Title: MERMAID data format

Doc. no: QWG-MER-MERMAID-DF-02

Issue: 2

Revision: 3

Date: 22/03/2012

	Name	Company	Function	Signature	Date
Prepared by:	C. Mazeran	ACRI-ST	W.P. Manager		22/03/2012
	C. Lerebourg	ACRI-ST	Research Engineer		22/03/2012
Verified by:					22/03/2012
Validated by:					22/03/2012

Change Record

Issue	Revision	Date	Description	Approval
1	0	11/03/2009	Initial data format on March 2009	
2	0	22/01/2010	Major update of the format with respect to new extracted quantities, including in-situ atmospheric parameters and correction of in-situ water marine reflectance for solar illumination	
2	1	30/08/2010	Update of MQC and PQC flags section	
2	2	28/02/2012	Typo correction in some units	
2	3	22/03/2012	upade of output header definition (table 2)	

Contents

1 General Information

The MERMAID project aims at making available an easy-to-use centralised database of merged in-situ optical measurements with concurrent MERIS acquisitions to Ocean Colour researchers involved in the MERIS mission. Details on the project and access to the database are available on http://hermes.acri.fr/mermaid.

This document describes the data format of the extraction generated by the web interface.

The output of the extraction process generates four semi-colon separated files and a text file:

1. **extraction.csv** contains data of the whole block of pixels for each matchup (either 1, 3x3 or 5x5 pixels).
2. **extractionAvg.csv** contains averaged data of the block of pixels, computed upon the *flag acceptance* and *statistical screening* options.
3. **stats.csv** contains statistics comparing MERIS water-leaving reflectance $\rho_w(\lambda)$ to in-situ measurements. Statistics are computed per individual sites occuring in the extraction as well as for the whole dataset.
4. **uncertainties.csv** contains the in-situ water reflectance uncertainties provided by the PI, when available.
5. **parameter.txt** is a record of the extraction criteria.

Depending on the user's choice, extraction may also include RGB of Level 1 scenes (in directory **RGB**) and plots to compare MERIS water reflectance retrieval to in-situ data (scatter plots, histograms in directory **plots**).

The MERIS quantities are either Level 1b, Level 2, auxiliary data or intermediary data. The first three kinds of data are detailed in the MERIS product handbook available at http://envisat.esa.int/pub/ESA_DOC/ENVISAT/MERIS/meris.ProductHandbook.2_1.pdf.

The intermediary data are detailed in the MERIS Detailed Processing Model (DPM) available at http://earth.esa.int/pub/ESA_DOC/ENVISAT/MERIS/MERIS-DPM-L2-i8r0B.pdf.

Additional informations about the models used in MERIS and MERMAID can be found in the Reference Model Document (RMD) available at http://envisat.esa.int/instruments/meris/rfm/rmd_third_reprocessing_ocean.pdf.

Further details on the oceanic in-situ data are given on the web site through the MERIS Optical Measurement Protocols document (http://hermes.acri.fr/mermaid/dataproto/CO-SCI-ARG-TN-0008_MERIS_Optical_Measurement_Protocols_Issue2_Aug2011.pdf), or can be requested from the corresponding PI, whose affiliation and contact email are displayed on the MERMAID Data Policy page (http://hermes.acri.fr/mermaid/policy/policy.php).

The atmospheric in-situ data come from the AERONET web site http://aeronet.gsfc.nasa.gov/new_web/ocean_color.html.

2 Extracted files format

The data format of extracted files **extraction.csv** and **extractionAvg.csv** is detailed in table 2.

Units and type definitions are given in table 3 and 4. The time format UTC iso 860 is defined as yyyymmddThhmmssZ where yyyymmdd is year, month, day and hhmmss is UTC time expressed as hour, minute, second.

Spectral quantities are extracted for three different band sets (see table 5 for MERIS band definition):

- The Level 1 TOA radiance and reflectances are extracted on all 15 bands;
- the NIR residual marine signal computed by BPAC is extracted at only B9, B12, B13 and B14;
- all other Level 2 signals are extracted on 13 bands: B1 to B10, B12, B13, B14.

For AERONET atmospheric data, two measurements bracketing the satellite overpath are available: the closest atmospheric dataset before and after satellite overpath (with suffix "_1" and "_2" respectively).

Note that some of the variables, although defined as integers, are printed out as floating point numbers with no decimal numbers in order to support the "NaN" value. For example, the aerosol model indexes (parameters **iaer1** and **iaer2**) are integers and printed as floating points ("f3.0"). They therefore appear in the csv files with decimal points.

In file **extractionAvg.csv**, note that quantities **iaer1**, **iaer2** and **aer_mix** are set to "NaN" because their mean would be meaningless. On the contrary, the mean of the detector index is computed for reference, yet it might be a floating point value because of duplicated detectors in the macro-pixel.

Table 2: Columns of extracted files.

	Header ID	Unit	Type	Description
Meta Data	MATCHUP_ID	-	S	Matchup identification
Meta Data	Site	-	S	Site name
Meta Data	PI	-	S	Principal investigator
Meta Data	Lat_IS	deg	F	In-situ latitude
Meta Data	Lon_IS	deg	F	In-situ longitude
	TIME_IS	UTC	S	In-situ measurement date and time
	PQC	-	S	Processing Quality Control Flag (cf Protocol doc.)
	MQC	-	S	Measurement Quality Control Flag (cf Protocol doc.)
	land_dist_IS	km	F	Distance to coast
	thetas_IS	deg	F	In-situ derived Solar zenith angle
	thetav_IS	deg	F	In-situ sensor zenith angle
	dphi_IS	deg	F	In-situ azimuth angle
	rho_wn_IS_b	dl	F	In-situ normalised water reflectance at band b
	rho_wn_ISME_b	dl	F	rho_wn_IS corrected for solar illumination at band b
	Es_IS_b	$mW.m^{-2}.nm^{-1}$	F	Sea level solar illumination at band b
	Kd_IS_b	m^{-1}	F	Diffuse attenuation coefficient for downwelling irradiance at bands b
	KPAR_IS	m^{-1}	F	Diffuse attenuation coefficient for PAR
	PARzX%_IS	m	F	Depth of X% light level of surface PAR
	a_IS_b	m^{-1}	F	Total absorption coefficient at band b. $a = a_w + a_g + a_p$
	a_det_IS_b	m^{-1}	F	Detrital absorption at band b. $a_{det} = a_p - a_{ph}$
	a_p_IS_b	m^{-1}	F	Particulate absorption coefficient at band b. $a_p = a_{ph} + a_{det} = a - a_w - a_g$
	a_ph_IS_b	m^{-1}	F	Algal pigment absorption coefficient at band b. $a_{ph} = a_p - a_{det}$
	a_g_IS_b	m^{-1}	F	Colored Dissolved Organic Matter absorption coefficient at band b. $a_g = a - a_w - a_p$
	bb_IS_b	m^{-1}	F	Total backscattering coefficient at band b. $b_{bb} = b_{bw} + b_{bp}$
	TSM_IS	$g.m^{-3}$	F	Total Suspended Matter. $TSM\_IS = OSM\_IS + MSM\_IS$
	OSM_IS	$g.m^{-3}$	F	Organic Suspended Matter
	MSM_IS	$g.m^{-3}$	F	Mineral Suspended Matter
	POC_IS	$gC.m^{-3}$	F	Particulate Organic Carbon
	HPLC_chla_TOTAL_IS	$mg.m^{-3}$	F	Total Chla derived from HPLC pigment analysis. Sum of HPLC chla, div. chla, chlide-a + phaeopigments

continued on next page

Table 2 continued

		Header ID	Unit	Type	Description
In situ data		HPLC_chla_ONLY_IS	$mg.m^{-3}$	F	Chla only derived from HPLC pigment analysis
		SPECT_chla_IS	$mg.m^{-3}$	F	Chla derived from spectrophotometric procedures
		Fluor_chla_IS	$mg.m^{-3}$	F	Chla derived from fluorimetric procedures
		AERONET_chla_IS	$mg.m^{-3}$	F	In-situ Chlorophyll concentration
	AERONET	AOT_865_IS	dl	F	SIMBADA aerosol optical thickness at 865 nm
		time_IS_1	UTC	S	First AERONET measurement time
		AOT_IS_1_*	dl	F	First AERONET aerosol optical thickness at band b (table 6)
		alpha_NIR_IS_1	dl	F	First AERONET aerosol Angstrom exponent
		time_IS_2	UTC	S	Second AERONET measurement time
		AOT_870_IS_*	dl	F	Second AERONET aerosol optical thickness at band b (table 6)
		alpha_NIR_IS_2	dl	F	Second AERONET aerosol Angstrom exponent
		PROCESSING_VERSION	-	S	MERIS Ground Segment processing version
MERIS data	Pixel id & geometry	TIME	UTC	S	MERIS measurement date and time
		ORBIT	-	S	Orbit number
		RESOLUTION	-	S	Image resolution (FR/RR)
		DETECTOR	-	I	Detector index
		LAT	deg	F	Pixel latitude
		LON	deg	F	Pixel longitude
		SUN_ZENTIH	deg	F	Pixel solar zenith angle
		VIEW_ZENITH	deg	F	Pixel viewing zenith angle
		DELTA_AZIMUTH	deg	F	Pixel azimuthal difference between satellite and sun
		SCATT_ANGLE	deg	F	Pixel scattering angle
	Meteo	WINDM	$m.s^{-1}$	F	ECMWF wind speed modulus
		PRESS_ECMWF	hPa	F	ECMWF atmospheric pressure
		OZONE_ECMWF	DU	F	ECMWF ozone content
		VAPOUR_ECMWF	$\%$	F	ECMWF relative humidity
		LAND	-	I	Land
		CLOUD	-	I	Cloud
		ICE_HAZE	-	I	Ice or high aerosol load
		WHITE_SCATTERER	-	I	presence of white scatterer in water
		HIGH_GLINT	-	I	High (uncorrected) glint
		MEDIUM_GLINT	-	I	Medium glint
		PCD_1_13	-	I	Uncertain normalized surface reflectance
		PCD_14	-	I	Uncertain total water vapour content
		PCD_15	-	I	Uncertain algal pigment index 1 or cloud top pressure or TOA vegetation index

continued on next page

Table 2 continued

		Header ID	Unit	Type	Description
MERIS data	Flags	PCD_16	-	I	Uncertain yellow substance or total suspended matter or reftified reflectances
		PCD_17	-	I	Uncertain algal pigment index 2 or bottom of atmosphere vegetation index
		PCD_18	-	I	Uncertain PAR or cloud albedo or land surface pressure
		PCD_19	-	I	Uncertain aerosol type and optical thickness or cloud optical thickness
		OADB	-	I	Aerosol model is out of aerosol database
		ABSOA_DUST	-	I	Dust-like absorbing aerosol selected for atmospheric correction
		BPAC_ON	-	I	Bright pixel atmospheric correction activated and successful
		CASE2_S	-	I	Turbid (sediment dominated Case 2) water
		CASE2_ANOM	-	I	Anomalous scattering water
	Water param.	ALTITUDE	m	F	Local depth (positive downward)
		CHL1	$mg.m^{-3}$	F	Chlorophyll estimate - case 1 algorithm
		CHL2	$mg.m^{-3}$	F	Chlorophyll estimate - case 2 algorithm
		SPM	$g.m^{-3}$	F	Total Suspended Matter estimate - case 2 algorithm
		ODOC	m^{-1}	F	Yellow Substances estimate - case 2 algorithm
	Atmospheric param.	VAPR	$g.cm^{-2}$	F	Water vapour
		TAU_AER_05	dl	F	Aerosol optical thickness estimate at 560nm
		AOT_AER_13	dl	F	Aerosol optical thickness estimate at 865nm
		ALPHA	dl	F	Aerosol Angstrom exponent
		IAER_1	-	I	Index of bracketing aerosol model 1
		IAER_2	-	I	Index of bracketing aerosol model 2
		AER_MIX	dl	F	Aerosol mixing ratio
	Radiometric data	RN_b	dl	F	MERIS water reflectance at band b
		RHO_WN_b	dl	F	MERIS Normalised water reflectance at band b
		TOAR_b	LU	F	Top Of Atmosphere radiance at band b
		RHO_TOA_b	dl	F	Top Of Atmosphere reflectance at band b
		RHO_GC_b	dl	F	Gas (ozone, water, oxygen), smile and glint corrected Top Of Atmosphere reflectance at band b
		RHO_RAY0_b	dl	F	Pressure corrected Rayleigh reflectance at band b
		RHO_AER_b	dl	F	Aerosol reflectance at band b defined as rho_path-rho_ray where rho_path is the total atmospheric path reflectance retrieved by atmospheric correction

continued on next page

Table 2 continued

		Header ID	Unit	Type	Description
MERIS data	Radiometric data	T_DOWN_b	dl	F	Downward total transmittance (Rayleigh + aerosol) at band b
		T_UP_b	dl	F	Upward total transmittance (Rayleigh + aerosol) at band b
		T_RHO_W_C2_b	dl	F	NIR residual marine signal TOA computed by BPAC at band b

Table 3: Units and notation

Unit symbol	**Name**
dl	dimensionless
EU or $mW.m^{-2}.nm^{-1}$ (EU being Energy Unit)	Spectral irradiance
LU or $mW.m^{-2}.nm^{-1}.sr^{-1}$	Spectral radiance
Deg	Degree angle
UTC	Coordinated Universal Time
hPa	hectoPascal
DU	Dobson Unit ($=10^3$ atm.cm)

Table 4: Type and notation

Type Symbol	**Name**
S	String
I	Integer
F	Float

Table 5: MERIS Spectral bands

Band	Wavelength (nm)	Bandwidth (nm)
1	412.5	10
2	442.4	10
3	490.0	10
4	510.0	10
5	560.0	10
6	620.0	10
7	665.0	10
8	681.25	7.5
9	708.75	10
10	753.75	7.5
11	761.875	3.75
12	778.75	15
13	865.0	20
14	885.0	10
15	900.0	10

Table 6: AERONET AOT spectral bands

Band	Wavelength (nm)
1	412
2	443
3	490
4	555
5	667
6	675
7	870

3 Statistical data file

The data contained in **stats.csv** are statistical parameters derived from **extractionAvg.csv**. The estimators are summarised in table 7 below, where x_i stands for the reference in-situ normalized water-leaving reflectance and y_i for the MERIS'one. The overline stands for the mean over the N available points.

Note that two references are possible: directly the in-situ reflectance provided by the PI (and normalised) **rho_wn_IS**, or its correction for MERIS solar illumination **rho_wn_ISME**, see section 6. The choice is given to the user in the interface. No matchup will be returned if the latter is chosen and PI did not provide in-situ solar irradiance.

Table 7: Format of file **stats.csv**

Header	**Variable**	**Formula**
lambda	Wavelength (nm)	
N	Number of points	
RPD	Average Relative (signed) percent difference	$RPD = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i - x_i}{x_i}$
|RPD|	Average Relative (unsigned) percent difference	$RPD = \frac{1}{N}\sum_{i=1}^{N}\frac{\|y_i - x_i\|}{x_i}$
MAD	Mean Arithmetic Difference	$MAD == \frac{1}{N}\sum_{i=1}^{N} y_i - x_i$
RMSE	Root Mean Squared Error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2}$
slope	Linear fit slope	$slope = \frac{\overline{x_i y_i} - \overline{x_i}\,\overline{y_i}}{\overline{x_i^2} - (\overline{x_i})^2}$
intercept	Linear fit offset	$intercept = \overline{y_i} - slope * \overline{x_i}$
r^2	Coefficient of determination	$\frac{(\overline{x_i y_i} - \overline{x_i}\,\overline{y_i})^2}{\left(\overline{x_i^2} - (\overline{x_i})^2\right)^2 \left(\overline{y_i^2} - (\overline{y_i})^2\right)^2}$

4 Uncertainty file

The **uncertainties.csv** file contains, when available, the in-situ water reflectance percentage uncertainties provided by the PI's.

5 MQC and PQC flags

Flags are required to describe the level of quality control (QC) applied to the in-situ data. Two flags are defined in MERMAID: PQC and MQC, also defined in the MERIS Optical Measurement Protocols.

5.1 MQC: Measurement Quality Control

The MQC flag defines the quality control checks made by the PI, prior to submission to MERMAID. It pertains to the provision, or not, of a clearly defined measurement and processing protocol by the in-situ data provider.

Table 8: MQC flag criteria definition. Flag position is counted from the first numeric character after the leading "M". Unless otherwise specified: 0 = No / Not done, 1 = Yes / done/ provided, 2 = Unknown / not available / Not applicable (N/A).

Flag ID	Flag position	String options	Conditions and criteria
MQC	1	0 1	Protocol provided by PI
	2	0 1 2	Avoidance of self-shading
	3	0 1 2	Correction for straylight (2 = N/A)
	4	0 1 2	Made dark measurements (and used in processing)
	5	0 1 2	Measured immersion coefficients (and used in processing)
	6	0 1	Instrument calibration history provided
	7	0 1	Data processed to MERIS band characterisation
	8	0 1	Hyperspectral integration done
	9	0 1	Error budget provision
	10	0 1 2 (L1.5)	In-situ data filtering (PI's QC checks)
	11	0 1	In-situ ρ_w already normalised or f/Q and $\Re$ corrected
	12	0 1 2	Tilt measurement made
	13	0 1 2	Calibration of tilt sensor
	14	0 1 2 3	Type of E_s : E_s or $E_d(0^+)$ (0 = N/A, 1 = E_s measured in-situ, 2 = $E_d(0^+)$ measured in-situ/derived in-situ, 3 = E_s computed.)
	15	0 1	E_s tilt corrected
	16	0 1 2	Type of L_u: L_w or $L_u(0^-)$ (and extrapolated to $L_w(0^+)$. 0 = N/A
	17	0 1	L_u tilt corrected
	18	0 1 2 (L1.5)	(AERONET-OC only) Data quality level: 0 = N/A, 1 = L1.5, 2 = 2.0 (see AERONET website http://aeronet.gsfc.nasa.gov, for more details)

5.2 PQC: Processing Quality Control

The PQC flag defines the post-submission quality control performed on the in-situ data, and includes information on the normalisation procedure. The flag provides information on the MERMAID in-situ processing for both the optical data received (e.g. ρ_w) and the atmospheric parameters (e.g. α) newly added to the database.

Table 9: PQC flag criteria definition. Flag position is counted from the first numeric character after the leading "P". Unless otherwise specified: 0 = No / Not done, 1 = Yes / done/ provided, 2 = Unknown / not available.

Flag ID	Flag position	String options	Conditions and criteria
PQC	1	0 1	Passed in-situ ρ_w QC
	2	0 1 2	Hyperspectral integration
	3	0 1	Case 1 Normalisation by MERMAID *
	4	0 1	Case 2 Normalisation
	5	0 1	Band shifted correction (AERONET-OC data only; presently only AAOT)
	6	0 1	Nearest neighbour (refer to MERIS Optical Measurement Protocols): 0 = data at bands greater than ± 5nm from MERIS 1 = data at bands less than ± 5nm from MERIS ****NOMAD only**: Flag is 0 when data is at 560 nm and 1 when at 555 nm
	7	0 1	AlphaNIR (1 & 2) derived from 870-675 nm
	8	0 1	AlphaNIR (1 & 2) derived from 870-667 nm

* See MQC flag #11 to check if normalisation has already been performed by PI.

6 Correction of in-situ reflectance for solar irradiance

Because the solar illumination used in the computation of the in-situ reflectance might be different from that of the MERIS processing, MERMAID contains a complementary in-situ water reflectance **rho_wn_ISME**, consistent with the MERIS formulation of E_s. This quantity is for instance used in the vicarious calibration of MERIS. The in-situ solar illumination provided by PI may come from measurement or models, depending on the sites. It is not directly stored in the MERMAID extraction. Details on this quantity can be found in the MERIS Optical Measurement Protocol.

6.1 Calculation of MERIS solar irradiance

For a given detector indexed by i, we have from DPM Level 2 step 2.1.4 and RMD

$$\begin{aligned}\rho_{TOA}(\lambda) &= \frac{\pi TOAR(\lambda) seasonal\_fact}{\cos\theta_s^{MERIS} F_0(\lambda, i)} \\ &= \frac{\pi TOAR(\lambda)}{\cos\theta_s F_0(\lambda, i) d^2}\end{aligned}$$

where ρ_{TOA} is the Top of Atmosphere reflectance, $TOAR$ is the TOA radiance, θ_s^{MERIS} is the sun zenith angle at measurement time, $F_0(i)$ is the extra-terrestrial sun irradiance at reference date for MERIS detector i and $seasonal\_fact = 1/d^2$ the correction factor for seasonal variation of sun irradiance.

It follows that the MERIS Solar irradiance can be directly deduced from the MERMAID extraction by

$$F_0 d^2(\lambda, i) = \frac{\pi TOAR(\lambda)}{\cos\theta_s^{MERIS} \rho_{TOA}(\lambda, i)}.$$

6.2 Correction

Given the previous $F_0 d^2$, a MERIS-like total downward irradiance at sea-level for any solar zenith angle θ writes

$$E_s^{MERIS}(\lambda, \theta) = T_d^{MERIS}(\lambda, \theta)\cos(\theta) F_0 d^2,$$

where $T_d^{MERIS}(\lambda, \theta)$ is the total downward transmittance as computed in the MERIS processing, which includes the Rayleigh, aerosol and ozone contribution. In particular, this transmittance may require auxiliary data like aerosol models and meteorological quantity of the current pixel. It is thus dependent on the MEGS version.

Then, starting from the in-situ total downward irradiance E_S^{IS} provided by the PI, the correction simply writes

$$rho\_wn\_ISME(\lambda) = rho\_wn\_IS(\lambda) \frac{E_s^{IS}(\lambda)}{E_s^{MERIS}(\lambda, \theta_s^{IS})}.$$