 1.5 and 2.10.1.

Sensor Key	Source
avhrr-n06	AVHRR NOAA06

avhrr-n07	AVHRR/2 NOAA07
avhrr-n08	AVHRR NOAA08
avhrr-n09	AVHRR/2 NOAA09
avhrr-n10	AVHRR NOAA10
avhrr-n11	AVHRR/2 NOAA11
avhrr-n12	AVHRR/2 NOAA12
avhrr-n14	AVHRR/2 NOAA14
avhrr-n15	AVHRR/3 NOAA15
avhrr-n16	AVHRR/3 NOAA16
avhrr-n17	AVHRR/3 NOAA17
avhrr-n18	AVHRR/3 NOAA18
avhrr-n19	AVHRR/3 NOAA19
avhrr-m02	AVHRR/3 METOP A
avhrr-ma	AVHRR/3 METOP A
avhrr-m01	AVHRR/3 METOP B
avhrr-mb	AVHRR/3 METOP B

Variable Name	Description
ch1	Channel 1 reflectance
ch1_earth_counts	Channel 1 photon counts earth view (only v2.10.1)
ch1_noise	Channel 1 noise estimate (only v01.4-cspp, 1.5, v2.10.1)
ch2	Channel 2 reflectance
ch2_earth_counts	Channel 2 photon counts earth view (only v2.10.1)
ch2_noise	Channel 2 noise estimate (only v01.4-cspp, 1.5, v2.10.1)
ch3	Channel 3 reflectance (only AVHRR/2)
ch3_earth_counts	Channel 3 photon counts earth view (only v2.10.1) (only AVHRR/2)
ch3_noise	Channel 3 noise estimate (only v01.4-cspp, 1.5) (only AVHRR/2)
ch3a	Channel 3a reflectance (only AVHRR/2 AVHRR/3)
ch3a_earth_counts	Channel 3a photon counts earth view (only v2.10.1) (only AVHRR/2 AVHRR/3)
ch3a_noise	Channel 3a noise estimate (only v01.4-cspp, 1.5) (only AVHRR/2 AVHRR/3)
ch3b	Channel 3b brightness temperature
ch3b_bbody_counts	Channel 3b smoothed black body counts (only v2.10.1)
ch3b_earth_counts	Channel 3b photon counts earth view (only v2.10.1)
ch3b_nedt	Channel 3b noise estimate (only v01.4-cspp, 1.5)
ch3b_space_counts	Channel 3b photon counts space view (only v2.10.1)
ch4	Channel 4 brightness temperature
ch4_bbody_counts	Channel 4 smoothed black body counts (only v2.10.1)
ch4_earth_counts	Channel 4 photon counts earth view (only v2.10.1)
ch4_nedt	Channel 4 noise estimate (only v01.4-cspp, 1.5)
ch4_space_counts	Channel 4 photon counts space view (only v2.10.1)
ch5	Channel 5 brightness temperature (only AVHRR/3)
ch5_bbody_counts	Channel 5 smoothed black body counts (only v2.10.1)
ch5_earth_counts	Channel 5 photon counts earth view (only v2.10.1)
ch5_nedt	Channel 5 noise estimate (only v01.4-cspp, 1.5)
ch5_space_counts	Channel 5 photon counts space view (only v2.10.1)
cloud_mask	CLAVR-X cloud mask

cloud_probability	CLAVR-X cloud probability
dtime	Scanline time difference from start time
ict_temp	Temperature of internal calibration target
l1b_line_number	Level 1b line number
lat	Latitude coordinates
lon	Longitude coordinates
orbital_temperature	Average orbital temperature (t_inst) (only v2.10.1)
orbital_temperature_nlines	Number of lines over which the average orbital temperature was calculated (only v2.10.1)
prt_1	PRT1 smoothed temperatures (only v2.10.1)
prt_2	PRT2 smoothed temperatures (only v2.10.1)
prt_3	PRT3 smoothed temperatures (only v2.10.1)
prt_4	PRT4 smoothed temperatures (only v2.10.1)
qual_flags	Quality Flags
relative_azimuth_angle	Relative Azimuth Angle (not v2.10.1)
satellite_azimuth_angle	satellite azimuth angle (only v2.10.1)
solar_azimuth_angle	solar azimuth angle (only v2.10.1)
satellite_zenith_angle	Satellite zenith angle
solar_zenith_angle	Solar zenith angle

3.4.7 AVHRR FCDR

AVHRR FCDR data from FIDUCEO project in NetCDF format. Some variables need to be transferred to the MMD using a post-processing as their dimensionality does not match the core engine requirements. Please refer to chapter 4.1.5.1.14.

Sensor Key	Source
avhrr-n06-fcdr	AVHRR NOAA06 FCDR
avhrr-n07-fcdr	AVHRR/2 NOAA07 FCDR
avhrr-n08-fcdr	AVHRR NOAA08 FCDR
avhrr-n09-fcdr	AVHRR/2 NOAA09 FCDR
avhrr-n10-fcdr	AVHRR NOAA10 FCDR
avhrr-n11-fcdr	AVHRR/2 NOAA11 FCDR
avhrr-n12-fcdr	AVHRR/2 NOAA12 FCDR
avhrr-n14-fcdr	AVHRR/2 NOAA14 FCDR
avhrr-n15-fcdr	AVHRR/3 NOAA15 FCDR
avhrr-n16-fcdr	AVHRR/3 NOAA16 FCDR
avhrr-n17-fcdr	AVHRR/3 NOAA17 FCDR
avhrr-n18-fcdr	AVHRR/3 NOAA18 FCDR
avhrr-n19-fcdr	AVHRR/3 NOAA19 FCDR
avhrr-ma-fcdr	AVHRR/3 METOP A FCDR

Variable Name	Description
Ch1	Channel 1 reflectance
Ch2	Channel 2 reflectance
Ch3a	Channel 3a reflectance
Ch3b	Channel 3b brightness temperature

Ch4	Channel 4 brightness temperature
Ch5	Channel 5 brightness temperature
Time	Acquisition time in seconds since 1970-01-01 00:00:00
data_quality_bitmask	bitmask for quality per pixel
latitude	Geolocation latitude
longitude	Geolocation longitude
quality_pixel_bitmask	Flag mask
quality_scanline_bitmask	bitmask for quality per scanline
relative_azimuth_angle	Viewing geometry azimuth angle
satellite_zenith_angle	Viewing geometry zenith angle
scanline_map_to_origl1bfile	Indicator for mapping each line to its corresponding original level 1b file. See global attribute 'source' for the filenames. 0 corresponds to 1st listed file, 1 to 2nd file.
scanline_origl1b	Original scan line numbers from corresponding l1b records
solar_zenith_angle	Solar zenith angle
u_common_Ch1	common uncertainty per pixel for channel 1
u_common_Ch2	common uncertainty per pixel for channel 2
u_common_Ch3a	common uncertainty per pixel for channel 3a
u_common_Ch3b	common uncertainty per pixel for channel 3b
u_common_Ch4	common uncertainty per pixel for channel 4
u_common_Ch5	common uncertainty per pixel for channel 5
u_independent_Ch1	independent uncertainty per pixel for channel 1
u_independent_Ch2	independent uncertainty per pixel for channel 2
u_independent_Ch3a	independent uncertainty per pixel for channel 3a
u_independent_Ch3b	independent uncertainty per pixel for channel 3b
u_independent_Ch4	independent uncertainty per pixel for channel 4
u_independent_Ch5	independent uncertainty per pixel for channel 5
u_structured_Ch1	structured uncertainty per pixel for channel 1
u_structured_Ch2	structured uncertainty per pixel for channel 2
u_structured_Ch3a	structured uncertainty per pixel for channel 3a
u_structured_Ch3b	structured uncertainty per pixel for channel 3b
u_structured_Ch4	structured uncertainty per pixel for channel 4
u_structured_Ch5	structured uncertainty per pixel for channel 5

3.4.8 AVHRR FRAC

AVHRR full resolution L1B data in NOAA format. Reader supports gzip compressed products.

Sensor Key	Source
avhrr-frac-ma	AVHRR full resolution on MetopA

Variable Name	Description
radiance_[1, 2, 3a, 3b, 4, 5]	TOA radiances
reflec_[1, 2, 3a, 3b, 4, 5]	TOA reflectances
flags	Quality flags

cloudFlag	Cloud flags
sun_zenith	Solar zenith angle
view_zenith	Viewing zenith angle
delta_azimuth	Azimuth angle difference
latitude	
longitude	

3.4.9 CALIOP CLay

CALIOP Lidar L2 Cloud Layer Product in NetCDF file format.

Sensor Key	Source
caliop_clay-cal	CALIOP Clay Standard V4.10 on CALIPSO

Variable Name	Description
Profile_ID	The ID of the vertical profile
Latitude	Latitude in degrees of the laser footprint
Longitude	Longitue in degrees of the laser footprint
Profile_Time	Time expressed in International Atomic Time (TAI)
Profile_UTC_Time	Time expressed in Coordinated Universal Time (UTC)
Spacecraft_Position_x	ECR coordinate system X
Spacecraft_Position_y	ECR coordinate system Y
Spacecraft_Position_z	ECR coordinate system Z
Number_Layers_Found	The number of cloud layers found in this coulumn
Column_Feature_Fraction	The fraction of the 5-km horizontally averaged profile, which has been identified as containing a feature (i.e., either a cloud, an aerosol, or a stratospheric layer)
FeatureFinderQC	A set of feature finder QC flags
Feature_Classification_Flags	See https://www-calipso.larc.nasa.gov/resources/calipso_users_guide/data_summaries/vfm/index.php
ExtinctionQC_532	cloud extinction is only reported at 532 nm
CAD_Score	The cloud-aerosol discrimination score
Layer_IAB_QA_Factor	=1 - $F(\gamma'_{above})$, where $F(\gamma'_{above})$ is the cumulative probability of measuring a complete column IABequal to γ'_{above}.
Opacity_Flag	An opacity flag value of 1 indicates an opaque layer; values of 0 indicate transparent layers
Ice_Water_Path	The integral, from layer top to layer base, of the ice-water content profile within any ice cloud layer
Ice_Water_Path_Uncertainty	
Feature_Optical_Depth_532	Layer optical depth computed according to the procedures outlined in the CALIOP extinction retrieval ATBD, cloud optical depth is reported only for the 532 nm measurements
Feature_Optical_Depth_Uncertainty_532	
Layer_Top_Altitude	

Layer_Base_Altitude	

3.4.10 CALIOP VFM

CALIOP Lidar L2 Vertical Feature Product in NetCDF file format.

Sensor Key	Source
caliop_vfm-cal	CALIOP VFM Standard V4.10 on CALIPSO

Variable Name	Description
Profile_ID	The ID of the vertical profile
Latitude	Latitude in degrees of the laser footprint
Longitude	Longitue in degrees of the laser footprint
Profile_Time	Time expressed in International Atomic Time (TAI)
Profile_UTC_Time	Time expressed in Coordinated Universal Time (UTC)
Day_Night_Flag	Indicates the lighting conditions at an altitude of ~24 km above mean sea level. 0 = day, 1 = night
Land_Water_Mask	a 30 arc second resolution land/water mask provided by the SDP toolkit
Spacecraft_Position_x	ECR coordinate system X
Spacecraft_Position_y	ECR coordinate system Y
Spacecraft_Position_z	ECR coordinate system Z

3.4.11 HIRS

HIRS L1C data converted to NetCDF using tools developed at University of Reading.

Sensor Key	Source
hirs-tn	HIRS/2 on Tiros-N
hirs-n06 – n12	HIRS/2 on NOAA 6 to NOAA 12
hirs-n14	HIRS/2 on NOAA 14
hirs-n15 – n17	HIRS/3 on NOAA 15 to NOAA 17
hirs-n18 – n19	HIRS/4 on NOAA 18 to NOAA 19
hirs-ma	HIRS/4 on Metop-A
hirs-mb	HIRS/4 on Metop-B

Variable Name	Description
time	Acquisition time per scanline in seconds since 1970
lat	Latitude of pixel
lon	Longitude of pixel
bt_ch01 – ch19	Brightness temperatures for channel 1 to 19
radiance_ch01 – ch19	TOA radiances for channel 1 to 19
counts_ch01 – ch19	Raw measurement counts for channels 1 to 19
lza	Local zenith angle
scanline	Original scanline number
scanpos	Original scan position number

scanline_type	Scanline type flag

3.4.12 HIRS FCDR

HIRS FCDR data from FIDUCEO project in NetCDF format. Some variables need to be transferred to the MMD using a post-processing as their dimensionality does not match the core engine requirements. Please refer to chapter TODO.

Sensor Key	Source
hirs-n06 – n12-fcdr	HIRS/2 on NOAA 6 to NOAA 12
hirs-n14-fcdr	HIRS/2 on NOAA 14
hirs-n15 – n17-fcdr	HIRS/3 on NOAA 15 to NOAA 17
hirs-n18 – n19-fcdr	HIRS/4 on NOAA 18 to NOAA 19
hirs-ma-fcdr	HIRS/4 on Metop-A
hirs-mb-fcdr	HIRS/4 on Metop-B

Variable Name	Description
latitude	Latitude of pixel
longitude	Longitude of pixel
bt_ch01 – ch19	Brightness temperature, NOAA/EUMETSAT calibrated for channel 1 to 19
satellite_zenith_angle	platform_zenith_angle
satellite_azimuth_angle	sensor_azimuth_angle
solar_zenith_angle	solar_zenith_angle
solar_azimuth_angle	solar_azimuth_angle
time	Acquisition time in seconds since 1970-01-01 00:00:00
quality_scanline_bitmask	Scanline quality indicator
scanline_map_to_origl1bfile	Indicator for mapping each line to its corresponding original level 1b file. See global attribute 'source' for the filenames. 0 corresponds to 1st listed file, 1 to 2nd file
scanline_origl1b	Original scan line numbers from corresponding l1b records
quality_channel_bitmask_ch01 – ch19	Channel quality indicator for channel 1 to 19
u_independent_ch01 – ch19	Uncertainty from independent errors for channel 1 to 19
u_structured_ch01 – ch19	Uncertainty from structured errors for channel 1 to 19
u_common_ch01 – ch19	Uncertainty from common errors for channel 1 to 19

3.4.13 IASI

IASIS L1C data in Eumetsat format following the specification in AD 1. Supported is data using an MDR of class 8, subclass 2, version 4 and version 5.

The reader implements a virtual x/y raster on a per-EFOV basis, all per scan or per snapshot variables are extended to this raster.

IMPORTANT: this reader per-se does not supply the IASI spectrum, this has to be added using a post-processing step (see chapter 4.1.5.1.7).

Sensor Key	Source
iasi-ma	IASI on MetopA
iasi-mb	IASI on MetopB

Variable Name	Description
DEGRADED_INST_MDR	Quality of MDR has been degraded from nominal due to an instrument degradation
DEGRADED_PROC_MDR	Quality of MDR has been degraded from nominal due to a processing degradation
GEPSIasiMode	Instrument mode
GEPSOPSProcessingMode	Processing mode
OBT	On Board Time (Coarse time + Fine time)
OnboardUTC	Date of IASI measure (on board UTC)
GEPSDatIasi	Date of IASI measure (corrected UTC)
GEPS_CCD	Corner Cube Direction for all observational targets
GEPS_SP	Scan position for all observational targets
GQisFlagQualDetailed	Detailed quality flag for the system (*not in v4*)
GQisQualIndex	System-IASI general quality index
GQisQualIndexIIS	IIS imager quality index inside 1c product
GQisQualIndexLoc	Geometric quality index for sounder product
GQisQualIndexRad	Radiometric quality index for sounder product
GQisQualIndexSpect	Spectral quality index for sounder product
GQisSysTecIISQual	System-TEC quality index for IIS
GQisSysTecSondQual	System-TEC quality index for sounder
GGeoSondLoc_Lon	Location of pixel centre in geodetic coordinates for each sounder pixel (lon)
GGeoSondLoc_Lat	Location of pixel centre in geodetic coordinates for each sounder pixel (lat)
GGeoSondAnglesMETOP_Zenith	Measurement angles for each sounder pixel (zenith)
GGeoSondAnglesMETOP_Azimuth	Measurement angles for each sounder pixel (azimuth)
GGeoSondAnglesSUN_Zenith	Solar angles at the surface for each sounder pixel (zenith)
GGeoSondAnglesSUN_Azimuth	Solar angles at the surface for each sounder pixel (azimuth)
EARTH_SATELLITE_DISTANCE	Distance of satellite from Earth centre
IDefSpectDWn1b	Sample width of IASI 1C spectra (same as 1B)
IDefNsfirst1b	Number of the first sample of IASI 1C spectra (same as 1B)
IDefNslast1b	Number of the last sample of IASI 1C spectra (same as 1B)
GCcsRadAnalNbClass	Radiance Analysis: Number of identified classes in the sounder FOV
IDefCcsMode	Radiance Analysis: Image used is from AVHRR or IIS imager (degraded cases)

GCcsImageClassifiedNbLin	Radiance Analysis: Number of useful lines
GCcsImageClassifiedNbCol	Radiance Analysis: Number of useful columns
GCcsImageClassifiedFirstLin	First line of the classified image (number in the Avhrr raster, as per section 2.5)
GCcsImageClassifiedFirstCol	First column of the classified image (number in the Avhrr raster, as per section 2.5)
GIacVarImagIIS	Variance of IIS image (*not in v4*)
GIacAvgImagIIS	Average of IIS image (*not in v4*)
GEUMAvhrr1BCldFrac	Cloud fraction in IASI FOV from AVHRR 1B in IASI FOV (*not in v4*)
GEUMAvhrr1BLandFrac	Land and Coast fraction in IASI FOV from AVHRR 1B (*not in v4*)
GEUMAvhrr1BQual	Quality indicator. If the quality is good, it gives the coverage of snow/ice. (*not in v4*)

Note: The MMD will always generate variables according to the MDR_1C specification version 5, when encountering a file containing MDR_1C version 4 data, the additional variables will be “fake” variables that contain only the appropriate fill value. The variables are marked as “not in v4” in the table above.

3.4.14 MHS

MHS L1c data from NOAA CLASS storage, converted to HDF using “atovin” and “convert_to_hdf5”.

Sensor Key	Source
mhs-n18	MHS on NOAA18
mhs-n19	MHS on NOAA19
mhs-ma	MHS on MetopA
mhs-mb	MHS on MetopB

Variable Name	Description
btemps_ch1 … ch5	Brightness temperatures for channel 1 to 5
chanqual_ch1 … ch5	Channel quality flags
instrtemp	Instrument temperature in Kelvin
qualind	Quality indicator
scanqual	Scan line quality
scnlin	Scan line number
scnlindy	Scan line acquisition day of year
scnlintim	Scan line acquisition seconds of day
scnlinyr	Scan line acquisition year
Latitude	Pixel latitude in degrees
Longitude	Pixel longitude in degrees
Satellite_azimuth_angle	Satellite azimuth angle
Satellite_zenith_angle	Satellite zenith angle
Solar_azimuth_angle	Sun azimuth angle
Solar_zenith_angle	Sun zenith angle

All per-scan line variables are extended to cover the full pixel grid.

3.4.15 MODIS MxD06

Modis MxD06 L2 cloud products from Terra and Aqua platform, standard NASA HDF file format.

This reader only uses the low-resolution data. High resolution data is re-mapped to the low resolution data by picking the centre pixel of a 5x5 1km pixel acquisition.

Sensor Key	Source
mod06-te	Modis cloud data from Terra platform
myd06-aq	Modis cloud data from Aqua platform

Variable Name	Description
Latitude	Pixel latitude in degrees
Longitude	Pixel longitude in degrees
Scan_Start_Time	TAI time at start of scan replicated across the swath
Solar_Zenith	Solar Zenith Angle, Cell to Sun
Solar_Zenith_Day	Solar Zenith Angle, Cell to Sun, Day Data Only
Solar_Zenith_Night	Solar Zenith Angle, Cell to Sun, Night Data Only
Solar_Azimuth	Solar Azimuth Angle, Cell to Sun
Solar_Azimuth_Day	Solar Azimuth Angle, Cell to Sun, Day Data Only
Solar_Azimuth_Night	Solar Azimuth Angle, Cell to Sun, Night Data Only
Sensor_Zenith	Sensor Zenith Angle, Cell to Sensor
Sensor_Zenith_Day	Sensor Zenith Angle, Cell to Sensor, Day Data Only
Sensor_Zenith_Night	Sensor Zenith Angle, Cell to Sensor, Night Data Only
Sensor_Azimuth	Sensor Azimuth Angle, Cell to Sensor
Sensor_Azimuth_Day	Sensor Azimuth Angle, Cell to Sensor, Day Data Only
Sensor_Azimuth_Night	Sensor Azimuth Angle, Cell to Sensor, Night Data Only
Surface_Temperature	Surface Temperature from Ancillary Data
Surface_Pressure	Surface Pressure from Ancillary Data
Cloud_Height_Method	Index Indicating MODIS Bands Used for Cloud Top Pressure Retrieval
Cloud_Top_Height	Geopotential Height at Retrieved Cloud Top Pressure Level (rounded to nearest 50 m)
Cloud_Top_Height_Nadir	Geopotential Height at Retrieved Cloud Top Pressure Level for Sensor Zenith (View) Angles <=32 Degrees (rounded to nearest 50 m)
Cloud_Top_Height_Nadir_Day	Geopotential Height at Retrieved Cloud Top Pressure Level for Sensor Zenith (View) Angles <=32 Degrees, Day Data Only (rounded to nearest 50 m)
Cloud_Top_Height_Nadir_Night	Geopotential Height at Retrieved Cloud Top Pressure Level for Sensor Zenith (View) Angles <=32 Degrees, Night Data Only (rounded to nearest 50 m)
Cloud_Top_Pressure	Cloud Top Pressure Level (rounded to nearest 5 mb)
Cloud_Top_Pressure_Nadir	Cloud Top Pressure Level for Sensor Zenith (View) Angles <= 32 Degrees (rounded to nearest 5 mb)

Cloud_Top_Pressure_Night	Cloud Top Pressure Level, Night Data Only (rounded to nearest 5 mb)
Cloud_Top_Pressure_Nadir_Night	Cloud Top Pressure Level for Sensor Zenith (View) Angles <= 32 Degrees (rounded to nearest 5 mb), Night Data Only
Cloud_Top_Pressure_Day	Cloud Top Pressure Level, Day Only (rounded to nearest 5 mb)
Cloud_Top_Pressure_Nadir_Day	Cloud Top Pressure Level for Sensor Zenith (View) Angles <= 32 Degrees (rounded to nearest 5 mb), Day Data Only
Cloud_Top_Temperature	Temperature from Ancillary Data at Retrieved Cloud Top Pressure Level
Cloud_Top_Temperature_Nadir	Temperature from Ancillary Data at Retrieved Cloud Top Pressure Level for Sensor Zenith (View) Angles <= 32 Degrees
Cloud_Top_Temperature_Night	Temperature from Ancillary Data at Retrieved Cloud Top Pressure Level, Night Only
Cloud_Top_Temperature_Nadir_Night	Temperature from Ancillary Data at Retrieved Cloud Top Pressure Level for Sensor Zenith (View) Angles <= 32 Degrees, Night Data Only
Cloud_Top_Temperature_Day	Temperature from Ancillary Data at Retrieved Cloud Top Pressure Level, Day Only
Cloud_Top_Temperature_Nadir_Day	Temperature from Ancillary Data at Retrieved Cloud Top Pressure Level for Sensor Zenith (View) Angles <= 32 Degrees, Day Data Only
Tropopause_Height	Tropopause Height from Ancillary Data
Cloud_Fraction	Cloud Fraction in Retrieval Region (5x5 1-km Pixels) from 1-km Cloud Mask
Cloud_Fraction_Nadir	Cloud Fraction in Retrieval Region (5x5 1-km Pixels) from 1-km Cloud Mask for Sensor Zenith (View) Angles <= 32 Degrees
Cloud_Fraction_Night	Cloud Fraction in Retrieval Region (5x5 1-km Pixels) from 1-km Cloud Mask, Night Only
Cloud_Fraction_Nadir_Night	Cloud Fraction in Retrieval Region (5x5 1-km Pixels) from 1-km Cloud Mask for Sensor Zenith (View) Angles <= 32 Degrees, Night Data Only
Cloud_Fraction_Day	Cloud Fraction in Retrieval Region (5x5 1-km Pixels) from 1-km Cloud Mask, Day Only
Cloud_Fraction_Nadir_Day	Cloud Fraction in Retrieval Region (5x5 1-km Pixels) from 1-km Cloud Mask for Sensor Zenith (View) Angles <= 32 Degrees, Day Data Only
Cloud_Effective_Emissivity	Cloud Effective Emissivity from Cloud Top Pressure Retrieval
Cloud_Effective_Emissivity_Nadir	Cloud Effective Emissivity from Cloud Top Pressure Retrieval for Sensor Zenith (View) Angles <= 32 Degrees
Cloud_Effective_Emissivity_Night	Cloud Effective Emissivity from Cloud Top Pressure Retrieval, Night Only
Cloud_Effective_Emissivity_Nadir_Night	Cloud Effective Emissivity from Cloud Top Pressure Retrieval for Sensor Zenith (View) Angles <= 32 Degrees, Night Data Only

Cloud_Effective_Emissivity_Day	Cloud Effective Emissivity from Cloud Top Pressure Retrieval, Day Only
Cloud_Effective_Emissivity_Nadir_Day	Cloud Effective Emissivity from Cloud Top Pressure Retrieval for Sensor Zenith (View) Angles <= 32 Degrees, Day Data Only
Cloud_Top_Pressure_Infrared	Cloud Top Pressure from IR Window Retrieval
Radiance_Variance	Band 31 Radiance Standard Deviation
Cloud_Phase_Infrared	Cloud Phase from 8.5 and 11 um Bands
Cloud_Phase_Infrared_Night	Cloud Phase from 8.5 and 11 um Bands, Night Only
Cloud_Phase_Infrared_Day	Cloud Phase from 8.5 and 11 um Bands, Day Only
Cloud_Phase_Infrared_1km	Cloud Phase at 1-km resolution from 8.5- 11 um BTDs and cloud emissivity ratios (12/11, 8.5/11, and 7.2/11 um)
IRP_CTH_Consistency_Flag_1km	Indicates Cloud_Phase_Infrared_1km results changed to ice from water when cloud_top_method_1km reports valid band 36/35 CO2-slicing result (1=change)
os_top_flag_1km	Upper Tropospheric/Lower Stratospheric (UTLS) Cloud Flag at 1-km resolution - valid from -50 to +50 Degrees Latitude
cloud_top_pressure_1km	Cloud Top Pressure at 1-km resolution from LEOCAT, Cloud Top Pressure Level rounded to nearest 5 mb
cloud_top_temperature_1km	Cloud Top Temperature at 1-km resolution from LEOCAT, Temperature from Ancillary Data at Retrieved Cloud Top Pressure Level
cloud_emissivity_1km	Cloud Emissivity at 1-km resolution from LEOCAT Cloud Top Pressure Retrieval
cloud_top_method_1km	Index Indicating the MODIS Band(s) Used to Produce the Cloud Top Pressure Result
surface_temperature_1km	Surface Temperature for Each 1-km MODIS Pixel Interplated from Ancillary Data
cloud_emiss11_1km	11 micron Cloud Emissivity at 1-km resolution from LEOCAT for All Clouds
cloud_emiss12_1km	12 micron Cloud Emissivity at 1-km resolution from LEOCAT for All Clouds
cloud_emiss13_1km	13.3 micron Cloud Emissivity at 1-km resolution from LEOCAT for All Clouds
cloud_emiss85_1km	8.5 micron Cloud Emissivity at 1-km resolution from LEOCAT for All Clouds
Cloud_Effective_Radius	Cloud Particle Effective Radius two-channel retrieval using band 7(2.1um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Effective_Radius_PCL	Cloud Particle Effective Radius two-channel retrieval using band 7(2.1um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Effective_Radius_16	Cloud Particle Effective Radius two-channel retrieval using band 6(1.6um) and either band 1(0.65um), 2(0.86um), or

	5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Effective_Radius_16_PCL	Cloud Particle Effective Radius two-channel retrieval using band 6(1.6um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Effective_Radius_37	Cloud Particle Effective Radius two-channel retrieval using band 20(3.7um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Effective_Radius_37_PCL	Cloud Particle Effective Radius two-channel retrieval using band 20(3.7um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Optical_Thickness	Cloud Optical Thickness two-channel retrieval using band 7(2.1um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Optical_Thickness_PCL	Cloud Optical Thickness two-channel retrieval using band 7(2.1um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Optical_Thickness_16	Cloud Optical Thickness two-channel retrieval using band 6(1.6um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Optical_Thickness_16_PCL	Cloud Optical Thickness two-channel retrieval using band 6(1.6um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Optical_Thickness_37	Cloud Optical Thickness two-channel retrieval using band 20(3.7um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Optical_Thickness_37_PCL	Cloud Optical Thickness two-channel retrieval using band 20(3.7um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges

Cloud_Effective_Radius_1621	Cloud Particle Effective Radius two-channel retrieval using band 7(2.1um) and band 6(1.6um)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Effective_Radius_1621_PCL	Cloud Particle Effective Radius two-channel retrieval using band 7(2.1um) and band 6(1.6um)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Optical_Thickness_1621	Cloud Optical Thickness two-channel retrieval using band 7(2.1um) and band 6(1.6um)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Optical_Thickness_1621_PCL	Cloud Optical Thickness two-channel retrieval using band 7(2.1um) and band 6(1.6um)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Water_Path	Column Water Path two-channel retrieval using band 7(2.1um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Water_Path_PCL	Column Water Path two-channel retrieval using band 7(2.1um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Water_Path_1621	Column Water Path two-channel retrieval using band 7(2.1um) and band 6(1.6um)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Water_Path_1621_PCL	Column Water Path two-channel retrieval using band 7(2.1um) and band 6(1.6um)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Water_Path_16	Column Water Path two-channel retrieval using band 6(1.6um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Water_Path_16_PCL	Column Water Path two-channel retrieval using band 6(1.6um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Water_Path_37	Column Water Path two-channel retrieval using band 20(3.7um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from best points: not failed in any way, not marked for clear sky restoral
Cloud_Water_Path_37_PCL	Column Water Path two-channel retrieval using band 20(3.7um) and either band 1(0.65um), 2(0.86um), or 5(1.2um) (specified in Quality_Assurance_1km)from

	points identified as either partly cloudy from 250m cloud mask test or 1km cloud edges
Cloud_Effective_Radius_Uncertainty	Cloud Effective Particle Radius (from band 7(2.1um)) Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m
Cloud_Effective_Radius_Uncertainty_16	Cloud Effective Particle Radius (from band 6(1.6um) Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m
Cloud_Effective_Radius_Uncertainty_37	Cloud Effective Particle Radius (from band 20(3.7um)) Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m
Cloud_Optical_Thickness_Uncertainty	Cloud Optical Thickness Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m based on the Cloud_Optical_Thickness and Cloud_Effective_Radius results
Cloud_Optical_Thickness_Uncertainty_16	Cloud Optical Thickness Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m based on the Cloud_Optical_Thickness_16 and Cloud_Effective_Radius_16 results
Cloud_Optical_Thickness_Uncertainty_37	Cloud Optical Thickness Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m based on the Cloud_Optical_Thickness_37 and Cloud_Effective_Radius_37 result
Cloud_Water_Path_Uncertainty	Cloud Water Path Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m based on the Cloud_Water_Path result
Cloud_Effective_Radius_Uncertainty_1621	Cloud Effective Particle Radius Relative Uncertainty (Percent) using band 7(2.1um) and band 6(1.6um)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m
Cloud_Optical_Thickness_Uncertainty_1621	Cloud Optical Thickness Relative Uncertainty (Percent) using band 7(2.1um) and band 6(1.6um)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m
Cloud_Water_Path_Uncertainty_1621	Cloud Water Path Relative Uncertainty (Percent) using band 7(2.1um) and band 6(1.6um)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m
Cloud_Water_Path_Uncertainty_16	Cloud Water Path Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m using the VNSWIR-1.6um retrieval

Cloud_Water_Path_Uncertainty_37	Cloud Water Path Relative Uncertainty (Percent)from both best points and points identified as cloud edge at 1km resolution or partly cloudy at 250m using the VNSWIR-3.7um retrieval
Above_Cloud_Water_Vapor_094	Above-cloud water vapor amount from 0.94um channel, ocean only, tau > 5.
IRW_Low_Cloud_Temperature_From_COP	Low Cloud Temperature from IR Window retrieval using cloud emissivity based on cloud optical thickness
Cloud_Phase_Optical_Properties	Cloud Phase Determination Used in Optical Thickness/Effective Radius Retrieval
Cloud_Multi_Layer_Flag	Cloud Multi Layer Identification From MODIS Shortwave Observations
Cirrus_Reflectance	Cirrus Reflectance
Cirrus_Reflectance_Flag	Cirrus_Reflectance_Flag
Cloud_Mask_5km	First Byte of MODIS Cloud Mask Plus Additional Stats for L3 (2nd Byte) **NOTE: both bytes are merged to one short value, first byte as higher byte of the short**
Quality_Assurance_5km_03	Quality Assurance at 5x5 Resolution - Number of cloudy pixels
Quality_Assurance_5km_04	Quality Assurance at 5x5 Resolution - Number of clear pixels
Quality_Assurance_5km_05	Quality Assurance at 5x5 Resolution - Number of missing pixels
Quality_Assurance_5km_09	Quality Assurance at 5x5 Resolution - Cloud height Category

3.4.16 SLSTR

SLSTR L1B data in SAFE format. This reader reads only stripe A data, both nadir and oblique view. The overall resolution is 1km. The high-resolution variables are resampled to the 1km Raster:

- Data variables: averaging of all non-fill-value acquisitions
- Flag variables: or-aggregation of all non-fill-value pixels

The reader supports directory based and zip compressed product types. Due to performance isues, the SLSTR data is completely cached, so it is advised to run SLSTR matches with high memory (> 8GB) and to reduce the reader cache size, see 4.1.1.

Sensor Key	Source
slstr-s3a	SLSTR on Sentinel 3A

Variable Name	Description

latitude_tx	Latitude of detector FOV centre on the earth's surface
longitude_tx	Longitude of detector FOV centre on the earth's surface
sat_azimuth_tn	Satellite azimuth angle (nadir view)
sat_zenith_tn	Satellite zenith angle (nadir view)
solar_azimuth_tn	Solar azimuth angle (nadir view)
solar_zenith_tn	Solar zenith angle (nadir view)
sat_azimuth_to	Satellite azimuth angle (oblique view)
sat_zenith_to	Satellite zenith angle (oblique view)
solar_azimuth_to	Solar azimuth angle (oblique view)
solar_zenith_to	Solar zenith angle (oblique view)
S(1…6)_radiance_an	TOA radiance for channel S(1…6) (A stripe grid, nadir view) (500 m – resampled to 1km)
S(1…6)_exception_an	toa_radiance_status_flag (500 m – resampled to 1km)
S(7…9)_BT_in	Gridded pixel brightness temperature for channel S(7…9) (1km TIR grid, nadir view) (1 km)
S(7…9)_exception_in	toa_brightness_temperature_status_flag (1 km)
confidence_in	Status flag data (1km)
pointing_in	Status flag data (1km)
bayes_in	Status flag data (1km)
cloud_in	Status flag data (1km)
S(1…7)_radiance_ao	TOA radiance for channel S(1…7) (A stripe grid, oblique view) (500 m – resampled to 1km)
S(1…7)_exception_ao	toa_radiance_status_flag (500 m – resampled to 1km)
S(7…9)_BT_io	Gridded pixel brightness temperature for channel S(7…9) (1km TIR grid, oblique view) (1 km)
S(7…9)_exception_io	toa_brightness_temperature_status_flag (1 km)
bayes_io	Status flag data (1km)
cloud_io	Status flag data (1km)

3.4.17 SSMT2

SSMT2 L1C data in NetCDF format.

Sensor Key	Source
ssmt2-f11	SSMT/2 on DMSP F11
ssmt2-f12	SSMT/2 on DMSP F12
ssmt2-f14	SSMT/2 on DMSP F14
ssmt2-f15	SSMT/2 on DMSP F15

Variable Name	Description
ancil_data_Year_1	Scan information – year

ancil_data_DayofYear_1	Scan information – day of year
ancil_data_SecondsofDay_1	Scan information – seconds of day
ancil_data_SatLat	Satellite nadir latitude
ancil_data_SatLong	Satellite nadir longitude
ancil_data_SatAlt	Satellite altitude
ancil_data_SatHeading	Satellite heading
ancil_data_Year_2	Scan information – year
ancil_data_DayofYear_2	Scan information – day of year
ancil_data_SecondsofDay_2	Scan information – seconds of day
tb_ch1 … tb_ch5	Brightness temperatures in channels 1 to 5
lon	Pixel longitude
lat	Pixel latitude
channel_quality_flag_ch1 … ch5	Channel quality flags for channels 1 to 5
gain_control_ch1 … ch5	Gain control for channels 1 to 5
counts_to_tb_gain_ch1 … ch5	Conversion scale factors for channel 1 to 5
counts_to_tb_offset_ch1 … ch5	Conversion offset values for channels 1 to 5
Temperature_misc_housekeeping_thermistorcount01 … 18	Thermistor status information for thermistors 1 to 18
warm_counts_ch1_cal1 … ch5_cal4	Warm calibration counts for channels 1 to 5 and views 1 to 4
cold_counts_ch1_cal1 … ch5_cal4	Cold calibration counts for channels 1 to 5 and views 1 to 4
Satellite_zenith_angle	Artificial variable containing the satellite zenith angle derived from the scan position and the platform orbit altitude

3.5 Supported Insitu Data

This chapter lists the input in-situ data source types available for matchup processing. In subsequent chapters, the sensors, the data format and the available variables are described in more detail.

3.5.1 GRUAN reference files

Specific ASCII data format developed by University of Leicester that maps geolocation and time to an original GRUAN measurement file. Must be used with the prost-processing that adds the full reference file name, see 4.1.5.1.13.

Sensor Key	Source
gruan-uleic	GRUAN reference files

Variable Name	Description
time	Acquisition time in seconds since 1970-01-01
lat	Latitude of measurement position
lon	Longitude of measurement position
source_file_path	GRUAN original measurement file reference path

3.5.2 OceanRain

In-situ data collected at the University of Hamburg within "the Ocean Rain And Ice-phase precipitation measurement Network". (http://oceanrain.cen.uni-hamburg.de)

ASCII dataset, 64 characters per line, one measurement per minute.

Sensor Key	Source
ocean-rain-sst	The OceanRAIN SST dataset

Variable Name	Description
time	Acquisition time in seconds since 1970-01-01
lat	Latitude of measurement position
lon	Longitude of measurement position
sst	SST in degrees Celsius

3.5.3 SST-CCI

In-situ data collected for SST-CCI, sea surface temperature. Formatted to NetCDF following the project specifications (RD 4). Supported is data of processing versions v03.3 and v04.0.

Sensor Key	Source
drifter-sst	Drifter based SST data

ship-sst	Ship based SST data
gtmba-sst	Global Tropical Moored Buoy Array SST data
radiometer-sst	Radiometer based SST data
argo-sst	Argo floater based SST data
xbt-sst	Expendable Bathythermograph SST data
mbt-sst	MBT SST data
ctd-sst	CTD SST data
animal-sst	Animal based SST measurements
bottle-sst	Bottle SST data

Variable Name (v03.3)	**Description**
insitu.time	Acquisition time in seconds since 1978-01-01
insitu.lat	Latitude of measurement position
insitu.lon	Longitude of measurement position
insitu.sea_surface_temperature	SST in degrees Celsius
insitu.sst_uncertainty	Measurement uncertainty in degrees Celsius
insitu.sst_depth	Measurement depth in meters
insitu.sst_qc_flag	In situ sea surface temperature MOHC QC flag
insitu.sst_track_flag	In situ sea surface temperature MOHC track flag
insitu.mohc_id	Unique ID from MOHC database
insitu.id	Unique ID generated from acquisition year, month and mohc_id

Variable Name (v04.0)	**Description**
insitu.collection	HadIOD observation identifier
insitu.lat	Latitude of measurement position
insitu.lon	Longitude of measurement position
insitu.mohc_id	Unique ID from MOHC database
insitu.prof_id	HadIOD profile identifier
insitu.qc1	Observation quality flag 1
insitu.qc2	Observation quality flag 2
Insitu.sea_surface_temperature	SST in Celsius
insitu.sst_depth	Measurement depth in meters
insitu.sst_depth_corr	Measurement depth correction
insitu.sst_plat_corr	Measurement platform correction
insitu.sst_plat_corr_unc	Uncertainty of measurement platform correction
insitu.sst_random_uncertainty	Measurement random uncertainty in degrees Celsius
insitu.sst_type_corr	Measurement platform type correction
insitu.sst_type_corr_unc	Uncertainty of measurement platform type correction
insitu.sst_uncertainty	Measurement uncertainty in degrees Celsius
insitu.subcol1	HadIOD data sub-collection code 1
insitu.subcol2	HadIOD data sub-collection code 2
insitu.time	Acquisition time in seconds since 1978-01-01 00:00:00
insitu.id	Unique ID generated from acquisition year, month and mohc_id

4 Configuration

The Fiduceo MMS is a highly configurable software system that allows customisation wherever applicable.

4.1 Configuration Files

This chapter lists the available parameters in various configuration files that are used by the software. All configuration files should be located in a single configuration directory.

4.1.1 System Configuration

The system configuration file “system-config.xml” contains global settings that affect all components of the MMS. An example XML configuration is displayed below; each tag is explained later in the document.

```
<system-config>
  <geometry-library name="S2"/>
  <reader-cache-size>64</reader-cache-size>
  <temp-directory>/tmp</temp-directory>

  <archive>
    <root-path>
      /usr/local/data/fiduceo
    </root-path>

    <rule sensors="drifter-sst, ship-sst, gtmba-sst, …">
      insitu/SENSOR/VERSION"
    </rule>
  </archive>
</system-config>
```

geometry-library: defines the geometry library to be used for geometric operations. Available values are

- “S2” – use the Google S2 spherical library – the default and recommended value
- “JTS” – use the Java Topology Suite library – experimental

reader-cache-size: defines the number of product-readers which are cached with open files. It depends on the size of the input sensor products how many readers can be cached. Default value is 8.

temp-directory: defines the system directory for temporary file storage. When omitted to set the software uses the standard directory as defined by the system property “java.io.tmpdir”. All temporary files created by the system will be deleted on process exit, if files cannot be cleaned an associated entry in the log-file states this.

archive: defines all settings related to the satellite data archive.

root-path: defines the root directory containing the input datasets used for matchup processing by default, the archive is organised as

<root>/<sensor-platform>/<version>/<year>/<month>(/<day>)

rule: this tag allows to define specific archiving rules that may differ from the default organisation. Any number of rules may be added to the configuration. A rule can apply to one or many sensors, defined by the "*sensors*" attribute of the rule tag. Multiple sensors are written as comma-separated list of sensor - keys. All sensor-keys defined in chapter 3.4 are valid entries for this tag.

A rule defines a general path into the archive. The elements of the path must be separated by a forward slash "/". Path elements can be either selected from a list of pre-defined entities or defined as custom elements. The pre-defined entities will be expanded dynamically during runtime; the custom elements are treated as constant path elements. Pre-defined path elements are:

- SENSOR – will be expanded to the sensor key being processed
- VERSION – will be expanded to the data version being processed
- YEAR – will be expanded to the currently processed year
- MONTH - will be expanded to the currently processed month
- DAY - will be expanded to the currently processed day
- DAY_OF_YEAR - will be expanded to the currently processed day of year

4.1.2 Database Configuration

The database configuration file "database.properties" defines the connection properties for the MMS database. It is a java properties file containing unstructured key/value pairs.

The following settings are available:

driverClassName: defines the database connection driver class name. The following values are available:

- "mongodb": use the MongoDb database driver class – the default and recommended value
- "org.postgres.Driver": use the PostGIS JDBC database driver
- "org.h2.driver": use the H2 in memory database driver
- "com.mysql.jdbc.driver": use the MySQL JDBC database driver – experimental, do not use

url: defines the database connection URL. The format depends on the driverClassName selected, please consult the documentation of the specific database vendor for details.

username: defines the database user name that is submitted during the database connection process.

password: defines the password that is submitted during the database connection process.

4.1.3 MMD Writer Configuration

The MMD output format and content can be configured using this file. In addition, it allows fine-tuning of the writer performance and behaviour. The configuration file is an XML file named "mmd-writer-config.xml" and located in the standard configuration directory.

An example configuration is displayed below; each tag is explained later in the document.

```
<?xml version="1.0" encoding="UTF-8"?>
<mmd-writer-config>
  <overwrite>false</overwrite>
  <cache-size>2048</cache-size>
  <reader-cache-size>8</reader-cache-size>
  <netcdf-format>N4</netcdf-format>
  <variables-configuration>
    …
  </variables-configuration>
</mmd-writer-config>
```

overwrite: this tag defines the MMD writer behaviour when encountering an already existing file with the same name in the target directory.

- false: software raises an exception and stops
- true: software deletes the existing file and writes the new one.

cache-size: defines the cache size in number of matchups that is used by the mmd-writer. Default value is 2048. Reduce this number when experiencing OutOfMemory errors on large subset window sizes.

reader-cache-size: defines the reader cache size in number of open file readers. The reader cache keeps frequently used input files open to increase the MMD writing performance. This setting defines the number of input files that are kept simultaneously open.

netcdf-format: defines the NetCDF file format version used for the MMD files. Available values are:

- "N3" - write MMDs in NetCDF version 3 format
- "N4" - write MMDs in NetCDF version 4 format – the default value

variables-configuration: allows to rename or exclude variables or attributes. A detailed description is given in the following chapter.

4.1.3.1 Variables Configuration

The variables configuration section of the MMD writer configuration allows fine-tuning the content of the final matchup data files. It is possible to

- rename a sensor
- define a separator per sensor namespace for name concatenation

- exclude variables
- rename variables
- rename attributes of variables

For a detailed description of the content and structure of an MMD file and the naming convention used, please refer to chapter 0.

4.1.3.1.1 Renaming a sensor

In the standard configuration the sensors reference are named as they appear in the descriptions in chapters 3.4 and 3.5. This behaviour can be changed using the sensor-rename tag. Multiple sensor-rename tags are allowed inside a variables configuration.

Example:

```
<mmd-writer-config>
  <variables-configuration>
    <sensor-rename source-name="amsub-n16" target-name="AMSU-B_NOAA16" />
  </variables-configuration>
</mmd-writer-config>
```

4.1.3.1.2 Define a separator per sensor namespace

The MMS stores input variables to the MMD using the pattern <sensor-name><separator><variable-name>, for example “ssmt2-f14_tb-ch4”. This configuration aspect allows to change the separator character, which is set to the underscore character ‘_’ by default. Separators can be associated to one or many sensors, which have to be added as a comma-separated list.

Example:

```
<mmd-writer-config>
  <variables-configuration>
    <separator sensor-names="sen1, sen2, ..." separator = "." />
  </variables-configuration>
</mmd-writer-config>
```

4.1.3.1.3 Exclude variables

It is possible to exclude variables of the input data from being written to the MMD file. The exclude tag allows to exclude any number of variables from a given sensors variables list. The names appearing in the exclude tag must exactly match the variable names of the input products as listed in chapters 3.4 and 3.5.

Example:

```
<mmd-writer-config>
  <variables-configuration>
```

```
    <sensors names="hirs-n17, hirs-n16">
      <exclude source-name="not_needed" />
      <exclude source-name="useless" />
    </sensors>
  </variables-configuration>
</mmd-writer-config>
```

4.1.3.1.4 Rename variables

Using this tag, it is possible to rename variables, i.e. to define a different name for the variable in the MMD as it originally appeared in the input data. This tag only changes the variable name, the sensor and separator parts of the final composite name are not affected, these can be changed separately. The names of the source variables appearing in the rename tag must exactly match the variable names of the input product as listed in chapters 3.4 and 3.5.

Example:

```
<mmd-writer-config>
  <variables-configuration>
    <sensors names="hirs-n17, hirs-n16">
      <rename source-name="stupid" target-name="cool" />
      <rename source-name="boring" target-name="exciting" />
    </sensors>
  </variables-configuration>
</mmd-writer-config>
```

4.1.3.1.5 Rename attributes of variables

It is possible to adapt the names of attributes of any variable in the MMD. The names of the source variables and attribute names appearing in the rename-attribute tag must exactly match the variable and attribute names of the input product as listed in chapters 3.4 and 3.5.

```
<mmd-writer-config>
 <variables-configuration>
  <sensors names="hirs-n17, hirs-n16">
   <rename-attribute source-name="attN1" target-name="attN1_r" />
   <rename-attribute variable-names="var1" source-name="attN2" target-name="attN2_r" />
   <rename-attribute variable-names="var1, var2" source-name="attN3" target-name="attN3_r" />
  </sensors>
 </variables-configuration>
</mmd-writer-config>
```

4.1.4 UseCase Configuration

This configuration file defines the parameter-set to be used for a matchup processing. It is an XML file that has to be located in the configuration directory used for processing.

An example configuration is displayed below; each tag is explained later in the document.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<use-case-config name="mmd02">
  <sensors>
    <sensor>
      <name>avhrr-n17</name>
      <primary>false</primary>
    </sensor>
    <sensor>
      <name>avhrr-n18</name>
      <primary>true</primary>
    </sensor>
  </sensors>

  <dimensions>
    <dimension name="avhrr-n17">
      <nx>5</nx>
      <ny>5</ny>
    </dimension>
    <dimension name="avhrr-n18">
      <nx>5</nx>
      <ny>5</ny>
    </dimension>
  </dimensions>

  <location>
    <lon>-12.7</lon>
    <lat>33.24</lat>
  </location>

  <write-distance>true</write-distance>
  <random-sampling>
    <points-per-day>20000</points-per-day>
    <distribution>true</distribution >
  </random-sampling>
  <output-path>/tmp/fiduceo_test/usecase-02</output-path>

  <conditions>
    <time-delta>
      <time-delta-seconds>10800</time-delta-seconds>
    </time-delta>
    <spherical-distance>
      <max-pixel-distance-km>3.0</max-pixel-distance-km>
    </spherical-distance>
  </conditions>

  <screenings>
    <angular>
      <primaryVZAVariable name="satellite_zenith_angle" />
```

```
      <secondaryVZAVariable name="satellite_zenith_angle" />
      <maxAngleDelta>10.0</maxAngleDelta>
    </angular>
  </screenings>

</use-case-config>
```

Each usecase configuration file starts with an XML version and character encoding statement. The enclosing <use-case-config name="mmd02"> tag defines the content of the configuration. Each usecase configuration has to be given a name that is set in the “name” attribute as shown above. This name is used during MMD file generation to create the result file names.

sensors: this tag encloses a list of sensors to be processed. Each sensor element in this list must contain a valid <name> tag that defines the senor name. The sensor names must match one name of the list of supported sensors, see chapters 3.4 and 3.5. One (and only one) of the sensors must be tagged as being the primary (i.e. reference) sensor in this list of sensors. This is accomplished by adding a primary tag with the value “true”:

```
<primary>true</primary>
```

The primary tag with a value of “false” can be omitted, the software automatically assumes this when the tag is missing.

Optionally, sensor data can be restricted to a single data/processing version. This can be expressed by adding a version tag:

```
<data-version>v2.08</data-version>
```

When the version tag is omitted, the MMS uses all available data versions for this sensor for processing.

dimensions: enclosed by this tag is a list of dimension tags. These define the output window size for each matchup, i.e. the size of the data extract stored to the MMD with the matchup x/y location in the centre. It is required that for each sensor defined in the sensor list, an associated dimension is being defined. The association is realised by matching the names, stored as dimension attribute “name”.

The tags “nx” and “ny” define the extent of each dimension, they define the complete width and height of the extract. Following the requirement of a central matchup pixel, both values have to be odd numbers.

location: optional tag to define a matchup processing that extracts input sensor pixels that contain the location defined by the tags **lon** and **lat**. If set, this setting supersedes the standard matchup processing and allows overflight detection for a fixed ground location. All geodesic matches within the time-interval and sensor(s) configured are considered as valid matchups.

write-distance: specifies whether the MMD file shall contain an additional variable that stores the matchup centre distance in kilometres. Valid values are “true” and “false”, defaults to “false”.

output-path: defines the target directory for the processing. All MMD files generated will be stored in the directory denoted by this tag.

conditions: this tag encloses a list of condition plugin configurations. The values in each configuration are depending on the plugin and are explained in more detail in chapter 4.1.4.1.

screenings: this tag encloses a list of screening plugin configurations. The values in each configuration are depending on the plugin and are explained in more detail in chapter 4.1.4.2.

random-sampling: specifies to use a random sampling point matchup strategy. The random “Seed Point Matchup Strategy” is helpful to avoid that all the pixels of the entire observations are matchups, e.g. if the primary and secondary observation geometries are congruent, because they are on the same platform.

points-per-day: Defines the number of global generated pseudo random seed points (sobol generated points) per day. The points are nearly equidistant distributed on the globe and nearly equidistant distributed in the time range defined by the matchup tool command line arguments. (See 3.2 Matchup Tool)

distribution: defines the random distribution properties. The time and longitude axes are always equally distributed. The latitude axis can be scaled using the distribution tag; valid values are:

- FLAT: equally distributed – which increases the probability of polar matches because it generates an equal-angle sampling.
- COSINE_LAT: the latitude distribution is weighted by cos(latitude). This generates an equal area distribution by good approximation
- INV_TRUNC_COSINE_LAT; the latitude distribution is weighted by $\frac{1}{3} * \min(\frac{1}{\cos(latitude)}, 3)$. This weighting increases the probability of polar matches even more than FLAT.

The old syntax:

<random-points-per-day>20000</random-points-per-day>

is still supported by the software and generates a cosine-weighted distribution.

4.1.4.1 Conditions

This chapter lists the configuration settings for all condition plugins available for the MatchupTool. A condition plugin serves for checking that a matchup conforms to the condition defined. A condition plugin operates solely on the raw matchup information, i.e.

- Primary and secondary pixel
    - Geo-location longitude and latitude
    - Pixel raster location x and y
    - Pixel acquisition time in seconds since epoch
- Primary and secondary product raster width and height
- Primary and secondary extraction window sizes
- The processing time-range

When one or more of the subsequent plugin configurations are present in the configuration, the corresponding plugin is loaded and configured using the values supplied. During processing, the chain of plugins is applied to the matchups. The order of the condition configurations defines the processing order.

4.1.4.1.1 Border Distance Condition

This condition ensures that the matchup location within the data file raster is within a defined distance from the raster borders. The condition can check primary and multiple secondary acquisition positions using different parameters, according to the configuration. If the pixel is outside the defined boundaries, the matchup is rejected.

Example with one secondary sensor:

```
<border-distance>
  <primary>
    <nx>22</nx>
    <ny>6</ny>
  </primary>

  <secondary>
    <nx>5</nx>
    <ny>8</ny>
  </secondary>
</border-distance>
```

“primary”: if tag is present, the primary (i.e. reference) sensor position is checked

“secondary”: if tag is present, the secondary sensor position is checked

“nx” defines the minimal distance of the matchup centre location x coordinate to the left or right data raster boundary.

“ny” defines the minimal distance of the matchup centre location y coordinate to the upper or lower data raster boundary.

Example with three secondary sensors:

```
<border-distance>
  <primary>
    <nx>22</nx>
    <ny>6</ny>
  </primary>

  <secondary names=”sensor-A”>
    <nx>5</nx>
    <ny>8</ny>
  </secondary>

  <secondary names=”sensor-B, sensor-C”>
```

```
    <nx>3</nx>
    <ny>3</ny>
  </secondary>
</border-distance>
```

A “secondary” tag can be used twice or more times. Each of the “secondary” tags must have assigned a “names” attribute containing one or more secondary sensor names. Border distance calculations will be applied to all of these sensors.

It is not allowed to use a secondary sensor name twice.

It is not allowed to use more than one “secondary” tag without names.

It is not allowed to mix “secondary” tag without name and “secondary” tag with name.

4.1.4.1.2 Spherical Distance Condition

This condition ensures that the primary and secondary matchup location are within a given spherical distance. If the matchup pixel centre locations (in lon/lat) are further apart than the threshold value, the matchup is rejected. Example:

```
<spherical-distance>
  <max-pixel-distance-km>6</max-pixel-distance-km>
</spherical-distance>
```

4.1.4.1.3 Time Delta Condition

This condition verifies that the pixel acquisition time difference of the contributing primary and secondary (named or not named) products do not exceed the configured threshold. All pixels with an acquisition time difference above the threshold are rejected.

Example Case: <u>primary and one secondary sensor</u>

```
<time-delta>
  <time-delta-seconds>10800</time-delta-seconds>
</time-delta>
```

Example Case: <u>primary and two or more secondary sensors</u>

```
<time-delta>
  <time-delta-seconds names="name1, name2, ..."
          secondaryCheck="true">  <!-- default is false -->
    300
  </time-delta-seconds>
</time-delta>
```

“names” attribute contain comma separated secondary sensor names to which this condition applies
“secondaryCheck” attribute

- If “true” the check also applies between the named secondaries
- If “false” (default) the check apllies NOT between secondaries

Example Case: extra time delta for each sensor combination

```
<time-delta>

  <!-- time delta between primary and name1 sensor -->
  <time-delta-seconds names="name1">
    300
  </time-delta-seconds>

  <!-- time delta between primary and name2 sensor -->
  <time-delta-seconds names="name2">
    250
  </time-delta-seconds>

  <!-- time delta between named secondary sensors -->
  <time-delta-seconds names="name3,name4,..."
              primaryCheck="false"    <!-- default is true -->
              secondaryCheck="true">  <!-- default is false -->
    400
  </time-delta-seconds>
  ...
</time-delta>
```

“primaryCheck” attribute

- If “true” (default) the check applies between primary and seconday
- If “false” NOT primary check will be applied

4.1.4.1.4 Overlap Remove Condition

This condition plugin removes overlapping extraction windows. The plugin can operate either on the primary or on the secondary sensor extracts. It checks all matchup pixels for overlapping extraction windows and rejects all appropriate, so that the remaining list of matchups does not have overlapping areas. Example:

```
<overlap-remove>
  <reference>PRIMARY</reference>
</overlap-remove>
```

“reference”: denotes the sensor that shall be used for overlap removal. Valid choices are “PRIMARY” and “SECONDARY”.

If more than one secondary sensors defined in use case configuration, a “names” attribute must be applied to the “SECONDARY” reference tag to define to which secondary sensor the overlap remove should be applied.

```
<overlap-remove>
  <reference names=”name-a, name-b, ...”>
    SECONDARY
  </reference>
</overlap-remove>
```

4.1.4.1.5 UniqueSamples Condition

This condition plugin ensures that duplicate matches are removed from the MMD. Duplicate matches can e.g. occur when a single satellite measurement is associated to two or more in-situ measurements because the acquisition interval of the in-situ instrument is shorter than the matchup time criterion. The sample with the smallest geodetic distance to the reference is kept, all other matches are rejected.

Example:

```
<unique-samples>
  <reference-sensor>sensor_a</reference-sensor>
  <associated-sensor>sensor_b</associated-sensor>
</unique-samples>
```

The reference sensor is the sensor whose matchups shall be kept, the associated sensor is the one to be checked for multiple references. Both tags must contain a valid sensor key, as listed in section 3.4 and 3.5.

4.1.4.2 Screenings

This chapter lists the configuration settings for all screening plugins available for the MatchupTool. A screening plugin serves for checking that a matchup conforms to the condition defined. A screening plugin operates on the raw matchup information (see chapter 4.1.4.1) and additionally has access to the product data readers.

When one or more of the subsequent plugin configurations are present in the configuration, the corresponding plugin is loaded and configured using the values supplied. During processing, the chain of plugins is applied to the matchups. The order of the screening configurations defines the processing order.

4.1.4.2.1 Angular Screening

This plugin performs a number of screenings on the satellite or viewing zenith or azimuth angles. The names of the variables can be configured, as well as the screenings that shall be applied and the associated thresholds. A screening is executed only when the associated tag is present in the configuration. Example:

```
<angular>
  <primary-vza-variable name=”zenith_angle”/>
  <secondary-vza-variable name=”satellite_zenith_angle”/>
  <max-primary-vza>10.0</max-primary-vza>
  <max-secondary-vza>15.0</max-secondary-vza>
  <max-angle-delta>17.0</max-angle-delta>
</angular>
```

“primary-vza-variable” denotes the variable name containing the angle to be screened for the primary sensor.

“secondary-vza-variable” denotes the variable name containing the angle to be screened for the secondary sensor.

“max-primary-vza” denotes the threshold for the primary angular variable in decimal degrees. All matchups with a value higher than the threshold value are rejected. Requires “primary-vza-variable” to be set. If tag is not present, the screening is not applied.

“max-secondary-vza” denotes the threshold for the secondary angular variable in decimal degrees. All matchups with a value higher than the threshold value are rejected. Requires “secondary-vza-variable” to be set. If tag is not present, the screening is not applied.

“max-angle-delta” denotes the maximal angular difference absolute value in decimal degrees. Matchups with an angular difference higher than the threshold are rejected. Requires “primary-vza-variable” and “secondary-vza-variable” to be set. If tag is not present, the screening is not applied.

4.1.4.2.2 Angular Cosine Proportion Screening

This plugin performs a screening on satellite viewing zenith angles as defined by the University of Hamburg (see RD 1), following the equation:

$$\left|\frac{\cos(primaryVZA)}{\cos(secondaryVZA)} - 1\right| < \varepsilon_1$$

All matchups with a proportion value exceeding the threshold are rejected. Example configuration:

```
<angular-cosine-proportion>
   <primary-variable name=”zenith_angle”/>
   <secondary-variable name=”satellite_zenith_angle”/>
   <threshold>0.02</threshold>
</angular-cosine-proportion>
```

“primary-variable” denotes the variable name containing the angle to be screened for the primary sensor.

“secondary-variable” denotes the variable name containing the angle to be screened for the primary sensor.

“threshold” denotes the threshold value (i.e. the epsilon$_1$ in the equation above).

4.1.4.2.3 HIRS LZA Angular Screening

The LZA variable in the HIRS L1C data describes the local zenith angle for each scanline x-position. It does not distinguish left and right nadir positions. To implement a correct angular difference screening, the scanposition dataset has to be taken into account.

This plugin implements a correct angular difference screening for HIRS data following the algorithm:

$sigma_{primary} = scanpos_{primary} < 28 ? -1.0 : 1.0$

$sigma_{secondary} = scanpos_{secondary} < 28 ? -1.0 : 1.0$

$angDelta = abs(sigma_{primary} * lza_{primary} – sigma_{secondary} * lzas_{econdary})$

All matchups with an angDelta above the “max-lza-delta” threshold are rejected. Values are in decimal degrees. Example configuration:

```
<hirs-lza-delta>
   <max-lza-delta>10.0</max-lza-delta>
</hirs-lza-delta>
```

4.1.4.2.4 Buehler Cloud Screening

This plugin performs a cloud screening for microwave sensors based on the algorithm described in RD 2 and RD 3. It is applicable to AMSU-B, MHS and SSM/T-2. It can be applied either to primary, secondary or both sensor data. The screening is only executed when all three configuration values for the sensors are set. To switch the screening for a sensor off, remove the associated tags from the configuration. All matchups detected as cloudy are rejected.

Example configuration:

```
<buehler-cloud>
   <primary-narrow-channel name=”btemp_ch18”/>
   <primary-wide-channel name=”btemp_ch20”/>
   <primary-vza name=”Satellite_zenith_angle”/>

   <secondary-narrow-channel name=”btemp_ch3”/>
   <secondary-wide-channel name=”btemp_ch4”/>
   <secondary-vza name=”Satellite_zenith_angle”/>
</buehler-cloud>
```

“primary-narrow-channel” denotes the variable name containing the narrow bandwidth channel for the primary sensor.

“primary-wide-channel” denotes the variable name containing the wide bandwidth channel for the primary sensor.

“primary-vza” denotes the variable name containing the satellite zenith angle values for the primary sensor.

"secondary-narrow-channel" denotes the variable name containing the narrow bandwidth channel for the secondary sensor.

"secondary-wide-channel" denotes the variable name containing the wide bandwidth channel for the secondary sensor.

"secondary-vza" denotes the variable name containing the satellite zenith angle values for the secondary sensor.

4.1.4.2.5 Pixel Value Screening

This screening plugin allows selecting matchups using a mathematical expression consisting of any combination of input variables. The plugin allows entering different expressions for primary and secondary sensor. Mathematical expressions can be arbitrarily complex but must evaluate to a Boolean value. The mathematical expressions are evaluated using the matchup location only. All matchups where either the primary or the secondary expression evaluate to false are rejected.

Mathematical expressions can be composed of variables and functions.

Variables available are all variables contained in the input satellite or in-situ data using the name as it appears in the file and as documented in section chapters 3.4 and 3.5, e.g. "btemps_ch5" for AMSU-B brightness temperature in channel 5. In addition to this, the constants "PI" and "E" are available denoting pi and e. And of course any number typed into the expression is treated as numerical constant.

Mathematical operations and functions available are:

Function	Description
+	Addition
-	Subtraction
*	Multiplication
/	Division
==	Equality comparison
!=	Inequality comparison
>	Greater than
>=	Greater than or equal
<	Less than
<=	Less than or equal
abs	Calculates absolute value of argument
acos	Arc cosine of argument
ampl	Calculates amplitude of arguments ($a^2 + b^2$)
asin	Arc sine of argument
atan	Arc tangent of argument
atan2	The theta component of the point (r, theta) in polar coordinates that corresponds to the point (x, y) in Cartesian coordinates. (https://docs.oracle.com/javase/7/docs/api/java/lang/Math.html#atan2(double,%20double))
avg	Mean value of arguments

bit_set	Checks whether a bit is set or not, first argument is the value to check, second argument is the bit index (e.g. bit_set(cloud_flags_fward , 1))
ceil	Round upwards to next integer value
cos	Cosine of argument
cosech	Hyperbolic cosecant of argument
cosh	Hyperbolic cosine of argument
deg	Converts argument from radians to decimal degrees
distance	Calculates the distance (sqrt of sum of squares)
exp	Exponential of argument
exp10	Ten to the power of argument
feq	Fuzzy equality (1e-6 accuracy) or when used with three arguments, third argument supplies maximal delta.
floor	Round downwards to next integer value
fneq	Fuzzy inequality (1e-6 accuracy) or when used with three arguments, third argument supplies maximal delta.
inf	Check argument for infinity. Returns 1 if argument is positive or negative infinite, returns 0 otherwise
log	Natural logarithm of argument
log10	Logarithm to the base of 10 of argument
max	Maximum value of arguments
min	Minimum value of arguments
nan	Checks argument for “not a number”. Returns 1 if argument is not a number, 0 otherwise
phase	Calculates phase of arguments (atan2(b,a))
pow	General power function a^b
rad	Converts argument in decimal degrees to radians
random_gaussian	Generates a random number from a gaussian distributed process
random_uniform	Generates a random number from a uniformly distributed process
round	Round to closes integer value
sech	Hyperbolic secant of argument
sign	Compute signum, returns +1 for positive, -1 for negative and 0 for zero argument
sin	Sine of argument
sinh	Hyperbolic sine of argument
sq	Square of argument
sqrt	Square root of argument
tan	Tangent of argument
tanh	Hyperbolic tangent of argument

All matchups where the mathematical expression for either primary or secondary sensor evaluates to false are rejected. Example configuration:

```
<pixel-value>
  <primary-expression>btemp_ch4 >= 180.5</primary-expression>
  <secondary-expression>cloud_probability < 0.25</secondary-expression>
</pixel-value>
```

4.1.4.2.6 Window Value Screening

This screening plugin allows selecting matchups using a mathematical expression consisting of any combination of input variables. The plugin allows entering different expressions for primary and secondary sensor. Mathematical expressions can be arbitrarily complex but must evaluate to a Boolean value. The mathematical expressions are evaluated for each pixel of a window defined by the matchup location and an interval defined by the use case configuration.

Additional to the mathematical expression two more properties must be defined. The percentage and the evaluate property. The evaluate property decide whether pixels with fill value should be ignored or all pixels of the window are used to calculate the percentage pixel count.

All matchups where either the primary or the secondary percentage evaluate to less pixels are valid are rejected.

Mathematical expressions can be composed of variables and functions.

Variables available are all variables contained in the input satellite or in-situ data using the name as it appears in the file and as documented in section chapters 3.4 and 3.5, e.g. “btemps_ch5” for AMSU-B brightness temperature in channel 5. Satellite variables are always scaled to geophysical entities before evaluated, if the original data is stored as scaled integer data.

In addition to this, the constants “PI” and “E” are available denoting pi and e. And of course any number typed into the expression is treated as numerical constant.

Mathematical operations and functions available are the same as in 4.1.4.2.5.

Example case 1: Primary and only one secondary sensor defined in use case configuration

```
<window-value>
  <primary>
    <expression>btemp_ch4 >= 180.5</expression>
    <percentage>22.3</percentage>
    <evaluate>EntireWindow</evaluate>  // second option is “IgnoreNoData”
  </primary>
  <secondary>
    <expression> cloud_probability < 0.25</expression>
    <percentage>50.0</percentage>
    <evaluate>IgnoreNoData</evaluate>
  </secondary >
</window-value>
```

Example case 2: Screening for one primary and multiple secondaries

```
<window-value>
  <primary>
    <expression>btemp_ch4 >= 180.5</expression>
    <percentage>22.3</percentage>
    <evaluate>EntireWindow</evaluate>  // second option is “IgnoreNoData”
  </primary>
```

```
  <secondary names=”name-A”>
    <expression> cloud_probability < 0.25</expression>
    <percentage>50.0</percentage>
    <evaluate>IgnoreNoData</evaluate>
  </secondary >
  <secondary names=”name-B, name-C”>
    <expression>CH_wavelength233 < 0.33</expression>
    <percentage>75.0</percentage>
  </secondary >
</window-value>
```

4.1.5 Post Processing Configuration

The post processing configuration file stores the parameters and configuration used by the post-processing tool. It is an XML file that has to be located in the general configuration files directory, see chapter 4.1. It contains the main tag *<post-processing-config>* and a list of *<post-processings>* which contain the configurations of the plugins. A list of implemented post processing plugins is described in 4.1.5.1.

Also the configuration file contains information whether the input mmd files should be overwritten in place

```
<overwrite>
```

or new extended files should be written to the output directory

```
<create-new-files>
  <output-directory><!—the existing output directory path --></output-directory>
</create-new-files>
```

<u>Example Post Processing Configuration File:</u>

```
<post-processing-config>
  <create-new-files>
    <output-directory><!—the existing output directory path --></output-directory>
  </create-new-files>
  <post-processings>
    ... <!-- List of plugin configurations -->
  </post-processings>
</post-processing-config>
```

Plugins can be chained and are processed in the order they appear in the configuration file.

4.1.5.1 Post Processing Plugins

Each post processing plugin has its own XML signature.

4.1.5.1.1 Spherical Distance Plugin

The spherical distance plugin is a post processing plugin, which can be used to add a spherical distance variable to a matchup dataset i.e. the distance of the matchup acquisitions for primary and additional sensor, calculated as geodesic distance in kilometers.

Requirements:

- The matchup dataset must contain latitude and longitude information for primary and secondary sensor.
- This geo information can be organized in 1 by 1, 3 by 3, 5 by 5 or each other odd manner.
- The matchup dataset must not contain a variable with the name, defined in `<var-name>` tag.

```
<spherical-distance>
  <target>
    <data-type>Float</data-type>
    <var-name>post_dist</var-name>
    <dim-name>matchup_count</dim-name>
  </target>

  <primary-lat-variable scaleAttrName="Scale">amsub-n16_Latitude</primary-lat-variable>
  <primary-lon-variable scaleAttrName="Scale">amsub-n16_Longitude</primary-lon-variable>
  <secondary-lat-variable>ssmt2-f14_lat</secondary-lat-variable>
  <secondary-lon-variable>ssmt2-f14_lon</secondary-lon-variable>
</spherical-distance>
```

If the geo information variable contains values which must be scaled to get the real lat/lon information, in the plugin XML the names of scaling (*scaleAttrName*) and offset (*offsetAttrName*) attributes have to be supplied. The plugin scales the variables accordingly before calculating the geodesic distance. If scaling or offset attributes are defined but not available, a runtime exception will be thrown.

The primary and secondary lat/lon value will always be extracted from the center pixel.

4.1.5.1.2 SST Insitu Timeseries Plugin

This plugin extracts a time series from in the situ data files and appends the data to the mmd file.

Requirements:

- An input mmd file must contain variables containing “_insitu.” in its name. The part before “_insitu.” is interpreted as the insitu sensor type.
  e.g. “gtmba-sst_insitu.id” resolves to insitu sensor type “gtmba-sst”
- An input mmd file must contain a variable named
  “<insitu-sensor-type>_file_name”
  which contains the original insitu data file name per matchup
- An input mmd file must contain a dimension named “file_name” which defines the field size of a file_name field in the “<insitu-sensor-type>_file_name variable.

- An input mmd file must contain a dimension named “matchup_count” which defines the number of matchups in this file.

Example sst time series post processing plugin xml part:

```
<sst-insitu-time-series>
  <version>v03.3</version>
  <time-range-in-seconds>129600</time-range-in-seconds>
  <time-series-size>220</time-series-size>
  <secondary-sensor-matchup-time-variable>
    amsre.acquisition_time
  </secondary-sensor-matchup-time-variable>
  <insitu-sensor>ship-sst</insitu-sensor>
  <file-name-variable>ship-sst.file_name</file-name-variable>
  <y-variable>ship-sst.y</y-variable>
</sst-insitu-time-series>
```

- version: defines the processing version of the insitu data files. Also encodes the data archive path.
- time-range-in-seconds: defines the complete time range extract range from the insitu acquisitons
- time-series-size: defines the size of the extract – should match the maximum number of acquisitions possible within the extraction time range
- secondary-sensor-matchup-time-variable: variable name for the reference times
- insitu-sensor: name of the insitu sensor (optional). This configuration parameter can be skipped if the MMD to process uses standard variable notation. If variable names or naming rules have been updated you can set the correct sensor name here.
- file-name-variable: name of the insitu file name variable in the input MMD (optional). This configuration parameter can be skipped if the MMD to process uses standard variable notation. If variable names or naming rules have been updated you can set the correct file name variable namne here.
- y-variable: name of the insitu y-positin variable in the input MMD (optional). This configuration parameter can be skipped if the MMD to process uses standard variable notation. If variable names or naming rules have been updated you can set the correct file y variable name here.

The plugin expects that the insitu file archive is organized without year/month/day paths. This means that all the insitu files of same sensor type and version should be in one directory.

Example archive snippet from system-config.xml

```
<system-config>
  . . .
  <archive>
    <root-path><!-- archive root path --></root-path>
    <rule sensors="animal-sst, gtmba-sst">insitu/SENSOR/VERSION</rule>
  </archive>
  . . .
</system-config>
```

Time series variables created by the plugin:

- insitu.dtime [int] (time delta between secondary matchup time and insitu time)
- insitu.id [long] (this a unique id generated by combining YEAR, MONTH and mohc_id)
- insitu.lat [float]
- insitu.lon [float]
- insitu.mohc_id [int]
- insitu.sea_surface_temperature [float]
- insitu.sst_depth [float]
- insitu.sst_qc_flag [short]
- insitu.sst_track_flag [short]
- insitu.sst_uncertainty [float]
- insitu.time [int]
- insitu.y [int]

4.1.5.1.3 NWP plugin

This plugin adds numerical weather prediction (NWP) data to matchup datasets. The NWP data is extracted from ERA-Interim data (http://www.ecmwf.int/en/research/climate-reanalysis/era-interim). The plugin allows two different processes to be executed, a time series extraction for single pixel matchups (like insitu data) or an n x m pixel extract with profile data for atmospheric parameters for satellite-sensor extracts. The plugin relies on the publicly available CDO software (https://code.zmaw.de/projects/cdo) to perform the time-series aggregation and projection tasks.

The parametrization for the time-series and sensor-extracts is configured in specific sub-tags, see example below.

Configuration example:

```
<nwp>
    <cdo-home>/usr/local/bin/cdo</cdo-home>
    <nwp-aux-dir>/the/auxiliary/files</nwp-aux-dir>
    <delete-on-exit>true</delete-on-exit>

    <time-series-extraction>
      <analysis-steps>19</analysis-steps>
      <forecast-steps>33</forecast-steps>

      <time-variable-name>acquisition-time</time-variable-name>
      <longitude-variable-name>animal-sst_inistu.lon</longitude-variable-name>
      <latitude-variable-name>animal-sst_inistu.lat</latitude-variable-name>

      <analysis-center-time-name>matchup.nwp.an.t0</analysis-center-time-name>
      <forecast-center-time-name>matchup.nwp.fc.t0</forecast-center-time-name>

      <an-ci-name>an_sea-ice-fraction</an-ci-name>
      <an-sstk-name>an_sea-surface-temperature</an-sstk-name>
      <fc-sstk-name>fc_sea-surface-temperature</fc-sstk-name>
      <an-u10-name>an_east_wind</an-u10-name>
      <an-v10-name>an_north_wind</an-v10-name>
```

```
    <fc-u10-name>10m_east_wind_component</fc-u10-name>
    <fc-v10-name>10m_north_wind_component</fc-v10-name>
    <fc-msl-name>mean_surface_pressure</fc-msl-name>
    <fc-t2-name>2m-temperature</fc-t2-name>
    <fc-d2-name>2m_dew_point</fc-d2-name>
    <fc-tp-name>total_precipitation</fc-tp-name>
    <an-clwc-name>cloud_liquid_water_content</an-clwc-name>
    <fc-clwc-content-name>cloud_liquid_water_content</fc-clwc-name>
    <an-tcwv-name>an_total_column_water_vapour</an-tcwv-name>
    <fc-tcwv-name>fc_total_column_water_vapour</fc-tcwv-name>
    <fc-sshf-name>fc_surface_sensible_heat_flux</fc-sshf-name>
    <fc-slhf-name>surface_latent_heat_flux</fc-slhf-name>
    <fc-blh-name>boundary_layer_height</fc-blh-name>
    <fc-ssrd-name>fc_downward_surface_solar_radiation</fc-ssrd-name>
    <fc-strd-name>fc_downward_surface_thermal_radiation</fc-strd-name>
    <fc-ssr-name>fc_surface_solar_radiation</fc-ssr-name>
    <fc-str-name>fc_surface_thermal_radiation</fc-str-name>
    <fc-ewss-name>fc_turbulent_stress_east_component</fc-ewss-name>
    <fc-nsss-name>fc_turbulent_stress_north_component</fc-nsss-name>
    <fc-e-name>fc_evaporation</fc-e-name>
</time-series-extraction>

<sensor-extraction>
    <time-variable-name>acquisition_time</time-variable-name>
    <longitude-variable-name>amsre.longitude</longitude-variable-name>
    <latitude-variable-name>amsre.latitude</latitude-variable-name>

    <x-dimension>5</x-dimension>
    <x-dimension-name>nwp_x</x-dimension-name>
    <y-dimension>5</y-dimension>
    <y-dimension-name>nwp_y</y-dimension-name>
    <z-dimension>60</z-dimension>
    <z-dimension-name>nwp_z</z-dimension-name>

    <an-ci-name>seaice_fraction</an-ci-name>
    <an-asn-name>snow_albedo</an-asn-name>
    <an-sstk-name>sea_surface_temperature</an-sstk-name>
    <an-tcwv-name>total_column_water_vapour</an-tcwv-name>
    <an-msl-name>mean_sea_level_pressure</an-msl-name>
    <an-tcc-name>total_cloud_cover</an-tcc-name>
    <an-u10-name>10m_east_wind_component</an-u10-name>
    <an-v10-name>north_wind</an-v10-name>
    <an-t2-name>2m_temperature</an-t2-name>
    <an-d2-name>2m_dew_point</an-d2-name>
    <an-al-name>albedo</an-al-name>
    <an-lnsp-name>log_surface_pressure</an-lnsp-name>
    <an-skt-name>skin_temperature</an-skt-name>
    <an-t-name>temperature_profile</an-t-name>
    <an-q-name>water_vapour_profile</an-q-name>
```

```
    <an-o3-name>ozone_profile</an-o3-name>
    <an-clwc-name>cloud_liquid_water</an-clwc-name>
    <an-ciwc-name>cloud_ice_water</an-ciwc-name>
    <an-tp-name>total_precipitation</an-tp-name>
  <sensor-extraction>

</nwp>
```

Configuration parameters are explained in detail below:

- “cdo-home”: denotes the absolute path to the root directory of the binary CDO operators installation.
- “nwp-aux-dir”: defines the directory where the ERA-Interim auxiliary files are located.
- “delete-on-exit”: Set whether to delete all temporary files after processing or not. Default value is: true

Configuration parameter for the time-series extraction process:

- “analysis-steps”: Defines the number of time steps around the matchup time for NWP analysis data (6 hr time resolution). Default is: 17.
- “forecast-steps”: Defines the number of time steps around the matchup time for NWP forecast data (3 hr time resolution). Default is: 33.
- “time-variable-name”: Defines the name of the time variable to use as reference time. Time variables are expected to store data in seconds since 1970 format.
- “longitude-variable-name”: Defines the name of the longitude variable to use as reference. Data is expected in decimal degrees east.
- “latitude-variable-name”: Defines the name of the longitude variable to use as reference. Data is expected in decimal degrees north.
- “analysis-center-time-name”: Defines the name of the variable for analysis center times. Values are in seconds since 1970-01-01. Default: matchup.nwp.an.t0.
- “forecast-center-time-name”: Defines the name of the variable for forecast center times. Values are in seconds since 1970-01-01. Default: matchup.nwp.fc.t0.
- “an-ci-name”: Defines the name of the target variable for analysis sea-ice-fraction (CI). Default: matchup.nwp.an.sea_ice_fraction.
- “an-sstk-name”: Defines the name of the target variable for analysis sea surface temperature (SSTK). Default: matchup.nwp.an.sea_surface_temperature.
- “fc-sstk-name”: Defines the name of the target variable for forecast sea surface temperature (SSTK). Default: matchup.nwp.fc.sea_surface_temperature.
- “an-u10-name”: Defines the name of the target variable for analysis 10m east wind component (U10). Default: matchup.nwp.an.10m_east_wind_component.
- “an-v10-name”: Defines the name of the target variable for analysis 10m north wind component (V10). Default: matchup.nwp.an.10m_north_wind_component.
- “fc-u10-name”: Defines the name of the target variable for forecast 10m east wind component (U10). Default: matchup.nwp.fc.10m_east_wind_component.
- “fc-v10-name”: Defines the name of the target variable for forecast 10m north wind component (V10). Default: matchup.nwp.fc.10m_north_wind_component.
- “fc-msl-name”: Defines the name of the target variable for forecast mean sea level pressure (MSL). Default: matchup.nwp.fc.mean_sea_level_pressure.
- “fc-t2-name”: Defines the name of the target variable for forecast 2m temperature (T2). Default: matchup.nwp.fc.2m_temperature.

- “fc-d2-name”: Defines the name of the target variable for forecast 2m dew point (D2). Default: matchup.nwp.fc.2m_dew_point.
- “fc-tp-name”: Defines the name of the target variable for forecast total precipitation (TP). Default: matchup.nwp.fc.total_precipitation.
- “an-clwc-name”: Defines the name of the target variable for analysis cloud liquid water content (CLWC). Default: matchup.nwp.an.cloud_liquid_water_content
- “fc-clwc-name”: Defines the name of the target variable for forecast cloud liquid water content (CLWC). Default: matchup.nwp.fc.cloud_liquid_water_content
- “an-tcwv-name”: Defines the name of the target variable for analysis total column water vapour (TCWV). Default: matchup.nwp.an.total_column_water_vapour
- “fc-tcwv-name”: Defines the name of the target variable for forecast total column water vapour (TCWV). Default: matchup.nwp.fc.total_column_water_vapour
- “fc-sshf-name”: Defines the name of the target variable for forecast surface sensible heat flux (SSHF). Default: matchup.nwp.fc.surface_sensible_heat_flux
- “fc-slhf-name”: Defines the name of the target variable for forecast latent sensible heat flux (SLHF). Default: matchup.nwp.fc.surface_latent_heat_flux.
- “fc-blh-name”: Defines the name of the target variable for forecast boundary layer height (BLH). Default: matchup.nwp.fc.boundary_layer_height.
- “fc-ssrd-name”: Defines the name of the target variable for forecast downward surface solar radiation (SSRD). Default: matchup.nwp.fc.downward_surface_solar_radiation.
- “fc-strd-name”: Defines the name of the target variable for forecast downward surface thermal radiation (STRD). Default: matchup.nwp.fc.downward_surface_thermal_radiation.
- “fc-ssr-name”: Defines the name of the target variable for forecast surface solar radiation (SSR). Default: matchup.nwp.fc.surface_solar_radiation.
- “fc-str-name”: Defines the name of the target variable for forecast surface thermal radiation (STR). Default: matchup.nwp.fc.surface_thermal_radiation.
- “fc-ewss-name”: Defines the name of the target variable for forecast turbulent stress east component (EWSS). Default: matchup.nwp.fc.turbulent_stress_east_component.
- “fc-nsss-name”: Defines the name of the target variable for forecast turbulent stress north component (NSSS). Default: matchup.nwp.fc.turbulent_stress_north_component.

Configuration parameter for the sensor extraction process:

- “time-variable-name”: Defines the name of the time variable to use as reference time. Time variables are expected to store data in seconds since 1970 format.
- “longitude-variable-name”: Defines the name of the longitude variable to use as reference. Data is expected in decimal degrees east.
- “latitude-variable-name”: Defines the name of the longitude variable to use as reference. Data is expected in decimal degrees north.
- “x-dimension”: defines the extract window extent in x direction.
- “x-dimension-name“: defines the name of the NetCDF dimension for the x direction extent of the NWP extract.
- “y-dimension”: defines the extract window extent in y direction.
- “y-dimension-name“: defines the name of the NetCDF dimension for the y direction extent of the NWP extract.
- “z-dimension”: defines the extract window extent in z direction (i.e. atmospheric layers).

- "z-dimension-name": defines the name of the NetCDF dimension for the z direction extent of the NWP extract.
- "an-ci-name": Defines the name of the target variable for analysis sea-ice-fraction (CI). Default: amsre.nwp.seaice_fraction.
- "an-asn-name": Defines the name of the target variable for analysis snow albedo (ASN). Default: amsre.nwp.snow_albedo.
- "an-sstk-name": Defines the name of the target variable for analysis sea surface temperature (SSTK). Default: amsre.nwp.sea_surface_temperature.
- "an-tcwv-name": Defines the name of the target variable for analysis total column water vapour content (TCWV). Default: amsre.nwp.total_column_water_vapour.
- "an-msl-name": Defines the name of the target variable for analysis mean sea level pressure (MSL). Default: amsre.nwp.mean_sea_level_pressure.
- "an-tcc-name": Defines the name of the target variable for analysis total cloud coverage (TCC). Default: amsre.nwp.total_cloud_cover.
- "an-u10-name": Defines the name of the target variable for analysis 10m east wind component (U10). Default: amsre.nwp.10m_east_wind_component.
- "an-v10-name": Defines the name of the target variable for analysis 10m north wind component (V10). Default: amsre.nwp.10m_north_wind_component.
- "an-t2-name": Defines the name of the target variable for analysis 2m temperature (T2). Default: amsre.nwp.2m_temperature.
- "an-d2-name": Defines the name of the target variable for analysis 2m dew point (D2). Default: amsre.nwp. 2m_dew_point.
- "an-al-name": Defines the name of the target variable for analysis albedo (AL). Default: amsre.nwp.albedo.
- "an-lnsp-name": Defines the name of the target variable for analysis logarithmic surface pressure (LNSP). Default: amsre.nwp.log_surface_pressure.
- "an-skt-name": Defines the name of the target variable for analysis skin temperature (SKT). Default: amsre.nwp.skin_temperature.
- "an-t-name": Defines the name of the target variable for analysis temperature profile (T). Default: amsre.nwp.temperature_profile.
- "an-q-name": Defines the name of the target variable for analysis water vapour profile (Q). Default: amsre.nwp.water_vapour_profile.
- "an-o3-name": Defines the name of the target variable for analysis ozone profile (O3). Default: amsre.nwp.ozone_profile.
- "an-clwc-name": Defines the name of the target variable for analysis cloud liquid water (CLWC). Default: amsre.nwp.cloud_liquid_water.
- "an-ciwc-name": Defines the name of the target variable for analysis cloud ice water (CIWC). Default: amsre.nwp.cloud_ice_water.
- "an-tp-name": Defines the name of the target variable for analysis total precipitation (TP). Default: amsre.nwp.total_precip.

4.1.5.1.4 Cloud Flagging for HIRS

This plugin computes and adds cloud flag data to a matchup dataset containing HIRS data.

Currently the software is not able to differentiate the cases "water" and "ice covered water". So "ice covered water" now is calculated like "water".

Example plugin xml part:

```
<hirs-l1-cloudy-flags>
   <hirs-sensor-name>hirs-n18</hirs-sensor-name>
   <hirs-var-name-source-file-name>hirs-n18_file_name</hirs-var-name-source-file-name>
   <hirs-var-name-source-x>hirs-n18_x</hirs-var-name-source-x>
   <hirs-var-name-source-y>hirs-n18_y</hirs-var-name-source-y>
   <hirs-var-name-processing-version>hirs-n18_processing_version</hirs-var-name-processing-version>
   <hirs-var-name-source-11_1-um>bt_ch08</hirs-var-name-source-11_1-um>
   <hirs-var-name-cloud-flags>hirs-n18_cloudy_flags</hirs-var-name-cloud-flags>
   <hirs-var-name-latitude>hirs-n18_lat</hirs-var-name-latitude>
   <hirs-var-name-longitude>hirs-n18_lon</hirs-var-name-longitude>
   <hirs-var-name-11_1-um>hirs-n18_bt_ch08</hirs-var-name-11_1-um>
   <hirs-var-name-6_5-um>hirs-n18_bt_ch12</hirs-var-name-6_5-um>
   <distance-product-file-path>/path/to/the/distance-NetCDF-file.nc</distance-product-file-path>
</hirs-l1-cloudy-flags>
```

The plugin differentiates three matchup scenes:

- "water"
- "ice covered water"
- "land"

In the case of a "water" scene, it is needed to reopen the source orbit and extract a 45 x 45 pixels area around the center position of the current matchup.

Therefor this list of parameters is needed:

- <hirs-sensor-name>
  The sensor name, which was used to ingest the input product to the database.
- <hirs-var-name-source-file-name>
  The name of the MMD variable, which contains the source filename of the input file
- <hirs-var-name-source-x>
  The name of the MMD variable, which contains the source center pixel **x** of the matchup
- <hirs-var-name-source-y>
  The name of the MMD variable, which contains the source center pixel **y** of the matchup
- <hirs-var-name-processing-version>
  The processing version, which was used to ingest the input product to the database.
- <hirs-var-name-source-11_1-um>
  The name of the 11,1µm brightness temperature variable of the source file

In the cases "ice covered water" or "land" the plugin can compute the flags without extraction of data from the source product.

The following list of parameters are effective for all cases.

- <hirs-var-name-cloud-flags>
  The name of the cloud flags variable, which will be attached to the MMD file.

- <hirs-var-name-11_1-um>
  The name of the variable containing the brightness temperature values of 11,1µm.
- <hirs-var-name-6_5-um>
  The name of the variable containing the brightness temperature values of 6,5µm.

Since HIRS input data currently does not contain quality flags, the only way to detect “land” or “water” scene is to fetch a distance to land value from a verified map file.

To detect “land” case currently the distance file: Globolakes-static_distance_to_land_Map-300m-P5Y-2005-ESACCI_WB-fv1.0_RES120.nc is used.

- Input: lat and lon position of the matchup center pixel
- Output: distance to land value (unit km)
- Evaluation: “land” = distance to land value is less than 0.3 km

Therefor the following list of parameters is needed.

- <hirs-var-name-latitude>
  The name of the variable containing the latitude values.
- <hirs-var-name-longitude>
  The name of the variable containing the longitude values.
- <distance-product-file-path>
  The path to the distance to land map file.

4.1.5.1.5 ElevationToSolzen

This plugin converts viewing geometry angles stored as elevation angles (i.e. angles defined wrt. the normal plane) to zenith angles (i.e. angles defined wrt. the nadir line). It optionally allows to remove the original elevation angle variables.

Example configuration:

```
<elevation-to-solzen-angle>
  <convert source-name = "elevation_angle_1"
        target-name = "zenith_angle_1"
        remove-source = "false"/>
  <convert source-name = "elevation_angle_2"
        target-name = "zenith_angle_2"
        remove-source = "true"/>
</elevation-to-solzen-angle>
```

The configuration consists of a list of “convert” elements that describe a single angle conversion process. Any number of “convert” elements can be added to the configuration. Each of the element has three attributes:

- “source-name”: defines the source elevation variable name. Must appear exactly as stored I the MMD.
- “target-name”: defines the target zenith angle variable name.
- “remove-source”: Flag to keep/remove the source variable. If set to “true”, the source variable is remove fro the MMD, if set to “false”, the variable is kept. Default value when not set is “true”.

4.1.5.1.6 AddAmsrAngles

Post processing plugin that adds solar zenith and solar azimuth angle variables to AMSR-E and AMSR -2 data that follow the (A)ATSR angles convention. These are calculated following the equations:

$$solarZenithAngle = SunElevation + EarthIncidence$$

$$solarAzimuthAngle = (EarthAzimuth - SunAzimuth + 180.0)\ \%\ 360.0$$

Example configuration:

```
<add-amsr-solar-angles>
   <sun-elevation-variable name = "Sun_Elevation" />
   <sun-azimuth-variable name = "Sun_Azimuth" />
   <earth-incidence-variable name = "Earth_Incidence" />
   <earth-azimuth-variable name = "Earth_Azimuth" />
   <sza-target-variable name = "amsre.solar_zenith_angle" />
   <saa-target-variable name = "amsre.solar_azimuth_angle" />
</add-amsr-solar-angles>
```

The configuration parameter are:

- sun-elevation-variable: name of the sun elevation input variable
- sun-azimuth-variable: name of the sun azimuth input variable.
- earth-incidence-variable: name of the earth incidence input variable
- earth-azimuth-variable: name of the earth azimuth input variable
- sza-target-variable: name of the target variable containing the solar zenith angle
- saa-target-variable: name of the target variable containing the solar azimuth angle

4.1.5.1.7 AddAmsr2ScanQualityData

This post processing plugin adds the AMSR2 Scan_Data_Quality variable to the MMD. It copies the matching original 512 bytes data array to the MMD data.

Example configuration:

```
<add-amsr2-scan-data-quality>
   <filename-variable name = "amsr2-gcw1_file_name" />
   <processing-version-variable name = "amsr2-gcw1_processing_version" />
   <y-variable name = "amsr2-gcw1_y" />
```

```
   <target-variable name = "amsr2-gcw1_Scan_Data_Quality" />
</add-amsr2-scan-data-quality>
```

The configuration parameter are:

- filename-variable: name of the file name input variable
- processing-version-variable: name of the processing version input variable.
- y-variable: name of the original scanline input variable
- target-variable: name of target variable

4.1.5.1.8 AddIasiSpectrum

This plugin extends MMDs that have been generated using the IASI reader (see chapter 3.4.12) to contain the full IASI hyper-spectral data. This step is implemented as a post-processing because the target variable is a 4-dimensional dataset that cannot per-se be handled by the MMS core engine. The variable generated by this plugins has the dimensions matchup/y/x/spectrum.

Example configuration:

```
<add-iasi-spectrum>
   <target-variable name = "GS1cSpect" />
   <reference-variable name = "OnboardUTC" />
   <x-variable name = "iasi-ma_x" />
   <y-variable name = "iasi-ma_y" />
   <file-name-variable name = "iasi-ma_file_name" />
   <processing-version-variable name = "iasi-ma_processing_version" />
 </add-iasi-spectrum>
```

The configuration parameter are:

- target-variable: name of the target variable
- reference-variable: name of reference input variable. This is used to detect the target dimensions.
- x-variable: name of the x-position input variable containing the matchup raster x position
- y-variable: name of the x-position input variable containing the matchup raster y position
- file-name-variable: name of the input variable that contains the filename where the matchup was taken from
- processing-version-variable: name of the input variable containing the file processing version

4.1.5.1.9 AddDistanceToLand

A post-processing plugin that adds a variable containing the distance to the closest landmass/coastline. The distance is calculated based on an auxiliary file generated in the GLOBOLAKES project with a spatial resolution of 0.0083 degrees. Land pixels contain a distance of 0.0 km.

Example configuration:

```
<add-distance-to-land>
   <aux-file-path>/data/fiduceo/aux/globolakes.nc</aux-file-path>
   <target-variable name = "distance-to-land" />
```

```
  <lon-variable name = "xyz_longitude" />
  <lat-variable name = " xyz_latitude " />
</add-distance-to-land>
```

The configuration parameter are:

- aux-file-path: absolute path to the GLOBOLAKES distance auxiliary file
- target-variable: name of the variable written. The datatype is float and the dimension of this variable is the same as the dimension of the longitude variable used as input.
- lon-variable: input variable name of the variable containing the geolocation longitudes
- lat-variable: input variable name of the variable containing the geolocation latitudes

4.1.5.1.10 CALIOP Feature Classification Flags Midpoints Extract

This plugin extracts horizontal midpoints from each vertical layer of the source CALIOP file from the variable “Feature_Classification_Flags” and adds the data to the mmd file to an user defined target variable.

Example configuration:

```
<post-processings>
    <caliop-level2-vfm-flags>
        <mmd-source-file-variable-name>
            caliop_vfm.file_name
        </mmd-source-file-variable-name>
        <mmd-processing-version-variable-name>
            caliop_vfm.processing_version
        </mmd-processing-version-variable-name>
        <mmd-y-variable-name>
            caliop_vfm.y
        </mmd-y-variable-name>
        <target-fcf-variable-name>
            caliop_vfm.Center_Feature_Classification_Flags
        </target-fcf-variable-name>
    </caliop-level2-vfm-flags>
</post-processings>
```

The configuration parameter are:

- <mmd-source-file-variable-name>
  The name of the MMD variable, which contains the source filename of the input file
- <mmd-processing-version-variable-name>
  The processing version, which was used to ingest the input product to the database.
- <mmd-y-variable-name>
  The name of the MMD variable, which contains the source center pixel **y** of the matchup.
  A center x variable is not needed because CALIOP products have a width of 1
- <target-fcf-variable-name>
  The name of the new variable which will be added.

4.1.5.1.11 CALIOP add L2 CLay variables

This plugin adds “caliop_clay-cal” (clay = cloud layer) product variables and data to an mmd file which already contains “caliop_vfm-cal“ data. (vfm = vertical feature mask)

Example configuration:

```
<post-processings>
    <caliop-sst-wp100-clay>
        <mmd-source-file-variable-name>
            caliop_vfm.file_name
        </mmd-source-file-variable-name>
        <processing-version>
            4.10
        </processing-version>
        <mmd-y-variable-name>
            caliop_vfm.y
        </mmd-y-variable-name>
        <target-variable-prefix>
            caliop_clay.
        </target-variable-prefix>
    </caliop-sst-wp100-clay>
</post-processings>
```

The configuration parameter are:

- <caliop-sst-wp100-clay>
  The main tag of this post processing plugin for caliop cloud layer product data
- <mmd-source-file-variable-name>
  The name of the MMD variable, which contains the VFM source filename. The VFM source file name is needed to create the correspding caliop_clay-cal filename.
- <processing-version>
  The processing version of caliop_clay-cal products, which was used to ingest the clay products to the database.
- <mmd-y-variable-name>
  The name of the MMD variable, which contains the VFM source center pixel **y** of the matchup. A center x variable is not needed because CALIOP products have a width of 1
- <target-variable-prefix>
  To insert the caliop cloud layer product variables to an mmd file a variable prefix is needed. e.g. “caliop-cal.”

4.1.5.1.12 AIRS add channel variables

This plugin adds channel data from AIRS L1B source products

Example configuration:

```
<post-processings>
    <add-airs-channel-data>
        <mmd-source-file-variable-name>
            airs-aq_file_name
```

```
            </mmd-source-file-variable-name>
            <mmd-processing-version-variable-name>
                airs-aq_processing_version
            </mmd-processing-version-variable-name>
            <mmd-x-variable-name>
                airs-aq_x
            </mmd-x-variable-name>
            <mmd-y-variable-name>
                airs-aq_y
            </mmd-y-variable-name>
            <mmd-variable-name-cut-out-reference>
                airs-aq_cutOutRef
            </mmd-variable-name-cut-out-reference>
            <target-variable-name-radiances>
                airs-aq_radiances
            </target-variable-name-radiances>
            <target_variable_name_CalFlag>
                airs-aq_CalFlag
            </target_variable_name_CalFlag>
            <target_variable_name_SpaceViewDelta>
                airs-aq_spaceViewDelta
            </target_variable_name_SpaceViewDelta>
        </add-airs-channel-data>
</post-processings>
```

Configuration parameters are:

- < add-airs-channel-data >
  The main tag of this post processing plugin for AIRS channel data
- <mmd-source-file-variable-name>
  The name of the MMD variable, which contains the AIRS product source filename. The source file name is needed to fetch channel data from the sam file as other data per matchup.
- <mmd-processing-version-variable-name>
  The processing version of AIRS products, which was used to ingest the clay products to the database.
- <mmd-x-variable-name>
  The name of the MMD variable, which contains the AIRS source center pixel **x** of the matchup.
- <mmd-y-variable-name>
  The name of the MMD variable, which contains the AIRS source center pixel **y** of the matchup.
- <mmd-variable-name-cut-out-reference>
  Name of reference input variable. This is used to detect the target dimensions.
- <target-variable-name-radiances>
  Target variable name for radiances channel data.
  e.g. “airs-aq_radiances”
- <target_variable_name_CalFlag>
  Target variable name for CalFlag channel data.
  e.g. “airs-aq_CalFlag”

- <target_variable_name_SpaceViewDelta>
  Target variable name for SpaceViewDelta channel data.
  e.g. “airs-aq_SpaceViewDelta”

4.1.5.1.13 AddGruanSource

Post processing that adds the GRUAN reference file path to the MMD.

Example configuration:

```
<add-gruan-source>
  <target-variable name = "gruan-source-file" />
  <y-variable name = "gruan-uleic_y" />
  <file-name-variable name = "gruan-uleic_file_name" />
  <processing-version-variable name = "gruan-uleic_processing_version" />
</add-gruan-source>
```

Configuration parameters are:

- <target-variable>: the name of the variable in the target containing the file references
- <y-variable>: The name of the variable containing the GRUAN y matchup location
- <file-name-variable>: the name of the variable containing the Gruan reference file in-situ file name
- <processing-version-variable>: the name of the variable containing the input file version.

4.1.5.1.14 AddAvhrrCorrCoeffs

Post processing that adds the AVHRR FIDUCEO FCDR correlation variables to the MMD. This processing adds a variable that contains one set of correlation coefficients for each input file. The data dimension is therefore different from the number of matchups. Instead an additional dimension “input_files” is created that contains for each dimension the associated input file name.

Example configuration:

```
<add-avhrr-corr-coeffs>
  <file-name-variable name = "input-file-name-variable" />
  <processing-version-variable name = "input-processing-version" />
  <target-x-elem-variable name = "cross_element_correlation_coefficients" />
  <target-x-line-variable name = "cross_line_correlation_coefficients" />
</add-avhrr-corr-coeffs>
```

Configuration parameters are:

- <file-name-variable>: the name of the variable containing the AVHRR FCDR input file name.
- <processing-version-variable>: the name of the variable containing the input file version.
- <target-x-elem-variable>: the name of the target variable for the cross element correlation coefficients variable
- <target-x-line-variable>: the name of the target variable for the cross line correlation coefficients variable

5 MMD file format

TODO

5.1 Global Attributes

This section lists all global attributes of the MMD that are added by the MMS.

5.1.1 sensor-name

This attribute contains a comma separated list of sensor names that have been used to generate the MMD. The first name in the list is the name of the primary (reference) sensor.

5.2 Extra MMD Variables

In addition to the original sensor variable extracts, the MMS adds a number of useful variables that facilitate traceability of the data. These are listed in the following chapters.

5.2.1 <sensor-name>_x, <sensor-name>_y

These variables contain the x and y position of the matchup location for the sensor *<sensor>* in the input data raster. The variables are vectors with one value per matchup layer containing the matchup centre location.

5.2.2 <sensor-name>_file_name

A variable that contains for each matchup-layer the name of the sensor input file for each sensor *<sensor>*.

5.2.3 <sensor-name>_processing_version

Contains the processing version of the original sensor input file for each sensor *<sensor>* as stated during the database ingestion of the sensor data.

5.2.4 <sensor-name>_acquisition_time

This variable contains for each pixel in the matchup-extraction window the normalized acquisition time in seconds since 1970-01-01 (Unix epoch). All sensor specific acquisition times are converted to this value with the highest precision possible.

5.3 CF Confomity

5.3.1 CF conform usage of “units” attribute

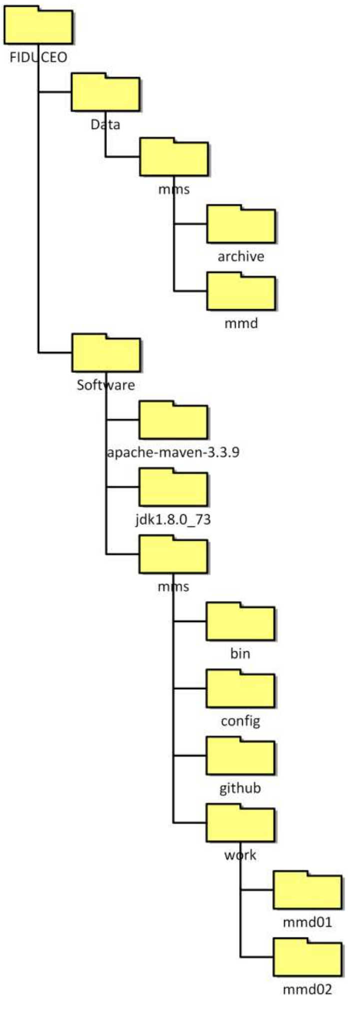
If the MMD-File-Writer detects, that an input variable contain a non CF conform “unit” attribute, the writer duplicates the atrribute with the CF conform name “units” instead of “unit”. Finally such a variable contain two unit attributes with the same value. One with a non CF confom attribute name and one with a CF conform attribute name.

6 Installation and Operation on CEMS

This chapter covers the installation and operation of the Fiduceo MMS on the parallel processing facility JASMIN/CEMS supplied by CEDA. The Fiduceo project has a specific virtual machine assigned on the CEMS system for operations.

6.1 Workspace Structure

This chapter describes the Fiduceo workspace directory structure and the way the MMS in embedded into it. The sketches omit all non-MMS-related directories for a more clear view of the workspace structure.



The directory branch below “Data” contains a subdirectory for the “mms” that contains MMS related data. The “archive” directory contains all sensor and in-situ input data ordered in the directory structure described in the MMS Implementation Plan (see RD 1) whereas the directory “mmd” serves as target directory for the processing results. Subdirectories for each use-case will be generated during the processing of matchups.

The directories below “Software” contain all software comprising the MMS and also all third-party software used to build and run the MMS. Subdirectories “apache-maven-3.3.9” and “jdk1.8.0_73” contain the apache maven build manager and the Java JDK/runtime required to compile and run the MMS.

The subdirectory “mms” contains all MMS related software, configuration, a build directory and a set of working directories. The current operating version of the MMS is contained in the directory “bin”. The configuration files used by the MMS are stored in the directory “config”.

The subdirectory “github” is the build directory. It contains the complete source code of the MMS and is configured to be able to fetch code changes from github and to build and deploy the software.

In the “work” directory contained are the working directories for processing. Any processing job needs to be operated from a dedicated sub-directory of “work”. Each subdirectory for a specific use-case processing contains status files, log-directories and all other files required by PMonitor to operate.

6.2 Building the MMS on CEMS

The build and deployment process is mostly automated. To build and deploy “master” branch (i.e. the latest development version) of the github repository, change to the directory “FIDUCEO/Software/mms/github/fiduceo” and invoke the script “build_and_install_on_cems.sh”. This script performs the following steps:

- Checkout the latest version of the code from github
- Invoke maven to build the code and run all tests
- Clean up the “bin” directory
- Invoke maven to build an assembly of the binaries and deploy it to the “bin” directory.

When these steps have been run successfully, the MMS software is ready to operate.

6.3 Operating the MMS on CEMS

This chapter covers all information with respect to operations of the MMS on the CEMS environment and specific configuration and set-up tasks.

6.3.1 Setting up the environment

Before the MMS is able to operate on the LOTUS cluster, a number of environment variables and paths need to be set. This is done by simply sourcing a pre-configured script.

Assuming that you are in a use-case subdirectory e.g. in “FIDUCEO/Software/mms/work/mmd08” sourcing the environment file is accomplished by executing:

```
[tblock01@lotus mmd08]$. ../../bin/mms-env.sh
```

The script file is automatically installed to the directory “FIDUCEO/Software/mms/bin” by the maven deployment script.

6.3.2 CEMS Specific Configuration

The CEMS environment requires a number of specific configuration settings that ensure the proper operation on the cluster. All configuration files accessed by the operational MMS are located in the directory “FIDUCEO/Software/mms/config”.

IMPORTANT: The configuration files in this directory will not be updated by the MMS build and deploy procedure. Any changes in the file content, e.g. additional fields or field renaming, have to be inserted manually into the operational configuration.

6.3.2.1 System Configuration

The MMS system configuration is set up to pick up the correct input data archive path for the CEMS installation. The following setting is mandatory on CEMS:

archive-root: /group_workspaces/cems2/fiduceo/Data/mms/archive

For a description of the system configuration file, please refer to chapter 4.1.1.

6.3.2.2 Database Configuration

The database installed on the Fiduceo workspace is a MongoDB server installation, version 3.2. This database requires a fixed set of access configuration parameters for the database configuration:

driverClassName: mongodb

url: mongodb://172.26.69.130:27017/FIDUCEO

username: xxxx

password: xxxx

Username and password are available on request. For a description of the database configuration file, please refer to chapter 4.1.2.

6.3.3 Operating the MMS

On the parallel processing environment on the LOTUS cluster, the MMS software is controlled by an additional Python layer, PMonitor. This software layer controls the deployment and parallel execution of MMS-Tool processes (Ingestion, Matchup, Post Processing) to the cluster nodes. It ensures that processing can be paused and re-invoked and that on e.g. hardware or network failures, large processing jobs can safely continue where they have been interrupted.

PMonitor is invoked from the console on the lotus VM using three shell scripts that are located in the directory “FIDUCEO/Software/mms/bin” – and that are automatically available on the path when the environment is set up correctly following the instructions in chapter 6.3.1.

pmstart.sh: this command is used to submit a parallel processing task to the cluster. It assumes a single cmd-line argument, the name of the workflow-file that contains the process description, e.g.

“pmstart.sh ingest_avhrr_n07.py”.

pmstop.sh: this command stops the execution of parallel tasks and kills the currently running processes on LOTUS. PMonitor ensures that the processing can later pick up processing at the point where the job has been stopped. It assumes a single cmd-line argument, the name of the workflow-file that contains the process description, e.g.

“pmstop.sh ingest_avhrr_n11.py”.

pmclean.sh: this command cleans up all temporary files that have been generated during the processing execution. It assumes a single cmd-line argument, the name of the workflow that has been processed, e.g.

“pmclean ingest_avhrr_m02”.

The shell scripts are designed in a way that it is safe to close the console after a processing job has been started. Processing will keep on running in the background.

6.3.4 Workflow Files

The definition of the jobs to be executed by PMonitor is stored in workflow files. These are Python files that create and configure a workflow object, which is using PMonitor to execute the parallel processing on the cluster. Workflow files are available from and have to be stored into the directory “FIDUCEO/Software/mms/bin/python”.

An example workflow-file is displayed below:

```
from workflow import Workflow

w = Workflow('usecase02_avhrr_m02_n18', 7, '/group_workspaces/cems2/fiduceo/Software/mms/config')

w.add_primary_sensor('avhrr-m02', '2006-10-30', '2015-12-31', 'v01.2')
w.add_secondary_sensor('avhrr-n18', '2006-10-30', '2015-12-31', 'v01.2')

w.set_usecase_config('usecase-02.xml')

w.run_matchup(hosts=[('localhost', 24)])
```

A workflow is constructed using three parameters:

- usecase: the name of the usecase the workflow is operating on – used to name log-files and other process control files
- time_slot_days: the time slicing applied to the processing period in days. Every time slice is processed as a separate job on the cluster.
- config_dir: the absolute path to the configuration directory

Subsequently, a number of sensors can be added to the workflow using the commands

- add_primary_sensor(…): adds the primary (i.e. reference) sensor to the workflow
- add_secondayr_sensor(…): adds a secondary (i.e. associated) sensor to the workflow.

Each workflow must have a reference sensor, the number of associated sensors is not restricted. The parametrisation for both methods is the same, using:

- name: the sensor name as listed in chapter 3.4.
- start_date: processing interval start date , formatted yyyy-MM-dd
- end_date: processing interval end date , formatted yyyy-MM-dd
- version: the sensor data version

The start and end dates of the sensors must not (although they can) match the processing interval. These tags are meant to narrow down the validity range of the data. PMonitor automatically calculates the largest possible processing interval.

The workflow must know the processing configuration file to be used. This is accomplished by the call to set_usecase_config(). The argument is the name of a configuration file stored in the configuration directory passed in to the constructor. Configuration files are described in chapters 4.1.4 and 4.1.5.

The final line sets the hostnames and the parallelization to be used to execute the job. On CEMS, the only possible host for submitting jobs is “localhost”, assuming the MMS is operated on one of the CEMS science nodes. The number behing the host name denotes the maximum number of parallel nodes used to execute the job.

IMPORTANT: the LSF scheduler on CEMS is overloaded when using more than approx. 150 parallel nodes.