

EPS-SG Generic Product Format Specification (GPFS)

Doc.No. : EUM/LEO-EPSSG/SPE/13/702108
Issue : v5 e-signed
Date : 27 May 2024
WBS : LEO-EPSSG-925010

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

Signature Table

	<i>Name</i>	<i>Function</i>	<i>Signature</i>	<i>Date</i>
Book-captain:	Carlos Mas	EPS SG Instrument Functional Chain Engineer		
Reviewed by:	Michela Sunda	EPS-SG Instrument Chain Team Leader		
Reviewed by:	Yago Andres	EPS-SG GS Technical Manager		
Reviewed by:	Georg Riker	EPS-SG Configuration Management Engineer		
Reviewed by:	Esteban Tomas	LEO Quality & Product Assurance Manager		
Approved by:	Davide Castellazzi	EPS-SG System Manager		

Distribution List

<i>Name</i>	<i>Organisation</i>
EPS-SG Team (mailing list)	EUMETSAT

Change Record

Version	Date of Version	Document Change Request (DCR) Number	Description of changes
<i>as on profile</i>		<i>if applicable</i>	
V1 Draft	05/12/2013	N/A	Initial version
V1C Draft	30/07/2014	N/A	Initial version following internal reviews
V1E	01/12/2014	N/A	Update following further internal review
V1H	17/11/2015	EUM/LEO-EPSSG/DCN/15/834841	Included updated schema file EPS-SG-GPFS.xsd with targetNamespace Section 1.7: Added acronyms Section 2.1: Added applicability of filename convention to other formats Section 3.2: Added type identifiers, added free_text field to filename, changed environment identifiers, changed RO identifiers Section 3.2.2: new section defining Product ID Added new section 4.3.3 on presentation of integer constants Section 4.3.2: added reference to CDL data types defined in [netCDFUsersGuide] Section 5: Clarified the mandatory structure of product formats and changed group names to lower case Section 5.2.3: Added type identifiers, changed environment identifiers, changed instrument attribute Section 8: updated schema and XML files included Section 3.2: - added description for productID - changed length of free description from 16 to 15
V2.0	18/07/2016	EUM-EPSSG-DCR-302	General: defined CF units for time variables and dimensions length column added. Section 1.5; added applicable document [BIPM] Section 1.6: Update of link to [CF] and [NACDD] Section 1.7: updated list of acronyms Section 3.2: changed product generation time to start of generation

Version	Date of Version	Document Change Request (DCR) Number	Description of changes
<i>as on profile</i>		<i>if applicable</i>	
			Section 3.2: - Updated text with “.nc” extension. - Added GHG, NAC and TWV product types. Section 4: added applicability of [BIPM] Section 4.3.1: text update to allow lower case. Section 4.3.3: error correction: Hx11 -> HxB Section 4.3.4.1: updated user precision specification text. Section 4.3.4.2: removed precision references and updated standard_name. Section 4.3.4.3: section on Time as Integer Data deleted. Section 5.1: updated title. Section 5.1.1: - Removed subgroup quality statistics in Figure 1 and sentence allowing empty groups. - Corrected typo in caption. Section 5.1.2: amended instrument status group description Section 5.2.3, Table 6: - Updated convention -> Convention, organisation -> institution and disposition mode. - Removed baseline and repeated type table. - updated comments of instrument param. - Added Doi, “absolute orbit” to orbit numbers and “with ms precision” to sensing times. Section 5.2.6.1, deleted Data from title and update of Table 8: - Deleted leap second, orbit and state vector dimensions. Section 5.2.6.2, deleted Data from title, text added and update of Table 9: - Added reference frame. - Deleted epoch_time_ut1 and tolerances. - Dimensions and comments updated. Section 5.2.7: added note on instrument modes definition.

Version	Date of Version	Document Change Request (DCR) Number	Description of changes
<i>as on profile</i>		<i>if applicable</i>	
			Section 5.2.8, added Table 13 with sub-group variables and update of Table 12 attributes: - Added baseline, format and atbd version. - Deleted generating facility. Section 5.2.9: Update of Table 14 and Table 16: - Changed parameters from time to duration. Section 8.1: added updated XML schema Section 8.2: added updated XML file Table 15: updated gap information definition
V2A	21/11/2016	EUM-EPSSG-DCR-451	General: minor formatting and text improvements. Section 1.3.1: corrected examples for enumerated and grouping Section 1.5: - updated link [netCDFUsersGuide] - added L0 PFS reference document Section 1.6: updated acronyms list Section 5.1.1: updated caption of Figure 1 Section 5.2.6.1: improved text comment for manoeuvre_items dimension Section 5.2.6.2: removed annotation for the manoeuvre type in Table 9 (manoeuvre_occurrence) and closed related open point. Section 5.2.8: clarified the scalar nature of the source attribute Section 5.2.9: - Table 14 updated comments to clarify applicability for specific L1/L2 attributes: - degraded_manoeuvre - degraded_instrument - Table 15 added quality Dimensions and updated description - Table 16 improved notes description for gap_start/end_time_utc variables Section 6: corrected XML schema file name Section 7: record format version updated Section 8 changed to Appendix A: - added updated XML schema - added updated XML format description file with a generic product example and clarifying text. Section 9 changed to Appendix B.
V3	11/07/2017	EUM-EPSSG-DCR-693	Section 1.6: added MoD in list of acronyms, removed TBC, TBD, TBW Section 5.1.1: added a note clarifying that the root group does not explicitly appear in the XML file. Section 5.2.6.1: changed possible range of manoeuvre_items

Version	Date of Version	Document Change Request (DCR) Number	Description of changes
	<i>as on profile</i>	<i>if applicable</i>	
			Section 5.2.6.2: changed possible values of manoeuvre_occurrence and added not that the manoeuvre variables appear only in case of manoeuvres Section 5.2.8: removed processor_full_name from source attribute, added possible names for processor_name Section 5.2.9: changed type of overall_quality_flag to NC-USHORT, removed other quality flags from table 14. Added explanation on setting of overall_quality_flag, and changed sub-sequent tables to refer to the overall_quality_flag Section 7: added new format version to Table 17 Appendix A1: updated XSD schema according to document update Appendix A2: updated XML file according to document update, and clarified this to be an example Appendix C: List of TBC and TBD: removed
V3A	27/11/2017	EUM-EPSSG-DCR-762	Section 1.5: Updated reference and link to CF conventions Section 4.3.4.2: Removed reference to CDL notation Appendix A1: Inclusion of new XSD schema file, aligned with XML format description. Appendix A2: removed notion of product example and clarified that the XML file describes the product format
V3B	22/01/2018	EUM-EPSSG-DCR-820	Section 7: Update of product format version control numbers Annex A2: Corrections and changes to XML file removal of version and date attributes change of valid_mean to -1.e-9 for all time variables correction of valid range for manoeuvre variable typo correction of missing_value of eccentricity
V3C	12/02/2019	EUM-EPSSG-DCR-1131	Section 7: Update of product format version control numbers Annex A.1: Annexed revised XML schema
V3D	07/01/2020	EUM-EPSSG-DCR-1515	Section 3.2: Change of file name location indicator to enforce capitals in part of it Update of signature table Section 7, Table 18: Update of version number
V4	22/09/2021	EUM-EPSSG-DCR-1830 EUM-EPSSG-DCR-2103	Section 3.2, Table 5, disposition_mode: addition of identifiers indicating last product of a dump, type: addition of new product types CHK and REP Section 7, Table 18: Update of version number Section 1.4: Addition of WMO Manual on GTS as applicable document

Version	Date of Version	Document Change Request (DCR) Number	Description of changes
<i>as on profile</i>		<i>if applicable</i>	
			Section 1.5: Removal of WMO Manual on GTS as reference document Section 3.2, Table 5, type: Introduction of new product type SMP, correction of freeformat field and its length Addition of Section 6: GTS Packaging Appendix A2: Changed attributes valid_min, valid_max of semi major axis, eccentricity and inclination in XML format file.
V5	27/05/2024	EUM-EPSSG-DCR-2782	Section 5.2.6: Introduced text to allow free definition of the range for angular variables at individual PFS level. Section 3.2: updated description of free text. Product Format Version Control updated. Removed V3E from the table as it was not published.

Table of Contents

Introduction	10
1.1 Purpose and Scope	10
1.2 Document Structure.....	10
1.3 Conventions and Terminology used in this Document.....	11
1.3.1 Conventions.....	11
1.3.2 Terminology	12
1.4 Applicable Documents.....	12
1.5 Reference Documents.....	13
1.6 Acronyms.....	14
2 EPS-SG Products Content	15
2.1 Overview.....	15
2.2 Global and Regional Products.....	15
2.3 Aggregation of Products and Aggregated Products.....	16
3 Generic Specifications	17
3.1 Introduction.....	17
3.2 Product File Naming Convention.....	17
3.2.1 EPS-SG Data Designators	23
3.2.2 Product ID.....	23
4 NetCDF and SI Conventions	24
4.1 Guiding Principles.....	24
4.2 The NetCDF-4 Enhanced Data Model	24
4.3 NetCDF Conventions and Best Practices	25
4.3.1 Naming Conventions for Variables, Groups, Dimensions and Attributes.....	25
4.3.2 Data Types.....	25
4.3.2.1 Common User-Defined Compound Data Types	26
4.3.3 Presentation of Integer Constants	26
4.3.4 Time in an EPS-SG Product.....	26
4.3.4.1 Time as a Character String Data Type.....	26
4.3.4.2 Time as a Real Data Type.....	27
4.3.5 Time as a Logical Compound.....	28
5 NetCDF Product Structure	29
5.1 Product groups in EPS-SG.....	29
5.1.1 root.....	29
5.1.2 Status.....	30
5.1.3 Data	31
5.1.4 Quality.....	31
5.2 Elements Generic to EPS-SG Products	31
5.2.1 Global Dimensions.....	31
5.2.2 Global Variables	31
5.2.3 Global Attributes	31
5.2.4 Groups	34
5.2.5 Common Attributes	34
5.2.6 satellite status group common to all EPS-SG Products	34
5.2.6.1 satellite status Dimensions.....	34
5.2.6.2 satellite status Variables.....	34
5.2.7 instrument status group common to all EPS-SG Products.....	37
5.2.8 processing status group common to all EPS-SG products	38
5.2.9 quality group	39
5.3 Missing Data.....	41
6 GTS Packaging.....	41
7 Specific Product Format Specifications (PFS)	42
8 Product Format Version Control.....	43

Appendix A	Product Format XML Description	45
Appendix B	Format Conventions.....	46

INTRODUCTION

1.1 Purpose and Scope

This is the Generic Product Format Specification (GPFS) for the products made by the EPS-SG Payload Data Processing (PDP) Facilities, including products owned by the SAF and generated centrally, and the various test tools. It specifies the aspects that are common to all of the EPS-SG Products. This document addresses the native format of the products generated in the EPS-SG ground segment. User formats such as BUFR, HDF, or others will be specified elsewhere.

The text in this document shall be considered as specifying requirements applicable to all EPS-SG products in native format.

The document may be updated to reflect common fields identified during the development of individual product format specifications. It is complemented by the EPS-SG Product Format Specifications of Level 0, Level 1, and Level 2 products from each of the product processing facilities in the EPS-SG system.

1.2 Document Structure

Section 1 is this introduction, which also addresses conventions and terminology used in the rest of the document and lists applicable and reference documents.

Section 2 introduces EPS-SG Level 0, Level 1, and Level 2 products and some of the issues with regard to their formatting. It also addresses the issue of granularity of products.

Section 3 sets out conventions that are applicable to all EPS-SG products. This includes the naming convention used for these products.

Section 4 sets out the guiding principles behind the use of netCDF-4 in EPS-SG products, and describes the encoding of time in the products.

Section 5 describes the NetCDF product structure and the elements generic to all products.

Section 6 sets out the guiding principles behind the use of any formats used to encode EPS-SG products.

Section 7 provides guidance with respect to the product format version control.

Section 8 describes the generic parts of the products in XML.

Section 9 addresses format conventions.

1.3 Conventions and Terminology used in this Document

1.3.1 Conventions

Table 1: Conventions used in this document

Operation	Meta character(s)	Explanation	Example
Enumerated		A vertical bar separates alternatives	“IAS” “VII” “RO” “MWS” “SCA” “3MI” “MWI” “ICI” “SN5” designates one of “IAS”, “VII”, “RO”, “MWS”, “SCA”, “3MI”, “MWI”, “ICI”, “S5”
Grouping	()	Parentheses define scope and precedence	SGA1_SND_(IASx MWSx)_E_2 designates either SGA1_SND_IASx_E_2 or SGA1_SND_MWSx_E_2
Character class	[]	Square brackets designate a set of possible characters to choose from	[1-3] matches 1, 2, or 3
Quantifiers	{min,max}	Curly brackets designate the minimum and maximum number of repetitions of the preceding group	(abcd){2,3} designates abcdabcd, or abcdabcdabcd
	?	The question mark indicates <i>zero or one</i> of the preceding element	(a)(x)? Designates a or ax
	*	The asterisk indicates <i>zero or more</i> of the preceding element	(a)(x)* designates a, ax, axx, axxx, ...
	+	The plus sign indicates <i>one or more</i> of the preceding element.	(a)(x)+ designates ax, axx, axxx, ...
String	“”	Exact string	“Unidata Dataset Discovery v1.0” designates the string within the matched quotes
Character	‘’	Exact character	‘,’ designates the comma character

Operation	Meta character(s)	Explanation	Example
Range	< > ≤ ≥	The left and right angle brackets indicate “greater than” or “less than”, respectively. The other two symbols represent “greater than or equal to” and “less than or equal to”, respectively. These symbols may be used to indicate a valid range for a variable.	If X is an integer, then: $0 < X < 500$ designates a range for X of any integer value from 1 to 499. $0 \leq X \leq 500$ designates a range for X of any integer value from 0 to 500. $0 \leq X$ designates a range for X of any positive integer or 0. $X < 0$ designates a range for X of any negative integer If Y is a real value, then: $0 \leq X \leq 500$ designates a range for X of any value between 0 and 500.

1.3.2 Terminology

Term	Explanation
Product	A product is generated by a product generation function in the Ground Segment.
Product Granule	A product whose size is optimised to meet the NRT timeliness requirements for its generation and dissemination to the users.
Field	A field contains a data type or an array of data types
Data Type	A data type is a way to encode information following a given data model (description of objects represented by a computer system)

1.4 Applicable Documents

Table 2: Applicable documents

ID	Reference Number	Title
[CONV]	EUM/PEPS/STD/09/0160	EPS-SG Mission Conventions and Standards Document
[netCDFUsersGuide]	http://www.unidata.ucar.edu/software/netcdf/docs/user_guide.html	The netCDF Users' Guide, netCDF version 4.1.3, 10 June 2011
[BIPM]	http://www.bipm.org/en/publications/si-brochure/	BIPM SI Brochure
[WMO-386]	https://library.wmo.int/index.php?lvl=notice_display&id=21811	Manual on the Global Telecommunication System: Annex III to the WMO Technical Regulations (2015 edition, updated in 2020)

1.5 Reference Documents

Table 3: Reference documents

ID	Reference Number	Title
[CF]	http://cfconventions.org/Data/cf-conventions/cf-conventions-1.7/cf-conventions.pdf	netCDF Climate and Forecast (CF) Metadata Conventions: Version 1.7, 18/07/2017
[L0PFS]	EPS-SG L0 Product Format Specification	EUM/LEO-EPSSG/SPE/13/703928
[NACDD]	http://wiki.esipfed.org/index.php?title=Category:Attribute_Conventions_Dataset_Discovery	NetCDF Attribute Convention for Dataset Discovery
[OGSRD]	EUM/LEO-EPSSG/REQ/13/725156	EPS-SG Overall Ground Segment Requirements Document
[SRD]	EUM/LEO-EPSSG/SPE/13/735903	EPS-SG System Requirements Document (SRD)
[WMO-306]	http://www.wmo.int/pages/prog/www/WMOCodes.html	WMO Manual on Codes, Common Table C. 2012 Edition.
[NcML]	http://www.unidata.ucar.edu/software/netcdf/ncml/v2.2/AnnotatedSchema.html	Annotated Schema for NcML-2.2

1.6 Acronyms

Acronym	Meaning
AOI	Area Of Interest
ASCII	American Standard Code for Information Interchange
BIPM	Bureau International des Poids et Mesures
BLO	Binary Large Object, also known as BLObs of data
BUFR	Binary Universal Form Representation
CDL	Common data form Description Language
CF	Climate and Forecast metadata convention
DCS	Data Collection System
EPS-SG	EUMETSAT Polar System – Second Generation
GPFS	Generic Product Format Specification
GRIB	GRIdded Binary
GS	Ground Segment
HDF	Hierarchical Data Format
IAS	Infra-red Atmospheric Sounding
ICI	Ice Cloud Imaging
IR	Infra-Red
ISP	Instrument Source Packet
JPEG	Joint Photographic Experts Group
L0	Level 0
L1	Level 1
L2	Level 2
MoD	Mean of Date
MW	Micro-Wave
MWI	Micro-Wave Imaging
MWS	Micro-Wave Sounding
NcML	netCDF Markup Language
netCDF	Network Common Data Form
NIR	Near Infra-Red
NRT	Near Real Time
OSI SAF	SAF on Ocean and Sea Ice
O3M SAF	SAF on Ozone Monitoring and Atmospheric Chemistry
PDP	Payload Data Processing
PGS	Product Generation Specification
PFS	Product Format Specification
PNG	Portable Network Graphics
RO	Radio Occultation
SAF	Satellite Application Facility
SCA	Scatterometry
SI	International System of Units
SN5	Sentinel 5
SWR	Short-Wave infra-red Radiance
UTC	Universal Time Coordinated
UVR	Ultraviolet Visible Radiance
VII	Visible Infra-red Imaging
WMO	World Meteorological Organization
XML	eXtensible Markup Language
XSD	XML Schema Definition
3MI	Multi-viewing Multi-channel Multi-polarisation Imaging

2 EPS-SG PRODUCTS CONTENT

2.1 Overview

All EPS-SG products are structured using the enhanced netCDF-4 data model.

In the L0 products, which are netCDF-4 datasets that make no bit-level alterations whatsoever to the instrument/spacecraft data delivered, the down-linked instrument source packets or telemetry packets are stored as variable-length byte arrays.

At higher product levels, besides being provided in the netCDF-4 as native format, the EPS-SG products to be delivered to users are formatted using a number of different encodings, for example netCDF-4, GRIB-2, HDF-5, BUFR, JPEG, PNG. The list of possible product formats to be provided to users will be decided and described elsewhere. This document describes only the native format. The filename conventions described in this document are applicable to other formats as well, subject to change of the filename extensions (“type” and “compression” as laid out in [WMO-386]).

2.2 Global and Regional Products

Each EPS-SG product type is generated at the EUMETSAT Headquarters in NRT with *mission type* “global” or “regional”, respectively, in the context of the EPS-SG global and regional missions.

Global products are in addition archived in the EUMETSAT Data Centre.

Each global product does not include a full orbit of data, and equivalently each regional product does not cover the complete regional Area of Interest (AOI). To cope with the timeliness constraints for dissemination of global and regional L1/L2 products to the users, the EPS-SG Ground Segment will generate and disseminate both global and regional products in the form of *product granules* of limited duration/size. A product granule is a product whose size has been optimised to meet the NRT timeliness requirements for its generation and dissemination to the users.

Therefore, multiple global *product granules* will be required to cover an entire orbit, and multiple regional *product granules* from different orbits will be required to cover the complete regional AOI.

The duration of *product granules* may also differ for global and regional products of the same mission (e.g. for the same instrument), but also for different product types (for different instruments or type of data e.g. NAVATT and HKTM). The size of *product granules* will be defined during the design phase of the Ground Segment.

Global and Regional *product granules* of the same *type* (e.g. Global VII L0 products and Regional VII L0 products) may therefore also differ for their duration. The duration of L0 *product granules* of a given type is expected to be fixed and not varying along the orbit. See [SRD] for the definition of “Global” and “Regional”.

2.3 Aggregation of Products and Aggregated Products

Aggregation may occur at several points in the EPS-SG ground segment and its product processing functions. The structure of the products, however, independent of size, shall conform to the requirements in this document and the applicable product format specification document.

3 GENERIC SPECIFICATIONS

3.1 Introduction

This section defines conventions that are applicable to all EPS-SG products in native format.

3.2 Product File Naming Convention

The product name is designed to be embedded in the name of the file containing the product.

The name of the file will follow the WMO file naming convention [WMO-386].

The “type” and “compression” parts listed in that convention will not be part of the product file name: they will rather be part of the filename.

The “type” part is only “nc”. The accepted “compression” values are set out in Table 4 reproduced from [WMO-386] Attachment II-15.

Table 4: Accepted compression values and methods

Compression	Meaning
Z	The file has been compressed using the Unix COMPRESS technique
zip	The file has been compressed using the PKWare zip technique
gz	The file has been compressed using the gzip technique
bz2	The file has been compressed using the bzip2 technique

The product file name will thus be as follows:

**(pflag) ‘_’ (productidentifier) ‘_’ (oflag) ‘_’ (originator) ‘_’
(YYYYMMDDhhmmss) ‘_’ (freeformat)**

Where freeformat contains a number of product name fields separated by the underscore symbol “_”.

The order of the fields is mandatory.

The contents of the fields are as follows:

Table 5: Product file name fields

Product Name Field	Description	Size in Characters
pflag	“W” <i>Note: “W” stands for “WMO Product Identifier”</i>	1
productidentifier	(locationindicator) ‘,’ (datadesignator) ‘,’ (freedescription)	41
locationindicator	(country)’-(organisation)’-(location) where (country) is the ISO 3166 standard 2 letter code or “XX” for international organisations (organisation) is the name of the organisation (location) is the production centre location Examples: “XX-EUMETSAT-Darmstadt” for products generated at EUMETSAT central facility “fr-meteofrance-lannio” for products generated by the OSI SAF at Météo France in Lannion Note: Capitals have to be used in the locationindicator for XX-EUMETSAT and first letter of Darmstadt	21
datadesignator	“SAT” Table C-13 of [WMO-306]	3
freedescription	(spacecraft) ‘-’ (productID)’	Variable (maximum 15 chars)
spacecraft	Metop-SG-A satellites: (“SGA”)[1-3] Metop-SG-B satellites: (“SGB”)[1-3]	4
productID	(instrument)’-(processing_level)’-(type)	10

instrument	“IAS” “MWS” “VII” “SCA” ”3MI” ”MWI” ”RO_” ”ICI” ”SN5” “DCS” “MSP” “NAV” “HKT” Instruments: IAS = Infra-red Atmospheric Sounder MWS = Micro-Wave Sounder VII = Visible Infra-red Imager SCA = Scatterometer 3MI = Multi-viewing Multi-channel Multi-polarisation Imager RO_ = Radio Occultation receiver MWI = Micro-Wave Imager ICI = Ice Cloud Imager SN5 = Sentinel 5 DCS = Advanced Data Collection System Other: MSP=Multi-Sensor Product (single main instrument cannot be assigned) NAV = NAVATT L0 data HKT=House-Keeping Telemetry data	3								
processing_level	(“00” ”1A” ”1B” ”1C” ”1D” ”1F” ”02”) “00” = Raw data/ source packets “1A” = for SCA: internal product “1B” = Calibrated and geolocated science data “1C” = for IAS: calibrated, apodised and geolocated radiance spectra with VII imagery added; for 3MI: calibrated and geolocated Stokes vectors “1D” = for IAS: principal component scores of IAS level 1C radiance spectra “1F” = Fundamental Climate Data Records “02” = geolocated geophysical variables	2								
type	Character string providing an indication of the type of product: <i>Examples (not exhaustive list) of Product Type ID:</i> <table><tr><th>type</th><th>Meaning</th></tr><tr><td>AER</td><td>aerosol</td></tr><tr><td>ALH</td><td>aerosol layer height</td></tr><tr><td>AMV</td><td>atmospheric motion vector</td></tr></table>	type	Meaning	AER	aerosol	ALH	aerosol layer height	AMV	atmospheric motion vector	3
type	Meaning									
AER	aerosol									
ALH	aerosol layer height									
AMV	atmospheric motion vector									

EPS-SG Generic Product Format Specification (GPFS)

	AUI	aerosol index	
	AOD	aerosol optical thickness	
	BND	bending angle	
	CHK	IAS L0 check mode	
	CH4	methane	
	CLA	auxiliary cloud	
	CLD	cloud	
	CO_	carbon monoxide	
	CTP	cloud top pressure (O ₂ A band)	
	ENG	Engineering product	
	FDY	formaldehyde	
	GHG	green-house gases	
	GLY	glyoxal	
	GSC	GSICS statistics	
	HET	scene heterogeneity	
	ICM	Internal Cloud Mask	
	IRR	irradiance	
	LIW	Liquid/ice water paths	
	NAC	nitric acid column	
	NIR	NIR radiance	
	NO2	nitrous oxide	
	OCA	Optimal cloud analysis	
	O3_	ozone	
	O3P	ozone profile	
	PCS	Principal component scores	
	PRE	Pre-processed internal product	
	QPR	water-vapour profile	
	RAD	radiance	
	REP	IAS L0 repeated source packets	
	RFL	reflectance	
	SFC	surface variables	
	SFT	surface temperature	
	SFW	surface wind	
	SFE	surface emissivity	
	SMP	S-5 small pixel pixel product (level 0)	
	SND	Sounding product	
	SRC	source (level 0)	
	SO2	sulphur dioxide	
	SZF	sigma zero full resolution	
	SZR	sigma zero re-sampled	
	SSM	soil moisture	
	SWR	SWIR radiance	
	TPR	temperature profile	

EPS-SG Generic Product Format Specification (GPFS)

	TRA	atmospheric trace gas									
	TPW	Water-vapour column (MW)									
	TWV	temperature/water vapour profiles									
		type	Meaning								
	UVR	UV-VIS radiance									
	UV_	UV radiance									
	VER	Verification and check data (level 0)									
	WVV	Water-vapour column (VIS-NIR)									
WVI	Water-vapour column (IR)										
oflag	“C” <i>Note: Currently, “C” is the only acceptable value (as imposed by [WMO-386]); future use foresees this flag to indicate how to decode the “originator” field</i>		1								
originator	“EUMT”		4								
YYYYMMDDhhmmss	Is the UTC time of the start of the product generation, defined as the creation time of the product file <u>Abbreviated Generalised Time Format</u> , where: YYYY: 20[0-9]{2,2} : 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		14								
freeformat	(mission_type)’_’ (environment)’_’ (sensing_start)’_’ (sensing_end)’_’ (disposition_mode)’_’ (processing_mode)’_’ (free_text)		41								
mission_type	“G “ ”R “ “L” <table><tr><th>Identifier</th><th>Meaning</th></tr><tr><td>G</td><td>global product type</td></tr><tr><td>R</td><td>regional product type</td></tr><tr><td>L</td><td>local product type</td></tr></table>		Identifier	Meaning	G	global product type	R	regional product type	L	local product type	1
Identifier	Meaning										
G	global product type										
R	regional product type										
L	local product type										

environment	[“O” “V” “D” “I” “E”]	1
sensing_start	UTC Time of start of Sensing Data formatted in Abbreviated Generalised Time format (see above).	14
sensing_end	UTC Time of end of Sensing Data formatted in Abbreviated Generalised Time format (see above).	14
disposition_mode	Identification of the type of processing (“T” “t” “C” “c” “O” “o” “V” “v”) T, t = test data C, c = produced during commissioning O, o = routine operations V, v = during validation of a new processor version during routine operations The ‘mode’ of disposition is related to the suitability of the data for various kinds of uses, and hence the use that should be made of it and the destination to which it should (or should not) be sent. Lower-case letters indicate the last product of a dump.	1
processing_mode	Identification of the mode of processing (“N” “B” “R”) N = Near Real Time (NRT) B = Backlog (only for Level 0 products) R = Reprocessing	1
free_text	For arbitrary use, e.g. to discriminate product names that would otherwise be identical For RO L1 products: The GNSS satellite used for the occultation is encoded in this 3 chars (either letters, digits or symbols), e.g.: in the case of GPS satellite with PRN 20 is G20. Other GNSS systems would be marked with the letter E for Galileo, R for GLONASS, and C for COMPASS. For all the other products “T__” for products generated in TGPS “__” for all the other products	3

Example product file name (for illustrative purposes only):

**W_XX-EUMETSAT-Darmstadt,SAT,SGA1-VII-1B-
RAD_C_EUMT_20220101121212_G_O_20220101103000_20220101104000_C_N____**

This is a global L1b product, generated in the context of the EPS-SG Global mission, for the METImage instrument (VII mission) embarked on the Metop-SG/A1 satellite (SGA1).

The product was created on the 01 January 2022 at 12:12:12 hours, with a sensing start date of 01 January 2022 at 10:30:00 hours and a sensing end date of 01 January 2022 at 10:40:00 hours. The file was generated in the Ground Segment operational (O) environment during commissioning I in NRT (N) processing mode.

The corresponding product file name using netCDF formatting will carry an extension “.nc” as shown below:

**W_XX-EUMETSAT-Darmstadt,SAT,SGA1-VII-1B-
RAD_C_EUMT_20220101121212_G_O_20220101103000_20220101104000_C_N____.nc**

The file is in netCDF-4 format and is not compressed in any other way than possibly internally.

3.2.1 EPS-SG Data Designators

The field “datadesignator” specifies the type of data with reference to the categories and subcategories defined in the Common Table C-13 of [WMO-306]. The category for satellite data should always be “SATELLITE”. However, following the practice of existing EUMETSAT programmes, to limit this field length “SATELLITE” is truncated to “SAT”.

As the use of sub-categories is optional, it has been suppressed, since the full information on the satellite, instrument, product level, and type is contained in the “freedescription” field.

3.2.2 Product ID

For the sake of identifying products generically, the instrument, processing_level, and type identifiers are grouped to present a unique Product ID in the form

(instrument)’-‘(processing_level)’-‘(type)

Examples: IAS-00-SRC, VII-1B-RAD, MSP-02-LIW

4 NETCDF AND SI CONVENTIONS

4.1 Guiding Principles

All EPS-SG products are formatted following the netCDF-4 data model as described in [netCDFUsersGuide]. As noted in [netCDFUsersGuide], 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 EPS-SG 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.

EPS-SG products should where practicable conform to the [CF] convention extended where appropriate to satellite data, and should where practicable conform to the “Highly Recommended” aspects of the [NACDD]. It is known that [CF] is not fit for including all aspects of EPS-SG yet and that this convention is evolving.

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

4.2 The NetCDF-4 Enhanced Data Model

The enhanced netCDF-4 data model [netCDFUsersGuide] supports **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.

- **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 in it.
- **Variables** have a name, a ‘shape’ – a rank defined by their dimensions –, and a data type. They can be composed of one of several different predefined or user defined data types. Predefined types include – but are not limited to – byte, int, short, unsigned 64-bit int, string, etc. User-defined types include arrays of variable-length arrays, opaque Binary Large Objects (BLOBs) of bytes, or compound types (analogous to C structs).

(No matter where a type is defined in a netCDF-4 file, its definition is visible throughout it.). 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.
- **Groups** are used to organise 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.

The data in a netCDF-4-structured EPS-SG product are arranged in nested groups, as set out in the following sections.

4.3 NetCDF Conventions and Best Practices

The EPS-SG products should where practicable follow the conventions specified in the global attributes Conventions and Metadata_conventions fields (Section 5.2.3).

In general, the Conventions will be the latest version of “The Climate and Forecast (CF) Metadata Conventions” [CF] and the Metadata_conventions will be “The NetCDF Attribute Convention for Dataset Discovery” [NACDD].

4.3.1 Naming Conventions for Variables, Groups, Dimensions and Attributes

The convention for the standard names used in [CF] is:

- Standard names consist of 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 EPS-SG products.

4.3.2 Data Types

The data model of netCDF-4 provides a number of pre-defined atomic data types, the so-called CDL data types listed in [netCDFUsersGuide]; 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.

Beyond the atomic data types, other data types can be defined by users that allow for arrays of integer-string pairs, compound as a collection of one or more atomic or user-defined types. Furthermore, a collection of objects of a known size but unknown contents (so-called “blobs of data”) can be stored under the netCDF-4 data model.

4.3.2.1 Common User-Defined Compound Data Types

The netCDF-4 enhanced data model allows the definition of user-defined compound data types. Currently, there are no compound data types that are common to all products. Compound data types will be added in later versions of this document if needed.

4.3.3 Presentation of Integer Constants

By default, integer constants are to be interpreted as decimal numbers. Alternative presentations are binary, octal, or hexadecimal numbers, to be indicated by prefixes Bx, Ox, Hx, respectively. For example, the decimal number 11 can be represented as Bx1011, Ox13, HxB.

4.3.4 Time in an EPS-SG Product

Time may be encoded as either

- An ASCII character string
- A real data type
- An integer data type
- A compound data type.

Unless otherwise specified, or required to meet operational needs, time shall be encoded as a real data type (see section 4.3.4.2).

4.3.4.1 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 [CF].

The format is:

(YYYY)"-"(MM)"-"(DD)" "(hh)":(mm)":(ss)".(ddd)

where:

YYYY:	20[0-9]{2,2}	: 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
ddd:	[0-9][0-9][0-9]	: milliseconds

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

2011-07-27 21:19:27.099

Note: This allows for a precision of one millisecond and these precisions shall be specified by the user. The milliseconds, including the “.”, can be omitted if of no significance to the user.

4.3.4.2 Time as a Real Data Type

Rationale

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

When time is encoded as a real or integer data type in an EPS-SG 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.

Specification of time encoding using real data type

When time is encoded as a real data type in an EPS-SG product, it shall be as an NC_DOUBLE data type with units of seconds and a reference time of “2020-01-01 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.

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

Example in XML/NcML of standard time encoding (only relevant attributes shown for clarity):

XML/NcML:

```
<dimension name="time" length="100" />
<variable name="dtime" shape="time" type="double">
  <attribute name="standard_name" type="String" value="Time" />
  <attribute name="units" type="String"
    value="seconds since 2020-01-01 00:00:00.000" />
</variable>
```

This example shows an array of 100 times, stored in an array of NC_DOUBLE data types. The times are stored as seconds since midnight on 1 January 2020. The data stored have a precision of 1 millisecond.

Note 1: When expressed as a NC_DOUBLE with a reference of "seconds since 2020-01-01 00:00:00.000", the time variable can be expressed with 15 significant figures which provide a precision of 1 microsecond until approximately the year 2050.

Note 2: For the use with test data prior to 2020 negative times will occur.

If greater precision is required or the same precision is required beyond this period, then a logical compound shall be used (see below).

4.3.5 Time as a Logical Compound

In cases where higher precision is needed, e.g. for the Radio Occultation mission, it will be necessary to define a logical compound that includes a time reference together with a time increment. This logical compound is made up of the NC_INT carrying the days since the reference date 2020-01-01 00:00:00.000 and an NC_DOUBLE carrying the seconds elapsed since start of the day.

XML/NcML:

```
<dimension name="time" length="10" />

<variable name="jtime" shape="time" type="int">
  <attribute name="long_name" type="String" value="Days since reference date" />
  <attribute name="units" type="String" value="days since 2020-01-01" />
  <attribute name="precision" type="String" value="1 day" />
</variable>

<variable name="seconds" shape="time" type="double">
  <attribute name="long_name" type="String" value="Seconds since start of day" />
  <attribute name="units" type="String" value="seconds since start of day" />
  <attribute name="precision" type="String" value="1 ps" />
</variable>
```

This example shows an array of 10 times, stored in an array of (NC_INT, NC_DOUBLE). The times are stored as days since 1 January 2020 with a precision of 1 day and the seconds elapsed since the start of the day with a precision of one picosecond.

5 NETCDF PRODUCT STRUCTURE

The groups described below, and their contents, can appear in any EPS-SG Product. However, it is not mandatory that any product contains all of them. Only the top two levels of the hierarchy are common to all product formats.

5.1 Product groups in EPS-SG

5.1.1 root

The root level of an EPS-SG product contains three groups. In addition, it contains sets of global attributes, including search metadata, dimensions, variables, and user-defined types.

The groups separate the product data into:

- **status** conveying information about the status of the satellite, the instrument, and the processing chain. It includes also information such as On-Board Time (OBT) versus UTC correlation, orbital elements, and leap seconds.
- **data** containing the core measurements, such as source packets at level 0, radiances, backscatter coefficients or bending angles at level 1, geophysical variables and level2.
- **quality**, providing overall quality information, represented by statistics about the science data as far as not being part of the science data themselves.

The global attributes, which conform as far as is practicable to the netCDF-4 Attribute Convention for Dataset Discovery [NACDD], provide information identifying the instrument(s), product type, processing level, instrument model(s), and spacecraft.

Note: The “root” group does not explicitly appear in the XML file attached in Annex A.2, it is implicitly represented by the NetCDF file.

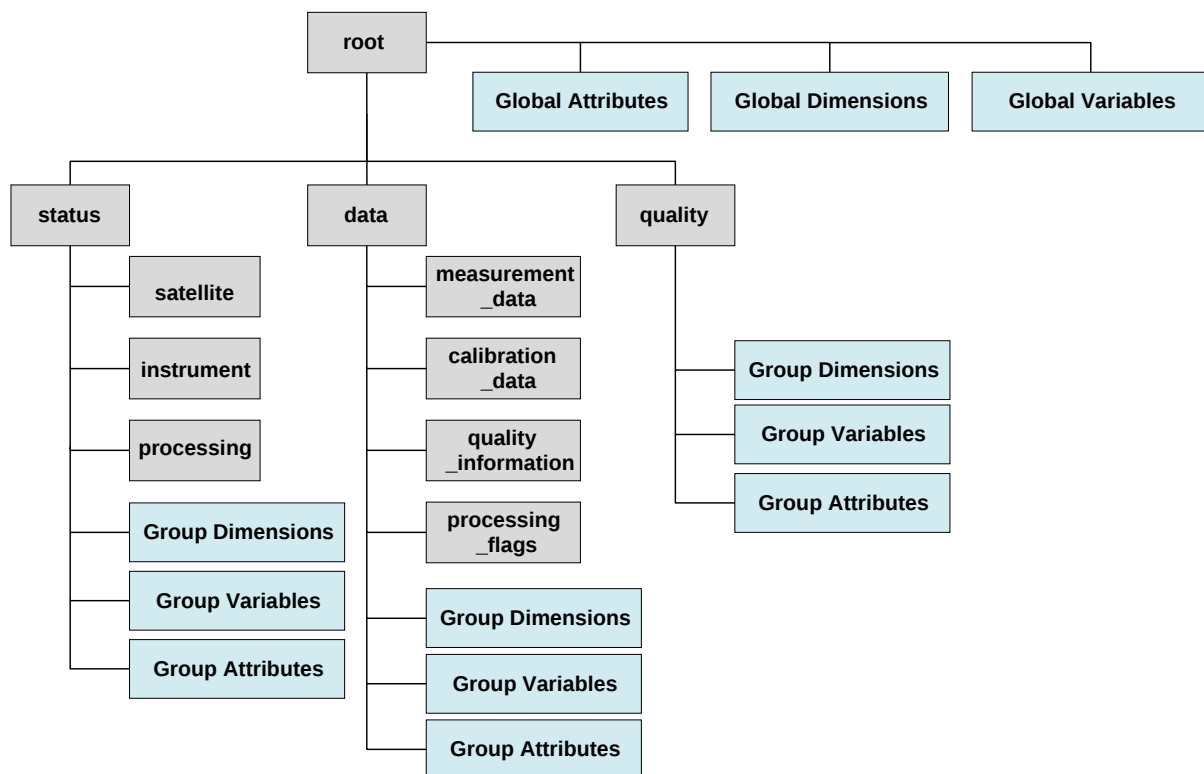


Figure 1: Sample overall structure and sections of EPS-SG products. Variables, dimensions, and attributes are shown in blue boxes. Groups are shown in grey boxes. Groups can be nested to any depth. Any group – not just the first two levels of groups shown here – can contain dimensions, variables and attributes. Groups can be omitted if not meaningful.

5.1.2 Status

The status data can in turn be divided into three groups particular to the satellite, the instrument, and the processing system.

- **satellite:** The satellite status information includes information sufficient to reconstruct the position and attitude of the spacecraft throughout the time during which the sensing data in the product were acquired, including:
 - Orbit characteristics (e.g., semi-major axis, eccentricity, inclination, argument of perigee, right ascension, mean anomaly),
 - Orbital state vector for selected points in time,
 - Spacecraft attitude for selected points in time,
 - Sub-satellite latitude and longitude for selected points in time,
 - Leap seconds.
 - Information on manoeuvres.
- **instrument:** The instrument status includes flags and information indicative of the instrument status during the sensing period, as deduced from instrument data. This sub-group is common to all products and further entries can be specified in specific PFSs,

beyond the ones given here, or removed if not meaningful (e.g. multi-mission products). In case of NAVATT or HKTM products this sub-group is to be omitted.

- **processing:** The processing status covers details of the product's generation, such as variants of retrieval methods, which are not present as global attributes.

5.1.3 Data

The data group contains the information appropriate to the processing level. At Level 0 they contain instrument or telemetry source packets. At Level 1b, for example, they might contain calibrated spectral radiances, calibration coefficients, and geolocation, viewing and sun angle geometry. At level 2, they might contain retrieved specific humidity.

If data compression is used internal to the product the corresponding specific PFS shall provide the necessary details.

Contents and sub-grouping of the data group will be specified in the specific PFS.

5.1.4 Quality

The product quality group statistically summarises the quality of the whole product. From this information the user will deduce the utility of the product in general.

Contents and sub-grouping of the quality group is partly covered in this document and can be expanded as necessary in the specific PFS.

5.2 Elements Generic to EPS-SG Products

This section specifies elements generic to EPS-SG products. First it lists those at the global (root) level. Then it lists those common elements that can occur in different levels and groups in the product, as detailed in the specific PFS.

5.2.1 Global Dimensions

No common global dimensions are currently envisaged.

5.2.2 Global Variables

No common global variables are currently envisaged.

5.2.3 Global Attributes

The following global attributes are common to all EPS-SG products:

Table 6: Global attributes for all EPS-SG products

Attribute name=	Data Type type=	Meaning and/or value
Conventions	NC_STRING	e.g. “CF-1.6”
metadata_conventions	NC_STRING	e.g. “Unidata Dataset Discovery v1.0”
product_name	NC_STRING	Product name formatted as set out in section 3.2
title	NC_STRING	Short description of the product
summary	NC_STRING	A summary as defined in the relevant product format specification.
doi	NC_STRING	Digital Object Identifier Note: used only for reprocessing campaigns and not applicable for NRT. Only applicable to L1/L2 products.
keywords	NC_STRING	As defined in the relevant product format specification.
history	NC_STRING	(“original generated product” “aggregated product” “sub-setted product”)
institution	NC_STRING	“EUMETSAT” <i>Note: This field may be extended with other values should products be generated in other locations.</i>
spacecraft	NC_STRING	Metop-SG A satellites: (“SGA”[1-3]) Metop-SG B satellites: (“SGB”[1-3])
instrument	NC_STRING	Instrument or product identifier (“IAS” “MWS” “VII” “RO_” “3MI” “MWI” “ICI” “SN5” “SCA” “MSP” “NAV” “HKT”)
product_level	NC_STRING	Product processing level (“0” “1A” “1B” “1C” “1D” “1F” “2”) “0” = Raw data/ source packets “1A” = internal SCA product “1B” = Calibrated and geolocated science data “1C” = for IAS: calibrated, apodised and geolocated radiance spectra with VII imagery added; for 3MI: calibrated and geolocated Stokes vectors “1D” = for IAS: principal component scores of IAS level 1C radiance spectra “1F” = Fundamental Climate Data Records “2” = geolocated geophysical variables

EPS-SG Generic Product Format Specification (GPFS)

Attribute name=	Data Type type=	Meaning and/or value
type	NC_STRING	Character string providing an indication of the meaning and type of product as per type in Table 5.
mission_type	NC_STRING	(“Global” “Regional” “Local”)
disposition_mode	NC_STRING	Identification of the type of processing (“Test” “Commissioning” “Operational” “Validation”) Test = Test data Commissioning = Produced during commissioning Operational = expected quality as per requirements based on fully performed validation Validation = During validation of a new processor version during routine operations The ‘mode’ of disposition is related to the suitability of the data for various kinds of uses, and hence the use that should be made of it and the destination to which it should (or should not) be sent.
sensing_start_time_utc	NC_STRING	UTC time of start of sensing data formatted in CF date and time format with ms precision.
sensing_end_time_utc	NC_STRING	UTC time of end of sensing data formatted in CF date and time format with ms precision.
environment	NC_STRING	(“Operational” “Validation” ” Integration & Verification” “Development” ”Engineering”)
references	NC_STRING	“www.eumetsat.int” <i>Note: It is intended that users of the product can access published, web-based references describing the data and the methods used to produce it at this address.</i>
orbit_start	NC_UINT	Absolute orbit number at sensing_start_time_utc
orbit_end	NC_UINT	Absolute orbit number at sensing_end_time_utc

5.2.4 Groups

The following root groups are common to all EPS-SG products

Table 7: Root groups for all EPS-SG products

Group Name
status
data
quality

5.2.5 Common Attributes

The attributes in this section are in addition to the attributes conforming to the applicable conventions, such as long_name, units, valid_min, valid_max, valid_range, missing_value, etc. as defined in [netCDFUsersGuide].

No further common attributes identified at this time.

If a product is a subset of another product, then the amount of science data is reduced. All metadata fields in a subset or superset product shall be consistent with the sub-/superset. That is, they shall be reduced/enlarged and adapted (e.g. start/end times) correspondingly.

5.2.6 satellite status group common to all EPS-SG Products

The following group “satellite” within the status group is common to all products.

5.2.6.1 satellite status Dimensions

Table 8: satellite status Dimensions

Dimension name=	Comment	Dimension length=
manoeuvre_items	Number of manoeuvres occurring between product start and end. Note: Only applicable to L1/L2 products.	"" $0 \leq N$

5.2.6.2 satellite status Variables

Orbital elements, state vector, sub-satellite points are derived either from live orbit data provided by the satellites (if used at the time of the processing) or from on-ground predicted orbit information.

Manoeuvres parameters and attitude errors are only applicable to L1/L2 products and only set if live attitude data provided by the satellites is available and used at the time of the processing.

Variables that contain angular information (i.e. the units are degrees) can be expressed according to different conventions for the range, i.e. between 0 and 360, or between -180 and

180 degrees. The convention is explicit in the product as per the variables' attributes `valid_min` and `valid_max`. The convention to be used in each individual product is left to be defined at the individual PFS.

Table 9: satellite status Variables

Variable name=	Data Type type=	Attribute name="long_name" value=	Attribute name="units" value=
Orbit Parameters shape=1			
epoch_time_utc	NC_DOUBLE	"Epoch time in UTC of the orbital elements"	"seconds since 2020-01-01 00:00:00.000"
semi_major_axis	NC_DOUBLE	"Semi major axis of the orbit at epoch time [TOD]"	"m"
eccentricity	NC_DOUBLE	"Eccentricity of the orbit at epoch time [TOD]"	""
inclination	NC_DOUBLE	"Inclination of the orbit at epoch time [TOD]"	"degrees"
perigee_argument	NC_DOUBLE	"Argument of perigee of the orbit at epoch time [TOD]"	"degrees"
right_ascension	NC_DOUBLE	"Right ascension of the orbit at epoch time [TOD]"	"degrees"
mean_anomaly	NC_DOUBLE	"Mean anomaly of the orbit at epoch time [TOD]"	"degrees"
earth_sun_distance_ratio	NC_DOUBLE	"Ratio of current Earth-Sun distance to Mean Earth-Sun distance" Note: Only applicable to L1/L2 products.	""
Location Summary shape=1			
subsat_latitude_start	NC_DOUBLE	"Latitude of sub-satellite" point at start of the product"	"degrees_north"
subsat_longitude_start	NC_DOUBLE	"Longitude of sub-satellite point at start of the product"	"degrees_east"
subsat_latitude_end	NC_DOUBLE	"Latitude of sub-satellite point at end of the product"	"degrees_north"
subsat_longitude_end	NC_DOUBLE	"Longitude of sub-satellite point at end of the product"	"degrees_east"
State Vector and Attitude Parameters shape=1			
state_vector_time_utc	NC_DOUBLE	"Time of the state vector and attitude items"	"seconds since 2020-01-01 00:00:00.000"
x_position	NC_DOUBLE	"X position of the orbital state vector [EARTH+FIXED]"	"m"

EPS-SG Generic Product Format Specification (GPFS)

y_position	NC_DOUBLE	“Y position of the orbital state vector [EARTH+FIXED]”	“m”
z_position	NC_DOUBLE	“Z position of the orbital state vector [EARTH+FIXED]”	“m”
x_velocity	NC_DOUBLE	“X velocity of the orbital state vector [EARTH+FIXED]”	“m/s”
y_velocity	NC_DOUBLE	“Y velocity of the orbital state vector [EARTH+FIXED]”	“m/s”
z_velocity	NC_DOUBLE	“Z velocity of the orbital state vector [EARTH+FIXED]”	“m/s”
yaw_error	NC_DOUBLE	“Yaw attitude error” Note: Only applicable to L1/L2 products.	“degrees”
roll_error	NC_DOUBLE	“Roll attitude error” Note: Only applicable to L1/L2 products.	“degrees”
pitch_error	NC_DOUBLE	“Pitch attitude error” Note: Only applicable to L1/L2 products.	“degrees”
Leap Second Information shape=1			
leap_second_time_utc	NC_DOUBLE	“UTC time of occurrence of a leap second in this product (if leap second occurred in the product time window); it represents the time after the leap second occurrence (i.e. midnight of day after the leap second; no leap second results in 0)”	“seconds since 2020-01-01 00:00:00”
leap_second_value	NC_SHORT	“Value of leap second in product (1, 0, or -1)” 1 = increment -1 = decrement	“s”
Manoeuvre Information (included only in case of manoeuvre) shape=manoeuvre_items			
manoeuvre_occurrence	NC_BYTE	“Occurrence of manoeuvres between start and end times of the product (1 or 2)” 1 = in-plane manoeuvre occurred 2 = out-of-plane manoeuvre occurred Note: Only applicable to L1/L2 products.	“”
manoeuvre_start_time_utc	NC_DOUBLE	“UTC time of start of manoeuvre” Note: Only applicable to L1/L2 products.	“seconds since 2020-01-01 00:00:00.000”

manoeuvre_end_time_utc	NC_DOUBLE	“UTC time of end of manoeuvre” Note: Only applicable to L1/L2 products.	“seconds since 2020-01-01 00:00:00.000”
------------------------	-----------	--	---

5.2.7 instrument status group common to all EPS-SG Products

The following group “instrument” within the status group is common to all products that are related to single instruments. In case of NAVATT or HKTM data as well as multi-mission products this sub-group is to be omitted. Further entries beyond the ones given here can be specified in specific PFSs.

Table 10: instrument status Dimensions

Dimension name=	Comment	Dimension length=
mode_items	Number of modes the instrument assumed during product duration	”” $1 \leq N$

Table 11: instrument status Variables

Variable name=	Data Type type=	Attribute name="long_name" value=	Attribute name="units" value=
Instrument Modes shape=mode_items			
mode_start_time_utc	NC_DOUBLE	"Start time of the mode"	"seconds since 2020-01-01 00:00:00.000"
mode_end_time_utc	NC_DOUBLE	"End time of the mode"	"seconds since 2020-01-01 00:00:00.000"
instrument_mode	NC_STRING	"Name of the instrument mode assumed" Note: valid modes in [L0PFS].	"

5.2.8 processing status group common to all EPS-SG products

The following group "processing" within the Status Group is common to all products. Further entries beyond the ones given here can be specified in specific PFSs.

Table 12: processing status Attributes

Attribute name=	Data Type type=	Meaning / Value
processor_name	NC_STRING	Name of the product processor
		Product Generation processor_name
		Function
		L0
		3MI L1B
		3MI L1C
		3MI L2
		MAP
		IASI-NG L1C
		IASI-NG L1D
		IASI-NG L2
		ICI L1B
		MWI L1B
		MWI-ICI L2
		MWS L1B
		MWS L2
		RO L1B
		SCA L1B
		S5 L1B
		S5 L2
		VII L1B
		VII L2
		VII L2 Cloud Mask
		VII L2_CM

EPS-SG Generic Product Format Specification (GPFS)

processor_version	NC_STRING	Version number of the processor
processing_mode	NC_STRING	(“NRT” “Backlog” “Reprocessing”) Processing mode in which the product was generated Note: “Backlog” only for Level 0.
format_version	NC_STRING	Product format version control number.
pgs_reference_and_version	NC_STRING	Reference and version of the PGS
pfs_reference_and_version	NC_STRING	Reference and version of the PFS
atbd_reference_and_version	NC_STRING	“Reference and version of the ATBD” Note: Only applicable to L1/L2 products.
baseline (optional)	NC_STRING	“Climate data record collection version in reprocessed data – optional attribute” Note: Only applicable to L1/L2 products.
source	NC_STRING	A scalar string as particularised in the relevant product format specification, containing an array of bracketed strings of the form specified as follows: (AUXILIARY_DATA_NAME)* (INPUT_PRODUCT_NAME)* where the asterisks indicate zero or more instances <i>Note 1: it is intended that users of the product can determine from the source attribute the version of the processing software and algorithm and the configuration data used to create the product, as well as the products that were inputs to its creation.</i> <i>Note 2: In some cases (e.g. RO level 1) the inclusion of a full list of auxiliary data files could be excessively long; the corresponding specific PFS shall ensure that the included information remains at a useful level.</i>

Table 13: processing status Variables

Variable name=	Data Type type=	Attribute name=“long_name” value=	Attribute name=“units” value=
Creation Time Information shape=1			
creation_time_utc	NC_DOUBLE	“UTC time of the start of the product creation”	“seconds since 2020-01-01 00:00:00”

5.2.9 quality group

The following quality group is common to all products. Further entries beyond the ones given here can be specified in specific PFSs. The variables quantifying the overall product quality

are indicative and can be altered according to the specific products as needed. Particularly, for the L0 products there is no evaluation of information inside the instrument source packets.

Table 14: quality Attributes

Attribute name=	Data Type type=	Meaning /Value
overall_quality_flag	NC_USHORT	“0” if overall quality is OK Individual bits of the flag are set to indicate degraded conditions, the first four bits are set in case of: Bit 0: Missing input product(s) Bit 1: Data gap(s) Bit 2: Corrupted input product(s) Bit 3: Instrument anomaly Bit 4: missing or degraded auxiliary data Bits 5 to 15 can be set according to specifications in individual PFSs

Further quality flags can be defined in the specific PFS, as appropriate for the product described.

Table 15: quality Dimensions

Dimension name=	Comment	Dimension length=
gap_items	Number of gaps identified during product duration. Note: it will not appear in the Product if overall_quality_flag bit 1 equals 0.	”” $1 \leq N$

Table 16: quality Variables

Variable name=	Data Type type=	Attribute name=“long_name” value=	Attribute name=“units” value=
Product Duration shape=1			
duration_of_product	NC_DOUBLE	“Entire duration of the product”	“s”
duration_of_data_present	NC_DOUBLE	“Amount of data present in the product”	“s”
duration_of_data_missing	NC_DOUBLE	“Amount of data missing in the product”	“s”
duration_of_data_degraded	NC_DOUBLE	“Amount of data degraded in product”	“s”
Gaps Information shape=gap_items			

gap_start_time_utc	NC_DOUBLE	“Gap start time in UTC” Note: will not appear in the Product if overall_quality_flag bit 1 equals 0. CF date and time format.	seconds since 2020-01-01 00:00:00.000
gap_end_time_utc	NC_DOUBLE	“Gap end time in UTC” Note: will not appear in the Product if overall_quality_flag bit 1 equals 0. CF date and time format.	seconds since 2020-01-01 00:00:00.000

Table 16 gives an indication of what information could be included to summarise the overall product quality. It might not be applicable in certain cases (e.g. RO, where the information should rather be based on the number of occultations and not on the duration of a product) and can be replaced by summaries as described in the specific PFS.

5.3 Missing Data

Following [netCDFUsersGuide] missing data in variables are to be filled with values given by the attribute *missing_value*. This attribute can be scalar or vector containing values indicating missing data. These values must be outside the valid range of the respective variable and be defined in the specific PFS.

6 GTS PACKAGING

Products will be disseminated via the WMO Global Telecommunications System, which requires packaging of BUFR products as described in [WMO-386]. The packaging in GTS structure is depicted in Figure 2.

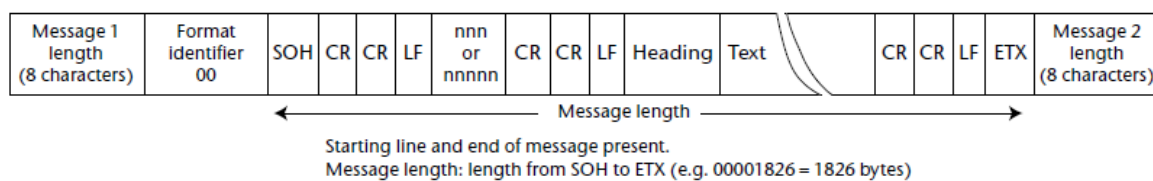


Figure 2: Structure of a GTS message

The transmission sequence number, nnn, is a three-digit group giving the transmission sequence of messages from one centre over a particular channel to the receiving centre on that channel. Numbers 000 to 999 inclusive must be used in a cyclic manner.

The field labelled Heading contains a character array of size 19, holding the abbreviated heading as defined in [WMO-386] section 2.3.2.

The field labelled Text contains the product in BUFR format.

The following simple source code in C may help explain what is needed to build the required structure as per [WMO-386] at pages 136-137:

```
/*
*****
/* write data to file */
(void)fprintf(fpOutFile, "%8.8d00\1\r\r\n000\r\r\n%s\r\r\n",
    *uliTotBUFRLength + 35, azABH);
if (1 != fwrite(acIntBuf, (size_t)*uliTotBUFRLength, 1, fpOutFile))
{
    (void)printf("Could not append bulletin %s data to output file.\n", azABH);
}
(void)fprintf(fpOutFile, "\r\r\n3");

*****
*/
```

In this snippet:

- fpOutFile is a file pointer to the output file
- uliTotBUFRLength is an unsigned long int (uint32_t) containing the length of the BUFR message, from start of “BUFR” to end of “7777”
- azABH is a char array of size 19, holding the abbreviated bulletin header (eg “ISZX01 EUMS 220930” and the null terminator)
- acIntBuf is a char (int8_t) array holding the BUFR message, from start of “BUFR” to end of “7777”

7 SPECIFIC PRODUCT FORMAT SPECIFICATIONS (PFS)

Product tables of the Specific Product Format Specifications (PFS) for each product shall be written in eXtensible Markup Language (XML), with mark-ups as defined by the XML schema definition file EPS-SG-GPFS.xsd as attached to this document (see section A.1).

8 PRODUCT FORMAT VERSION CONTROL

The product format version number should be updated whenever there is a change in the format or contents of a product that requires an update to software that has to read the product or has to check if the product is assembled correctly from the component records. This could be a change in the format itself (record field deleted, added, resized, retyped), a change in the contents of a field (e.g. scale factor change) or a change in the way that a field has to be interpreted. Any such record update requires the record format version number to be incremented. So, the updating of a record necessarily implies an updating of the format of any product that utilises the record, necessitating an update of the product format version number.

In addition, if a product no longer contains a certain record that was once compulsory, or adds a new record, or changes the way in which a record is used, then there should also be a new product format version number.

To summarise, the product format version number is updated:

1. when any record format version number of a record used in that product changes,
2. when a compulsory record is removed from a product,
3. when a new record is added to the product,
4. when the use of a record changes.

A recommended way to use *major.minor* versions of the product format version number is to issue both minor updates for a change resulting from a PFS update, and major updates for a change resulting from GPFS updates that affect all products. Then a GPFS update would reset all products back to a new major of (say) 12.0, and then 12.1, 12.2 etc. versions would indicate PFS-only updates.

Table 17 provides the *Product Format Version Control Numbers* of the GPFS.

Table 17: Record Format Version Numbers

Product Format Version Control Number (format_version)	Generic Product Format Specification Issue (gpfs_reference_and_version)
0.0	1E
1.0	2.0
1.1	2A
3.0	3
3.0	3A
4.0	3B
5.0	3D
5.1	4
5.2	5

Appendix A

PRODUCT FORMAT XML DESCRIPTION**A.1 EPS-SG Product XML schema definition**

The EPS-SG Product XML schema definition file is attached in the file EPS-SG-GPFS.xsd:



EPS-SG-GPFS_V3C.x
sd

A.2 EPS-SG Generic Product Format Specification XML file

An EPS-SG Product XML description file is attached in the file EPS-SG-GPFS.xml. Note that it does not represent a real EPS-SG product but the product format description in XML:



EPS-SG-GPFS-41.xml

Appendix B

FORMAT CONVENTIONS

Products to be encoded with various formats (e.g., NetCDF, BUFR, GRIB) 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, including those in Section 3.