Jason-CS/Sentinel-6 Generic Product Format Specification (GPFS)

EUMETSAT
Eumetsat-Allee 1, D-64295 Darmstadt, Germany
Tel: +49 6151 807-7
Fax: +49 6151 807 555
http://www.eumetsat.int

Doc.No. : **EUM/LEO-JASCS/SPE/17/897975**
Issue : v4 e-signed
Date : 19 November 2019
WBS/DB : LEO-JCS-200000

Document Signature Table

	Name	Function	Signature	Date
Prepared by:	Cristina Martin-Puig	Ocean Surface Topography Expert		
Reviewed by:	Remko Scharroo	EUMETSAT Project Scientist		
Reviewed by:	Paola Colagrande	Jason-CS IFC System Engineer		
Reviewed by:	Julia Figa Saldana	Jason-CS Ground Segment and PDAP Manager		
Reviewed by:	Conrad Jackson	Jason-CS System Manager		
Reviewed by:	Antonio Fanigliulo	Quality Assurance Engineer		
Reviewed by:	David Antonelli	Configuration Management Engineer		
Approved by:	Manfred Lugert~~Francois Parisot~~	Jason-CS Programme Manager		

Formatted: Font: (Default) Arial, 10 pt

Distribution List

Name	Organisation
As per signature table	EUMETSAT

Document Change Record

Version	***Date of Version*** *as on profile*	***Document Change Request (DCR) Number*** *if applicable*	***Description of changes***
1 Draft	23/01/2017		First Draft
1A	07/02/2017		Initial version for internal review
1B	14/02/2017		Version released for System PDR
1C	12/04/2017		Version for pre-releasing the PDAP ITT. This version is also addressing the System PDR review RIDS: RID_024 is now addressed. The new SAFE 2.x format is described in section 5.
1D	15/05/2017		Version for the PDAP ITT including minor editorial changes.
1E	06/11/2017		Added internal review comments. SAFE 2.x was de-scoped during system PDR close out. This document shall then provide a description of SENTINEL-SAFE format.
1F	16/01/2018		This version addresses internal review RIDS
2	24/01/2018	JCS_DCR_12	Document released for PDAP KO
2A	16/04/2018		Referenced the latest CF conventions document version 1.7 instead of 1.6. Removed duplicate reference. Added reference to Level 0 PFS. Included NetCDF groups description logic, included common Global Attributes for NetCDF files generated by the PDP. Included clarification on ISO 19115 standards, and their applicability. Added clarification on BUFR production. Removed units from Table 3-3 that are not used in any of the Product Format Specifications; corrected units for latitude and longitude. Minor editorial changes.
2B	19/04/2018	JCS_DCR_63	Version released for PDAP data package # 2.
2C	14/09/2018		Updated “title” attribute to add radiometer product. Changed “_meas_” to “_measurement_” [MPWG-23]. Changed “semi_major_ellipsoid_axis to “ellipsoid_semi_major_axis” [MPWG-23]. Added new Section 3.3.2 to clarify that what follows only applies to manifests.
3	24/10/2018	JCS_DCR_139	Version released for the System Check Point#2/CDR

Version	*Date of Version* *as on profile*	*Document Change Request (DCR) Number* *if applicable*	*Description of changes*
3A	10/04/2019	JCS_DCR_178	Corrected malformed Section link. Clarified “optional NetCDF compression”.
4	19/11/2019	JCS_DCR_313	Replaced Sentinel-6/Jason-CS by Jason-CS/Sentinel-6. Provided example of attribute “references” (Redmine issue #1275)

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

1.1 Purpose and Scope

This is the Generic Product Format Specification (GPFS) for the products made by the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Payload Data Processing (PDP) within the Payload Data Acquisition and Processing (PDAP) facilities at the EUMETSAT Mission Control Centre (MCC). It specifies the aspects that are common to all of the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Products. This document addresses the native format of the products generated in the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 PDP.

The text in this document shall be considered as specifying requirements applicable to all ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products in native format.

The document may be updated to reflect common fields identified during the development of individual product format specifications. The ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Product Format Specifications of Level 0, Level 1, and Level 2 products from each of the product processing facilities in the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 system complement this document.

1.2 Applicable Documents

AD- 1	PGSI-GSEG-EOPG-FS-05-0001	SAFE Control Book Volume 1 Core Specifications. (19-02-2015)
AD- 2	PGSI-GSEG-EOPG-FS-05-0002	SAFE Control Book Volume 2 Recommendations for Specifications (25/09/2014)
AD- 3	EUM/LEO-JASCS/SPE/17/912241	Sentinel-6_Jason-CS – Metadata Specification
AD- 4	http://www.unidata.ucar.edu/software/netcdf/docs/netcdf.pdf	The NetCDF Users' Guide (NetCDF version 4.1.3 from 10 June 2011)
AD- 5	EUM/LEO-JASCS/SPE/17/899011	~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 File Naming Convention
AD- 6	EUM/LEO-JASCS/SPE/12/0039	Sentinel-6 System Requirements Document

1.3 Reference Documents

RD- 1	http://cfconventions.org/Data/cf-conventions/cf-conventions-1.7/cf-conventions.html	NetCDF Climate and Forecast (CF) Metadata Conventions: Version 1.7, 1 August 2017
RD- 2	http://wiki.esipfed.org/index.php?title=Category:Attribute_Conventions_Dataset_Discovery	NetCDF Attribute Convention for Dataset Discovery
RD- 3	http://www.bipm.org/en/measurement-units/	Measurement units: the SI; from the Bureau International des Poids et Mesures
RD- 4	http://www.wmo.int/pages/prog/www/WMOCodes.html	WMO Manual on Codes, Common Table C. 2012 Edition.
RD- 5	EUM/LEO-JASCS/SPE/17/898270	~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Level 0 Product Format Specification (L0 PFS)
RD- 6	EUM/LEO-JASCS/DEF/13/695184	Sentinel-6 Glossary of Terms and Acronyms Document

RD- 7	EUM/LEO-JASCS/SPE/17/899450	~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Generic Auxiliary Data Specification
RD- 8	EUM/LEO-JASCS/SPE/17/899679	~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 ALT Level 1 Auxiliary Data Specification
RD- 9	EUM/LEO-JASCS/SPE.17/900471	~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 ALT Level 2 Auxiliary Data Specification

1.4 Acronyms

Acronym	*Meaning*
AIP	Archival Information Package
AMR	Advanced Microwave Radiometer
ANX	Ascending Node Crossing
ASCII	American Standard Code for Information Interchange
BUFR	Binary Universal Form Representation
EO	Earth Observation
GPFS	Generic Product Format Specifications
HKTM	House Keeping TeleMetry
ISP	Instrument Source Packets
LTDP	Long Term Data Preservation
MCC	Mission Control Centre
MJD	Modified Julian Days
NetCDF	Network Common Data Form
NRT	Near Real Time
NTC	Non Time Critical
OAIS	Open Archival Information System
PDAP	Payload Data Acquisition and Processing
PDP	Payload Data Processing
RO	Radio Ocultation
SAFE	Standard Achieve Format for Europe
STC	Short Time Critical
XFDU	XML Format Data Unit

1.5 Document Structure

Section 1 provides the introduction to this document. It includes the purpose, scope, applicable documents, reference documents and the acronyms.

Section 2 introduces the different products for the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 mission.

Section 3 sets out conventions that are applicable to all products.

Section 4 describes the NetCDF product structure and the elements generic to all products.

Section 5 addresses the packaging format. It details the format and contents of the packages for all levels.

Section 6 gives the metadata xml format.

Section 7 addresses the format conventions.

2 ~~SENTINEL-6/JASON-CS~~JASON-CS/SENTINEL-6 PRODUCTS CONTENT

2.1 Overview

The ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Payload Data Processing (PDP) produces its main user data products in only two native formats. The Level 0 products are formatted in binary, whereas the Level 1 and Level 2 are structured using the enhanced NetCDF-4 data model.

All ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products shall be compliant with the Sentinels Standard Archive Format for Europe (SENTINEL-SAFE[1]) [AD- 1][AD- 2].

SAFE is designed to act as a standard format for archiving and conveying Earth Observation (EO) data within the archiving facilities at ESA and cooperating agencies (e.g. EUMETSAT). SAFE has been design to be compliant with the Open Archival Information System (OAIS) standard. One of the SAFE features originating from OAIS is the use of XFDU packaging format (see section 5).

The ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products shall also be delivered to the users in a number of alternative encodings, for example BUFR (for ALT Level 2). This document only addresses the native format.

2.2 Product Levels

2.2.1 Level-0 Products

Level 0 Product	This term is used to describe reformatted, time-sorted and annotated Instrument Source Packets (ISPs) that belong to a single instrument. The Level 0 products for all instruments share a common format.

2.2.2 Poseidon-4 Products

Level 1A Product	This term is used to describe products containing Level 1 intermediate output of the HR processor. It includes geo-located bursts of Ku-band echoes (at ~140 Hz) with all instrument calibrations applied. This product includes the full rate complex waveforms input to the delay/Doppler or HR processor.
Level 1B Product	This term is used to describe the output of the LR and HR processors at Level 1. It includes geo-located, and fully calibrated multi-looked HR Ku-band waveforms.

[1] *http://earth.esa.int/SAFE/*

Level 2 Product	This term refers to products that contain the geophysical information derived from the altimeter: range, orbital altitude, time, and water vapour from the AMR-C, environmental and geophysical corrections, along with significant wave height, and wind-speed information.

2.2.3 AMR-C Products

Level 2 Product	This term refers to the AMR-C geophysical information derived from the radiometer.

2.2.4 RO Products

Level 1A Product	The L1A generally consists of pseudo-range, carrier phase and amplitude as measured by the RO instrument in its various measurements modes.
Level 1B Product	The primary content of Level 1B RO data are vertical bending angle profiles provided as function of the impact parameter, along with geo-referencing and some diagnostic data.

2.2.5 Ancillary/Auxiliary Products Definition

The Ancillary and Auxiliary data products format specification is described in a dedicated set of documents. This section only provides a brief description of these products.

Auxiliary Data	Generic term covering several data categories used in the products processing steps at the PDP, either of **static** or **dynamic** nature, from sources either internal or external to the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 System. We should find at least as auxiliary data: Processing parameters (e.g. in a classical ASCII file format for easy reading and modification) Characterization database (at instrument and satellite level): this database would be probably static after the launch. Calibration database: regularly updated through the calibration process. Orbit files: MOE and POE Climatological database (e.g. tide models): static Meteorological data (either forecast or a posterior analysis): dynamic Geographical database (e.g. land mask): static
Ancillary Data	At Level-0 we will also have ancillary annotation satellite data, and representation information. This is in detail described in the Level-0 PFS.

3 GENERIC SPECIFICATIONS

3.1 Introduction

This section defines conventions that are applicable to all ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products.

3.1.1 Terminology

A **Product Package** is defined as a physical collection of components that are gathered together, using a defined packaging scheme, into a single container. It includes a manifest file (XML format) describing the contained file or files, and the relationships among those files.

A product package is associated to the concept of PDP Product Category defined as an attribute of the Products Package, specifying the scope of a product. Two main categories are actually identified:

- **User Product** is a Product Package (EO product package, EO auxiliary package, EO collection package and Rep information package) generated for the operational dissemination to the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 users according to specific timeliness requirements, via subscription or Catalogue.
- **Internal Product** is a Product Package generated by the PDP used only internally for normal processing reasons (e.g. as an intermediate output/input product between two consecutive processing steps or an ancillary data product) thus it is not made available to the users. These products are not directly linked to a user dissemination timeliness requirement although their processing generation time inside the PDP is driven by these end-to-end requirements. However, these products may be disseminated when needed to identify special users (e.g. cal/val users, SALP, etc.). An Internal Product, depending on the evolution of the mission needs, might change its status into User Product during the mission lifetime.

Regarding timeliness, the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Products are categorized as:

- **Near Real Time (NRT)** products support operational objectives related to marine meteorology, air-sea interactions studies and real time operational oceanography. These products are generated using the auxiliary data available at the time of processing, which will be limited in quality (i.e. not the most precise orbit information and geophysical corrections). Their dissemination by the PDP is due in less than **3 hours** after data acquisition.
- **Short Time Critical (STC)** products support operational oceanography, i.e. improve ocean state analysis, forecasts and hind casts produced by numerical ocean prediction. Their accuracy is limited by the auxiliary data available at the processing time. STC products are based on preliminary orbit vectors and non-consolidated auxiliary data. Their dissemination by the PDP is due within **36 hours** after data acquisition.
- **Non-Time Critical (NTC)** products provide information on ocean topography and mean sea level in support of ocean and climate monitoring services. They rely on the availability of auxiliary of the highest quality. NTC products are consolidated products as they are generated using the most precise auxiliary information available. Their dissemination is due within **60 days** after data acquisition.

Within Product Format Specifications document the following terms are also used:

- **Product Type** refers to the fixed part of the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 PDP file naming convention [AD- 5], including the data source element the processing level and the product type ID. It consists of a total of 13 characters including the separator element ("_").
- **Product File name**, as defined by the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 PDP file naming convention [AD- 5], refers to the product package filename. It shall not be confused with the **Component File Name**, which refers to the filename of the physical components inside the package.
- **Sensing Start time and Sensing Stop time** is the period of time (in UTC) during which the data sensing occurred on board of the satellite, as calculated from the Satellite Binary Time counter.
- **Ascending Node Crossing (ANX)** refers to the latest crossing of the equator by the satellite going from South to North.
- **Orbit** is one full revolution of the satellite starting and ending at an ascending node. The satellite orbit is specified in two ways: absolute orbit and relative to a specific orbit cycle. The orbit numbers are specified at the sensing start time and sensing stop time of the product.
- **Pass** spans half an orbital revolution and is either ascending (South-North) or descending (North-South). This means that a pass always starts at the turnover point, i.e. the passing near the South or North Pole. The pass number represents the number of passes since the beginning of the mission (absolute) or since the beginning of the cycle (relative). Odd pass numbers are ascending, even are descending.
- **Cycle** is one full completion of the repeat period. A cycle starts at the equator when combined with orbit numbers, at the southern rollover point in when combined with pass numbers.
- **nssdcIdentifier** univocally identifies the mission according to standard defined by the World Data Centre for Satellite Information (WDCdlSI), available at http://nssdc.gsfc.nasa.gov/nmc/scdlquery.html

3.2 Product File Naming Convention

The file naming convention will be described in a separate document [AD- 4].

3.3 Data Representation

3.3.1 Data Types

The convention applies to the element data types are listed in the next table.

Table 3-1: Data Types

Variable Type	C Type	NetCDF type	Abbreviation	Range
character	char	byte	sc: signed char	-128 to 127
		ubyte	uc: unsigned char	0 to 255
2-byte integer	short	short	ss: signed short integer	-32768 to 32767
		ushort	us: unsigned short integer	0 to 65535
4-byte integer	long	int	sl: signed long integer	-2147483648 to 2147483647
		uint	ul: unsigned long integer	0 to 4294967295
8-byte integer	long long	int64	sll: signed long long integer	-9223372036854775808 to 9223372036854775807
		uint64	ull: unsigned long long integer	0 to 18446744073709551615
4-byte floating point	float	float	flt	3.4028e+38 (max) 1.17549e-38 (min)
8-byte floating point	double	double	dbl	1.79e+308 (max) 2.22e-308 (min)

The ANSI/IEEE 754-1985 is the chosen standard for storing real numbers.

3.3.2 Manifest data types

In the manifest file the following additional data type are also used to describe metadata.

Table 3-2: Additional Data Types used in manifests

Variable Type	C Type	NetCDF Type	Abbreviation
String	char[2]	char	S
String enumerative (string with limited number of possible values to be taken out of a set of predefined keywords)	char	char	E
String of the type yyyy-mm-ddThh:mm:ss.uuuuuu, representing the UTC date (year, month, day) and time (hours, minutes, seconds) separated by the character T, e.g. 1974-04-18T10:00:00.000000	char	char	UTC
Fixed String (only one value possible)	char	char	FS
Boolean (“0” for FALSE, “1” for TRUE)	char	char	B

For the metadata contained in XML files, the lexical representation that can be found in the xml schema part 2 follows data types (see http://www.w3.org/TR/2001/REC-xmlschema-2-20010502/).

[2] *It is of variable size, the number of bytes given in the Product Description Tables refers to the maximum number.*

In cases where a field is not fully filled by a value, placeholder values are used. For ASCII strings, the placeholder character is the ASCII blank space character. For numerical-values, the placeholder value is zero unless otherwise stated.

3.3.3 Bit/Byte numbering

For the purpose of identifying bits within a binary format, the numbering convention shown below is used (big endian rules). Byte 0 is the least significant byte. Within a byte, bit 0 is the least significant bit. Bytes and bits numbering always starts at 0.

- 1-byte structure

BYTE 0
7 6 5 4 3 2 1 0

- 2-byte structure (short)

BYTE 1	BYTE 0
7 6 5 4 3 2 1 0	7 6 5 4 3 2 1 0

- 4-byte structure (long):

BYTE 3	BYTE 2	BYTE 1	BYTE 0
7 6 5 4 3 2 1 0	7 6 5 4 3 2 1 0	7 6 5 4 3 2 1 0	7 6 5 4 3 2 1 0

- 4-byte structure (float):

Use IEEE standard (big endian)

- 8-byte structure (float):

Use IEEE standard (big endian).

3.3.4 Units

The convention applies to the element unit are listed in the next table. This table is in no way exhaustive.

Table 3-3: Units

Quantity	*Unit*	*NetCDF notation*
Time	jd or MJD or MJD2000	Jd or MJD or MJD2000
	10^{-6} s	µs
	s	s
Distance	m	m
	10^{3} m	km
	km	km
Velocity	$m.s^{-1}$	m/s or m.s-1

Quantity	*Unit*	*NetCDF notation*
Percentage	%	%
Voltage	V	V
Product Size	bytes	B
Temperature	K (degree Kelvin)	K
	C (degree Celsius)	C
Frequency	Hz	Hz
Resistance	Ohm	Ohm
Latitude	° (degree with positive values for North latitude)	degrees_north
Longitude	° (degree with positive values East of Greenwich meridian)	degrees_east
Angle	° (degree)	degrees
	rad	rad
Back scatter coefficient	dB	dB
Pressure	hPa	hPa
Total Column Water Vapour	$kg.m^{-2}$	kg/m^2 or kg.m^-2
Dimensionless	nc (numerical count) dl	count 1

4 NETCDF AND SI CONVENTIONS

4.1 Guiding Principles

All ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Level-1 and 2 products are formatted following the NetCDF-4 data model as described in [RD- 2]. As noted in [RD- 2], NetCDF-4 provides an abstraction – the data model – that:

"…supports a view of data as a collection of self-describing, portable objects that can be accessed through a simple interface. Array values may be accessed directly, without knowing details of how the data are stored. Auxiliary information about the data, such as what units are used, may be stored with the data. Generic utilities and application programs can access NetCDF datasets and transform, combine, analyze, or display specified fields of the data."

Specifications constraining aspects of the data models common to all ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products follow in the sections below.

This document does not address the physical file structure of the storage layer underlying the NetCDF-4 data model.

~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products should where practicable conform to the "Climate and Forcast (CF) Metadata Conventions" [RD- 1] extended where appropriate to satellite data, and should where practicable conform to the "Highly Recommended" aspects of the "NetCDF Attribute Convention for Dataset Discovery" [RD- 2]. It is known that [RD- 1] is not fit

for including all aspects of ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 yet (e.g. data grouping) and that this convention is evolving.

All representation of numbers and units shall adhere to the International System of Units as defined in [RD- 3].

4.2 The NetCDF Enhanced Data Model

The enhanced NetCDF-4 data model [AD- 4] supports **groups, dimensions, variables, and attributes**. In addition, it supports named groups of dimensions, variables, and attributes, as well as user-defined types. **Groups** can be nested hierarchically. The product itself constitutes the **'root' group**.

- **Groups** are used to organize large amounts of data. Groups also define namespaces; this means that within a group, all the variables, types, and sub-groups must have unique names. In addition, groups define the scope of a dimension: as noted above, a dimension declared in one group is visible within all the sub-groups below it.
- **Dimensions** are used to define array variables that are used to store the bulk of the data in a product. Dimensions declared at the level of a particular group are visible within all the groups nested within it.
- **Variables** have a name, a 'shape' – a rank defined by their dimensions –, and a data type. Within the Sentinel-6 project we only use the predefined data types, such as byte, int, short, double, string, etc. The preferred use of variables is as vector or array rather than as scalar.
- **Attributes** store information about variables and groups. Every attribute is associated to a variable or group. Global attributes are considered attached to the root group. Every attribute has a name, a value, a data type, and a length. Attributes are dynamic and weakly typed: they can be declared with one value, data type, and length, and then have these changed later. Naming Conventions for Groups, Variables, Dimensions and Attributes

The convention for the standard names used in [RD- 1] is: Standard names consist of lowercase letters, digits and underscores, and begin with a letter.

This convention is also widely used as a NetCDF naming convention for non-standard names and should be adopted for the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products for the naming of groups, variables, dimensions, and attributes.

4.2.1 NetCDF Groups

The altimeter NetCDF products are organized in two layers of nested groups: one defines the posting rate of the data (140 Hz, 20 Hz or 1 Hz) or whether the data applies to the entire product and has no time dimension, the other the frequency band (Ku or C).

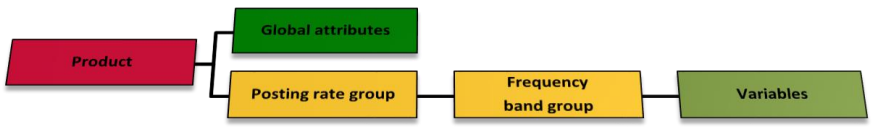

Figure 4-1 ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 product structure block diagram

The posting rate group *global* includes all the variables that apply to the entire product and have no time dimension. The posting rate group *data_140* is for the L1A products, whose posting rate is at 140 Hz. The posting rate groups *data_01* and *data_20* contain 1-Hz and 20-Hz variables that are well known in altimetry products, and they are used both in Level 1 and Level 2 products.

The frequency band groups *ku* and *c* contain the variables that apply to Ku-band and C-band data (or their corrections).

4.2.2 Dimensions

Dimensions are specified within the groups to which they are unique. If dimensions are common to multiple groups, they are defined in a layer that captures all groups in which those dimensions are used, possibly the root of the NetCDF product.

For example, in a Sentinel-6 Level 2 ALT product the timing of 20-Hz Ku-band and 20-Hz C-band is not the same, so a time dimension is defined both in the *data_20/ku* and *data_20/c* groups. At the same time, the 1-Hz Ku-band and 1-Hz C-band measurements **do** share the same time stamps, so there will be a separate time dimension in the *data_01* group, that applies to both the *data_01/ku* and dat*a_01/c* group.

4.2.3 Data Type

The data model of NetCDF-4 provides a number of pre-defined atomic data types, the so-called CDL data types listed in [AD- 4] and Table 3-1; all of them may be prefixed by “NC_” or “nf_”, which does not alter their meanings. Libraries such as the NetCDF library are responsible for translating the NetCDF-4 data model types (also called external types) into the native types of the target computer.

~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 NetCDF products make use of various of these atomic data types; non-atomic or user-defined types are *not* used in ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 NetCDF products.

4.2.4 Variable Attributes

Commonly used attributes of variables are defined in [RD- 1]:

- long_name
- unit
- add_offset (omitted when = 0)
- scale_factor (omitted when = 1)
- comment
- _FillValue (when applicable)
- standard_name (when applicable)
- source (when applicable)
- flag_values (for flags)
- flag_meanings (for flags)
- flag_masks (for flags)

4.2.5 Common Global Attributes to all Sentinel-6 Products

All Sentinel-6 products will include the following list of common metadata NetCDF Global Attributes:

Global Attributes	*Description*
Convention	*Name of the conventions followed by the dataset.* e.g. "CF-1.7"
title	*A succinct description of what is in the dataset*. "Altimeter " + "L1A " or "L1B LR " or "L1B HR " or "CAL1 L1 " or "CAL2 L1 " or "ECHO CAL " or "L2 LR " or "L2 HR " + "Near Real Time" or "Non Time Critical" or "Short Time Critical" *or* "Radiometer L2" + "Near Real Time" or "Non Time Critical" or "Short Time Critical"
institution	*Specifies where the original data was produced* "EUMETSAT"
source	*The method of production of the original data. If it was model-generated, source should name the model and its version, as specifically as could be useful.* "Processing Baseline XXXX"
history	*Provides an audit trail for modifications to the original data. Well-behaved generic NetCDF filters will automatically append their name and the parameters with which they were invoked to the global history attribute of an input NetCDF file.* "YYYY-MM-DD HH:MM:SS : Creation"
references	*Published or web-based references that describe the data or methods used to produce it. This is a configurable value and depends on the product, e.g., "Jason-CS/Sentinel-6 ALT Level 1 Product Generation Specification (901031 v4), Jason-CS/Sentinel-6 ALT Level 1 Product Format Specification (899201 v4), Jason-CS/Sentinel-6 ALT Level 1 NetCDF Dump (947129 v4)"*
product_name	*product filename as in SAFE package [AD-5]* "MMM_SS_LL_TTTTTTT_<start_time>_<end_time>_<generation_time>_<instance id>_<source>_<environment>_<class id>.<extension>"
contact	"ops@eumetsat.int"
mission_name	"Sentinel-6A" or "Sentinel-6B"
altimeter_sensor_name	"Poseidon-4A" or "Poseidon-4B"[3]
radiometer_sensor_name	"AMR-C"[4]
doris_sensor_name	"DORIS"
gnss_sensor_name	"GNSS"

[3] *Not necessary for radiometer product*

[4] *To be completed with different options after discussion with NASA/JPL. Idem for DORIS and GNSS after discussions with CNES and ESA.*

Global Attributes	Description
acq_station_name	*name of the acquisition station*[5] "FBK" for Fairbanks / "KIR" for Kiruna / "SVL" for Svalbard or "TRO" for Tromsø
netcdf_version	*NetCDF version including release date*[6] e.g. "4.5.0 of Dec 25 2017 16:13:17"
cycle_number	*cycle number [int]*
pass_number	*pass number in the cycle [int]*
absolute_rev_number	*absolute orbit number [int]*
first_measurement_time	*UTC time stamp of the first valid measurement* "YYYY-MM-DDTHH:MM:SS.ssssssZ"
last_measurement_time	*UTC time stamp of the last valid measurement* "YYYY-MM-DDTHH:MM:SS.ssssssZ"
first_measurement_latitude	*latitude of the first valid measurement [double]*
last_measurement_latitude	*latitude of the last valid measurement [double]*
first_measurement_longitude	*longitude of the first valid measurement [double]*
last_measurement_longitude	*longitude of the last valid measurement [double]*
equator_time	*UTC time of latest equator crossing* "YYYY-MM-DDTHH:MM:SS.ssssssZ"
equator_longitude	*longitude of latest equator crossing [double]*
ellipsoid_flattening	*flattening coefficient of the reference ellipsoid [double]*
ellipsoid_semi_major_axis	*semi major axis of the reference ellipsoid [double]*

4.3 Time in a ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Product

In the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 NetCDF products time may be encoded as either:

- An ASCII character string
- A real data type

Unless otherwise specified, or required to meet operational needs, time shall be encoded as a real data type.

4.3.1 Time as a Character String Data Type

When time is encoded as a character string in a product it shall be encoded in an NC_STRING data type in date and time format [RD- 1].

The format is:

*YYYY-MM-DD***T***hh*:*mm*:*ss.ssssss***Z**

Where:

[5] *This is the full list of potential acquisition stations. The final list is likely to only include a few. The final list will be available at the closing of the Sentinel-6/Jason-CS station service contract ITT (foreseen end of 2018).*

[6] *Information provided by the NetCDF library nc_inq_libvers. An example is given.*

YYYY	20[0-9][0-9]	year
MM	[0-1][0-9]	month
DD	[0-3][0-9]	day of month
hh	[0-2][0-9]	hour
mm	[0-5][0-9]	minute
ss	[0-5][0-9]	second
ssssss	[0-9][0-9][0-9][0-9][0-9][0-9]	microseconds

For example, 21:19:27.099 on 27 July 2011 would be encoded as:

2011-07-27T21:19:27.099000Z

4.3.2 Time as Real Data Type

Encoding time as a character string is not memory-efficient.

When time is encoded as a real or integer data type in a ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 product, it *shall* be encoded as a numeric value or an array of numeric values with associated attributes that specify the units, the reference time, and the precision at which the data is stored.

When time is encoded as a real data type in a ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 product, it shall be as an NC_DOUBLE data type with units of seconds and a reference time of "2000-01-01 00:00:00.000" and a given precision unless a different reference time is required for reasons of enhanced precision or precision over a longer time period. In any case the reference time shall be provided within the attribute "units" associated with the respective time variable, e.g. "seconds since 2000-01-01 00:00:00".

This standard encoding is the default encoding for all times unless another encoding is specified (see below).

When expressed as a NC_DOUBLE with a reference of "seconds since 2000-01-01 00:00:00", the time variable can be expressed with 15 significant figures which provide a precision of 1 microsecond until approximately the year 2285.

5 PRODUCT PACKAGE OVERVIEW

SAFE format is designed as an Archival Information Package (AIP) aiming at preserving the data from a specific EO space mission and instrument.

Different objects specified following the preserved data set composition recommended by the Long Term Data Preservation (LTDP) [7] guidelines define the content information of the AIP.

This section describes the high-level format and structure of the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products. The term *product* is used to refer to the AIP.

The format of the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products is derived from the XML Formatted Data Unit (XFDU) by using a specialisation process that maximises the reuse of the namespaces and schemas as defined in the SAFE format.

The XFDU is the CCSDS standard for packaging data and metadata, into a single package called "Package Interchange File".

Further information on the XFDU and SAFE conformance classes is available at http://earth.esa.int/SAFE/basicinfo.html.

5.1 High Level Product Package Data Structure

The purpose of this section is to describe the high-level product structure of the ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products without requiring an extensive knowledge of the XML Schema nor NetCDF format details.

The different high-level physical structures of SAFE-SENTINEL product package shall be:

- ***EO Product Package*** - for ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 products. This shall include:
  - *EO Data product*
  - *EO Data product metadata (manifest)*
  - *EO Data representation information (for L0 only)*
- ***EO Auxiliary Package*** – for ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 auxiliary data files. This shall include:
  - *EO Auxiliary data*
  - *EO Auxiliary metadata*
  - *EO Auxiliary representation information*

Each package is composed of:

- *Measurement data file* (EO data product or auxiliary data)
- *Metadata file*, also referred hereafter as manifest
- *Representation information data* to detail binary files structure

[7] *https://www.earthobservations.org/documents/se/130_GEO_ltdp_guidelines.pdf*

The *measurement* data file can be of any format (this is applicable to Level 0 products, and auxiliary data; Level 1 products must be NetCDF; Level 2 products shall be NetCDF and optionally BUFR).

The *metadata* or manifest file can be split into three different parts at higher level:

- The *Information Package Map* contains the logical view of the package. It provides a hierarchical view of the content of the XFDU using a series of nested Content Unit elements. Content Units contain pointers to data objects and to the metadata associated with those data objects.
- The *Metadata Section* records all of the metadata for all items in the package. The ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 Product packages can contain Referenced Metadata (that points to a representation data file), Wrapped Metadata (information included directly in the manifest file) and data object pointer (that points to an external annotation data file). Regarding Wrapped Metadata, two classes of metadata have been individuated:
    - *Primary Metadata*: metadata common to all products, regardless of the instrument or the level of processing;
    - *Secondary Metadata*: metadata specific for each instrument and/or level of processing.
- The *Data Object Section* contains all the physical information needed to get the location of each file composing the package.

The representation data is meant to describe the structure of a measurement data file that has not been implemented in a well-known standard format.

5.2 Product Components

5.2.1 Manifest File

The manifest of any ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 product shall comply with [AD-3] and contains information relevant to the product data composition. It describes all the physical components and interconnects all objects of the product. It can be considered as the map of the product and there is one manifest for each product.

The manifest file is an XML file that provides all the relevant product information for cataloguing and archiving purpose. Implementation of the Sentinel-SAFE format [AD- 1] is handled via XML schemas that are included in the specific instrument's volumes. In this volume the manifest content is represented with tables that specify the sequential order of the data and their definition. The table describing the entire content of the Manifest file have four columns described below:

- *Name*: name of the Element or of the Object in the Manifest file, i.e. tag name in XML format
- *Description*: Free text that defines the Element or the Object
- *Type*: Data Element type
- *Occurrence*: Number of Elements or Objects

The Manifest file contains XFDU objects; the table described in this paragraph are used to describe those objects.

The structure of the Manifest element is given in Table 5-1.

Table 5-1: XFDU manifest main structure

Name	Description	Data Type	Reference	Occ
<?xml version="" encoding="">	Attribute containing the relative path for the xfdu.xsd XML schema corresponding to the product	string	-	1
<informationPackageMap>	Contains a high-level textual description of the product and references to all product components	informationPackageMapType	-	1
<metadataSection>	Contains the product Metadata	metadataSectionType	Section 6	1
<dataObjectSection>	Contains references to the physical information needed to get the location of each file composing the package	dataObjectSectionType	Section 6	1

Detailed information of the manifest contents within each of the structures is given in [AD- 3].

5.2.2 Measurement Data File

The *measurement data file* formats used in the context of ~~Sentinel-6/Jason-CS~~Jason-CS/Sentinel-6 are:

- EO data product
  - L0 products: file in pure binary format containing the ISPs
  - L1 products: NetCDF4 format[8].
  - L2 products: NetCDF4 format and optionally BUFR (for NRT products only).
- EO auxiliary product
  - Auxiliary data files described in [RD- 7 to RD- 9].

5.2.3 Representation Data and Annotation Data

The Representation Data Files are XML schemas describing the measurement component. L0 products or/and auxiliary files require of such information. The contents of the Representation Data File are described in the Level 0 Product Format Specification [RD- 5] and within the auxiliary data format specification documentation [RD- 7 to RD- 9].

ISP Annotation data shall also be available for Level 0.

[8] *The PDP can be configured to use NetCDF internal compression.*

6 METADATA XML DESCRIPTION

To avoid duplication, XML schemas are kept in separate document delivered along with the Generic Product Format Specification documents [AD- 3].

Sentinel-6 products XML Metadata (manifest) shall be compliant with ISO-19115 following the SRD requirement as in [AD- 6]:

R-S-04900 *Metadata shall be appended to archived data and products in line with [ISO-19115], to support the search for and selection of items*.

7 FORMAT CONVENTIONS

Products to be encoded with various formats (e.g., NetCDF, BUFR, GRIB, etc) shall conform to the extent possible to the relevant agreements, standards, and conventions as well as to the provisions applicable to them in this document.