



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Distributed Oceanographic Data Match-Up Service

AIST-14

PI: Shawn Smith/FSU

FSU Co-I: Mark Bourassa

JPL/Caltech Co-Is: Thomas Huang, Vardis Tsonetos, Ben Holt

NCAR Co-I: Steve Worley

Development Engineers: Jocelyn Elya/FSU, Adam Stallard/FSU, and Zaihua Ji/NCAR

Collaborators: Fredrick Bingham/UNCW & Zhjin Li/JPL/Caltech (SPURS),
and Simon Yueh/JPL/Caltech (SMAP)



Overview

- **2014 ROSES AIST proposal to develop a Distributed Oceanographic Matchup Service (DOMS)**
 - 2 year collaborative effort between FSU, JPL, and NCAR
 - To develop a distributed web service to match satellite and in-situ marine observations to support platform comparisons, cross-calibration, validation, and quality control
- **Distributed Oceanographic Match-up Service (DOMS)**
 - Cloud computing matchup service
 - Host at JPL where it can be physically close to the PO.DAAC archive
 - Publically accessible web portal and webservice endpoints
 - Generalized satellite-in situ data collocation web-service
 - Distributed data sources: Initially 3 providers affiliated with the project and select in-situ & satellite datasets

Provider/Collaborator	In-situ data collection	Satellite data collection (swath & gridded)
NCAR	ICOADS/ARGO	-
FSU COAPS	SAMOS	-
PO.DAAC	SPURS	Aquarius SSS, JPLv3 QuikSCAT winds, L4 MUR-SST



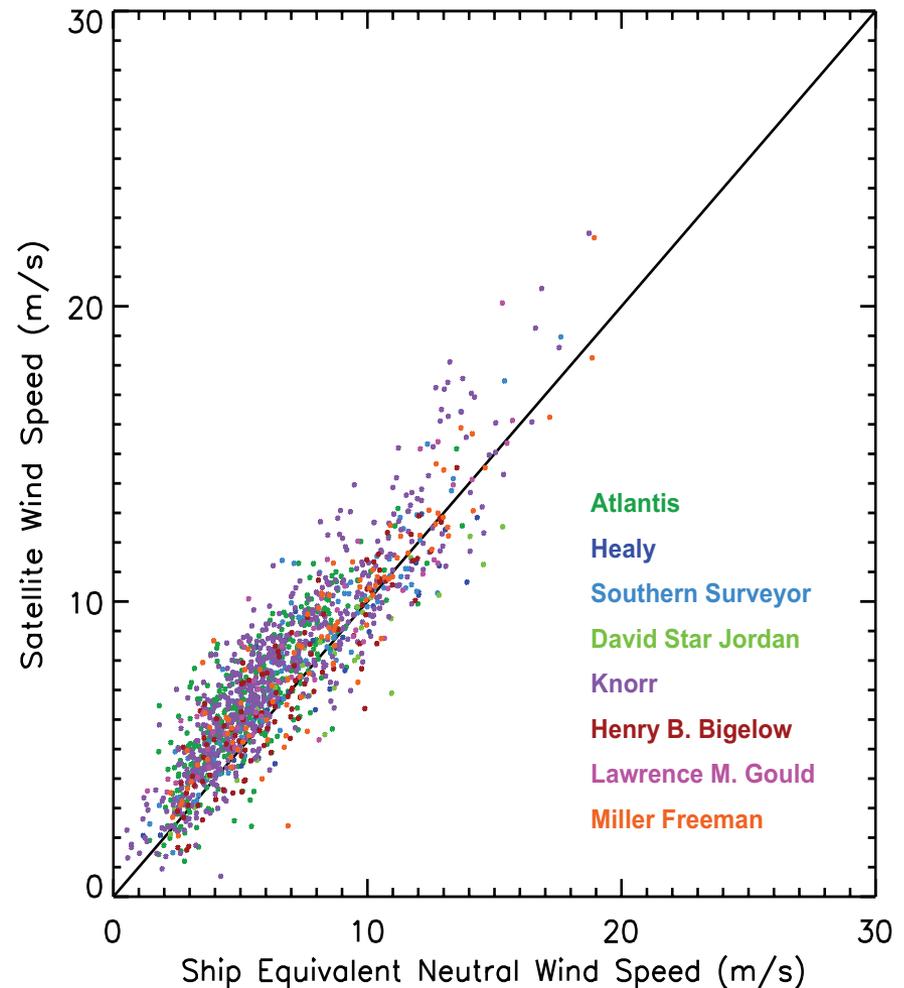
Overview (cont.)

- **Infuse common data access services** at FSU, NCAR and PO.DAAC, which consists of
 - **Extensible Data Gateway Environment (EDGE)** – a data aggregation service that supports OpenSearch, metadata export, and W10N protocol
 - **Pomegranate** – an implementation of the W10N specification
 - **Goal** - to improve handling and analysis of in-situ and satellite data
- **DOMS will allow users to input**
 - A series of geospatial references for satellite observations (e.g., footprint location, date, and time) and receive matched in-situ observations within a selectable temporal and spatial domain. Sufficient in situ metadata will also be received to improve the match interpretation, accounting for the heterogeneous observing systems.
 - In-situ geospatial data (e.g., positions of moorings, floats, or ships) and return corresponding satellite observations
- **DOMS is designed to be extensible:**
 - Incorporate other oceanographic data types
 - Integrate data from additional data providers
 - Webservice extended URL for machine-to-machine interactions
 - Web-form for interactive query definition by users
- **Primary advantage of DOMS** is the reduction in man hours required to match satellite/in situ data
 - Presently this is done using one-off programs at multiple institutions
 - Removes the need for satellite and in situ data to be collocated on a single server
 - Systematically recreate matchups if either in situ or satellite products are re-processed (new versions)



Why DOMS is Needed?

- Satellite product validation
 - Research ship wind observations co-located with winds from scatterometers
 - Ships provide necessary auxiliary parameters (air and sea temperature, pressure, humidity) and instrumental metadata (sensor height) to adjust observations to standard 10-m equivalent neutral wind.
 - Satellite wind speed is generally higher than ship winds, especially at higher wind speeds.

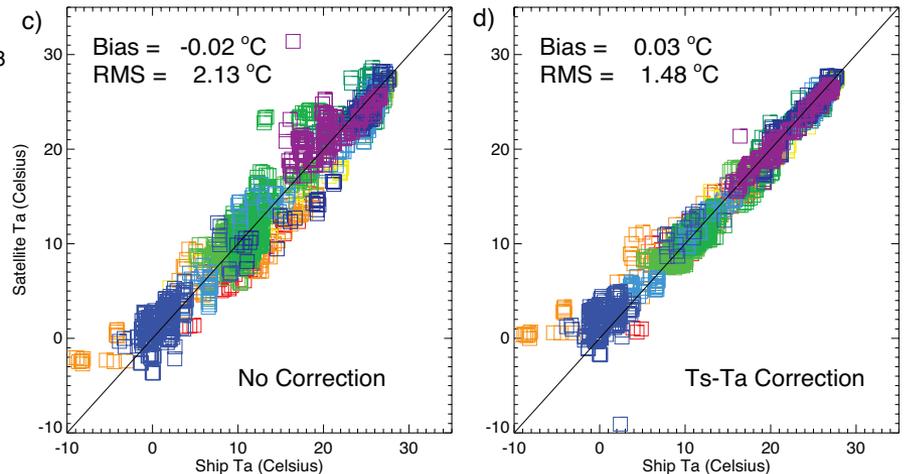
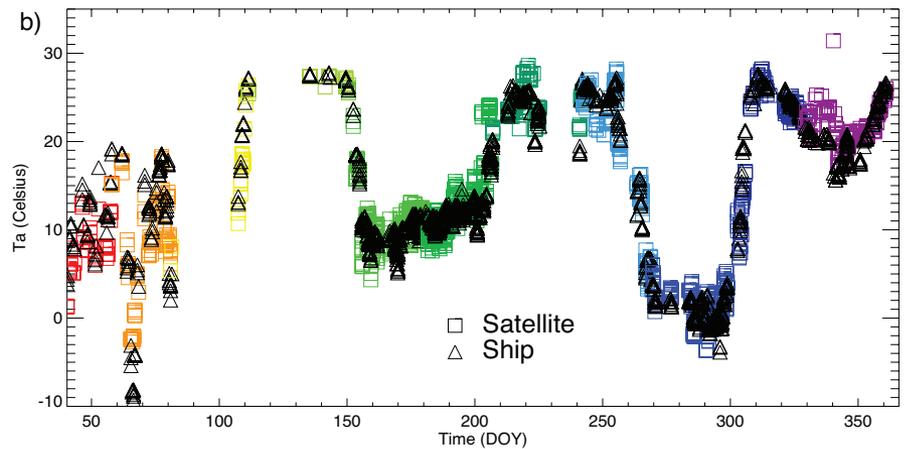
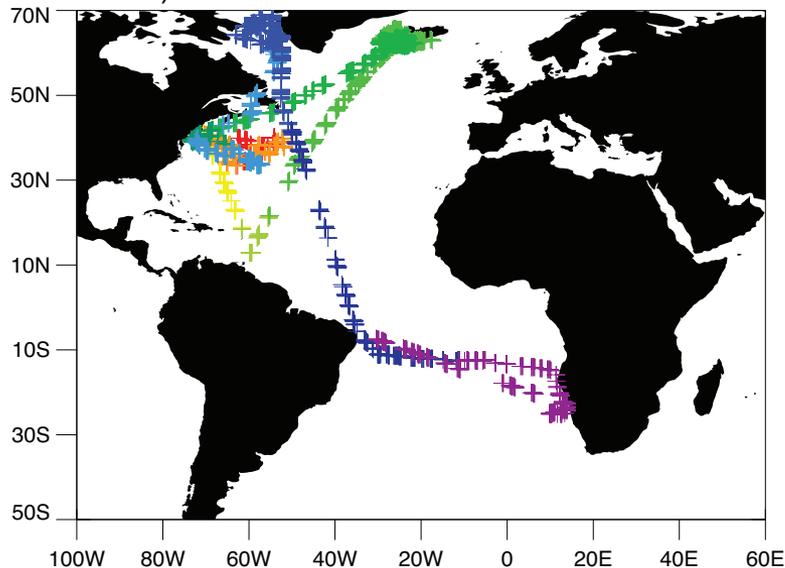




Why DOMS is Needed?

- Satellite algorithm development
 - Air temperature and humidity retrieval algorithms developed by Jackson et al. (2009)
 - Polynomial fit between collocated ship and satellite values provided correction that reduced air temperature RMS by 0.68°C .

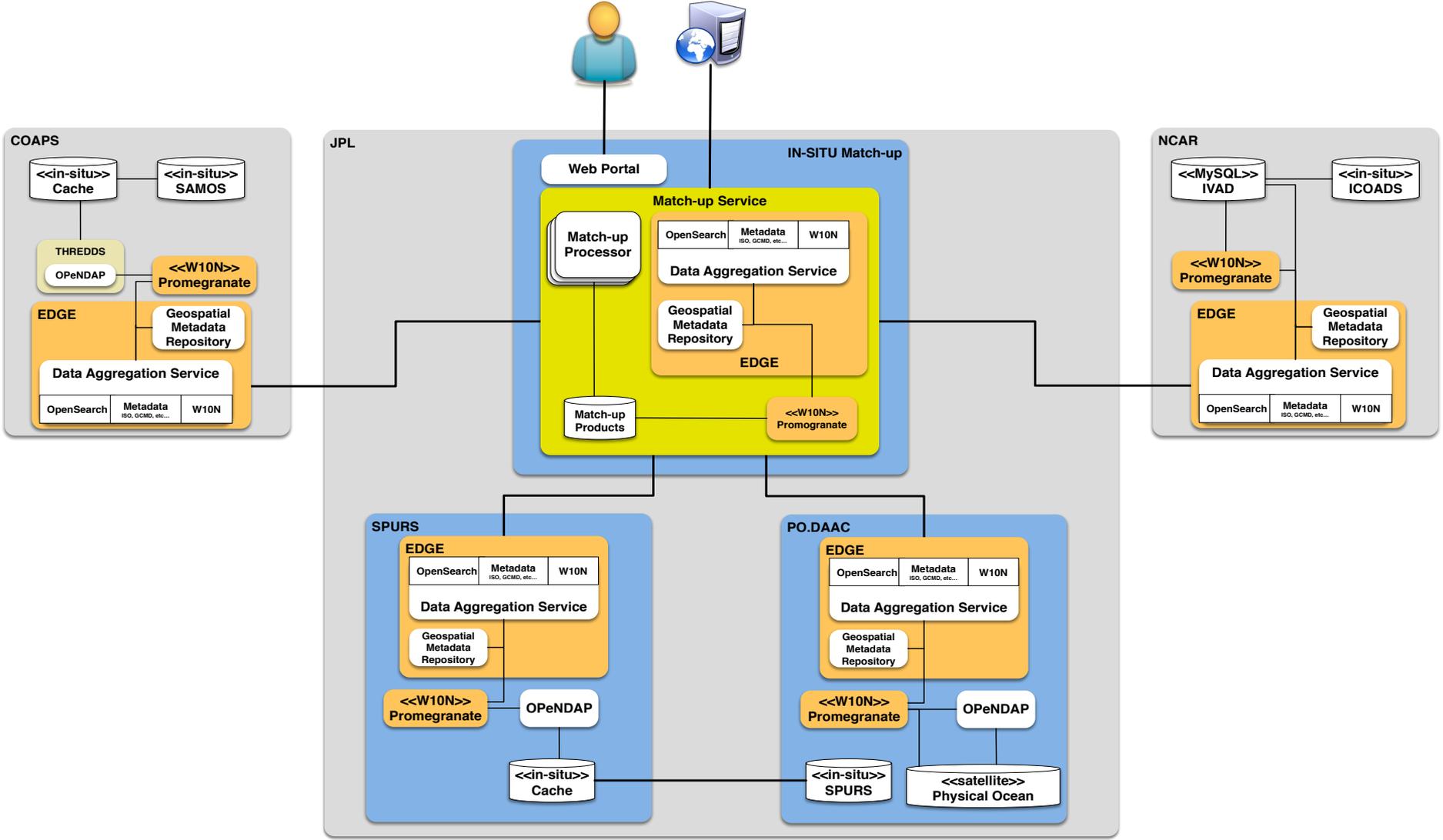
a) KNORR/Satellite Collocations for 2007



Adapted from Smith, Bourassa, and Jackson, *Sea Technology*, June 2012



DOMS' Distributed Architecture

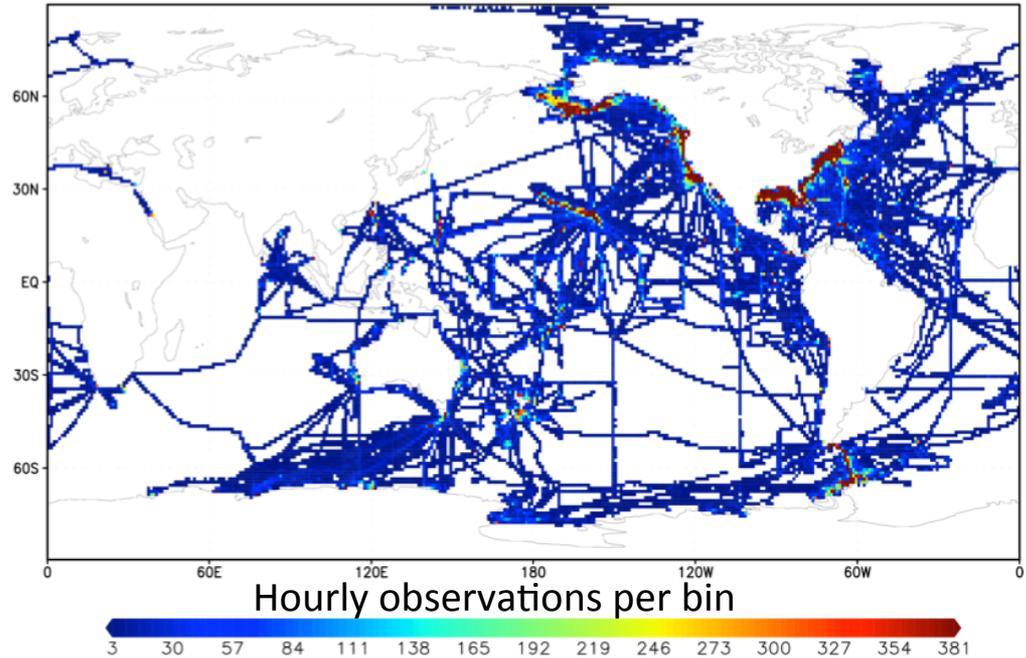




In-Situ Data: SAMOS

- **Shipboard Automated Meteorological and Oceanographic System (SAMOS)** initiative provides high-quality underway data from research vessels.
- Hosted at FSU/COAPS.
- ~30 vessels participating in FY2014
 - Vessels operated by WHOI, SIO, UH, UW, BIOS, NOAA, USCG, USAP, IMOS, SOI
 - ~30-40K one-minute observations/month/vessel

SAMOS Data Density: 2005-2014



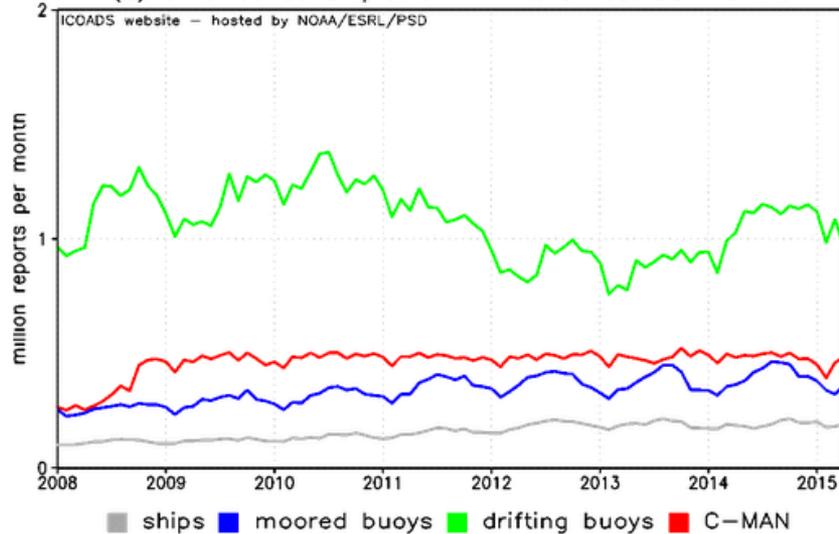
- Data include routine navigation (position, course, heading, speed), meteorology (wind, air temperature, humidity, pressure, rainfall, radiation), and oceanography (sea temperature and salinity).
- All data undergo scientific quality control.



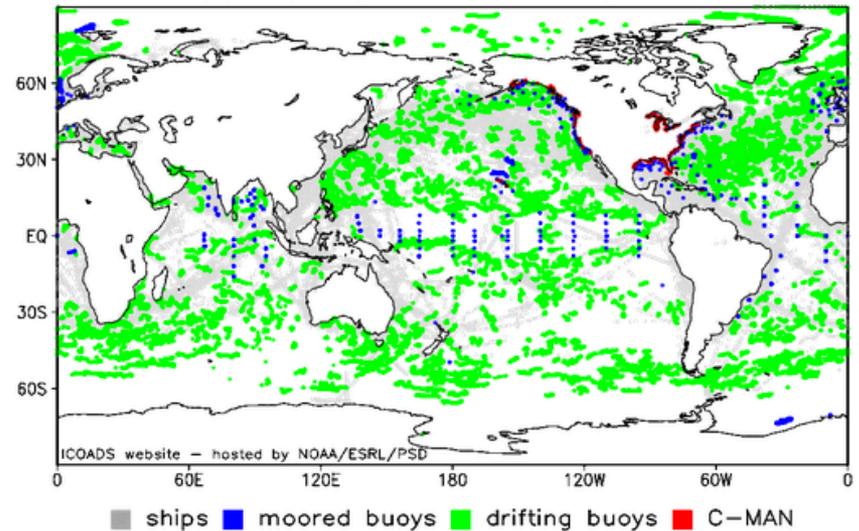


In-Situ Data: ICOADS

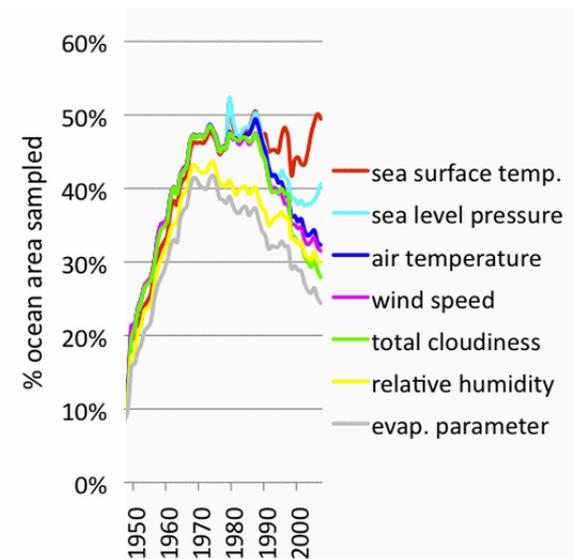
(a) time-series of platform mixture: last 88 months



(b) map of platform mixture: Apr 2015

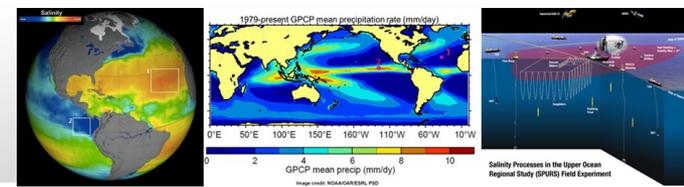


- **International Comprehensive Ocean-Atmosphere Data Set (ICOADS)**
- Global coverage from ocean observing systems (~ 3M/month)
 - VOS ships and R/Vs
 - Moored buoys: GTMBA and national systems
 - Drifting buoys: surface and ARGO
- % of ocean coverage varies by parameter
- Updated monthly with NCEP + NCDC GTS data streams
- Each record has UID and observing system tracking metadata





In-Situ Data: SPURS

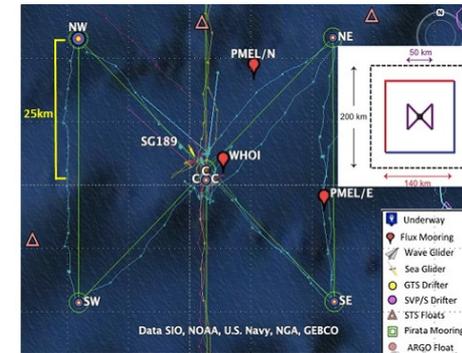


- NASA-funded oceanographic field campaigns/science salinity process studies:
 - SPURS-1: N. Atlantic (2012-13) : salinity max region
 - SPURS-2: Eastern Equatorial Pacific (15-16): high precipitation/low evaporation region

○ Aim:

- Elucidate key mechanisms responsible for near-surface salinity variations in the oceans
- Quantify the relative significance of circulation, evaporation, precipitation over a range of scales for representative areas of the open ocean.
- Surface Salinity observations valuable for Aquarius/SAC-D Cal/Val

Ship (cruise #)	Dates	Country	Chief Scientist
Thalassa	16-Aug - 13-Sep-2012	France	Reverdin
Knorr (209)	6-Sep - 9-Oct-2012	US	Schmitt
Endeavor-1 (522)	15-Mar - 15-Apr-2013	US	Schmitt
Sarmiento	14-Mar - 10-Apr-2013	Spain	Font
Endeavor-2 (533)	19-Sep - 13-Oct-2013	US	Fratantoni



○ SPURS-1 campaign

- Series of 5 cruises
- Advanced sampling technologies deployed in a nested design within a 900 x 800-mile square study area centered at 25°N, 38°W.
- 15 resulting datasets
- Natively heterogeneous formats converted to NODC NetCDF standard by SPURS-DMT
- Archived at the PO.DAAC, discoverable & distributed publicly as of 5/11/2015
 - PO.DAAC SPURS mission page: <http://podaac.jpl.nasa.gov/spurs>
 - SPURS-1 dataset catalog collection: <http://podaac.jpl.nasa.gov/datasetlist?ids=Collections&values=SPURS-I&view=list>





NASA PO.DAAC

http://podaac.jpl.nasa.gov

- The **NASA Physical Oceanographic Distributed Active Archive Center (PO.DAAC)** at Jet Propulsion Laboratory is an element of the **Earth Observing System Data and Information System (EOSDIS)**. The EOSDIS provides science data to a wide communities of user for NASA's Science Mission Directorate.
- Archives and distributes data relevant to the physical state of the ocean
- The mission of the **PO.DAAC** is to preserve NASA's ocean and climate data and make these universally accessible and meaningful.





Satellite Data: PO.DAAC



- **NASA DAACs with oceanography focus**

- Supports archival/distribution of extensive collection of satellite data by parameter:

SST (190), Ocean winds (68), Salinity (42), Altimetry (32), Ocean color (4), Circulation (3), etc.

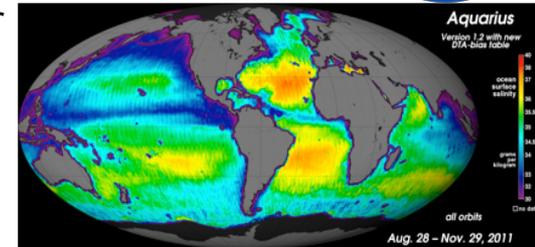
- Additionally supporting archival of select NASA-funded *in-situ* datasets
SPURS-1, SPURS-2, Ocean Melting Greenland (OMG)

- Datasets discoverable via:

Web-portal (<https://podaac.jpl.nasa.gov/>) , NASA ECHO/GCMD, P-WS

- Data holdings accessible via a range of tools & services:

ftp, OPeNDAP, W10n, CWS, THREDDS, LAS, HiTIDE, SOTO, etc.



- **DOMS will leverage a range satellite products available from PO.DAAC**

- Both Level 2 swath & higher resolution L3 and L4 gridded products

- Parameters and products of interest within the scope of DOMS:

- SST: L2 [GHRSSST-MODIS-A L2P](#), [GHRSSST-MODIS-T L2P](#) (1km res/ 2330km swath, 12hr repeat) L4 [MUR-SST](#) (1km, daily)

- SSS: L2 [Aquarius L2 v3.0](#), [CAP L2 v3.0](#) (100km res, 390km swath; 7day repeat)

- Winds: L2 [JPL Quikscat v3.0](#) (12.5km res, 1800km swath; 12hr repeat)



User Experience

- DOMS will provide a web portal interface for users to browse and to submit match-up requests interactively.
 - Planned to be hosted at JPL
 - Interface will allow users to “test/evaluate” searches by returning metadata only, creating visualizations, and then follow with full matched dataset
 - DOMS will provide flexible filtering and query specification by:
 - Instrument, sensor, parameter, provider
 - Matchup criteria: spatio-temporal domain (in x,y,z,t) and search radii/ tolerances)
- Additionally, DOMS will provide an underlying web service interface for machine-to-machine matchup operations to enable scalable data processing by external applications and services.
 - Tools will be provided to aid users in developing proper syntax for web service queries



Technical Challenges

- Ensuring that the match-up algorithms perform with sufficient speed to return desired information to the user
- Performing data matches using datasets that are distributed on the network
- Returning actual observations for the matches [e.g., salinity] with sufficient metadata so the value difference can be properly interpreted



Key Milestones

Distributed Oceanographic Match-up Service	Y1-Q1	Y1-Q2	Y1-Q3	Y1-Q4	Y2-Q1	Y2-Q2	Y2-Q3	Y2-Q4
DOMS hardware procurement and installation	■							
DOMS detailed use cases and concept of operation	■	◇						
EDGE and Webification integration and installation for SAMOS, ICOADS, and SPURS		■						
Design and demonstrate 3-way queries and subsetting by connecting to the 3 centers			■		◇			
DOMS detailed design and requirements				■		◇		
DOMS implementation and internal validation					■			◇
Upgrade ICOADS to Release 3.0						■		
DOMS Stress testing, project/peer reviews								■



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

THANKS

Questions, and more information

doms_team@coaps.fsu.edu