Standardized Output File Specifications for Match-Up Data

Maya DeBellis[1] (maya.debellis@jpl.nasa.gov), Edward M. Armstrong[1], Mark A. Bourassa[2], Thomas Cram[3], Jocelyn Lee Elya[2], Frank Greguska[1], Thomas Huang[1], Joseph Jacob[1], Shawn R. Smith[2], Vardis Tsontos[1], Steven J. Worley[3], Elizabeth Yam[1]

[1] NASA Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
[2] Center for Ocean-Atmospheric Prediction Studies, 2000 Levy Avenue, Building A, Suite 292, Tallahassee, FL 32306-2741, USA
[3] National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, USA

ESIP Winter Meeting 2019

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

[CL#19-0245]

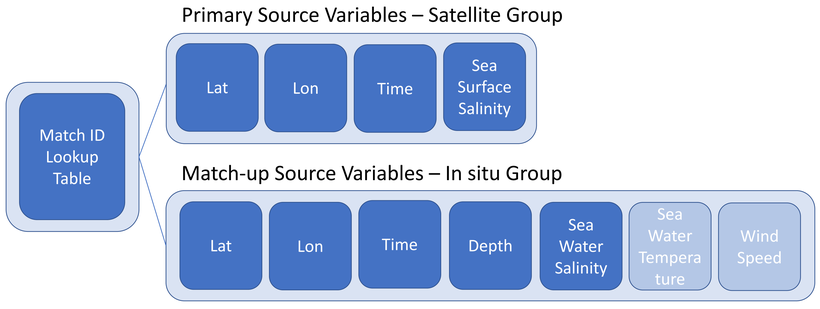
Abstract

Oceanographic applications increasingly rely on the integration and collocation of satellite and field observations providing complementary data coverage over a continuum of spatio-temporal scales. The Distributed Oceanographic Match-Up Service (DOMS) implements a technical infrastructure providing a generalized, publicly accessible and extensible data matching capability for remote in situ and satellite data stores in support of satellite mission cal/val and a range of research and operational applications. In situ datasets currently comprising DOMS include the SAMOS, ICOADS, and SPURS collections with a complementary set of satellite ocean wind, SST, and salinity datasets from the PO.DAAC. Developed initially under NASA/AIST14 support, the DOMS tool is now a component of the Apache Science Data Analytics Platform (SDAP) project, an open source project funded by AIST and colloquially known as OceanWorks. This follow-on project is a joint effort between the Jet Propulsion Laboratory, Florida State University, George Mason University, and National Center for Atmospheric Research.

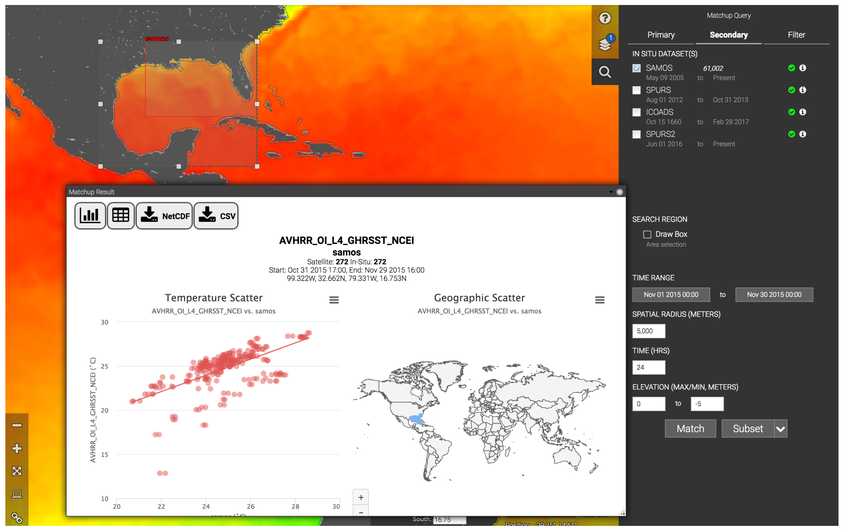
DOMS facilitates on-the-fly matchup queries both interactively via a web-based UI and programmatically via API through which users specify custom geospatial references and receive collocated satellite and field observations within the selected spatio-temporal domain and match-up window extent. Users can define the primary (satellite) and secondary (in situ) datasets for a match up operation, and can also specify whether they want all matching data within a user defined distance, depth and time radius or just nearest neighbor collocation data. Resulting match-up data can be exported as CSV or netCDF files.

An important aspect of the development work involved designing and implementing a standardized file and metadata format specification for the match-up output in accordance with Earth Science data interoperability standards. Here we describe features of this specification that support both netCDF4 and CSV output formats and implements a consistent set of relevant CF attributes but also DOMS specific metadata used to document query parameters fully. Emphasis here is placed on the DOMS nc4 output file specification, which leverages Group structures in a novel way to parsimoniously represent potentially even many-to-many relationships between matched satellite and in situ observations.

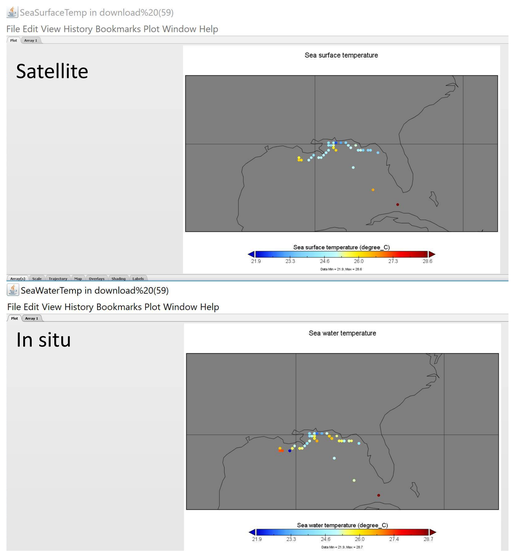
netCDF4 Match-up Structure

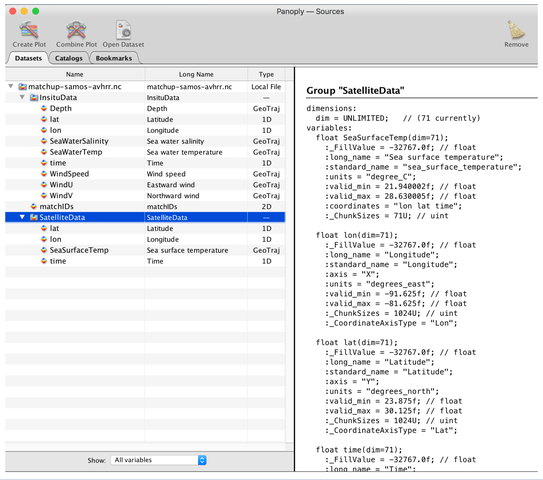


This figure represents the netCDF4 output file structure. The file contains two groups, each containing data for one source in the match-up pair. Also included in the file is an array of pairs of IDs representing the indexes of matched records between respective groups. The benefit of this match array structure is that it can support one-to-one through many-to-many relationships without any duplication of data. The netCDF is a more compact representation of the data as compared to the CSV, with smaller files sizes, because of this lack of data redundancy and the internal compression capabilities that nc4 affords.



netCDF4 Groups



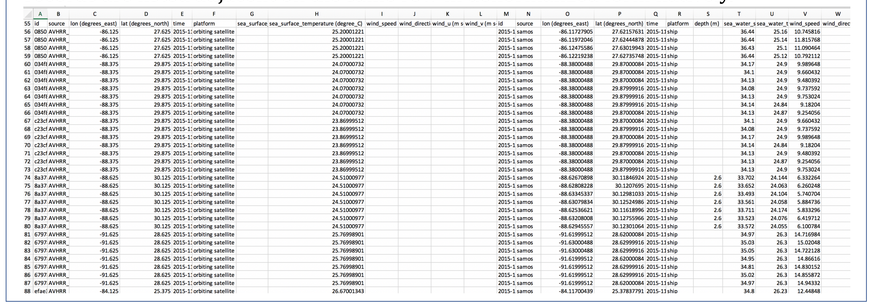


Since DOMS Group variables leverage existing CF standards, an off the shelf tool like Panoply that is CF-metadata aware can automatically interpret and map in situ and satellite group data independently just as if they were given in their separate CF-compliant source files.

Metadata is also included at the Group variable level. Both coordinate and measurement variable attributes such as valid_max/min, _FillValue, and units are included consistent with CF. In the future, a reader will be implemented to reconstruct the matches in each group using a join operation based on match IDs.

DOMS CSV File Structure

CSV and netCDF DOMS output formats are consistent to the extent possible, each supporting both singular or all neighboring point match up and with identical global variables. The nc4 variables are represented as columns in the CSV complete with key variable name and units attributes as column headers. The structure of the CSV data output is flat compared to a more normalized Group representation in the netCDF. Instead of an array of indices to match and reconstruct records, in the CSV matched records are listed adjacent on the same row in denormalized form natively.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
55	id	source	lon (degrees_east)	lat (degrees_north)	time	platform	sea_surface_sea_surface_temperature (degree_C)	wind_speed	wind_direction	wind_u (m s-1)	wind_v (m s-1)	id	source	lon (degrees_east)	lat (degrees_north)	time	platform	depth (m)	sea_water_salinity	sea_water_temperature	wind_speed	wind_direction
56	0850 AVHRR		-86.125	27.625	2015-11-orbiting satellite				25.2001221				2015-1 samos	-86.11727905	27.62157631	2015-11-ship			36.44	25.16	10.745816	
57	0850 AVHRR		-86.125	27.625	2015-11-orbiting satellite				25.2001221				2015-1 samos	-86.11972046	27.62444878	2015-11-ship			36.44	25.14	11.815768	
58	0850 AVHRR		-86.125	27.625	2015-11-orbiting satellite				25.2001221				2015-1 samos	-86.12475586	27.63019943	2015-11-ship			36.43	25.1	11.090464	
59	0850 AVHRR		-86.125	27.625	2015-11-orbiting satellite				25.2001221				2015-1 samos	-86.12219238	27.62735748	2015-11-ship			36.44	25.12	10.792112	
60	0349 AVHRR		-88.375	29.875	2015-11-orbiting satellite				24.0700732				2015-1 samos	-88.38000488	29.87000084	2015-11-ship			34.17	24.9	9.989648	
61	0349 AVHRR		-88.375	29.875	2015-11-orbiting satellite				24.0700732				2015-1 samos	-88.38000488	29.87000084	2015-11-ship			34.1	24.9	9.660432	
62	0349 AVHRR		-88.375	29.875	2015-11-orbiting satellite				24.0700732				2015-1 samos	-88.38000488	29.87000084	2015-11-ship			34.13	24.9	9.480392	
63	0349 AVHRR		-88.375	29.875	2015-11-orbiting satellite				24.0700732				2015-1 samos	-88.38000488	29.87999916	2015-11-ship			34.08	24.9	9.737592	
64	0349 AVHRR		-88.375	29.875	2015-11-orbiting satellite				24.0700732				2015-1 samos	-88.38000488	29.87999916	2015-11-ship			34.13	24.9	9.753024	
65	0349 AVHRR		-88.375	29.875	2015-11-orbiting satellite				24.0700732				2015-1 samos	-88.38000488	29.87999916	2015-11-ship			34.14	24.84	9.18204	
66	0349 AVHRR		-88.375	29.875	2015-11-orbiting satellite				24.0700732				2015-1 samos	-88.38000488	29.87000084	2015-11-ship			34.13	24.87	9.254056	
67	c23cf AVHRR		-88.375	29.875	2015-11-orbiting satellite				23.8699512				2015-1 samos	-88.38000488	29.87000084	2015-11-ship			34.1	24.9	9.660432	
68	c23cf AVHRR		-88.375	29.875	2015-11-orbiting satellite				23.8699512				2015-1 samos	-88.38000488	29.87999916	2015-11-ship			34.08	24.9	9.737592	
69	c23cf AVHRR		-88.375	29.875	2015-11-orbiting satellite				23.8699512				2015-1 samos	-88.38000488	29.87000084	2015-11-ship			34.17	24.9	9.989648	
70	c23cf AVHRR		-88.375	29.875	2015-11-orbiting satellite				23.8699512				2015-1 samos	-88.38000488	29.87999916	2015-11-ship			34.14	24.84	9.18204	
71	c23cf AVHRR		-88.375	29.875	2015-11-orbiting satellite				23.8699512				2015-1 samos	-88.38000488	29.87000084	2015-11-ship			34.13	24.9	9.480392	
72	c23cf AVHRR		-88.375	29.875	2015-11-orbiting satellite				23.8699512				2015-1 samos	-88.38000488	29.87000084	2015-11-ship			34.13	24.87	9.254056	
73	c23cf AVHRR		-88.375	29.875	2015-11-orbiting satellite				23.8699512				2015-1 samos	-88.38000488	29.87999916	2015-11-ship			34.13	24.9	9.753024	
74	8a37 AVHRR		-88.625	30.125	2015-11-orbiting satellite				24.51000977				2015-1 samos	-88.62670898	30.11866924	2015-11-ship	2.6	33.702	24.144	6.332064	
75	8a37 AVHRR		-88.625	30.125	2015-11-orbiting satellite				24.51000977				2015-1 samos	-88.62808228	30.1207605	2015-11-ship	2.6	33.652	24.063	6.260248	
76	8a37 AVHRR		-88.625	30.125	2015-11-orbiting satellite				24.51000977				2015-1 samos	-88.63345337	30.12981033	2015-11-ship	2.6	33.493	24.106	5.740704	
77	8a37 AVHRR		-88.625	30.125	2015-11-orbiting satellite				24.51000977				2015-1 samos	-88.63079834	30.12524986	2015-11-ship	2.6	33.561	24.058	5.884736	
78	8a37 AVHRR		-88.625	30.125	2015-11-orbiting satellite				24.51000977				2015-1 samos	-88.62536621	30.11618996	2015-11-ship	2.6	33.711	24.174	5.833296	
79	8a37 AVHRR		-88.625	30.125	2015-11-orbiting satellite				24.51000977				2015-1 samos	-88.63208008	30.12755966	2015-11-ship	2.6	33.523	24.076	6.419712	
80	8a37 AVHRR		-88.625	30.125	2015-11-orbiting satellite				24.51000977				2015-1 samos	-88.62945557	30.12301064	2015-11-ship	2.6	33.572	24.055	6.100784	
81	6797 AVHRR		-91.625	28.625	2015-11-orbiting satellite				25.76998901				2015-1 samos	-91.63999512	28.62000084	2015-11-ship			34.97	26.3	14.710984	
82	6797 AVHRR		-91.625	28.625	2015-11-orbiting satellite				25.76998901				2015-1 samos	-91.63000488	28.62999916	2015-11-ship			35.03	26.3	15.02048	
83	6797 AVHRR		-91.625	28.625	2015-11-orbiting satellite				25.76998901				2015-1 samos	-91.63000488	28.62999916	2015-11-ship			35.05	26.3	14.722128	
84	6797 AVHRR		-91.625	28.625	2015-11-orbiting satellite				25.76998901				2015-1 samos	-91.63999512	28.62000084	2015-11-ship			34.95	26.3	14.86616	
85	6797 AVHRR		-91.625	28.625	2015-11-orbiting satellite				25.76998901				2015-1 samos	-91.63999512	28.62999916	2015-11-ship			34.81	26.3	14.830152	
86	6797 AVHRR		-91.625	28.625	2015-11-orbiting satellite				25.76998901				2015-1 samos	-91.63999512	28.62999916	2015-11-ship			35.02	26.3	14.858872	
87	6797 AVHRR		-91.625	28.625	2015-11-orbiting satellite				25.76998901				2015-1 samos	-91.63999512	28.62000084	2015-11-ship			34.97	26.3	14.94332	
88	efae AVHRR		-84.125	25.375	2015-11-orbiting satellite				26.67001343				2015-1 samos	-84.11700439	25.37837791	2015-11-ship			34.8	26.23	12.44848	

Metadata

Consistency in Global Attributes between netCDF4 and CSV

Global Attributes

CF/ACDD Attributes

DOMS Attributes

```
:product_version = "1.0";
:Conventions = "CF-1.6, ACDD-1.3";
:title = "DOMS satellite-insitu machup output file";
:history = "Processing_Version = V1.0, Software_Name = DOMS, Software_Version = 1.03";
:institution = "JPL, FSU, NCAR";
:source = "doms.jpl.nasa.gov";
:standard_name_vocabulary = "CF Standard Name Table v27", "BODC controlled vocabulary";
:cdm_data_type = "Point/Profile, Swath/Grid";
:processing_level = "4";
:project = "Distributed Oceanographic Matchup System (DOMS)";
:keywords_vocabulary = "NASA Global Change Master Directory (GCMD) Science Keywords";
:creator_name = "NASA PO.DAAC";
:creator_email = "podaac@podaac.jpl.nasa.gov";
:creator_url = "https://podaac.jpl.nasa.gov/";
:publisher_name = "NASA PO.DAAC";
:publisher_email = "podaac@podaac.jpl.nasa.gov";
:publisher_url = "https://podaac.jpl.nasa.gov";
:acknowledgment = "DOMS is a NASA/AIST-funded project. NRA NNH14ZDA001N.";
:date_modified = "2018-12-18T21:09:09+0000";
:date_created = "2018-12-18T21:09:09+0000";
:time_coverage_start = "2015-11-01T00:00:00+0000";
:time_coverage_end = "2015-11-30T00:00:00+0000";
:time_coverage_resolution = "point";
:geospatial_lat_max = 21.0; // double
:geospatial_lat_min = 31.0; // double
:geospatial_lon_max = -79.0; // double
:geospatial_lon_min = -93.0; // double
:geospatial_lat_resolution = "point";
:geospatial_lon_resolution = "point";
:geospatial_lat_units = "degrees_north";
:geospatial_lon_units = "degrees_east";
:geospatial_vertical_min = 0.0; // double
:geospatial_vertical_max = 5.0; // double
:geospatial_vertical_units = "m";
:geospatial_vertical_positive = "down";
:URI_Matchup = "http://{webservice}/domsresults?id=1c1b4dd9-168e-43a4-9d84-
2e8a53bfd943&output=NETCDF";
:platform = "orbiting satellite, ship";
:_CoordSysBuilder = "ucar.nc2.dataset.conv.CF1Convention";
:DOMS_matchID = "1c1b4dd9-168e-43a4-9d84-2e8a53bfd943";
:DOMS_match_up = "samos";
:DOMS_num_matchup_matched = 607L; // long
:DOMS_num_primary_matched = 71L; // long
:DOMS_TimeWindow = 24L; // long
:DOMS_TimeWindow_Units = "hours";
:DOMS_SearchRadius = 1000.0; // double
:DOMS_SearchRadius_Units = "m";
:DOMS_ParameterPrimary = "sst";
:DOMS_platforms = "1,2,3,4,5,6,7,8,9";
:DOMS_primary = "AVHRR_OI_L4_GHRSST_NCEI";
:DOMS_time_to_complete = 79L; // long
:DOMS_time_to_complete_units = "seconds";
:DOMS_DatasetMetadata = "http://samos.coaps.fsu.edu/html/nav.php?s=2";
```