



# OceanSITES User's Manual

NetCDF Conventions and Reference Tables

Version 1.1

June 30<sup>th</sup>, 2008

# Table of contents

<b>1.OVERVIEW</b>	<b>4</b>
1.1.About OceanSITES	4
1.2.About this document	4
1.3.OceanSITES data management structure and data access	4
<b>2.OCEANSITES DATA FORMAT</b>	<b>4</b>
<b>3.DATA FILE DIMENSIONS</b>	<b>6</b>
3.1.Global attributes	7
3.2.Variables	11
3.2.1.Coordinate variables	11
3.2.2.Coordinate quality control variables	13
3.2.3.Data variables	14
<b>4.OCEANSITES METADATA FORMAT</b>	<b>17</b>
<b>5.REFERENCE TABLES</b>	<b>17</b>
5.1.Reference tables 1: data type and data code	17
5.1.1.Data type	17
5.1.2.Data code	17
5.2.Reference table 2: quality control flag scale	18
5.2.1.Reference table 2.1: quality control procedure indicator	18
5.2.2.Reference table 2.2: cell methods	19
5.3.Reference table 3: OceanSITES parameter dictionary	19
5.4.Reference table 4: Data Assembly Center Codes	22
5.5.Reference table 5: data mode	22
5.6.Reference table 6: OceanSITES sites catalog	23
<b>6.GDAC ORGANIZATION</b>	<b>24</b>
6.1.File naming convention	24
6.1.1.Data file naming convention	24
6.1.2.Metadata file naming convention	25
6.2.Index file for data files	25

## History

Version	Date	Comment
0.1	20/03/2003	TC: creation of the document
0.3	20/02/2004	TC: updates on locations, mooring name, data state indicator, parameters table, epic codes, history information
0.3.2	26/05/2004	N.G.: make more flexible, add dataset (metadata) file
0.4	01/06/2004	TC: separate data set description and data file, merge with Steve Hankins's straw man
0.6	28/06/2004	TC: updates from Nan Galbraith, Steve Hankins, Jonathan Gregory, Brian Eaton
0.7	23/05/2005	Maureen Edwards: NOCS data centre, new GF3 parameters
0.7	24/05/2005	Roy Lowry: physical parameters from BODC Data Markup Vocabulary
1.0	18/02/2006	TC: updates following OceanSITES data management meeting 2006, Hawai'i §2.1: LEVEL dimension replaces DEPTH to accomadate depth or pressure §2.2: QC_MANUAL field created §2.2: CONVENTION field removed §2.2: PLATFORM_CODE added §2.2: SITE_CODE added §2.2: WMO_PLATFORM_CODE added §2.3: DEPTH renamed DEPH to comply to GF3 §2.3: DATA_MODE set at measurement level §3: metadata file description transferred to "OceanSITES metadata proposal" until approval §5: file naming convention updated
1.0	19/02/2006	NG: data codes in chapter 4.1.2
1.0	28/04/2006	PF & NG: data mode optional
1.0	28/04/2006	TC & JG: §2.2 global attributes
1.1	April-May-June 2008	NG, MM,TC, ML: general revision based on OceanSITES 2008 meeting Epic codes removed Use ISO8601 for string dates Remove general attributes Update global attribute section for CF-1.1 compatibility New dimensions for DEPTH, LATITUDE, LONGITUDE Add an uncertainty attribute New presentation of the document

# 1. Overview

## 1.1. About OceanSITES

The OceanSITES program is the global network of open-ocean sustained timeseries sites, called ocean reference stations, being implemented by an international partnership of researchers. OceanSITES provides fixed-point timeseries of various physical, biogeochemical, and atmospheric variables at different locations around the globe, from the atmosphere and sea surface to the seafloor. The program's objective is to build and maintain a multidisciplinary global network for a broad range of research and operational applications including climate, carbon, and ecosystem variability and forecasting and ocean state validation.

All OceanSITES data are publicly available. More information about the project is available at <http://www.oceansites.org>

## 1.2. About this document

The main purpose of this document is to specify the format of the files that are used to distribute OceanSITES data, and to document the standards used therein. This includes naming conventions, or taxonomy, as well as metadata content.

## 1.3. OceanSITES data management structure and data access

The data flow within OceanSITES is carried out through three organisational units:

The **Principal Investigator (PI)**, typically a scientist at a research institution, maintains the observing platform and the sensors that deliver the data. He or she is responsible for providing the data and all auxiliary information to a **Data Assembly Center (DAC)**. The DAC assembles OceanSITES-compliant files from this information and delivers these to one of two **Global Data Assembly Centers (GDAC)**, where they are made publicly available. The user can access the data at either GDAC, cf. section "GDAC organization".

## 2. OceanSITES data format

OceanSITES uses the NetCDF (network Common Data Form) system, a set of software libraries and machine-independent data formats. Our implementation of NetCDF is based on the community-supported Climate and Forecast (CF) specification, which supplies a standard vocabulary and some metadata conventions.

OceanSITES has several more restrictions than the CF standard. These are intended to make it easier to share in-situ data, to make it simpler for the GDACs to aggregate data from multiple sites, and to ensure that the data can be created and understood by the basic NetCDF utilities.

- OceanSITES includes standard terms for the short name of both coordinate and data variables (measurements) .
- File names are created using a standard, described in section 6.2.

- Coordinate variables, which describe the dimensions of a data set, are limited to a single set of longitude, latitude, depth and time (X,Y,Z, and T) dimensions in any single file. If data from a reference station cannot all be put on to a single set of axes, then separate files are created for these data.

An OceanSITES data file contains measurements such as temperature and salinity, continuously performed at different levels on a platform (e.g. mooring), as well as meteorological or other parameters recorded at the site, derived variables associated with the site, and complete location, time, and provenance information.

The requirements are drawn almost exclusively from the NetCDF Style Guide:

- Units are compliant with CF/COARDS/Udunits ;
- The time parameter is encoded as recommended by COARDS and CF.
- Parameters are given standard names from the CF table
- Where time is specified as an attribute, the ISO8601 standard is used.

For more information on CF, COARDS, NetCDF, Udunits, and ISO8601 see:

- NetCDF: <http://www.unidata.ucar.edu/software/netcdf/docs/BestPractices.html>
- Udunits: <http://www.unidata.ucar.edu/software/udunits/>
- CF: <http://cf-pcmdi.llnl.gov/>
- COARDS: [http://www.ferret.noaa.gov/noaa\\_coop/coop\\_cdf\\_profile.html](http://www.ferret.noaa.gov/noaa_coop/coop_cdf_profile.html)
- ISO8601: [http://en.wikipedia.org/wiki/ISO\\_8601](http://en.wikipedia.org/wiki/ISO_8601)

### 3.Data file dimensions

NetCDF dimensions provide information on the size of the data variables. OceanSITES allows a single parameter for each of the data dimensions, i.e. time, depth, latitude and longitude. Requirements are described further in the section on coordinate variables. Standard names for OceanSITES dimensions should be in upper case.

Name	Example	Comment
TIME	TIME=unlimited	Number of time steps. Example: for a mooring with one value per day and a mission length of one year, TIME contains 365 time steps.
DEPTH	DEPTH=5	Number of depth levels. Example: for a mooring with measurements at 0.25, 10, 50, 100 and 200 meters, DEPTH=5.
LATITUDE	LATITUDE=1	Dimension of the LATITUDE coordinate variable. Examples: for a mooring without positioning system, LATITUDE=1. For a mooring with a GPS receiver, use LATITUDE of the same dimension as TIME and provide the actual location.
LONGITUDE	LONGITUDE=1	Dimension of the LONGITUDE coordinate variable. Examples: for a mooring without positioning system, LONGITUDE=1. For a mooring with a GPS receiver, use LONGITUDE of the same dimension as TIME and provide the actual location.
POSITION	POSITION=1	Dimension of the POSITION_QC variable.

### 3.1. Global attributes

The global attribute section of a NetCDF file contains metadata that describes the contents of the file overall, and allows for data discovery. All fields should be human-readable, and should be of character type, not numeric, even if the information content is a number. OceanSITES recommends that all of these attributes be used and contain meaningful information unless there are technical reasons rendering this impossible. However, files that do not at least contain the attributes listed as "mandatory" will not be considered OceanSITES-compliant. In OceanSITES, global attribute names are in lower-case letters.

Global attributes can be thought of as conveying five kinds of information:

- What: what are the data in this dataset;
- Where: the spatial coverage of the data;
- When: the temporal coverage of the data;
- Who: who produced the data;
- How: how were the data produced and made available.

The global attributes specification follows the recommendations of Unidata NetCDF Attribute Convention for Dataset Discovery, at <http://www.unidata.ucar.edu/software/netcdf/java/formats/DataDiscoveryAttConvention.html>

Name	Example	Definition
<b>WHAT</b>		
data_type	data_type="OceanSITES time-series data"	This field contains the type of data contained in the file. The list of acceptable data types is in reference table 1. Example: "OceanSITES time-series data". This attribute is mandatory.
format_version	format_version="1.1"	OceanSITES format version Example: "1.1". This attribute is mandatory.
platform_code	platform_code="CIS-1"	Platform unique code within OceanSITES project. Example: "CIS-1" mooring on CIS site (Central Irminger Sea). This attribute is mandatory.
date_update	date_update="2006-04-11T08:35:00Z"	File update or creation date (UTC). See note on time format below. This attribute is mandatory.
institution	institution="National Oceanographic Centre"	Specifies institution where the original data was produced.
site_code	site_code="CIS"	Name of the site within OceanSITES project. Example: "CIS" for Central Irminger Sea. The site codes are listed in reference table 6.
wmo_platform_code	wmo_platform_code="48409"	WMO (World Meteorological Organization) identifier. This platform number is unique within the OceanSITES project. Example: "48409" for CIS-1 mooring.
source	source="Mooring observation"	The method of production of the original data. For OceanSITES data, use one of the following: "Shipborne observation", "Mooring observation"

history	history= "2005-04-11T08:35:00Z data collected, A. Meyer.\n2005-04-12T10:11:00Z OceanSITES file with post- recovery data compiled and sent to DAC, A. Meyer."	Provides an audit trail for modifications to the original data. It should contain a separate line for each modification, with each line beginning with a timestamp, and including user name, modification name, and modification arguments. The time stamp should follow the format outlined in the note on time formats below.
data_mode	data_mode="R"	Indicates if the file contains real-time, post-recovery, or delayed-mode data. The list of valid data modes is in reference table 5.
quality_control_indicator	quality_control="6"	Level of quality control applied to data. The values are listed in reference table 2.1.
quality_index	quality_index="A"	A code value valid for the whole dataset: 0 unknown quality A excellent (no known problems, regular quality checking) B probably good (occasional problems, validation phase) C extremely suspect, frequent problems
references	references="http:// www.oceansites.org, http://www.noc.soton.ac.uk/ animate/index.php"	Published or web-based references that describe the data or methods used to produce it. Include a reference to OceanSITES and a project-specific reference if appropriate.
comment	comment="..."	Miscellaneous information about the data or methods used to produce it. Any free-format text is appropriate.
conventions	conventions="OceanSITES Manual 1.1, CF-1.1"	Name of the conventions followed by the dataset.
netcdf_version	netcdf_version="3.5"	Netcdf version used for the data set
title summary	title="CIS Mooring Data" summary="Oceanographic mooring data from CIS observatory in the Central Irminger Sea, North Atlantic, in 2005. Measured properties: temperature and salinity at ten depth levels."	Free-format text describing the dataset. The display of these two attributes together should allow data discovery for a human reader. "title": title of the dataset. Use the file name if in doubt. "summary": a longer description of the dataset. A paragraph of up to 100 words is appropriate.
naming_authority id	naming_authority="OceanSITES" id="OS_CIS-1_200502_TS"	The "id" and "naming_authority" attributes are intended to provide a globally unique identification for each dataset. For OceanSITES data, use: naming_authority="OceanSITES" and id=file name (without .nc suffix), which is designed to be unique.
cdm_data_type	cdm_data_type="Station"	The "cdm_data_type" attribute gives the Unidata CDM (common data model) data type used by THREDDS. E.g. "Point", "Trajectory", "Station", "Radial", "Grid", "Swath". Use "Station" for OceanSITES mooring data. More: <a href="http://www.unidata.ucar.edu/projects/THREDDS/CDM/CDM-TDS.htm">http://www.unidata.ucar.edu/projects/THREDDS/CDM/CDM-TDS.htm</a>
<b>WHERE</b>		
area	area="North Atlantic Ocean"	Geographical coverage. Try to compose of the following: North/Tropical/South Atlantic/Pacific/Indian Ocean, Southern Ocean, Arctic Ocean. For specific sea area, use the International Hydrographic Bureau sea areas available at : <a href="http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp">http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp</a> .
geospatial_lat_min	geospatial_lat_min="59.8"	The southernmost latitude, a value between -90 and 90 degrees.



geospatial_lat_max	geospatial_lat_max="59.8"	The northernmost latitude, a value between -90 and 90 degrees.
geospatial_lon_min	geospatial_lon_min="-41.2"	The westernmost longitude, a value between -180 and 180 degrees.
geospatial_lon_max	geospatial_lon_max="-41.2"	The easternmost longitude, a value between -180 and 180 degrees.
geospatial_vertical_min	geospatial_vertical_min="10.0"	Minimum depth for measurements
geospatial_vertical_max	geospatial_vertical_max="2000"	Maximum depth for measurements
<b>WHEN</b>		
time_coverage_start	time_coverage_start="2006-03-01T00:00:00Z"	Start date of the data in UTC. See note on time format below.
time_coverage_end	time_coverage_end="2006-03-05T23:59:29Z"	Final date of the data in UTC. See note on time format below.
<b>WHO</b>		
institution_references	institution_references="http://www.nocs.uk"	References to data provider institution, the place to find all information on the dataset (web-based, i.e. give URLs).
contact	contact="codac@nocs.uk"	Contact person's e-mail.
author	author="John Smith"	Name of the person responsible for the creation of the dataset.
data_assembly_center	data_assembly_center="EUROSITES"	Data Assembly Center (DAC) in charge of this data file. The data_assembly_center are listed in reference table 4.
pi_name	pi_name="Alice Juarez"	Name of the principal investigator in charge of the platform.
<b>HOW</b>		
distribution_statement	distribution_statement="Follows CLIVAR (Climate Variability and Predictability) standards, cf. <a href="http://www.clivar.org/data/data_policy.php">http://www.clivar.org/data/data_policy.php</a> . Data available free of charge. User assumes all risk for use of data. User must display citation in any publication or product using data. User must contact PI prior to any commercial use of data."	Statement describing data distribution policy. OceanSITES has adopted the CLIVAR data policy, which explicitly calls for free and unrestricted data exchange. Details at: <a href="http://www.clivar.org/data/data_policy.php">http://www.clivar.org/data/data_policy.php</a>
citation	citation="These data were collected and made freely available by the OceanSITES project and the national programs that contribute to it."	The citation to be used in publications using the dataset.
update_interval	update_interval="daily"	Update interval for the file, one of the following: "hourly", "daily", "yearly", "void". Use "void" for delayed-mode or archive data that do not need continuous updating.
qc_manual	qc_manual="OceanSITES User's Manual v1.1"	This field contains the name of the manual that describes the quality control procedure. As of now, there is no separate QC manual, so the user's manual is the appropriate reference.

### **Note on time formats**

Whenever time information is given in the global attributes, it ought to be a string of the format:

"YYYY-MM-DDT $hh:mm:ss$ Z" (i.e. year - month - day T hour : minute : second Z)

If higher resolution than seconds is needed, any number of decimal digits (".s") for the seconds is acceptable:

"YYYY-MM-DDT $hh:mm:ss.s$ Z"

In any case, the time must be in UTC. A capital "T" separates the date and the hour information. The string must end with a capital "Z", an old indication of UTC. These formats are two (of many) described by ISO8601.

Examples:

- 2005-10-24T08:00:00Z
- 2008-01-01T22:50:02.031Z

## **3.2. Variables**

NetCDF variables include data measured by instruments, parameters derived from the primary measurements, and coordinate variables, which may be nominal values, such as values for depth for instruments that do not directly record depth. The variable names are written in CAPITALIZED letters. Each variable has a specific set of attributes, some of which are mandatory.

### **3.2.1. Coordinate variables**

The coordinate variables orient the data in time and space. For this purpose, they have an "axis" attribute defining that they point in X, Y, Z, and T dimensions. The DEPTH variable may be positive in either upward or downward direction, which is defined in its "positive" attribute. All attributes in this section except the "comment" are mandatory, however "QC\_indicator" may be omitted for any parameter if there is a separate QC variable for that parameter.

Optionally, the Z axis may be represented as pressure, if, for example pressure is recorded directly by an instrument and the calculation of depth from pressure would cause a loss of information. Depth is strongly preferred, since it allows data to be used more directly; data with different coordinate variables must be recorded in separate files.

Type, Name, Dimension, Attributes	Comment
Double <b>TIME</b> (TIME); TIME:long_name = "time"; TIME:standard_name = "time"; TIME:units = "days since 1950-01-01T00:00:00Z"; TIME:_FillValue = 999999.0; TIME:valid_min = 0.0; TIME:valid_max = 90000.0; TIME:QC_indicator = <X>; TIME:QC_procedure = <Y>; TIME:uncertainty = <Z>; TIME:comment = "Optional comment..." TIME:axis = "T";	Date and time (UTC) of the measurement in days since midnight, 1950-01-01.  Example: Noon, Jan 2, 1950 is stored as 1.5. July 25, 2001, 19:14:00 is stored as 18833.8013889885.  <X>: Replaces TIME_QC if constant. Cf. note on quality control in data variable section, value from reference table 2. <Y>: Cf. note on quality control in data variable section, value from reference table 2.1. <Z>: Choose appropriate value.
Float <b>LATITUDE</b> (LATITUDE); LATITUDE:long_name = "Latitude of each location"; LATITUDE:standard_name = "latitude"; LATITUDE:units = "degrees_north"; LATITUDE:_FillValue = 99999.0; LATITUDE:valid_min = -90.0; LATITUDE:valid_max = 90.0; LATITUDE:QC_indicator = <X>; LATITUDE:QC_procedure = <Y>; LATITUDE:uncertainty = <Z>; LATITUDE:comment = "Optional comment..." LATITUDE:axis="Y";	Latitude of the measurements. Units: degrees north; southern latitudes are negative.  Example: 44.4991 for 44° 29' 56.76" N  <X>: Replaces LATITUDE_QC if constant. Cf. note on quality control in data variable section, value from reference table 2. <Y>: Cf. note on quality control in data variable section, value from reference table 2.1. <Z>: Choose appropriate value.
Float <b>LONGITUDE</b> (LONGITUDE); LONGITUDE:long_name = "Longitude of each location"; LONGITUDE:standard_name = "longitude"; LONGITUDE:units = "degrees_east"; LONGITUDE:_FillValue = 99999.0; LONGITUDE:valid_min = -180.0; LONGITUDE:valid_max = 180.0; LONGITUDE:QC_indicator = <X>; LONGITUDE:QC_procedure = <Y>; LONGITUDE:uncertainty = <Z>; LONGITUDE:comment = "Optional comment..." LONGITUDE:axis="X";	Longitude of the measurements. Unit: degrees east; western latitudes are negative.  Example: 16.7222 for 16° 43' 19.92" E  <X>: Replaces LONGITUDE_QC if constant. Cf. note on quality control in data variable section, value from reference table 2. <Y>: Cf. note on quality control in data variable section, value from reference table 2.1. <Z>: Choose appropriate value.
Float <b>DEPTH</b> (DEPTH); DEPTH:long_name = "Depth of each measurement"; DEPTH:standard_name = "depth"; DEPTH:units = "meters"; DEPTH:positive = "down"; DEPTH:_FillValue = -99999.0; DEPTH:valid_min = 0.0; DEPTH:valid_max = 12000.0; DEPTH:QC_indicator = <X>; DEPTH:QC_procedure = <Y>; DEPTH:uncertainty = <Z>; DEPTH:comment = "Optional comment..." DEPTH:axis="Z";	Depth of each measurement.  Example: 513 for a measurement 513 meters below sea surface.  Z axes may be positive="up" (atmospheric) or positive="down" (oceanic).  <X>: Replaces DEPTH_QC if constant. Cf. note on quality control in data variable section, value from reference table 2. <Y>: Cf. note on quality control in data variable section, value from reference table 2.1. <Z>: Choose appropriate value.

### 3.2.2.Coordinate quality control variables

The coordinate variables have the same quality control variables as the data variables. If the quality control values are constant, the information is given in attributes of the coordinate variables. For details, see <PARAM>\_QC in the section on data variables, and the note on quality control therein.

Type, Name, Dimension, Attributes	Comment
Byte TIME_QC(TIME);	Quality flag for each TIME value.
Byte POSITION_QC(PPOSITION);	Quality flag for each LATITUDE and LONGITUDE value.
Byte DEPTH_QC(DEPTH);	Quality flag for each DEPTH value.

### 3.2.3.Data variables

Data variables contain the actual measurements and indicators about their quality, uncertainty, and mode through which they were obtained. There are different options as to how the indicators are specified, whether in attributes or separate variables, which are outlined in the notes below the table. The variable names are standardized in reference table 3; replace <PARAM> with any of the names indicated there. Mandatory attributes are marked as such, however, OceanSITES requests that all other attributes be used and contain meaningful information unless technical reasons make this impossible.

Type, Name, Dimension, Attributes	Comment
Float <PARAM>(TIME, DEPTH); <PARAM>:long_name = "Y"; <PARAM>:standard_name = "<X>"; <PARAM>:units = "<Y>"; <PARAM>:_FillValue = <Y>; <PARAM>:QC_indicator = <X>; <PARAM>:QC_procedure = <X>; <PARAM>:valid_min = <Y>; <PARAM>:valid_max = <Y>; <PARAM>:comment = "<Y>"; <PARAM>:sensor_depth = <Y>; <PARAM>:ancillary_variables = "<Y>" ; <PARAM>:uncertainty = <Y>; <PARAM>:accuracy = <Y>; <PARAM>:precision = <Y>; <PARAM>:resolution = <Y>; <PARAM>: cell_methods = "<Y>"; <PARAM>:DM_indicator = "<Y>"	Contains data values. <PARAM> names are defined in reference table 3. Examples: TEMP, PSAL, DOXY.  These 3 attributes are mandatory : standard_name, units and _FillValue. The others attributes are highly recommended if available.  <X> : standardized attributes listed in reference tables <Y> : attributes whose value is set by the PI (Principal Investigator)  <b>long_name</b> : type char <b>standard_name</b> : type char, see reference. table 3 <b>units</b> : type char, see reference table 3 <b>_FillValue</b> : type float, see reference table 3 <b>QC_indicator</b> : type byte, see reference table 2 and note on quality control below <b>QC_procedure</b> type byte, see reference table 2.1 and note on quality control below <b>valid_min</b> : type float. Minimum value for valid data <b>valid_max</b> : type float. Maximum value for valid data <b>comment</b> . type char. Any free-format text with comments as appropriate. <b>sensor_depth</b> . type float. Nominal sensor depth(s) in meters, counting positive as per DEPTH:positive. <b>ancillary_variables</b> . type char. Other variables associated with <PARAM>, e.g. <PARAM>_QC. List as space-separated string. Example: TEMP:ancillary_variables="TEMP_QC TEMP_DM TEMP_UNCERTAINTY" <b>uncertainty</b> : type float. Overall measurement uncertainty, if constant. Cf. note on uncertainty below. <b>accuracy</b> : type float. Nominal sensor accuracy. Cf. note on uncertainty below. <b>precision</b> : type float. Nominal sensor precision. Cf. note on uncertainty below. <b>resolution</b> : type float. Nominal resolution of this data parameter. <b>cell_methods</b> : type char. Specifies cell method as per CF convention. Example: TEMP:cell_methods="TIME: point DEPTH: point LATITUDE: point LONGITUDE: point" Values are listed in table 2.2 <b>DM_indicator</b> : Type char. Data mode, if constant, as per reference table 5. Cf. note on data modes below.

<pre>Byte &lt;PARAM&gt;_QC(TIME, DEPTH); &lt;PARAM&gt;_QC:long_name = "quality flag"; &lt;PARAM&gt;_QC:conventions = "OceanSITES reference table 2"; &lt;PARAM&gt;_QC:_FillValue = -128; &lt;PARAM&gt;_QC:valid_min = 0; &lt;PARAM&gt;_QC:valid_max= 9; &lt;PARAM&gt;_QC:flag_values = 0, 1, 2, 3, 4, 5, 7, 8, 9; &lt;PARAM&gt;_QC:flag_meanings = "no_qc_performed good_data probably_good_data bad_data_that_are_potentially_correcta ble bad_data value_changed not_used nominal_value interpolated_value missing_value"</pre>	<p>Contains quality flags for values of associated &lt;PARAM&gt;. Flag scale is specified in reference table 2, and is included in the flag_meanings attribute.</p> <p><b>long_name:</b> type char. fixed value  <b>conventions.</b> type char. Required; fixed value  <b>_FillValue.</b> type byte. Required; fixed value  <b>valid_min.</b> type byte. Required; fixed value  <b>valid_max:</b> type byte. Required; fixed value  <b>flag_values</b> type byte. Required; fixed value  <b>flag_meanings:</b> type char. Required; fixed value</p>
<pre>Char &lt;PARAM&gt;_DM(TIME, DEPTH); &lt;PARAM&gt;_DM:long_name = "method of data processing "; &lt;PARAM&gt;_DM:conventions = " OceanSITES reference table 5"; &lt;PARAM&gt;_DM:flag_values = "R", "P", "D", "M"; &lt;PARAM&gt;_DM:flag_meanings = "real- time post-recovery delayed-mode mixed"; &lt;PARAM&gt;_DM:_FillValue = " ";</pre>	<p>Data mode. Indicates if the data point is real-time, delayed-, or post-recovery mode. It is included when the dataset mixes modes for a single variable.  See note on data modes below, and reference table 5.</p> <p><b>long_name:</b> type char. Required; fixed value:  <b>conventions:</b> type char. Required; fixed value:  <b>flag_values</b> type char. Required; fixed value:  <b>flag_meanings:</b> type char. Required; fixed value:  <b>_FillValue:</b> type char. Required; fixed value</p>
<pre>Float &lt;PARAM&gt;_UNCERTAINTY(TIME, DEPTH): &lt;PARAM&gt;_UNCERTAINTY:long_name = "uncertainty" &lt;PARAM&gt;_UNCERTAINTY:_FillValue= &lt;Y&gt; &lt;PARAM&gt;:units = "&lt;Y&gt;";</pre>	<p>Overall uncertainty of the data given in &lt;PARAM&gt;.</p> <p>Cf. note on uncertainty below.  <b>long_name:</b> type char. Required; fixed value  <b>_FillValue:</b> type float. Required.  <b>units:</b> type char. Required. Must be the same as &lt;PARAM&gt;:units.</p>

### Note on quality control (QC)

The quality of the data in a variable <PARAM> is described by the attribute <PARAM>:QC\_procedure, and one of the following: the attribute <PARAM>:QC\_indicator or the variable <PARAM\_QC>.

<PARAM>:QC\_procedure is mandatory and contains values from reference table 2.1, which describe what kind of quality control procedure has been applied.

It is mandatory to define one of <PARAM>:QC\_indicator or <PARAM\_QC>, both of which would contain values describing the data quality as per reference table 2. If the quality is the same for all <PARAM>, use <PARAM>:QC\_ind icator. Else, use <PARAM\_QC>.

### Note on uncertainty

If the overall measurement uncertainty for a variable <PARAM> is reasonably well-known, it must be provided in the attribute <PARAM>:uncertainty if it is constant, or in a variable of its own, <PARAM>\_UNCERTAINTY, if it is not constant. If uncertainty is given in either way, the attribute <PARAM>:accuracy is optional.

If it is impossible to estimate the overall measurement uncertainty, it is required to define at least the

attribute <PARAM>:accuracy with the nominal sensor accuracy.

The attributes <PARAM>:precision and <PARAM>:resolution are optional; they contain the sensor precision and resolution if defined.

### Note on data modes (DM)

Data mode may be represented as a global attribute "data\_mode" if all data is a single mode, or as an attribute to a variable <PARAM>:DM\_indicator if all data for <PARAM> is in a single mode. If a parameter contains a mixture of modes, these attributes should be set to "M" and the actual data modes should be represented by an extra variable, <PARAM>\_DM. The values for the data modes are explained in reference table 5.

### Example for sea temperature measurements and associated quality flags

```
Float TEMP(TIME, LEVEL);
TEMP:long_name = "sea water temperature in-situ ITS-90 scale";
TEMP:standard_name = "sea_water_temperature";
TEMP:units = "degree_Celsius";
TEMP:_FillValue = 99999.f;
TEMP:QC_indicator = 1;
TEMP:valid_min = -2.0f;
TEMP:valid_max = 40.f;
TEMP:comment = "";
TEMP:ancillary_variables = "TEMP_QC" ;
TEMP:resolution = 0.001f;
TEMP:uncertainty = 0.01f;
TEMP:error = 0.02f;
TEMP:cell_methods="TIME: point DEPTH: point LATITUDE: point
LONGITUDE: point";TEMP:DM_indicator="P";
```

```
Byte TEMP_QC(TIME, LEVEL);
TEMP_QC:long_name = "quality flag";
TEMP_QC:conventions = "OceanSITES reference table 2";
TEMP_QC:flag_values = 0, 1, 2, 3, 4, 5, 7, 8, 9;
TEMP_QC:flag_meanings = "no_qc_performed good_data probably_good_data
bad_data_that_are_potentially_correctable bad_data value_changed not_used nominal_value
interpolated_value missing_value"
TEMP_QC:_FillValue = 0;
```



## 4.OceanSITES metadata format

The OceanSITES data management team is developing a metadata format, probably using a SensorML implementation of XML, to allow metadata to be stored external to the NetCDF files.

## 5.Reference tables

### 5.1.Reference tables 1: data type and data code

#### 5.1.1.Data type

The data\_type global attribute should have one of the following valid values:

Data type
OceanSITES metadata
OceanSITES profile data
OceanSITES time-series data
OceanSITES trajectory data

#### 5.1.2.Data code

Data codes are used for file naming convention in chapter 6.1.1 and describe what kind of measurements the file contains.

They consist of the following letters:

Data Code	Meaning
C	Conductivity (electric conductivity of sea water)
M	Meteorological parameters (e.g. wind, air pressure, radiative fluxes, air temperature)
O	Oxygen concentration (in sea water)
S	Salinity (of sea water)
T	Temperature (of sea water)
V	Velocity (of sea water)

## 5.2. Reference table 2: quality control flag scale

The quality control flags indicate the data quality of the data values in a file, and are normally assigned after quality control procedures have been performed. These codes are used in the <PARAM>\_QC variables to describe the quality of each measurement, or in the attribute <PARAM>:QC\_indicator to describe the overall quality of the parameter.

Code	Meaning	Real-time comment	Post-recovery mode comment
0	No QC was performed	-	-
1	Good data	All real-time QC tests passed.	All post-recovery tests passed.
2	Probably good data	-	-
3	Bad data that are potentially correctable	These data are not to be used without scientific correction.	Possible re-calibration needed.
4	Bad data	Data have failed one or more tests.	Data have failed one or more tests.
5	Value changed	Data may be recovered after transmission error.	-
6	Not used	-	-
7	Nominal value	Data were not observed but reported; i.e. an anchor drop position or an instrument target depth, as opposed to a GPS position time series or a pressure sensor time series.	-
8	Interpolated value	Missing data may be interpolated from neighbouring data in space or time.	-
9	Missing value	-	-

### 5.2.1. Reference table 2.1: quality control procedure indicator

This table describes the quality procedures applied to all the measurement of a parameter. These values are used as an overall quality indicator (i.e. one summarizing all measurements) in the attributes of each variable <PARAM>.

Cf. <PARAM>:QC\_procedure attribute description in section 3.2.3.

Code	Meaning
0	No QC performed
1	Ranges applied, bad data flagged
2	Data interpolated
3	Sensor malfunctioning (data possibly useful)
4	Data missing
5	Data manually reviewed
6	Data verified against model or other contextual information

### 5.2.2. Reference table 2.2: cell methods

From NetCDF Climate and Forecast (CF) Metadata Conventions, Version 1.2, 4 May, 2008

In the *Units* column, *u* indicates the units of the physical quantity before the method is applied.

Cell Method	Units	Description
point	u	The data values are representative of points in space or time (instantaneous).
sum	u	The data values are representative of a sum or accumulation over the cell.
maximum	u	Maximum
median	u	Median
mid_range	u	Average of maximum and minimum
minimum	u	Minimum
mean	u	Mean (average value)
mode	u	Mode (most common value)
standard_deviation	u	Standard deviation
variance	u <sup>2</sup>	Variance

### 5.3. Reference table 3: OceanSITES parameter dictionary

#### Convention for parameter names, standard names and units

The parameter dictionary is now contained in an Excel file, which is available at [http://www.oceansites.org/parameters/...](http://www.oceansites.org/parameters/)

In the near future, it should be available in XML format.

- Parameter names should start with a 4 characters code based on SeaDataNet-BODC parameter discovery vocabulary.  
They are not strictly standardized, however.

When necessary, a parameter name has a suffix that designates secondary parameters<sup>1</sup>. The suffix starts with the character "\_".

- The NetCDF "standard\_name" attribute contains the standardized parameter name from CF conventions.
- The NetCDF "units" attribute are compliant with Uduunits as implemented in the CF/COARDS standards.

As the parameter names are not strictly standardized, one should use the standard\_name attribute to query a particular measurement from different data files.

### **Example**

On a mooring, sea temperature measured by a series of Microcat CTD is reported as TEMP, with a standard name of SEA\_WATER\_TEMPERATURE.

Secondary temperature measurement performed by an oxygen sensor is reported as DOXY\_TEMP with a standard name of temperature\_of\_sensor\_for\_oxygen\_in\_sea\_water.

For both measurements, the unit attribute is "degree\_Celsius".

### **References**

The OceanSITES standard names are taken from the CF standard names, available at:

- <http://cf-pcmdi.llnl.gov/documents/cf-standard-names/>

The parameter names are based on SeaDataNet-BODC parameter discovery vocabulary available at:

- [http://seadatanet.maris2.nl/v\\_bodc\\_vocab/welcome.aspx](http://seadatanet.maris2.nl/v_bodc_vocab/welcome.aspx)  
Select P021, " BODC Parameter Discovery Vocabulary"

The units are compliant with Uduunits, as implemented by the CF standard; definitions are available at:

- <http://www.unidata.ucar.edu/software/udunits>

The valid parameter names, standard names, long\_names and units are available at:

---

- <http://www.ifremer.fr/co/etc/oceansites/oceansites-user-manual-parameters.pdf>
- <http://www.ifremer.fr/co/etc/oceansites/oceansites-user-manual-units.pdf>

The following list an extract of the OceanSITES parameter dictionary for parameter name and standard name.

Parameter	Standard name
AIRT	air_temperature
ATMP	air_pressure
ATMS	air_pressure_at_sea_level
CDIR	direction_of_sea_water_velocity was sea_water_direction
CNDC	sea_water_electrical_conductivity
CSPD	sea_water_speed
DEPH	depth
DEWT	dew_point_temperature
DOX2	moles_of_oxygen_per_unit_mass_in_sea_water was dissolved_oxygen
DOXY	mass_concentration_of_oxygen_in_sea_water was dissolved_oxygen
DOXY_TEMP	temperature_of_sensor_for_oxygen_in_sea_water
DRYT	dry_bulb_temperature
DYNHT	dynamic_height
EWCT	eastward_sea_water_velocity
FLU2	fluorescence
HCSP	sea_water_speed
HEAT	heat_content
ISO17	isotherm_depth
LW	surface_downwelling_longwave_flux_in_air
NSCT	northward_sea_water_velocity
OPBS	optical_backscattering_coefficient
PCO2	surface_partial_pressure_of_carbon_dioxide_in_air was CO2_partial_pressure_in_dry/wet_gas
PRES	sea_water_pressure
PSAL	sea_water_salinity
RAIN	rainfall_rate
RAIT	thickness_of_rainfall_amount
RELH	relative_humidity
SDFA	surface_downwelling_shortwave_flux_in_air
SRAD	isotropic_shortwave_radiance_in_air
SW	surface_downwelling_shortwave_flux_in_air

TEMP	sea_water_temperature
TIME	time
UCUR	eastward_sea_water_velocity
UWND	eastward_wind
VAVH	sea_surface_wave_significant_height
VAVT	sea_surface_wave_zero_upcrossing_period
VCUR	northward_sea_water_velocity
VDEN	sea_surface_wave_variance_spectral_density
VDIR	sea_surface_wave_from_direction
VWND	northward_wind
WDIR	wind_to_direction
WSPD	wind_speed
WSPD	wind_speed

#### **5.4.Reference table 4: Data Assembly Center Codes**

Data Assembly Centers and institutions	
EUROSITES	EuroSites EU project
JAMSTEC	Jamstec, Japan
MEDS	MEDS, Canada
NDBC	National Data Buoy Center, USA
NOCS	National Oceanography Centre, Southampton
PMEL	PMEL, USA
SIO	SIO, Scripps, USA
WHOI	Woods Hole Oceanographic Institution, USA

#### **5.5.Reference table 5: data mode**

The values for the variables "<PARAM>\_DM", the global attribute "data\_mode", and variable attributes "<PARAM>:DM\_indicator" are defined as follows:

Value	Meaning
"R"	Real-time data. Data coming from the (typically remote) platform through a communication channel without physical access to the instruments, disassembly or recovery of the platform. Example: for a mooring with a radio communication, this would be data obtained through the radio.

"P"	Post-recovery data. Data obtained after the instruments or the platform have been recovered or serviced. Example: for instruments on a mooring, this would be data downloaded directly from the instruments after the mooring has been recovered on a ship.
"D"	Delayed-mode data. Data published after all calibrations and quality control procedures have been applied. This is the final status of the data.
"M"	Mixed. This value is only allowed in the global attribute "data_mode" or in attributes to variables in the form "<PARAM>:DM_indicator". It indicates that the file contains data in more than one of the above states. In this case, the variable(s) <PARAM>_DM specify which data is in which data mode.

## 5.6.Reference table 6: OceanSITES sites catalog

The OceanSITES sites catalog is available at [http://www.oceansites.org/parameters/..](http://www.oceansites.org/parameters/)

The following list is an extract from the catalog.

OceanSITES moorings catalogue based on data files available from GDAC				
site code	platform code	WMO code	DAC	contact
CIS	CIS-1	44478	EuroSites	NOCS
CIS	CIS-2	44478	EuroSites	NOCS
ESTOC	ESTOC-1	13471	EuroSites	NOCS
PAP	PAP-1	62442	EuroSites	NOCS
PAP	PAP-2	62442	EuroSites	NOCS
BTM	BTM	-	NDBC ?	UCSB
KEO	KEO32N145E	28401	PMEL	Cronin
MOVE	V404	41173	NDBC ?	SIO
MOVE	V405	-	NDBC ?	SIO
MOVE	V406	41174	NDBC ?	SIO
NTAS	NTAS	-	NDBC ?	Woods Hole
STRATUS	STRATUS	-	NDBC ?	Woods Hole
TAO	0n110w	32323	PMEL	McPhaden
TAO	0n140w	51311	PMEL	McPhaden
TAO	0n165e	52321	PMEL	McPhaden
TAO	0n170w	51010	PMEL	McPhaden
PIRATA	0n23w	15004	PMEL	McPhaden
RAMA	0n80.5e	23001	PMEL	McPhaden
PIRATA	10s10w	15001	PMEL	McPhaden
PIRATA	15n38w	13008	PMEL	McPhaden

## 6.GDAC organization

There are two GDACs (global data assembly centers) for redundancy, which are the users' access points for OceanSITES data. One GDAC is located in France (Coriolis, <http://www.coriolis.eu.org>), the other one in the USA (NDBC, National Data Buoy Center, <http://www.ndbc.noaa.gov>). The GDACs handle OceanSITES data, metadata, and index files on ftp servers. The servers at both GDACs are synchronized at least daily to provide the same OceanSITES data.

The user can access the data at either GDAC's ftp site:

- <ftp://ftp.ifremer.fr/ifremer/oceansites>
- <ftp://data.ndbc.noaa.gov/data/oceansites>

From these root directories of the GDACs downward, the organization of the directories and files is:

- `site/platform/OS_XXX_YYY_ZZZ<_PARTX>.nc`
- site: OceanSites site code
- platform: platform code

The site and platform codes are available in the "OceanSITES catalogue" document at either GDAC's root directory:

- `oceansites_catalogue.xml`

### 6.1. File naming convention

The OceanSITES file names use the following naming conventions:

#### 6.1.1. Data file naming convention

`OS_XXX_YYY_ZZZ<_PARTX>.nc`

- OS: OceanSITES prefix
- XXX: platform code from the OceanSITES catalogue
- YYY: configuration code
- ZZZ: list of "data codes" from reference table 1  
The data code is a list of the primary parameters measured. The data code will not list secondary parameters.
- <\_PARTX>: when an OceanSites data file size becomes excessive (eg: > 100Mb), it can be split in smaller parts: PART1, PART2, ... PARTN

Example:



- OS\_CIS-1\_200502\_TS.nc

This file contains temperature and salinity data from the CIS-1 mooring, from the Animate project, for the deployment performed in February 2005. Giving the deployment date (here: 200502) as the configuration code (YYY) is the recommended method if multiple site occupations exist, and if this makes identification unique.

### **6.1.2.Metadata file naming convention**

OS\_XXX\_YYY\_META.xml

- OS: OceanSITES prefix
- XXX: platform code
- YYY: configuration code

These metadata are for the corresponding data file(s) OS\_XXX\_YYY\_ZZZ<\_PARTX>.nc.

Example:

- OS\_CIS-1\_200502\_META.xml

This file contains the metadata of CIS-1 mooring, for the deployment performed in February 2005.

### **6.2.Index file for data files**

To allow for data discovery without downloading the data files themselves, an index file is created at the GDAC level, which lists all available data files and the location and time ranges of their data contents:

- The data index file is located at the root directory of the GDAC.
- The index file contains the list and a description of all data files available on the GDAC.
- There is a header section, lines of which start with # characters.
- The information sections are comma-separated values.
- Each line contains the following information:
  - FILE: the file name, beginning from the GDAC root directory
  - DATE\_UPDATE: the update date of the file, YYYY-MM-DDTHH:MI:SSZ
  - START\_DATE: first date for observations, YYYY-MM-DDTHH:MI:SSZ
  - END\_DATE: last date for observations, YYYY-MM-DDTHH:MI:SSZ

- SOUTHERN\_MOST\_LATITUDE
  - NORTHERN\_MOST\_LATITUDE
  - WESTERN\_MOST\_LONGITUDE
  - EASTERN\_MOST\_LONGITUDE
  - UPDATE\_INTERVAL: M monthly, D daily, Y yearly, V void
  - SIZE: the size of the file in megabytes
- The fill value is empty (eg. : ,).

GDAC data files index: oceansites\_files\_index.txt

```
# OceanSites FTP GLOBAL INDEX
# ftp://ftp.ifremer.fr/ifremer/oceansites
# Contact: http://www.oceansites.org
# Index update date YYYY-MM-DDTHH:MI:SSZ: 2008-03-30T18:37:46Z
#
#FILE,DATE_UPDATE,START_DATE,END_DATE,
SOUTHERN_MOST_LATITUDE,NORTHERN_MOST_LATITUDE,WESTERN_MOST_LONGITUDE,EASTER
N_MOST_LONGITUDE,UPDATE_INTERVAL,SIZE
TAO/0n170w/OS_0n170w_SW_LW_2m.nc,2008-04-12T08:05:00Z,2007-03-17T18:07:00Z,2008-04-
12T08:05:00Z,0,0,-170,-170,M,16.7
```