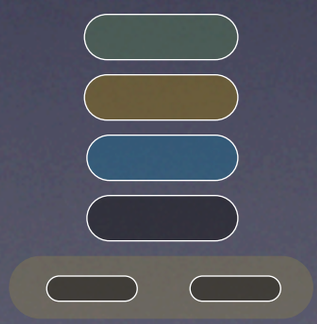
Comparison of Metadata Standards:

- This document compares the following metadata standards for the attribute groups used for describing geophysical datasets as an initial step to creating a NetCDF/CF based standard for use with CoastWatch satellite products:
  - CoastWatch HDF version 3.4 as described in the CoastWatch Utilities User's Guide
  - NetCDF Climate Forecast (CF) version 1.4, based on NetCDF and COARDS standards
  - Data Content Standard (DCS) for Remote Sensed Ocean Color Data version 1.0

File Information

CWHDF		CF-1.4	DCS
cwhdf_version	NC	Conventions	
satellite		source	sensorPlatform
sensor			sensor
data_source			
origin		institution	createInstitution
history	NC	history	createDateTime
processing_stream			inputFiles
orbit_type	NC	title	
station_code		references	
station_name		comment	

Legend:



- Required
- Required if applicable
- Recommended
- Optional
- Related attributes
- NetCDF/COARDS attribute

Time Coordinates

CWHDF	CF-1.4	DCS
pass_date		aquisitionStartDate
start_time		aquisitionEndDate
temporal extent		
composite	climatology	
pass_type	bounds	
	axis	
NC	long_name	
	standard_name	
NC	units	
	calendar	
	compress	
	leap_month	
	leap_year	
	month_lengths	
NC	valid_min	
NC	valid_max	
NC	valid_range	

Spatial Coordinates

CWHDF	CF-1.4	DCS
projection	grid_mapping_name	mapProjection
gctp_sys	semi_major_axis	ellipsoid
gctp_datum	semi_minor_axis	longOfCentralMeridian
gctp_zone	inverse_flattening	latOfTrueScale
gctp_parm	earth_radius	falseEasting
et_affine	standard_parallel	falseNorthing
rows	longitude_of_central_meridian	standardParallel
cols	latitude_of_projection_origin	geodeticDatum
	false_easting	pixelSize
	false_northing	
polygon_latitude	bounds	northernLatitude
polygon_longitude	NC valid_min	southernLatitude
	NC valid_max	westernLongitude
	NC valid_range	easternLongitude
region_code	standard_name	
region_name		
projection_type	axis	latitude
sensor_code	compress	longitude
sensor_parm	formula_terms	
	NC long_name	
	NC positive	
	NC units	

Data Values

CWHDF		CF-1.4	DCS
add_offset		add_offset	
add_offset_err	NC	scale_factor	
scale_factor			
scale_factor_err			
calibrated_nt			
_FillValue		_FillValue	
missing_value		missing_value	
	NC	valid_max	
	NC	valid_min	
	NC	valid_range	
long_name		long_name	observedProperty
	NC	standard_name	
direction_variable		ancillary_variables	processingFlags
quality_flag		flag_masks	ancillaryData
quality_mask		flag_meanings	
flag_bits		flag_values	

Data Values (continued)

CWHDF	CF-1.4	DCS
raster_type	cell_methods	
coordsys	coordinates	
	grid_mapping	
units	NC units	
atmospheric_correction	comment	observedPropertyAlgorithm
processing_algorithm	references	atmCorrection
	source	calibration
		processingLevel
		numberOfSpectralBands
		wavelengths
		bandwidths
format	cell_measures	
C_format	institution	
fraction_digits	standard_err_multiplier	
nav_affine		
direction_convention		

Notes:

- CF uses generic coordinate axes (ie: NetCDF variables) and associated metadata attributes for both temporal and spatial coordinates, where as CWHDF and DCS just use sets of attributes, one set for spatial and another for temporal. So there may *appear* to be gaps in metadata attribute groupings, where in fact the metadata is just stored differently.
- DCS does not specify how to deal with data file structure and there is no way to specify metadata for more than one "variable" to describe how it was produced (for example, chlorophyll concentration and turbidity in one file), where as CF and CWHDF distinguish between attributes attached to the whole file, and to specific variables, and allow for multiple variables.
- CWHDF allows for many map projections using a GCTP parameter array, where as CF and DCS have individual attributes for the map projection parameters of each projection system, and support a more limited set of projections.
- CF has a more sophisticated way of handling composite and climatology metadata than CWHDF. DCS does not handle composites or climatologies at all.
- CWHDF and DCS have relatively large sets of required attributes, where as CF has very few, most attributes are recommendations.

- No metadata standard explicitly addresses for what general purpose each attribute is required. We might propose a set of purposes for attributes to give them context, for example some purposes:
  - Correctly reading/interpreting variable data values.
  - Contacting the data provider for answering questions.
  - Helping archive users to quickly assess data suitability.
  - Allowing software to compute earth locations for each pixel.
  - Tracking changes to the data processing procedure.
  - Allowing certain types of further processing.
- These could be grouped and simplified into areas of concern:
  - Computation: fundamental use of the scientific data values themselves, needed for computational accuracy.
  - Information: for information about the data up to now, how it was produced, who produced it, when it was produced, what data sources were used, equations, how to distinguish it from other similar datasets, etc. Not needed for computation using the data.
  - Coordinate: for information about locating the data in space and time.
- Why have a general purpose for each attribute? Because metadata standards generally disagree on the *level of detail*, because the requirements for each standard are different. For example the metadata useful to an automatic data plotting program are different than those useful to an archive that needs to search data. What level of detail is needed for CoastWatch data?

References:

• *CoastWatch Software Library and Utilities User's Guide*, Version 3.2.4, Development Prerelease, Revised July 1, 2009.
• *NetCDF Climate and Forecast (CF) Metadata Conventions*, Version 1.4, February 27, 2009. (http://cf-pcmdi.llnl.gov/documents/cf-conventions/1.4/cf-conventions.html)
• *COARDS netCDF profile: Conventions for the standardization of NetCDF files*, Accessed November 25, 2009. (http://ferret.wrc.noaa.gov/noaa_coop/coop_cdf_profile.html)
• *Data Content Standard for Remotely Sensed Ocean Color Data*, IOOS Data Integration Framework - 2008, Version 1.0, July 7, 2008.
• *Mapping NCDDS Ocean Color DCS v1.0 to CoastWatch netCDF v0.61.*
• *The Recommended GHRSST Data Processing Specification Version 2.0, Revision 0.6 L2P Product Specification*, June 10, 2009.
• *The Recommended GHRSST Data Specification (GDS) L3 Product Specification*, Version 2.0, Revision 0.3.1, June 26, 2009.
• *The Recommended GHRSST Data Processing Specification GDS (Version 2.0 revision 0.1) L4 product specification, June 26, 2009.*