

NODC Submission Information Form

(V1.3, Revised 01/2009)

FORM APPROVAL PENDING

Following the definitions and principles of the *Open Archival Information System (OAIS) Reference Model (ISO 14721:2003)*, this form documents the mutual understanding between a *Producer*, defined as a person or organization who provides information to be preserved, and an *Archive*, defined as the organization that intends to preserve information for access and use over the long term. It should accompany all data submissions to the National Oceanographic Data Center (NODC) and be completed to the extent possible.

The information contained on this form may be used to:

1. Populate NODC's Accession Tracking Data Base and product-specific databases
2. Create metadata records conforming to the Content Standard for Digital Geospatial Metadata (CSDGM), Vers. 2 (FGDC-STD-001-1998) and/or ISO 19115
3. Generate a formal archive appraisal package, for submissions requiring management level approval
4. Develop a list of *Producer* requirements requested of the *Archive*

The information contained on this form is true and correct to the best understanding of the *Producer* and *Archive* at the time of its submission. In the future, this information may be amended, updated, or revised as necessary and some submissions may require management level approvals before archival services can be provided.

Instructions:

This form is divided into six sections. Section 1 provides only the most basic Tracking Information and is the only section absolutely required at the time of submission. **However, within 1 month all submissions must also include information for Section 2**, which provides basic Data Discovery and Usage Information, and all submissions should strive to provide information through Sections 3 and 4, which provide more comprehensive and detailed information on the data set and its requirements for long term stewardship. Section 5 is required for submissions to the *Archive* that are expected to be periodic or routine in nature, and supports automation of archival services. Section 6 is optional and is only required for submissions that are expected to need management level approval and a formal archive appraisal package. The *Archive* Point of Contact will provide guidance as needed on all of these questions and will work with the *Producer* to ensure both parties reach a mutual understanding.

When complete, please email the signed form (see the last page of the document) to the *Archive* Point of Contact with a copy to NODC.DataOfficer@noaa.gov. Closing the email with "Signed," followed by your name is an acceptable form of signature.

NODC Submission Information Form

Section 1 – Basic Tracking Information

All elements in this section are REQUIRED, and will enable the *Archive* to establish a unique and durable tracking number known as an NODC Accession Number for the submission. It also clearly establishes whether the *Archive* is able to freely redistribute the data, and if not, what the restrictions are. The *Producer* will be provided the Accession Number in a confirmation receipt, along with the web address where the *Producer* can access the data set. When Section 1 is complete, the submission is considered “Initialized”.

1. Date of submission of this form (or its update): **26 February 2009**
2. Describe the scope of this data submission information.

Long-term stewardship at NODC for CoastWatch Level 2 Ocean Color products, including data from SeaWiFS, MODIS and MERIS instruments, using the CLASS system.

3. What is the Data Set title? A useful title includes a listing of two or three of the observed variables, the name of one or two of the platforms used to collect data or the project responsible for the data collection activity, the location, and the range of observation dates. For example, “Temperature, salinity, and nutrient data from bottle casts from the *Akademic Korolev*, *Alpha Helix*, *Polar Star*, and *Surveyor* in the Bering and East Siberian Seas from 1987-1999.”

NOAA CoastWatch Level 2 Ocean Color Data from the SeaWiFS, MODIS and MERIS Instruments.

4. Primary Point of Contact for *Producer* – please provide name, organization, position, address, telephone, fax, and e-mail address.

Linda Stathoplos, Oceanographer
NESDIS/OSDPD/SSD/PIB
5200 Auth Road, WWBG Building Rm 510
Camp Springs, MD 20746-4325
Phone: (301)763-8142 x137
Fax: (301)899-9196
Linda.Stathoplos@noaa.gov

5. Primary Point of Contact for *Archive* – please provide name, organization, position, address, telephone, fax, and e-mail address.

Rick Vizbulis, CLASS Program Manager
NESDIS/Office of Systems Development
4231 Suitland Road, Rm 1160
Suitland, MD 20746
Phone: (301)817-4603
Fax: (301)817-4403

NODC Submission Information Form

Rick.Vizbulis@noaa.gov

6. Can NODC freely and openly redistribute this dataset? If no, list the *Producer's* constraints of the Data Set in the *Archive* for Users in terms of:
 - a. User access to the Data Set
MODIS: no restrictions.
MERIS data may not be distributed to anyone besides authorised ESA MERIS users. The updated list of authorised ESA users is available at <http://eopi.esa.int/search> (select the country, then search within list of ESA US PI's). CoastWatch users not yet registered by ESA can do it easily at <http://eopi.esa.int> (select registration).
SeaWiFS data may not be distributed to anyone besides users covered by the SeaWiFS CoastWatch contract. Richard Stumpf (see Section 4, Question 7) is the Point-of-Contact for information on those covered by the CoastWatch contract.
 - b. Uses of the Data Set by Users
Not intended for legal use. Data may contain inaccuracies due to clouded or mixed pixels.

NODC Submission Information Form

Section 2 – Basic Data Discovery and Usage Information

All elements in this section are REQUIRED if applicable and allow the *Archive* to enable users to find, access, and use the data described by this submission. It is strongly recommended that this information be provided at the time of the submission though it is acceptable if some elements are provided within **one month** of Initialization. When Section 2 is complete, the submission is considered “Active”.

1. What is the purpose for collecting this Data Set?

Biologists are using ocean color (e.g. chlorophyll-a) products to predict harmful algal blooms, manage living marine resources, and assess ocean climate effects. CoastWatch provides near real-time ocean color products from remote-sensing platforms in order to serve the efforts of these users, as well as the efforts of federal, state and local government environmental decision makers. Other users include recreational boaters, educators, resource managers, and researchers.

2. Provide a general descriptive abstract about the Data Set.

Research and monitoring programs are generally interested in live phytoplankton biomass and the suspended sediment concentration. For sediment concentrations, a measure of scatter provides the best information. At present, experimental backscattering algorithms exist, however, additional work is necessary to validate these algorithms. As an alternative, using the remote sensing reflectance for bands that are most sensitive to scatter (and least sensitive to absorption) provides an estimate of the live phytoplankton biomass in the surface layer and is derived from the amount of water-leaving radiances at various wavelengths.

CoastWatch provides near real-time ocean color products, e.g. chlorophyll-a, with data from NASA’s Earth Observing Satellite (EOS), Aqua, as well as GeoEye’s Orbview-2 satellite (via contractual purchase) and the European Space Agency’s (ESA) Envisat satellite. Each satellite-specific product has unique characteristics dependent on the sensor and algorithm applied.

3. What is the time period covered by the Data Set?

CoastWatch retains Level 2 ocean color products for 5-10 days, so the time period covered by the L2 products will be the day archival storage begins (*according to agreement*) minus 5-10 days. Archive will continue forward in time indefinitely.

4. What is the geospatial coverage of the Data Set (Easternmost longitude, westernmost longitude, northernmost latitude, southernmost latitude)? Note western longitudes and southern latitudes are negative, and use decimal degrees if possible.

CoastWatch ocean color products cover 13 regions (7 coastal, 6 open ocean) globally. SeaWiFS ocean color products cover 10 of these 13 regions; MODIS ocean color products cover all 13; and MERIS products cover 7 of the 13 regions. The following table defines the bounding coordinates for each of the 13 regions and identifies the ocean color product(s) available for each:

NODC Submission Information Form

Region	UL Longitude	UL Latitude	LR Longitude	LR Latitude	Available Products
Northeast	-82.0	46.0	-60.0	30.0	MODIS, SeaWiFS, MERIS
Southeast	-88.0	37.0	-72.0	22.0	MODIS, SeaWiFS, MERIS
Caribbean	-80.0	30.0	-60.0	8.0	MODIS, SeaWiFS, MERIS
Gulf of Mexico	-99.0	31.0	-79.0	17.0	MODIS, SeaWiFS, MERIS
Great Lakes	-93.0	51.0	-75.0	38.0	MODIS, SeaWiFS, MERIS
West Coast	-142.0	51.0	-112.0	29.0	MODIS, SeaWiFS, MERIS
Eastern Tropical Pacific	-142.0	30.0	-80.0	8.0	MODIS, SeaWiFS, MERIS
Alaska	-180.0	64.0	-126.0	50.0	MODIS, SeaWiFS
Hawaii	-167.0	29.0	-147.0	10.0	MODIS, SeaWiFS
Pacific Basin	-180.0	50.0	-140.0	0.0	MODIS, SeaWiFS
Equatorial Atlantic	-60.0	30.0	0.0	-10.0	MODIS
North Atlantic	-60.0	60.0	0.0	20.0	MODIS
Chesapeake Bay	-78.0	40.0	-74.0	36.0	MODIS

5. List the measured variables or parameters in the Data Set (e.g., Temperature, Salinity, etc.)

Ocean color (chlorophyll concentration).

6. List the platform(s) from which the Data Set is derived.

OrbView-2 satellite (SeaWiFS); Aqua and Terra* satellites (MODIS); Envisat satellite (MERIS).

***backup only**

7. List the instrument(s) used to derive the Data Set.

**Sea-viewing Wide Field-of-view Sensor (SeaWiFS)
Moderate Resolution Imaging Spectroradiometer (MODIS)
Medium Resolution Imaging Spectroradiometer (MERIS)**

8. List the observation types in the Data Set (e.g., Biological Data, Physical Data, etc.).

Biological data.

9. List the mission/project name(s) to which the Data Set contributes.

OrbView-2 (owned and operated by GeoEye), ongoing since 1 August 1997; NASA Aqua, ongoing since April 2002; NASA Terra, ongoing since February 2000; ESA Envisat, ongoing since March 2002; CoastWatch Ocean/Ocean Optics/Ocean Color.

NODC Submission Information Form

10. Give the expected size(s) in bytes and number of files in the submission.

File type	Format	Number of files	Total file size (MB)
SeaWiFS total	HDF, md5	39	5405.1
General_Purpose	HDF	7	1229.89
General_Purpose_(Stumpf)	HDF	6	1405.41
General_Purpose	HDF (CW)	7	1459.54
General_Purpose	md5	7	0.000348
General_Purpose_(Stumpf)	HDF (CW)	6	1310.25
General_Purpose_(Stumpf)	md5	6	0.00029
MODIS Aqua total	HDF, md5	87	9093.9
General_Purpose	HDF	29	5581.85
General_Purpose	HDF (CW)	29	3512
General_Purpose	md5	29	0.001711
MERIS total	HDF, md5	8	4800
General Purpose	HDF	4	4800
General Purpose	md5	4	0.002
All total	HDF, md5	134	19299

11. Give the file format and format version (e.g., netCDF-3, HDF-5, ASCII CSV, etc.).

Hierarchical Data Format (HDF-4).

12. Does this Data Set conform to any file-level data content or metadata content standards? (e.g., COARDS/CF, HDF-EOS, WOCE, GHRSSST)

CoastWatch Hierarchical Data Format (CW HDF) – combination metadata and file content standard (http://coastwatch.noaa.gov/cw_form_hdf.html). Metadata for these files are FGDC-STD-001-1998 compliant (see question 16 in this section).

13. Please describe the file contents. Include enough information to make these data understandable to future users. For example, a table containing as applicable: parameter definition, data type, byte size/length, scale factor, offset, precision, and units. This information is especially important for ASCII and other formats which are not self-describing like netCDF and HDF. If this information is already contained in a file or file headers included in this submission, please indicate the file name.

CoastWatch Ocean Color L2 product files may contain any or all of the following variables within a given file:

Variable	Description	Type	Units	Scale	Offset
chlor_a	Chlorophyll a concentration	float	mg m ⁻³	1	0
algal_1	Chlorophyll a concentration	float	mg m ⁻³	1	0
algal_2	Chlorophyll a concentration	float	mg m ⁻³	1	0

NODC Submission Information Form

K_490	Diffuse attenuation coefficient at 490 nm	int16	m ⁻¹	0.0002	0
Rrs_667/670	Remote sensing reflectance at 667/670 nm	float	sr ⁻¹	1	0
Rrs 412/885	Remote sensing reflectance at 412/885 nm	float	sr ⁻¹	1	0
Es_667/670	Extraterrestrial solar irradiance	float	mW cm ⁻² um ⁻¹	1	0
rhos_869	Surface reflectance	int16	count	0.0001	0
l2_flags	Quality flags	int32	-	1	0
nLw_412 – nLw_885	Directional normalized water-leaving radiance for various nm values; ranging from 412 to 885; available wavelengths vary with each instrument.	int16	mW cm ⁻² um ⁻¹ sr ⁻¹	0.001	0
sst	sea surface temperature – 11micron	int16	degrees celsius	0.005	0
latitude	latitude	float	degrees	1	0
longitude	longitude	float	degrees	1	0
sat_zenith	satellite zenith angle	int16	degrees	0.01	0
rel_azimuth	relative azimuth angle	int16	degrees	0.01	0
sun_zenith	solar zenith angle	int16	degrees	0.01	0
par	Photosynthetically Active Radiation	uint16	count	1	-32750
L1a band 1	SeaWiFS radiance count	uint16	count	1	0
L1a band 2	SeaWiFS radiance count	uint16	count	1	0
L1a band 3	SeaWiFS radiance count	uint16	count	1	0
L1a band 4	SeaWiFS radiance count	uint16	count	1	0
L1a band 5	SeaWiFS radiance count	uint16	count	1	0
L1a band 6	SeaWiFS radiance count	uint16	count	1	0
L1a band 7	SeaWiFS radiance count	uint16	count	1	0
L1a band 8	SeaWiFS radiance count	uint16	count	1	0
EV_Band8	Earth View Band 8 1KM reflective solar bands; scaled integers	uint16	count	1	0
EV_Band9	Earth View Band 9 1KM reflective solar bands; scaled integers	uint16	count	1	0
EV_Band10	Earth View Band 10 1KM reflective solar bands; scaled integers	uint16	count	1	0
EV_Band11	Earth View Band 11 1KM reflective solar bands; scaled integers	uint16	count	1	0
EV_Band12	Earth View Band 12 1KM reflective solar	uint16	count	1	0

NODC Submission Information Form					
	bands; scaled integers				
EV_Band13l o	Earth View Band 13low 1KM reflective solar bands; scaled integers	uint16	count	1	0
EV_Band13h i	Earth View Band 13high 1KM reflective solar bands; scaled integers	uint16	count	1	0
EV_Band14l o	Earth View Band 14low 1KM reflective solar bands; scaled integers	uint16	count	1	0
EV_Band14h i	Earth View Band 14high 1KM reflective solar bands; scaled integers	uint1	count	1	0
EV_Band15	Earth View Band 15 1KM reflective solar bands; scaled integers	uint16	count	1	0
EV_Band16	Earth View Band 16 1KM reflective solar bands; scaled integers	uint16	count	1	0

Dimensions for each of these variables (number of rows and columns) depends on the CoastWatch region covered, but will be the same for all variables within a given file.

14. Give the file-naming convention for the file(s) to be submitted, with the range/domain of each field value in the filename.

For CoastWatch HDF:

DataProduct [C]YYYYDDD_TTTT.hdf

DataProduct:

- a. SWFCW: SeaDAS standard SeaWiFS processed from L1
- b. SWRCW: Stumpf SeaDAS processed from L1
- c. MODSCW: SeaDAS standard MODIS processed from L1B
- d. MERCW: MERIS L2 ocean color file

[C]:

- S: SeaWiFS
- P: MODIS Aqua
- M: MERIS

YYYYDDD: four digit year + day of year

TTTT: four digit starting time (HHMM or hour + minutes) of granule

For Standard HDF (SeaWiFS and MODIS):

SeaWiFS:

YYYYDDDDHHMMSS.L2_SSS_[OC4,RS].hdf

YYYYDDDDHHMMSS: four digit year + day of year + hour + minute + second

SSS: HRPT station code

MODIS:

PYYYYDDDDHHMMSS.L2.hdf

YYYYDDDDHHMMSS: four digit year + day of year + hour + minute + second

NODC Submission Information Form

15. Please provide a list of existing reports, publications, user guides, web sites, or other supporting documentation relevant to the Data Set.

- a. **Okeanos System Interface Control Document (ICD), Version 2.02, June 2008**
- b. **Interface Control Document (ICD) Between CoastWatch Ocean Color and CLASS, Version 2.0, March 2006**
- c. **Producer-Archive Submission Agreement (SA) for Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Level 1A (L1A) Data between CLASS and CoastWatch Ocean Color, Version 1.0, March 2006**

16. What metadata exists for this Data Set? Is it in a standard format/can it be automatically translated into a standard format? Describe the granularity of this metadata (For example, is it collection level metadata? If not, to what file or grouping of files does it apply?)

Metadata for the products is available within the data set files (HDF), and is documented in the Okeanos System ICD. In addition, collection-level metadata is submitted for publishing and update to the NOAA Metadata Manager and Repository (NMMR). This is initiated by CoastWatch Ocean Color personnel. These are standard FGDC-compliant (FGDC-STD-001-1998) records which, with the use of a metaparser, are available in HTML, ASCII, SGML, XML, DTD, DIF, and plain text formats.

17. If applicable, describe the temporal resolution of the primary parameter(s) in the Data Set.

For SeaWiFS L1A and L2 data, a ‘granule’ is the entire portion of orbit visible by the ground station, which can be greater than 10 minutes. The SeaWiFS L2 products also collect the granules into daily and 61-day composites . MODIS-Aqua L1A files are received as individual 5-minute granules and processed to L2, and are also composited for an entire day (by region) and for 61 days (by region). The MODIS-Aqua Chesapeake Bay regional product is composited into calendar month averages instead of 61-day composites. MERIS products are composited for an entire day and for 61 days. MERIS N1 files are 2600 second granules (half orbit).

18. If applicable, describe the horizontal resolution of the primary parameter(s) in the Data Set.

**SeaWiFS and MODIS: Approximately 1.1 km per pixel at nadir.
MERIS: Approximately 1.2 KM (L2 reduced resolution)**

19. If applicable, describe the vertical resolution of the primary parameter(s) in the Data Set.

Not applicable.

20. If applicable, describe the projection grid or coordinate system used in the Data Set.

Original satellite projection.

21. If the Technical Contact for the *Producer* is different from the Primary Contact for the *Producer* (1.4), please provide name, organization, position, address, phone, fax, and email.

Same as 1.1.

NODC Submission Information Form

22. If the Metadata Contact for the *Producer* is different from the Primary Contact for the *Producer* (1.4), please provide name, organization, position, address, phone, fax, and email.

Heng Gu, Senior Programmer/Analyst
NOAA ORA (SPS Contractor)
5200 Auth Road
Camp Springs, MD 20746
Phone: (301)763-8209 x347
Heng.Gu@noaa.gov

23. If the Technical Contact for the *Archive* is different from the Primary Contact for the *Archive* (1.5), please provide name, organization, position, address, phone, fax, and email.

Jeremy Throwe, Contractor
NESDIS/OSD/GSD
Suitland Federal Center, NSOF Building, Rm 1449
Suitland, MD 20746
Phone: (301)817-4686
Jeremy.Throwe@noaa.gov

24. If the Metadata Contact for the *Archive* is different from the Primary Contact for the *Archive* (1.5), please provide name, organization, position, address, phone, fax, and email.

Tess Brandon, Oceanographer
NESDIS/National Oceanographic Data Center
1315 East West Highway
SSMC3 Rm 4716
Silver Spring, MD 20910
Phone: (301)713-3272 x181
Tess.Brandon@noaa.gov

NODC Submission Information Form

Section 3 – Detailed Data Processing and Quality Information

All elements in this section are **STRONGLY RECOMMENDED** and allow the *Archive* to enable more complete and thorough understanding of the data over the long term.

1. What is the overall completeness and quality of the Data Set?

The processing methods for NOAA CoastWatch ocean color products include several validation and quality control routines, which result in a high quality, complete set of products. Match-ups with in situ data and other sensors feed back into product development and sensor calibration. Regional comparisons are also performed. These validation routines are performed both during lower-level and Level 2 processing. Routine vicarious comparisons are performed as a means of near-real time quality control, and ad hoc checks and the incorporation of user feedback serve as retrospective quality control. Data quality in general is ensured through the program oversight of CoastWatch Program scientist Dr. Menghua Wang.

2. Describe the data processing level of the Data Set. For example, is the Data Set unprocessed or minimally processed, quality controlled or calibrated, etc.? For satellite data, is it Level 0, 1, 2, 3, or 4?

NOAA Level 2. Level 2 ocean color data is processed from Level 0 granules through Level 1, Level 1A and Level 1B.

3. Summarize the science algorithms(s) used to derive the Data Set.

Algorithms used for SeaWiFS:

For chlorophyll-a (OC4v4), SeaDAS4.6 for atmospheric correction; for chlorophyll-a (Stumpf), Stumpf Iter=2 + BMP for atmospheric correction (Southeast and Gulf of Mexico regions).

Algorithms used for MODIS:

Algorithm Theoretical Basis Documents (ATBD's) are found for MODIS through the NASA GSFC website at http://modis.gsfc.nasa.gov/data/atbd/ocean_atbd.php.

OC3-CB (also known as OC-SPIE): Chesapeake Bay regional algorithm is described in the following paper:

Werdell, P. Jeremy, et al., "Approach for the long-term spatial and temporal evaluation of ocean color satellite data products in a coastal environment," Proc. of SPIE Vol. 6680, 66800G (2007).

Algorithms used for MERIS:

Algorithm Theoretical Basis Documents (ATBD's) are found for MERIS through the ESA website at <http://envisat.esa.int/instruments/meris/atbd>.

4. Describe the steps taken to process the Data Set, including for each step the methodology, source data, and time/frequency, and listing any input data sets used to derive the Data Set.

NODC Submission Information Form

Level-2 processing is performed using the SeaDAS Multi-Sensor Level-1 to Level-2 (MSL12) code. This software is capable of retrieving oceanic optical properties and various derived products from the observed top-of-atmosphere (TOA) radiances collected by a variety of ocean remote sensing radiometers.. For MODIS, the Level-1B file and the geolocation file for each 5-minute granule are input into MSL12. For SeaWiFS, each Level 2 scene corresponds to a Level 1A scene. The Level-2 processing also makes use of meteorological (U and V winds, surface pressure, water vapor) and ozone information from ancillary sources. Before computing Level 2 data, pixels are eliminated if they contain clouds, sun glint, or other abnormalities. For pixels that pass these screens, an atmospheric correction is applied to subtract the atmospheric scattering components from the total radiance in order to obtain the water-leaving radiances.

For SeaWiFS:

NOAA produces CoastWatch SeaWiFS ocean color products daily in near-real time from data collected by two ground stations (East US and West US). Data is purchased from GeoEye under a NOAA Contract and therefore data is restricted to US civil marine operational applications. NOAA obtains Level-1A data from GeoEye and processes the data to Level-2 ocean products using NASA SeaDAS and NOAA algorithms. For a more detailed description of the processing steps, see http://oceancolor.gsfc.nasa.gov/DOCS/SW_proc.html.

For MODIS:

NOAA produces CoastWatch MODIS ocean color products daily in near-real time from global data collected by the NASA Goddard Space Flight Center (GSFC). These products are provisional and currently provided by NOAA on a “best effort” basis. NOAA obtains Level-0 data from NASA GSFC in 5-minute granules. These Level-0 data sets are processed to geolocated, calibrated radiances (L1b) and derived MODIS data products (Level 2) using NASA SeaDAS ‘msl12’ software. For a more detailed description of the processing steps, see http://oceancolor.gsfc.nasa.gov/DOCS/MODISA_processing.html.

For MERIS:

NOAA CoastWatch produces chlorophyll-a concentration products daily from the MERIS instrument aboard the Envisat spacecraft. Level 2 MERIS data are obtained from the European Space Agency. These data are converted to Level 2 CoastWatch HDF format.

5. Describe the Data Set’s dependency on other data (e.g. ancillary files), processing systems, software, or entities that are not to be submitted to the Archive.

SeaWiFS L2 files are dependent on L1A files delivered by the GeoEye SeaWiFS Data Server; MODIS L2 files are dependent on 5-minute L1A granule files and L1B geolocation files delivered by the NOAA-MODIS Data Server; MERIS L2 files are dependent on L2 files delivered by the ESA MERIS Data Server.

MODIS and SeaWiFS-derived products are dependent on the SeaDAS Multi-Sensor Level-1 to Level-2 (MSL12) software. Necessary ancillary data sets for input into MSL12 include meteorological (U and V winds, surface pressure, water vapor) and ozone information.

NODC Submission Information Form

6. ~~Detail any measures taken by the Producer to assess the quality of the Data Set, including data comparisons, and an assessment of the attribute accuracy. Give information about omissions, selection criteria, and other rules used to derive the Data Set.~~

As the data is processed, checks are made for different defined conditions. When certain tests and conditions are met for a given pixel, then a flag is applied to that pixel for that condition. This is done in the processing by setting the bit number assigned to that condition. If a certain flag exists for a pixel, it can be specified that a mask should be applied to it. During processing of Level 2, if a condition is set as a mask, then the flagged pixel will be set to zero and not valid. For a list of these flags, see question 7 below.

CoastWatch performs validation of the data through buoy match-ups or other in situ measurements for adjusting processing algorithms as necessary. A comparison is performed between the standard deviation, mean and median of the difference between science quality data and comparable NOAA products created using NCEP GFS ancillary data with <5% difference for chlorophyll and nLw values. Current CoastWatch products are <8% different compared with science quality products.

Chlorophyll concentrations are accurate to within +/- 35% in the range of 0.05 – 100 micrograms per liter.

7. List any quality assessment parameters included in the Data Set. For example, this may be an explanation of quality flags and their range/domain of values.

The table below shows the flags and masks that are operational in the Level 2 Ocean Color processing:

Bit	Name	Description
03	BADANC	Reduced quality of ancillary data
07	COASTZ	Pixel is in shallow water
08	NEGLW	Negative water-leaving radiance retrieved
12	TURBIDW	Turbid water detected
14	HITAU	High aerosol optical thickness
18	ABSAER	Possible absorbing aerosol (disabled)
19	TRICHO	Possible trichodesmium contamination
21	MODGLINT	Moderate sun glint contamination
24	DARKPIXEL	Rayleigh-subtraced radiances is negative
25	SEAICE	Possible sea ice contamination
28	SSTWARN	SST quality is reduced
29	SSTFAIL	SST quality is bad
30	HIPOL	High degree of polarization
31	spare	spare
32	OCEAN	not cloud or land

NODC Submission Information Form

Section 4 – Data Stewardship Information

All elements in this section are **STRONGLY RECOMMENDED** and enable the *Archive* to provide more comprehensive *data stewardship* over the long term. Data stewardship requires a more extensive set of functions than traditional long-term preservation of data and information, and includes activities such as monitoring the needs of user communities, compliance testing, quality assurance, and use of this Data Set in larger integrated product databases. Importantly, this section provides the *Producer* with an opportunity to request specific services from the *Archive*. This document does not imply that all of these services will be provided, but typically the *Archive* will work to meet them on a best-effort basis.

1. Please describe any quality control or quality assurance procedures the *Archive* should perform on this Data Set when it is submitted to the *Archive*.

None

2. How will the *Producer* provide updates to the *Archive* when changes occur in the Data Set, transmission mechanism, format, content, etc.? How often might such changes be expected to occur?

The Producer will provide updates to the Archive via the contacts listed in this document. It is presently unknown how often such changes might occur.

3. Does the *Producer* request reports on the *Archive*'s dissemination of the Data Set? If so, what statistics should be included? (Please note federal regulations strictly limit the amount and kind of information that can be recorded by federal agencies.)

Dissemination reports will be provided by the Archive upon request from the Producer. The details of these reports will be documented in an updated CoastWatch Ocean Color – CLASS ICD.

4. Does the *Producer* request standards compliance testing on the Data Set? For example, should the *Archive* verify data files are meeting netCDF Climate and Forecast (CF) conventions, or should metadata records be checked for adherence to the FGDC content standard? Will the *Producer* perform standards compliance testing prior to submission to the *Archive*?

No standards compliance testing is requested of the Archive. The Producer will ensure that files in the Data Set are compliant with the CoastWatch HDF standard.

5. Suggest action(s) for the *Archive* in the case of an error in transmission (e.g. missing data, duplicate data, incorrect file name or size, failure of compliance checks).

For FTP failure, retry FTP a configurable number of times over a configurable time period. For other errors including persistent FTP failures, contact Primary Contact for Producer (Section 1, Question 4).

NODC Submission Information Form

6. Please list any known NODC product databases (e.g. World Ocean Data Base) that this Data Set should become a part of.

None.

7. Please identify one or more Representative Users of the *Designated Community*. The Designated Community is defined in the OAIS Reference Model as the group of potential users who should be able to understand a Data Set over the long term. The *Archive* works specifically to preserve the data and information for this Designated Community.

a. Describe this user community and their requirements

Users in the NOAA National Ocean Service; require chlorophyll concentrations for use in coastal monitoring, habitat assessment, water quality assessment, harmful algal bloom (HAB) forecasts, carbon dioxide fluxes, eutrophication assessment, etc.

b. Provide Contact Information for a representative of this community - please provide name, organization, position, address, telephone, fax, and e-mail address

Richard Stumpf
NOAA NOS
1305 East West Highway
SSMC4 Rm 9115
Silver Spring, MD 20910-3278
Phone: (301)713-3028 x173
Fax: (301)713-4388
Richard.Stumpf@noaa.gov

8. List security requirements for dissemination of the Data Set from the *Archive* to the users.

Same as Section 1, Question 6a.

9. Once the Data Set is transferred to the *Archive*, how long should it take for it to become searchable? How long should it take to become accessible online?

2 hours.

10. Describe any preferred search criteria to be enabled for this Data Set in the *Archive* (e.g., search by time, search by geographic bounding box on a Polar Stereographic map, etc.)

No preference.

11. Describe any the preferred access mechanisms to be enabled for this Data Set in the *Archive* (e.g., OPeNDAP, Web Coverage Services, FTP, etc.)

Web browser, FTP.

NODC Submission Information Form

Section 5 – Logistics Information for Routine Transfers to the *Archive*

All elements in this section are REQUIRED for Data Sets that are expected to be routine, automated transmissions to the *Archive* from the *Producer*. This information is required for the *Archive* to establish and maintain the automated ingest and archive procedures. Questions 1 through 5 in Section 4 above are also required for automatic submissions.

1. Provide the mechanisms used to transfer digital data to the *Archive*. For routine, repeated submissions include the server, location, and protocol used.

FTP pull by Archive from Okeanos System ftp servers. Details will be described in an updated CoastWatch Ocean Color – CLASS ICD.

2. List any relevant Interface Control Document, Memorandum of Understanding, or other technical documents outlining how data will be transferred from *Producer* to *Archive*.

1084 CoastWatch Ocean Color to CLASS ICD – to be updated.

3. Describe the submission schedule in terms of starting/ending times and submission frequency for each submission session.

Daily; to be specified in ICD.

4. Give the volume of each submission session and the total anticipated volume per day or month in bytes.

Same as Section 2, Question 10.

5. List the steps in the transfer process from *Producer* to *Archive*.

For transfer to CLASS:

First, the Okeanos System pushes the Submission Information Package to a pre-designated CLASS directory on the Okeanos ftp server. Then it pushes the corresponding Archive Manifest to the same CLASS directory. Once the Archive Manifest is detected, CLASS begins processing the Archive Manifest and the associated Submission Information Package. For details regarding the Archive Manifest and this transfer process, see the ICD mentioned previously in this section.

6. List the *Producer's* preference for basic file validation routines (e.g. checksums, CRC32, MD5 or other).

MD5.

7. Does the *Producer* request a periodic record of receipt from the *Archive* for purposes of tracking the submitted data?

The transfer of records of receipts from the Archive to the Producer will be described in the ICD mentioned previously in this section.

NODC Submission Information Form

8. List any security requirements needed during submission from the *Producer* to the *Archive*.

Same as Section 1, Question 6a.

9. Is the content of each submission session considered by the Producer to be a continuation or new version of a previous submission, or is the content of each submission session considered by the Producer to be an independent or stand-alone collection of data?

The content of each submission is considered by the Producer to be a stand-alone collection of data.

NODC Submission Information Form

Section 6 – Archive Appraisal and Justification Information

Only in cases where a formal archive appraisal package is required by the *Archive* in order to gain management approval to provide archival services for this Data Set are these elements required. For these cases, also ensure the following questions have been answered: all of Section 2, and Section 3 questions 1 and 2. The Point of Contact for the *Archive* will provide additional guidance with this section.

1. What are the cost considerations for long-term maintenance of the Data Set? Are resources available for archiving and providing access to these records?

A formal cost assessment for the archive and long-term stewardship of NOAA CoastWatch Ocean Color products using the CLASS system has been performed by CLASS Program Manager Rick Vizbulis and has been included as an appendix to this document.

2. Has this Data Set ever physically resided at a scientific data center or center of data where stewardship was provided? Where does it reside now? What scientific expertise would best provide stewardship for this Data Set?

This data set has never physically resided at a scientific data center where stewardship was provided. It currently resides on the developmental Tethys system at STAR and on the operational Okeanos system at OSDPD. Science expertise in ocean optics would best provide stewardship for this data set.

3. Where does this Data Set fit within NOAA's mission?

Ecosystems: Ecosystem Research, Ecosystem Observations – The ocean color products provide information on phytoplankton biomass, primary productivity and water quality (e.g. Total Suspended Matter) for fisheries and coastal resource management, including the NOS HAB Bulletin.

Weather and Water: Coasts, Estuaries & Oceans (CEO); Science, Technology and Infusion – The ocean color products will support the CEO Program Coastal Storm Initiative.

Climate: Observations and Analyses – Ocean color data supports climate variability and change assessments.

Satellite Sub-Goal: Satellite Services – These ocean color products are created by processing lower-level satellite data to a more user-friendly package.

4. What is the value (scientific, public, government) of this Data Set in terms of current and anticipated future benefits?

Archived data will be used for new sensor/product development and evaluation. They will also be used for developing climatology products. They meet the user requests for the extended window of data accessibility.

5. Does the Data Set have legal mandates which require its archive at NOAA? Are there existing NARA disposition schedules that pertain to these records? If yes, please describe.

NODC Submission Information Form

~~Environmental Data Management at NOAA: Archiving, Stewardship and Access (2007):~~

“NOAA archives and provides access to a wide variety of data and products. These activities are based upon numerous legislative mandates. A few of many examples include:

- The National Climate Program Act of 1978 (15 USC CH29 PL 95-357), which calls for ‘...management and active dissemination of climateological data....’
- Magnuson-Stevens Fishery Conservation and Management Act (Public Law 94-265), which states that ‘The collection of reliable data is essential to the effective conservation, management, and scientific understanding of the fishery resources of the United States.’

NOAA must continue to archive and provide access to all data as required by law. This is a fundamental principle that NOAA will strictly adhere to, so this NRC study can be of most value to NOAA by focusing on the scientific and societal value of NOAA’s data and not the legislative mandates.”

--Committee on Archiving and Accessing Environmental and Geospatial Data at NOAA, National Research Council

6. Is the Data Set unique? If not, where else does it exist?

The data set is unique. It does not exist elsewhere.

7. Is the Data Set related to other records in a NOAA *Archive* (i.e. an extension, a new version, improved quality, etc.)? If yes, to what degree does this Data Set add value to other data sets held by NOAA or others?

This Data Set is related to CoastWatch L1A SeaWiFS Ocean Color data, which are archived using CLASS. This Data Set adds value to the L1A data in two principal ways. The L1A data is available only from the SeaWiFS instrument, whereas this Data Set contains data from three ocean color-sensing instruments. This allows for intercomparison and validation among the three sets of measurements. Also, the Data Set uses the L1A data along with ancillary data sets as input for processing up to L2.

8. Has the Data Set undergone user evaluation and/or scientific peer review, been used extensively in publications, and/or subjected to other appraisal processes? If yes, please describe.

The Data Set, part of CoastWatch, has undergone extensive user evaluation. Prior to 2001, CoastWatch in general had over 8000 registered users, and continues to provide data to thousands of users per month. A list of published references, including peer reviewed papers, proceedings, presentations, conference posters, and other documents that utilize or include NOAA CoastWatch data products exists at http://coastwatch.noaa.gov/cw_pub.html.

9. What is the current storage media for the Data Set? If in electronic format, does it still exist on other media (e.g. paper, film)? If yes, is it required to maintain copies on other media?

The current storage media for the Data Set is on electronic media. It is not necessary to maintain copies on other media.

10. Does appropriate hardware and software technology exist to enable usability of the Data Set? If yes, please describe.

NODC Submission Information Form

HDF is a widely used format among satellite products and is commonly used with scientific analysis software such as Matlab, IDL, Fortran, and GRADS. In addition to these tools, CoastWatch also produces several utilities designed for use specifically with CoastWatch products. These are available through the CoastWatch website at http://coastwatch.noaa.gov/cw_software.html.

11. Does the Data Set have intrinsic value? Intrinsic value implies that an object containing data has value beyond the data content in the object. For example, the original deck logs from the HMS *Beagle* have intrinsic value, but the digitized observations from those logs do not because the digitized files are easily copied viewed, and/or redistributed.

No.

The signatures below indicate the belief that the information contained on this form is true and correct to the best understanding of the *Producer* and *Archive*. These signatures also acknowledge that in the future, this information may be amended, updated, or revised as necessary and that some submissions may require management level approvals before archival services can be provided.

Point of Contact for the *Producer*
Printed Name and Date:

Point of Contact for the *Archive*
Printed Name and Date:

NODC Submission Information Form

Appendix: CLASS Cost Estimates for support of CoastWatch Products (OceanColor L2 and *Emiliana huxleyi*)

23 December 2008

1 Assumptions

- There will be two products.
 - Ocean Color L2
 - *Emiliana Huxleyi*
- Both products will be produced for 10 years
- Daily data volumes are
 - Ocean Color L2: 16.5GB/day
 - *Emiliana Huxleyi*: 24MB/day
- Yearly data volumes for both products are 6.2TB
 - Ocean Color L2: 6,023GB
 - *Emiliana Huxleyi*: 255GB
- Total data volumes (10 years) for both products: 62TB
- Data will be delivered from current CoastWatch or OceanColor servers
 - Interface between CLASS and both servers can be leveraged
- Existing Submission Agreement for Ocean Color L1 can be extended
- No formal reviews required
 - Requirements
 - Design
 - Operational Readiness
- No new functional requirements
 - Data can be ingested by file name
 - All search criteria are included in the file name
 - No need to create AIPs by combining various files
 - Data will be delivered to consumers as received.
- Data will be received on October 2009 or later
 - Allow for upgrade to LTO-4
 - Data will be stored in LTO-4 for five years and then in LTO-6
 - Earliest release that changes can be included is CLASS release 5.2
- Data stewardship costs are not included

2 Staffing Requirements

2.1 System Engineering

- Requirements gathering and definition, ICD and SA updates
 - One System engineer working for six months at 10% level
 - One web developer working 1 month at 10% level

2.2 Software Development and Database configuration

- One mid level software developers working for six months, at 15% level.

2.3 Hardware and network upgrades

- No updates needed

NODC Submission Information Form

2.4 Operations Support

- One operator working 10 years at 5% level
 - Data acquisition and delivery support
 - Help Desk support
- One DBA, Network Administrator and System Administrator working 10 years at a combined 5% level

3 Cost Estimate

Costs only include percent of time personnel are allocated to support this data campaign.

3.1 Pre-Operations activities: \$22,500

- Systems Engineering costs : \$10,000
 - Review Submission Agreements
 - Develop ICD
 - Transfer information to development team
 - End-to-end test preparations and support
 - Project Management
 - Security oversight
- Web Help pages and metadata updates: \$1,500
 - Interface with NMMR
 - Update help pages with new product information
- Software development and database configuration: \$11,000
 - Requirements analysis
 - Database configuration
 - Unit testing
 - System integration and testing
 - Configuration Management

3.2 Hardware Upgrade costs: \$0

No Hardware or network upgrades are need for support of this data

3.3 Operations and Maintenance: \$132,000

1. LTO tapes –
 - First five years at two locations-- (100 LTO-4 tapes): \$6,000
 - Full archive at two locations – (50 LTO-6 tapes): \$3,000
2. Operator costs (life of mission – 10 years-) \$33,000
3. DBA, System and Network Administrator (life of mission –10 years-): \$90,000