CL#17-2720

An Introduction To

OceanWorks

Ocean Science Platform

Thomas Huang
Jet Propulsion Laboratory
California Institute of Technology

JPL Team
Ed Armstrong, Joseph Jacob, Nga Quach, Vardis Tsontos, and Brian Wilson

Florida State University Team
Shawn Smith, and Mark A. Bourassa

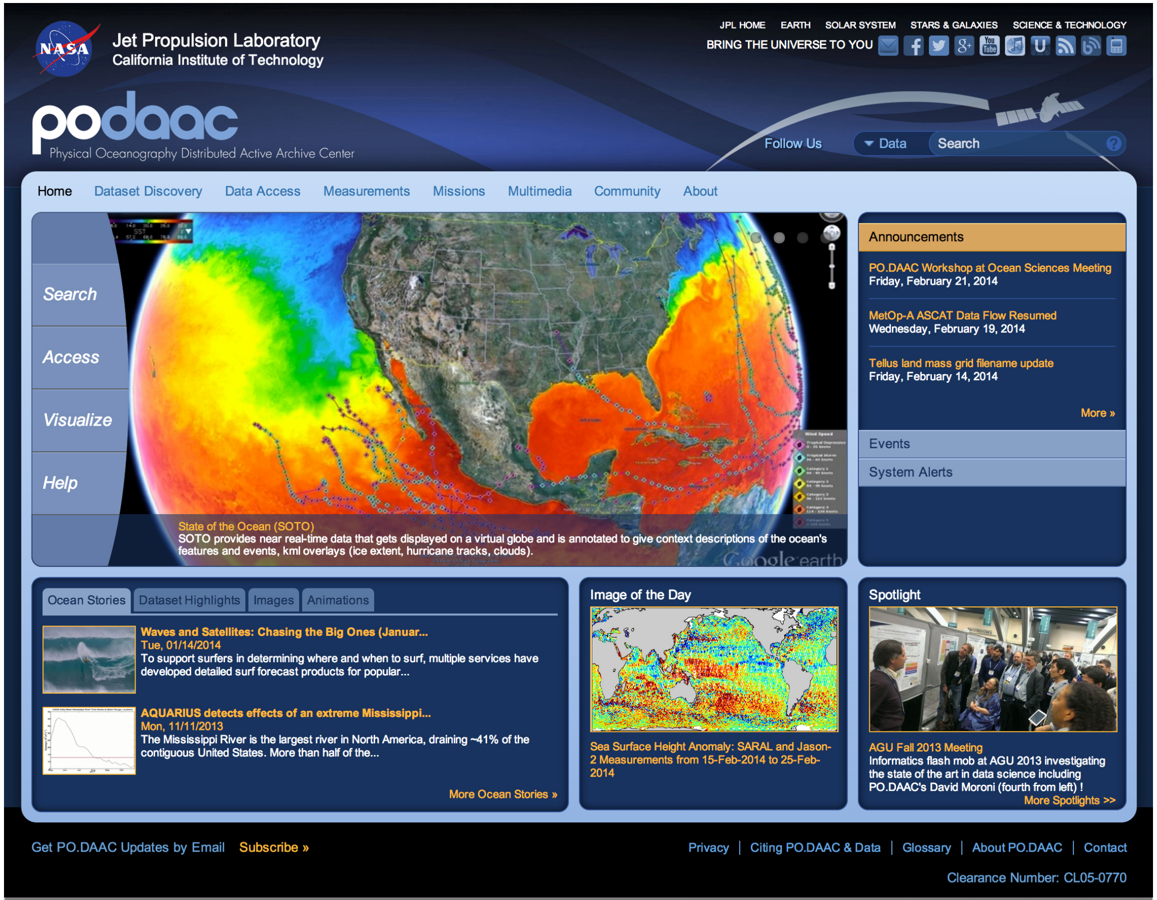
National Center for Atmospheric Research Team
Steve J. Worley

George Mason University Team
Chaowei (Phil) Yang, Yongyao Jiang, and Yun Li

NASA PO.DAAC

- 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.**

http://podaac.jpl.nasa.gov



State of the Ocean (SOTO)
SOTO provides near real-time data that gets displayed on a virtual globe and is annotated to give context descriptions of the ocean's features and events, kml overlays (ice extent, hurricane tracks, clouds).

Announcements

PO.DAAC Workshop at Ocean Sciences Meeting
Friday, February 21, 2014

MetOp-A ASCAT Data Flow Resumed
Wednesday, February 19, 2014

Tellus land mass grid filename update
Friday, February 14, 2014

More »

Events

System Alerts

Ocean Stories | Dataset Highlights | Images | Animations

Waves and Satellites: Chasing the Big Ones (Januar...
Tue, 01/14/2014
To support surfers in determining where and when to surf, multiple services have developed detailed surf forecast products for popular...

AQUARIUS detects effects of an extreme Mississippi...
Mon, 11/11/2013
The Mississippi River is the largest river in North America, draining ~41% of the contiguous United States. More than half of the...

More Ocean Stories »

Image of the Day

Sea Surface Height Anomaly: SARAL and Jason-2 Measurements from 15-Feb-2014 to 25-Feb-2014

Spotlight

AGU Fall 2013 Meeting
Informatics flash mob at AGU 2013 investigating the state of the art in data science including PO.DAAC's David Moroni (fourth from left) !
More Spotlights >>

Get PO.DAAC Updates by Email Subscribe »

Privacy | Citing PO.DAAC & Data | Glossary | About PO.DAAC | Contact

Clearance Number: CL05-0770

Big Data Reality

Reality

- With large amount of observational and modeling data, downloading to local machine is becoming inefficient
- Data centers are starting to provide additional services
  - Better searches – faceted, spatial, keyword, relevancy, etc.
  - Data subsetting – data reduction
  - Visualization – visual discovery

2015 NASA ESTO/AIST Big Data Study Roadmap: Moving from Data Archiving to Data Analytics

Increasing "big data" era is driving needs to

- Scale computational and data infrastructures
- Support new methods for deriving scientific inferences
- Shift towards integrated data analytics
- Apply computational and data science across the lifecycle

Scalable Data Management

- Capturing well-architected and curated data repositories based on well-defined data/information architectures
- Architecting automated pipelines for data capture

Scalable Data Analytics

- Access and integration of highly distributed, heterogeneous data
- Novel statistical approaches for data integration and fusion
- Computation applied at the data sources
- Algorithms for identifying and extracting interesting features and patterns

Overview

- **OceanWorks** is to establish an integrated data analytic center at the NASA PO.DAAC for Big Ocean Science. It focuses on technology integration, advancement and maturity
- **Collaboration between JPL, FSU, NCAR, and GMU**
- Bringing together PO.DAAC-related big data technologies
  - **AIST-14 OceanXtremes (PI: Huang/JPL) – TRL 4**

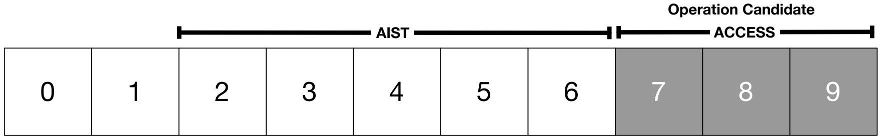
    Anomaly detection and ocean science
  - **NEXUS (PI: Huang/JPL) – TRL 6**

    Deep data analytic platform
  - **AIST-14 DOMS (PI: Smith/FSU) – TRL 4**

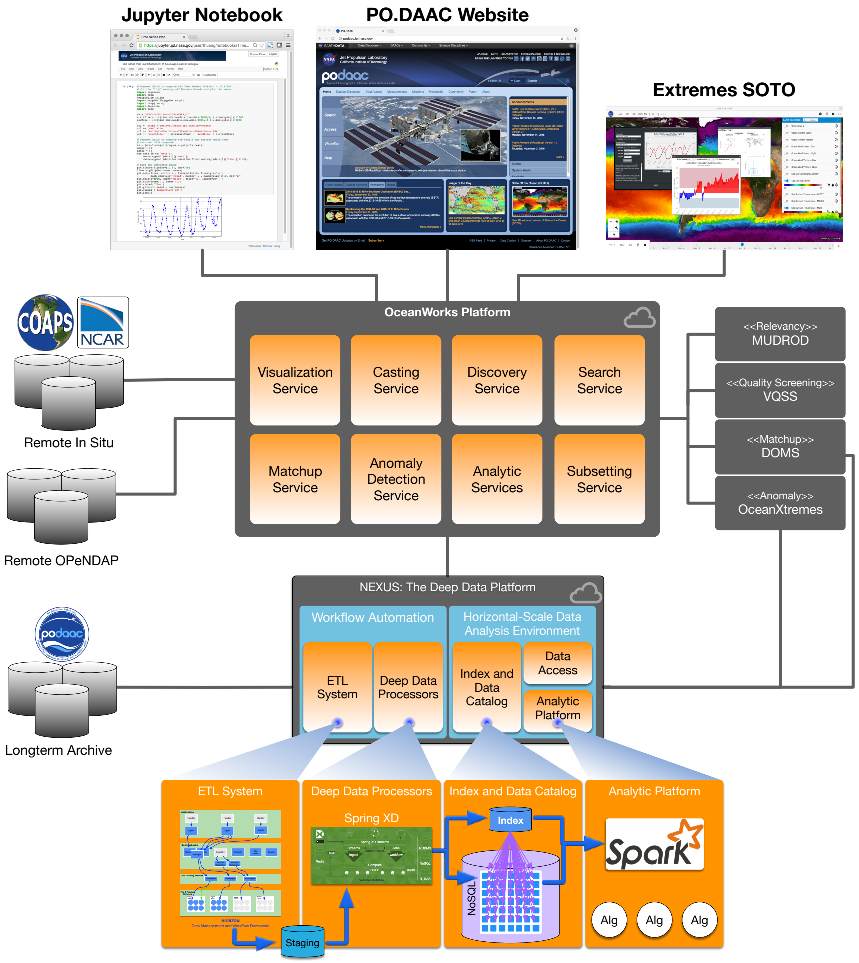
    Distributed in-situ to satellite matchup
  - **AIST-14 MUDROD (PI: Yang/GMU) – TRL 7**

    Search relevancy and discovery
  - **ACCESS-13 VQSS (PI: Armstrong/JPL) – TRL 7**

    Virtualized Quality Screening Service



Technology Readiness Level (TRL)



Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

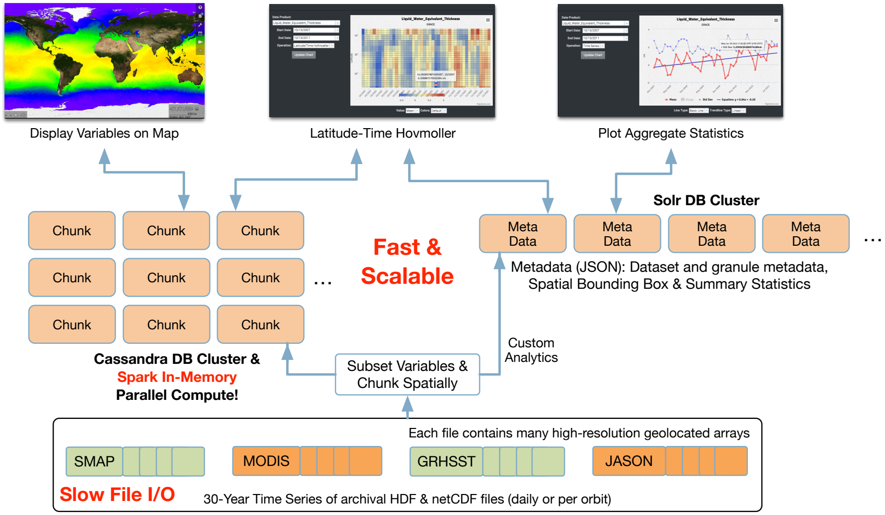
Objectives

- **Improve Data Discovery**: help users to discover relevant data, anomalies and events, quality and uncertainty information, and provide multi-observing system matchup data. **PO.DAAC Search** webservice will be developed to demonstrate new data search and discover related matches and phenomenon
- **Subset and Distribute Data**: high performance data subsetting that supports quality screening
- **Identify and Catalog Ocean phenomenon**: registry of detected oceanographic phenomenon and published using datacasting technology
- **Matchup between Satellite and In-Situ Observations**: advanced matchup services that support in-situ to satellite colocation. It provides a mechanism for users to input a series of geospatial references for satellite observations and receive the in-situ observations that matched to the satellite data within a selectable temporal and spatial search domain.
- **Analyze Satellite Observations**: long time-series, correlation map, time averaged map, Hovmöller, climatological map, etc
- **Visualize and analyze Satellite Observation on the Web**: high performance data visualizations with linkages to actual source measurements, events and analytics. A **prototype integration** with PO.DAAC's new State of the Ocean (SOTO) web application will be developed, called **Extremes SOTO**, to demonstrate this objective.
- **API Integration** with Jupyter notebook to demonstrate working directly with OceanWorks' webserivce platform

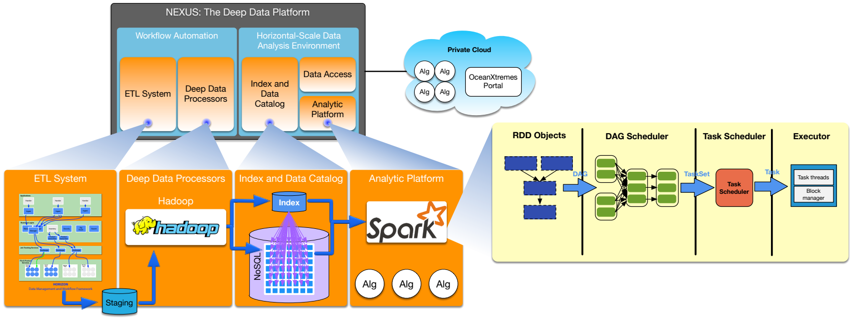
NEXUS: Deep Data Platform

- A data-intensive analysis solution using a new approach for handling science data to enable large-scale data analysis
- Streaming architecture for horizontal scale data ingestion
- Scales horizontally to handle massive amount of data in parallel
- Provides high-performance geospatial and indexed search solution
- Provides tiled data storage architecture to eliminate file I/O overhead
- A growing collection of science analysis webservices using Apache Spark: parallel compute, in-memory map-reduce framework
- Pre-Chunk and Summarize Key Variables
  - Easy statistics instantly (milliseconds)
  - Harder statistics on-demand using Spark (in seconds)
  - Visualize original data (layers) on a map quickly (Cassandra store)
- **Algorithms** – Time Series | Latitude/Time Hovmöller| Longitude/Time Hovmöller| Latitude/Longitude Time Average | Area Averaged Time Series | Time Averaged Map | Climatological Map | Correlation Map | Daily Difference Average

Open Source: Apache License 2
https://github.com/dataplumber/nexus


Display Variables on Map
Latitude-Time Hovmoller
Plot Aggregate Statistics
Chunk
Chunk
Chunk
Chunk
Chunk
Chunk
Chunk
Chunk
Chunk
...
Fast & Scalable
Solr DB Cluster
Meta Data
Meta Data
Meta Data
Meta Data
...
Metadata (JSON): Dataset and granule metadata, Spatial Bounding Box & Summary Statistics
Custom Analytics
Cassandra DB Cluster & Spark In-Memory Parallel Compute!
Subset Variables & Chunk Spatially
Each file contains many high-resolution geolocated arrays
SMAP
MODIS
GRHSST
JASON
Slow File I/O
30-Year Time Series of archival HDF & netCDF files (daily or per orbit)

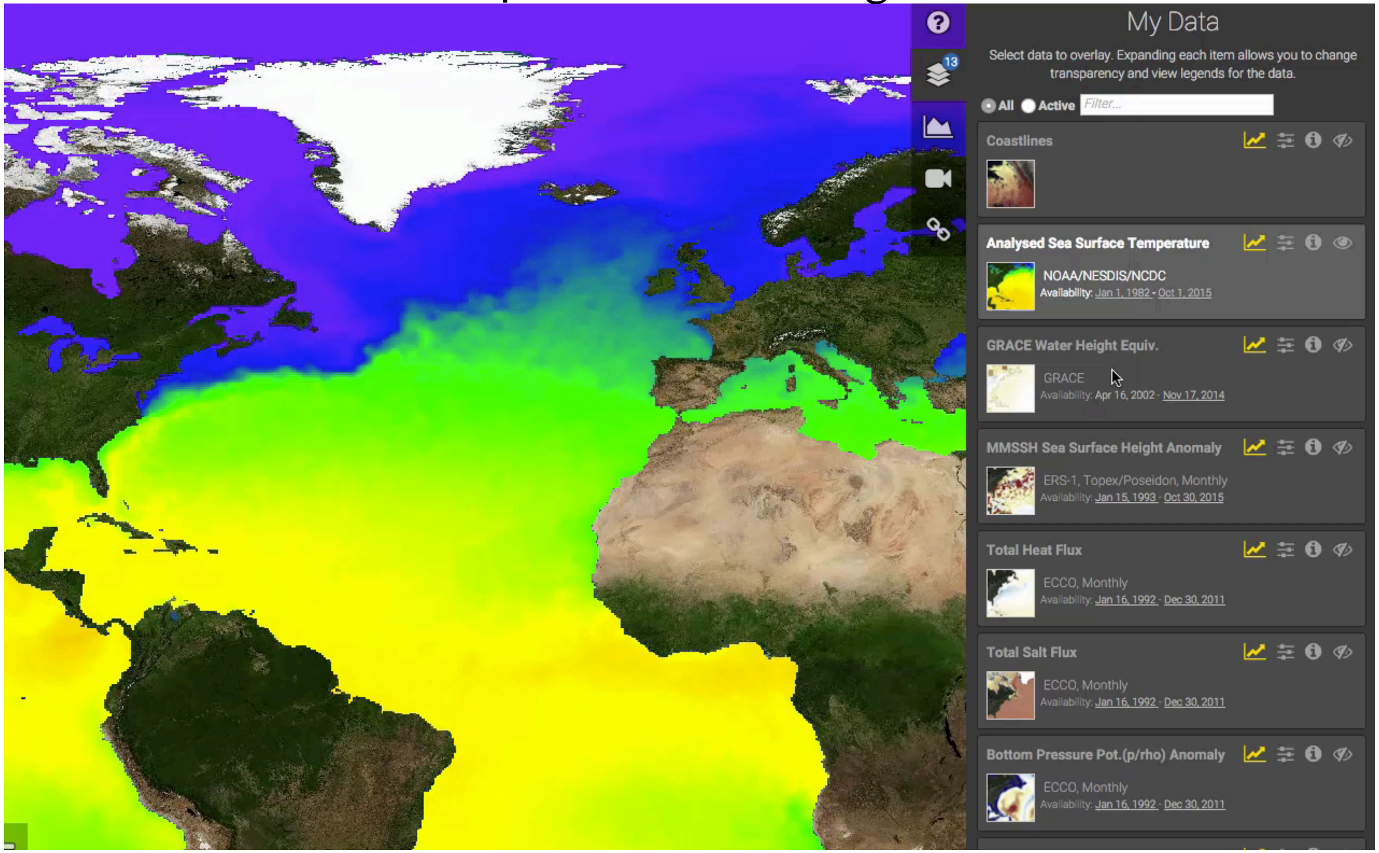
Two-Database Architecture



Deep Data Computing Environment (DDCE)

SEA LEVEL CHANGE
Observations from Space

Data Analysis Tool https://sealevel.nasa.gov



My Data

Select data to overlay. Expanding each item allows you to change transparency and view legends for the data.

○ All ○ Active Filter...

Coastlines

Analysed Sea Surface Temperature
NOAA/NESDIS/NCDC
Availability: Jan 1, 1982 - Oct 1, 2015

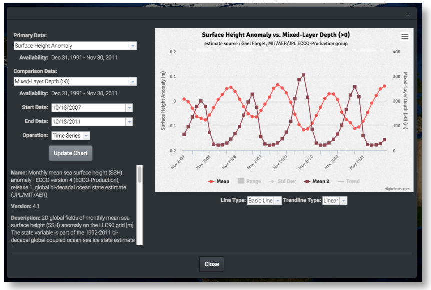
GRACE Water Height Equiv.
GRACE
Availability: Apr 16, 2002 - Nov 17, 2014

MMSSH Sea Surface Height Anomaly
ERS-1, Topex/Poseidon, Monthly
Availability: Jan 15, 1993 - Oct 30, 2015

Total Heat Flux
ECCO, Monthly
Availability: Jan 16, 1992 - Dec 30, 2011

Total Salt Flux
ECCO, Monthly
Availability: Jan 16, 1992 - Dec 30, 2011

Bottom Pressure Pot.(p/rho) Anomaly
ECCO, Monthly
Availability: Jan 16, 1992 - Dec 30, 2011



Primary Data: Surface Height Anomaly
Availability: Dec 31, 1991 - Nov 30, 2011
Comparison Data: Mixed-Layer Depth (>0)
Availability: Dec 31, 1991 - Nov 30, 2011
Start Date: 10/13/2007
End Date: 10/13/2011
Operation: Time Series

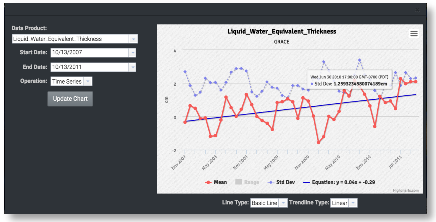
Name: Monthly mean sea surface height (SSH) anomaly - ECCO version 4 (ECCO-Production), release 1, global bi-decadal ocean state estimate (JPL/MIT/AER)

Version: 4.1

Description: 2D global fields of monthly mean sea surface height (SSH) anomaly on the LLC90 grid [m] The state variable is part of the 1992-2011 bi-decadal global coupled ocean-sea ice state estimate

Line Type: Basic Line Trendline Type: Linear

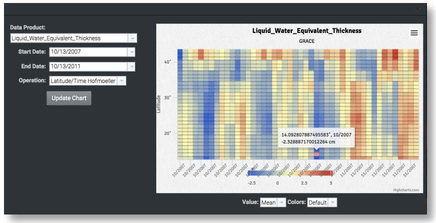
Update Chart
Close
Surface Height Anomaly vs. Mixed-Layer Depth (>0)
estimate source : Gael Forget, MIT/AER/JPL ECCO-Production group
Surface Height Anomaly (m)
Mixed-Layer Depth (>0) (m)
0.2
0.1
0
-0.1
-0.2
400
300
200
100
0
Nov 2007
May 2008
Nov 2008
May 2009
Nov 2009
May 2010
Nov 2010
May 2011
Mean
Range
Std Dev
Mean 2
Trend
Highcharts.com



Data Product: Liquid_Water_Equivalent_Thickness
Start Date: 10/13/2007
End Date: 10/13/2011
Operation: Time Series

Line Type: Basic Line Trendline Type: Linear

Update Chart
Liquid_Water_Equivalent_Thickness
GRACE
cm
4
2
0
-2
Nov 2007
May 2008
Nov 2008
May 2009
Nov 2009
May 2010
Nov 2010
Jul 2011
Wed Jun 30 2010 17:00:00 GMT-0700 (PDT)
Std Dev: 1.2593234580074189cm
Mean
Range
Std Dev
Equation: y = 0.04x + -0.29
Highcharts.com



Data Product: Liquid_Water_Equivalent_Thickness
Start Date: 10/13/2007
End Date: 10/13/2011
Operation: Latitude/Time Hofmoeller

Value: Mean Colors: Default

Facebook: 28K followers
Twitter: 22K followers

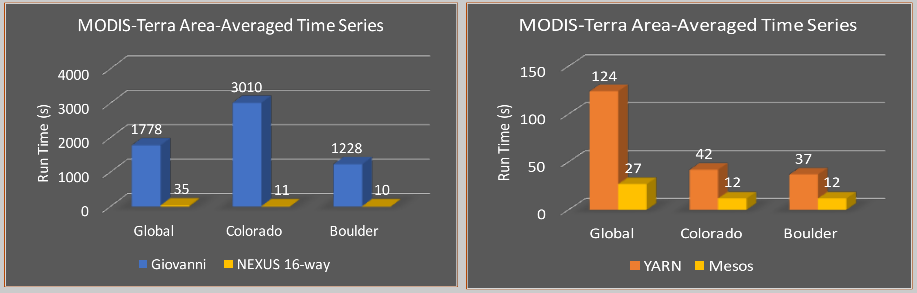
“NASA Sea Level Change Website Offers Everything You Need To Know About Climate Change”
http://www.techtimes.com/articles/147210/20160405/nasa-sea-level-change-website-offers-everything-need-know-climate.htm

“NASA’s new sea level site puts climate change papers, data, and tools online”
http://techcrunch.com/2016/04/04/nasas-new-sea-level-site-puts-climate-change-papers-data-and-tools-online/

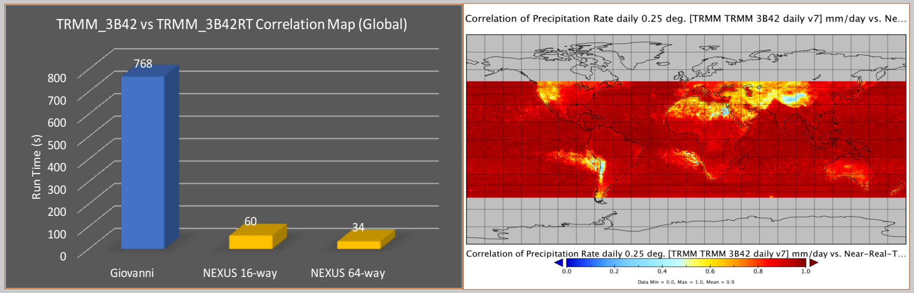
2016/2017 NEXUS Activities

- **2016 Earthdata prototyping**
  - Twelve factor application
  - Benchmarking against Giovanni
  - Amazon cloud deployment
- **2016 ESTO Data Container Study**
  - Look at a variety of technologies for reorganizing and storing Earth science data to make them more tractable to full-scale science analysis.
  - To understand the strong points and tradeoffs of the different approaches to large-scale analysis
  - NEXUS' plug-an-play data storage
    - Apache Cassandra – Java implementation
    - ScyllaDB – C++ implementation
  - Task 1 – long time series
    - Point Time series for Boulder, CO
    - Area-averaged time series for the state of Colorado
    - Area Averaged time series for the globe
  - Task 2 – climatological map
  - Document ETL process, elapsed time, compute and storage
- **2017 Earthdata prototyping**
  - Auto ingestion
  - NASAcompilant Generation Application Platform (NGAP) Infusion
  - Cloud Analysis Toolkit to Enable Earth Science (CATEE)
  - 2017 ESIP Summer Workshops

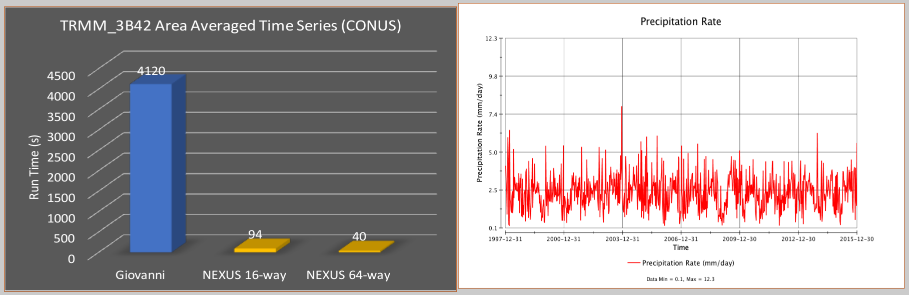
16-year area-averaged daily time series of MODIS-Terra Aerosol Optical Depth (AOD) 550 nm



14-year global correlation map of 0.25 degree TRMM daily precipitation rate

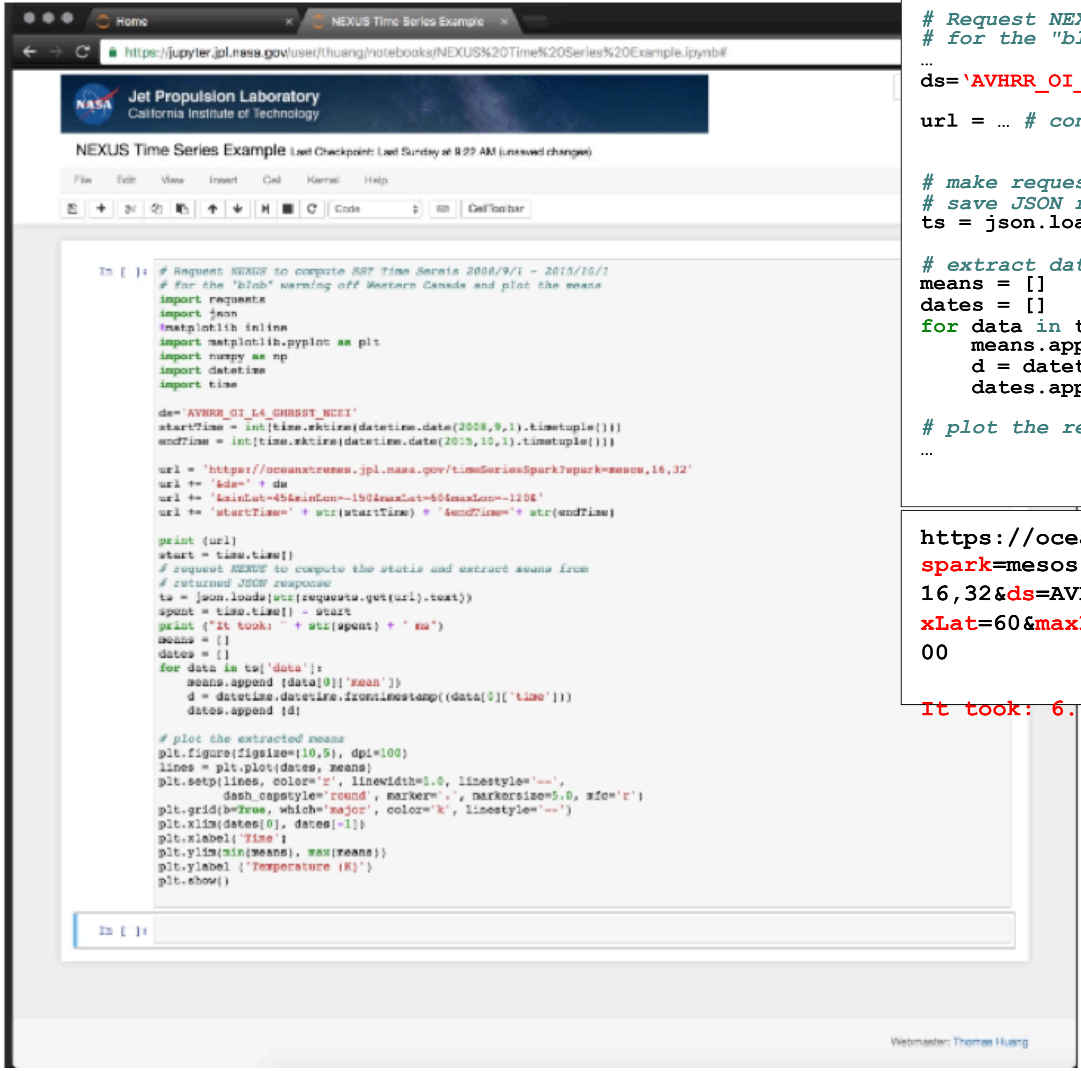


17-year area-averaged time series over the continental United States



Jupyter Notebook Integration

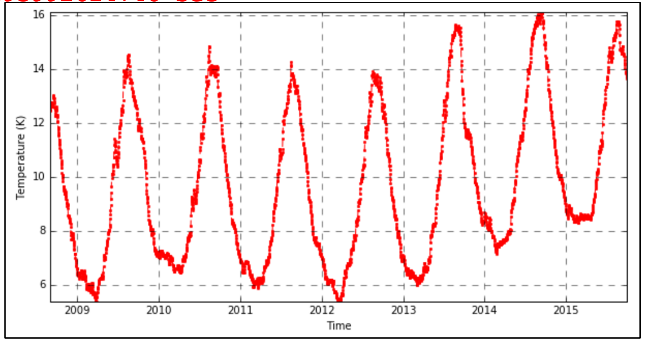
https://jupyter.jpl.nasa.gov


Home
NEXUS Time Series Example
https://jupyter.jpl.nasa.gov/user/thuang/notebooks/NEXUS%20Time%20Series%20Example.ipynb#
Jet Propulsion Laboratory
California Institute of Technology
NEXUS Time Series Example Last Checkpoint: Last Sunday at 9:22 AM (unsaved changes)
File Edit View Insert Cell Kernel Help
Code
CellToolbar
Webmaster: Thomas Huang

```
In [ ]: # Request NEXUS to compute SST Time Sereis 2008/9/1 - 2015/10/1
        # for the 'blob' warming off Western Canada and plot the means
        import requests
        import json
        %matplotlib inline
        import matplotlib.pyplot as plt
        import numpy as np
        import datetime
        import time

        ds='AVHRR_OI_L4_GHRSST_NCEI'
        startTime = int(time.mktime(datetime.date(2008,9,1).timetuple()))
        endTime = int(time.mktime(datetime.date(2015,10,1).timetuple()))

        url = 'https://oceanxtremes.jpl.nasa.gov/timeSeriesSpark?spark=mesos,16,32'
        url += '&ds=' + ds
        url += '&minLat=45&minLon=-150&maxLat=60&maxLon=-120&'
        url += 'startTime=' + str(startTime) + '&endTime='+ str(endTime)

        print (url)
        start = time.time()
        # request NEXUS to compute the statis and extract means from
        # returned JSON response
        ts = json.loads(str(requests.get(url).text))
        spent = time.time() - start
        print ("It took: " + str(spent) + " ms")
        means = []
        dates = []
        for data in ts['data']:
            means.append (data[0]['mean'])
            d = datetime.datetime.fromtimestamp((data[0]['time']))
            dates.append (d)

        # plot the extracted means
        plt.figure(figsize=(10,5), dpi=100)
        lines = plt.plot(dates, means)
        plt.setp(lines, color='r', linewidth=1.0, linestyle='--',
                 dash_capstyle='round', marker='.', markersize=5.0, mfc='r')
        plt.grid(b=True, which='major', color='k', linestyle='--')
        plt.xlim(dates[0], dates[-1])
        plt.xlabel('Time')
        plt.ylim(min(means), max(means))
        plt.ylabel ('Temperature (K)')
        plt.show()

In [ ]:
```

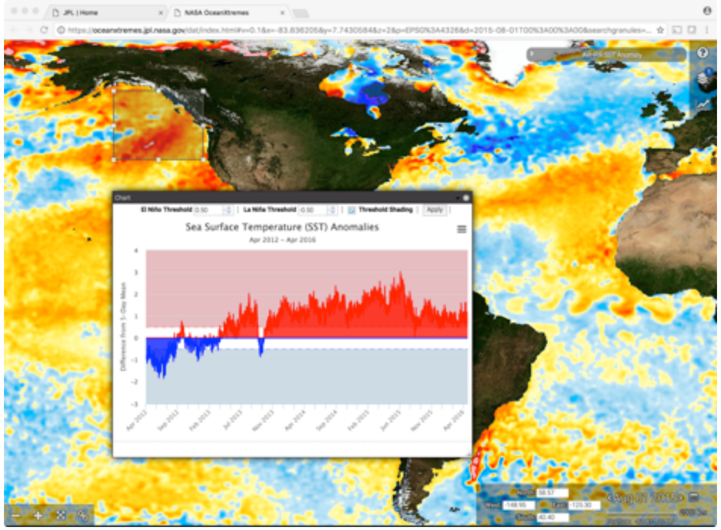
```
# Request NEXUS to compute SST Time Series 2008/9/1 - 2015/10/1
# for the "blob" warming off Western Canada and plot the means
…
ds='AVHRR_OI_L4_GHRSST_NCEI'

url = … # construct the webservice URL request


# make request to NEXUS using URL request
# save JSON response in local variable
ts = json.loads(str(requests.get(url).text))

# extract dates and means from the response
means = []
dates = []
for data in ts['data']:
    means.append (data[0]['mean'])
    d = datetime.datetime.fromtimestamp((data[0]['time']))
    dates.append (d)

# plot the result
…
```

```
https://oceanxtremes.jpl.nasa.gov/timeSeriesSpark?spark=mesos,
16,32&ds=AVHRR_OI_L4_GHRSST_NCEI&minLat=45&minLon=-150&maxLat=60&maxLon=-120&startTime=1220227200&endTime=1443657600

It took: 6.909193992614746 sec
```

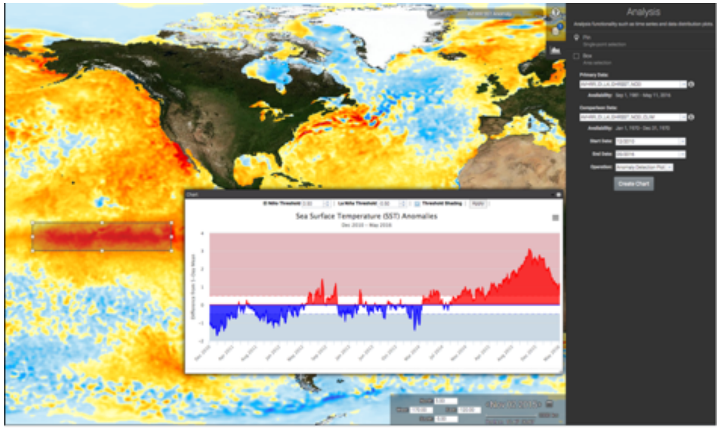


OceanXtremes: Oceanographic Data-Intensive Anomaly Detection and Analysis Portal

- https://oceanxtremes.jpl.nasa.gov
- An oceanographic data-intensive anomaly detection and analysis portal
- Cloud-based big data analytic platform for
  - Climatology generation
  - On-the-fly daily difference computation
  - Anomaly registry and publication
  - On-the-fly data analytics
- Recent highlights
  - **Recreated identification of ”The Blob”**
    - **The Blob** is the name given to a large mass of relatively warm water in the Pacific ocean off the coast of North America. It was first detected in late 2013 and continued to spread throughout 2014 and 2015.
    - SST anomaly = SST – SST Climatology at each location to compare with standard deviation - Dr. Chelle Gentemann, Senior Scientist at Earth & Space Research
  - **Recreated the El Niño 3.4 regional signal**
    - Dec. 2010 – May 2016
    - **El Niño** is a phenomenon in the equatorial Pacific Ocean characterized by a five consecutive 3-month running mean of sea surface temperature (SST) anomalies in the Niño 3.4 region that is above (below) the threshold of +0.5°C (-0.5°C). This standard of measure is known as the Oceanic Niño Index (ONI).



The Blob



El Niño 3.4 regional signal

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Hurricane Katrina Study – Using OceanXtremes

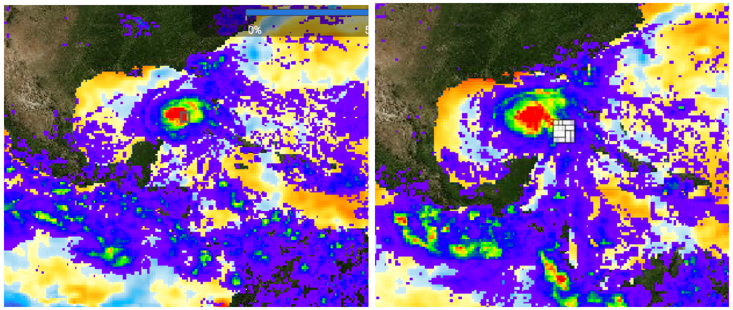
Hurricane Katrina passed to the southwest of Florida on Aug 27, 2005. The ocean response in a 1 x 1 deg region is captured by a number of satellites. The initial ocean response was an immediate cooling of the surface waters by 2 deg C that lingers for several days. Following this was a short intense ocean chlorophyll bloom a few days later. The ocean may have been "preconditioned' by a cool core eddy and low sea surface height.

The SST drop is correlated to both wind and precipitation data. The Chl-A data is lagged by about 3 days to the other observations like SST, wind and precipitation.

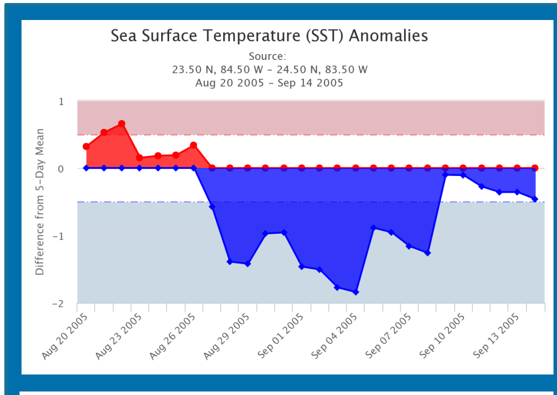
A study of a Hurricane Katrina–induced phytoplankton bloom using satellite observations and model simulations
Xiaoming Liu, Menghua Wang, and Wei Shi

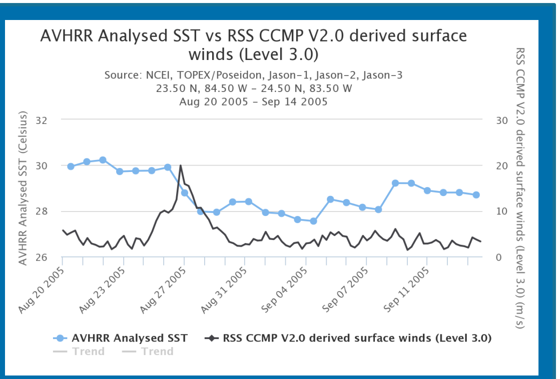
JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, C03023, doi:10.1029/2008JC004934, 2009

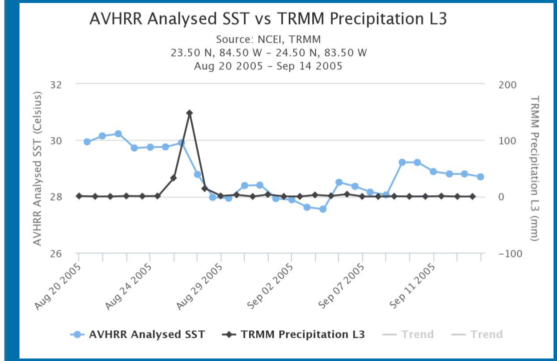
http://shoni2.princeton.edu/ftp/lyo/journals/Ocean/phybiogeochem/Liu-etal-KatrinaChlBloom-JGR2009.pdf

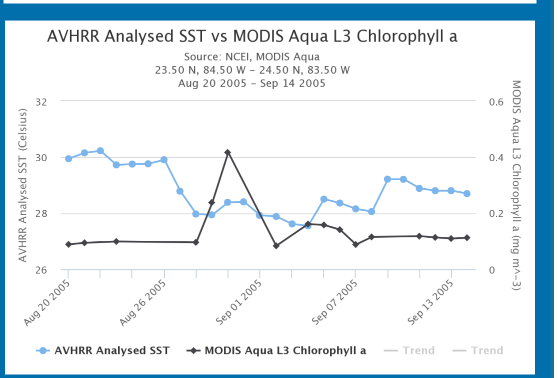


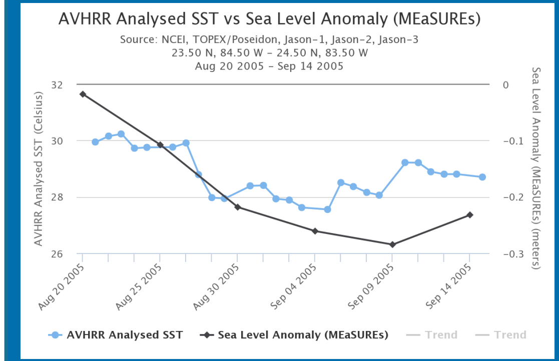
Hurricane Katrina
TRMM overlay SST Anomaly

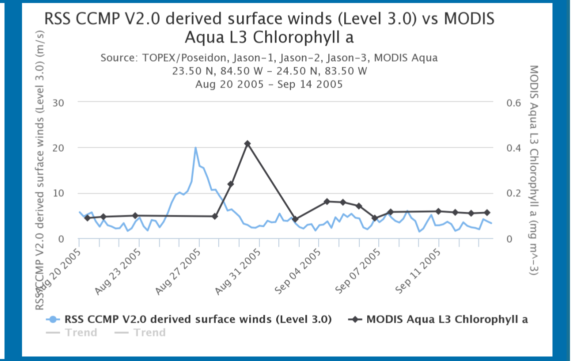










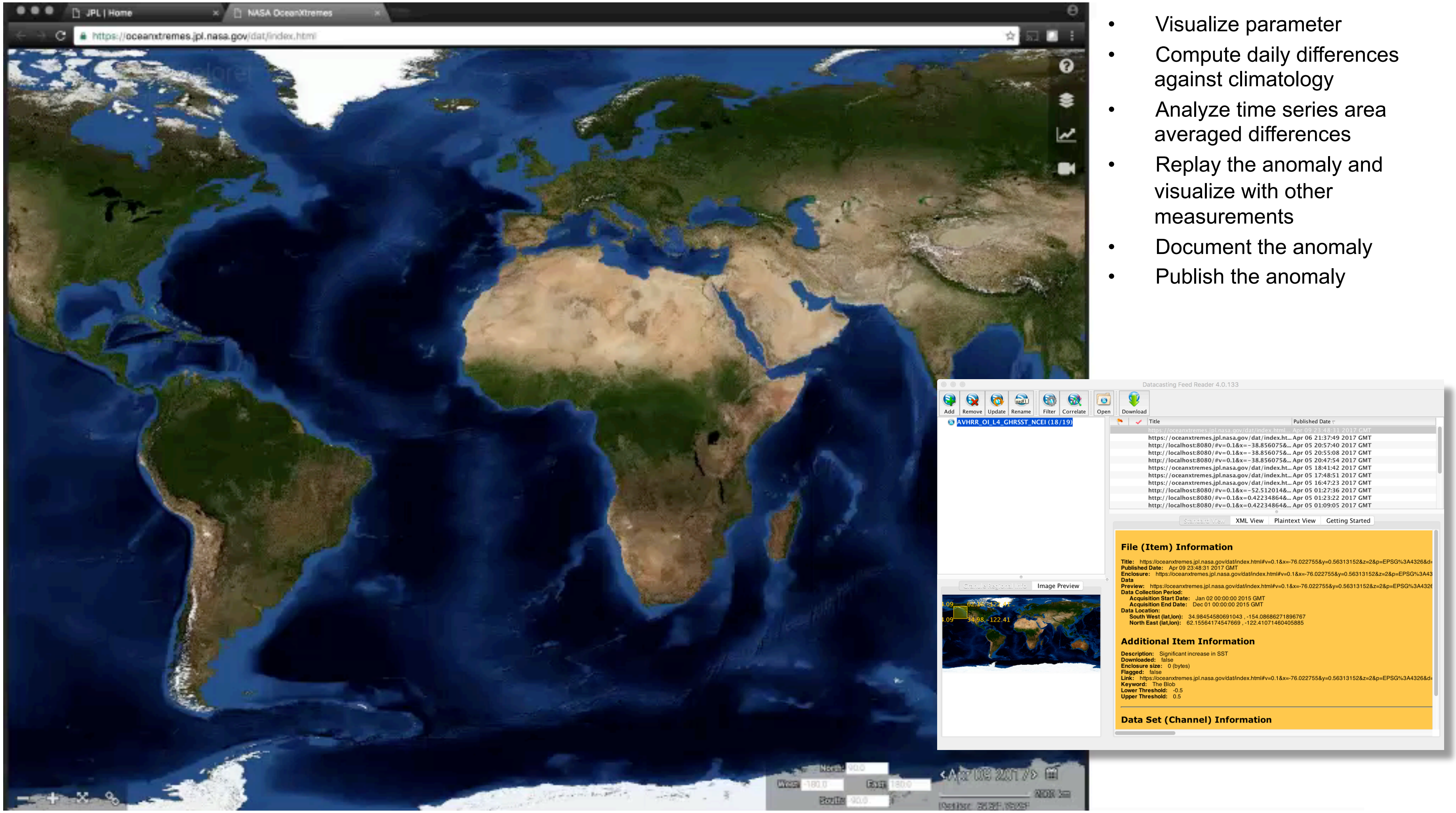


Powered By NEXUS

2017 ESTF

Earth Science Technology Office

OceanXtremes: Identify . Analyze . Share



Datacasting Feed Reader 4.0.133

AVHRR_OI_L4_GHRSST_NCEI (18/19)

Title	Published Date
https://oceanxtremes.jpl.nasa.gov/dat/index.html...	Apr 09 23:48:31 2017 GMT
https://oceanxtremes.jpl.nasa.gov/dat/index.ht...	Apr 06 21:37:49 2017 GMT
http://localhost:8080/#v=0.1&x=-38.856075&...	Apr 05 20:57:40 2017 GMT
http://localhost:8080/#v=0.1&x=-38.856075&...	Apr 05 20:55:08 2017 GMT
http://localhost:8080/#v=0.1&x=-38.856075&...	Apr 05 20:47:54 2017 GMT
https://oceanxtremes.jpl.nasa.gov/dat/index.ht...	Apr 05 18:41:42 2017 GMT
https://oceanxtremes.jpl.nasa.gov/dat/index.ht...	Apr 05 17:48:51 2017 GMT
https://oceanxtremes.jpl.nasa.gov/dat/index.ht...	Apr 05 16:47:23 2017 GMT
http://localhost:8080/#v=0.1&x=-52.512014&...	Apr 05 01:27:36 2017 GMT
http://localhost:8080/#v=0.1&x=0.42234864&...	Apr 05 01:23:22 2017 GMT
http://localhost:8080/#v=0.1&x=0.42234864&...	Apr 05 01:09:05 2017 GMT

File (Item) Information

Title: https://oceanxtremes.jpl.nasa.gov/dat/index.html#v=0.1&x=-76.022755&y=0.56313152&z=2&p=EPSG%3A4326&d=
Published Date: Apr 09 23:48:31 2017 GMT
Enclosure: https://oceanxtremes.jpl.nasa.gov/dat/index.html#v=0.1&x=-76.022755&y=0.56313152&z=2&p=EPSG%3A43
Data Preview: https://oceanxtremes.jpl.nasa.gov/dat/index.html#v=0.1&x=-76.022755&y=0.56313152&z=2&p=EPSG%3A4326
Data Collection Period:
Acquisition Start Date: Jan 02 00:00:00 2015 GMT
Acquisition End Date: Dec 01 00:00:00 2015 GMT
Data Location:
South West (lat,lon): 34.98454580691043 , -154.08686271896767
North East (lat,lon): 62.15564174547669 , -122.41071460405885

Additional Item Information

Description: Significant increase in SST
Downloaded: false
Enclosure size: 0 (bytes)
Flagged: false
Link: https://oceanxtremes.jpl.nasa.gov/dat/index.html#v=0.1&x=-76.022755&y=0.56313152&z=2&p=EPSG%3A4326&d=
Keyword: The Blob
Lower Threshold: -0.5
Upper Threshold: 0.5

Data Set (Channel) Information

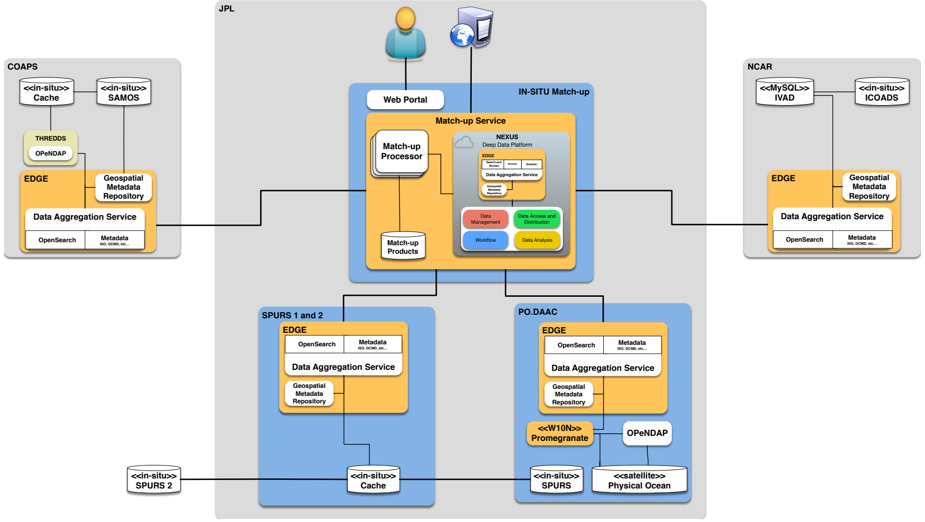
- Visualize parameter
- Compute daily differences against climatology
- Analyze time series area averaged differences
- Replay the anomaly and visualize with other measurements
- Document the anomaly
- Publish the anomaly

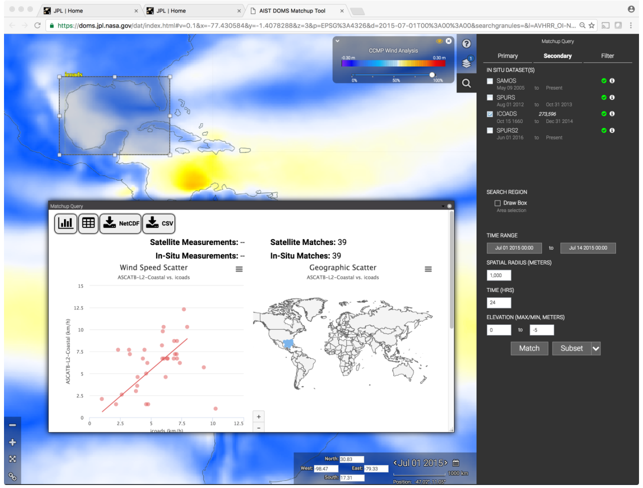
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

DOMS:
Distributed Oceanographic Matchup Service

- https://doms.jpl.nasa.gov
- Distributed Oceanographic Matchup Service
- Typically data matching is done using one-off programs developed at multiple institutions
- A primary advantage of DOMS is the reduction in duplicate development and man hours required to match satellite/in situ data
  - 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), i.e., matchup archives are always up-to-date.
- In situ data nodes at JPL, NCAR, and FSU operational.
- Provides data querying, subset creation, match-up services, and file delivery operational.
- Prototype graphical user interface (UI) and APIs accessible for external users.
- Plugin architecture for in situ data source using EDGE
  - Extensible Data Gateway Environment is an Apache License 2 open source technology
  - https://github.com/dataplumber/edge
- Defined specification for packaging matchup results. Working with Unidata and ESDSWG's data interoperability and standard groups




JPL | Home — JPL | Home — AIST DOMS Matchup Tool

https://doms.jpl.nasa.gov/dat/index.html#v=0.1&x=-77.430584&y=-1.4078288&z=3&p=EPSG%3A4326&d=2015-07-01T00%3A00%3A00&searchgranules=&l=AVHRR_OI-N...

Matchup Query

Primary | Secondary | Filter

IN SITU DATASET(S)

	Dataset		From		To
□	SAMOS		May 09 2005	to	Present
□	SPURS		Aug 01 2012	to	Oct 31 2013
☑	ICOADS	273,596	Oct 15 1660	to	Dec 31 2014
□	SPURS2		Jun 01 2016	to	Present

SEARCH REGION
□ Draw Box
Area selection

TIME RANGE
Jul 01 2015 00:00 to Jul 14 2015 00:00

SPATIAL RADIUS (METERS)
1,000

TIME (HRS)
24

ELEVATION (MAX/MIN, METERS)
0 to -5

Match | Subset

Matchup Query

NetCDF | CSV

Satellite Measurements: --
In-Situ Measurements: --
Satellite Matches: 39
In-Situ Matches: 39

North: 30.83
West: -98.47
East: -79.33
South: 17.31

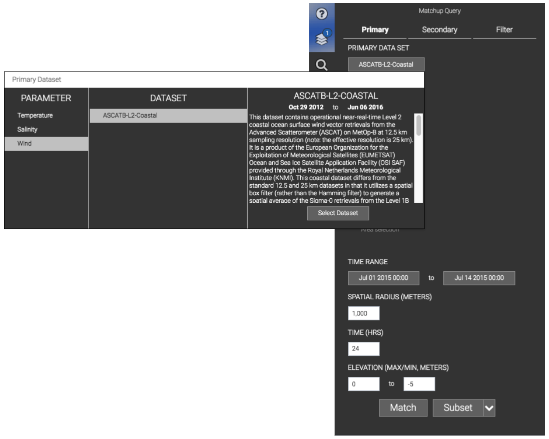
<Jul 01 2015>

Position: -47.02°, 11.05°

ESTO
Earth Science Technology Office

More than just matchup

Primary Dataset

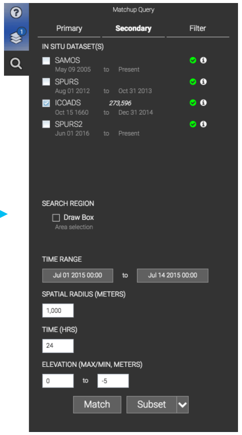


ASCATB-L2-COASTAL

Oct 29 2012 to **Jun 06 2016**

This dataset contains operational near-real-time Level 2 coastal ocean surface wind vector retrievals from the Advanced Scatterometer (ASCAT) on MetOp-B at 12.5 km sampling resolution (note: the effective resolution is 25 km). It is a product of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI SAF) provided through the Royal Netherlands Meteorological Institute (KNMI). This coastal dataset differs from the standard 12.5 and 25 km datasets in that it utilizes a spatial box filter (rather than the Hamming filter) to generate a spatial average of the Sigma-0 retrievals from the Level 1B

Match-up In-Situ



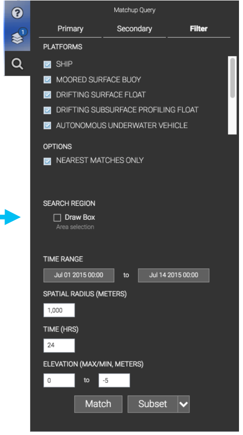
□ SAMOS — May 09 2005 to Present

□ SPURS — Aug 01 2012 to Oct 31 2013

☑ ICOADS 273,596 — Oct 15 1660 to Dec 31 2014

□ SPURS2 — Jun 01 2016 to Present

Optional Platform

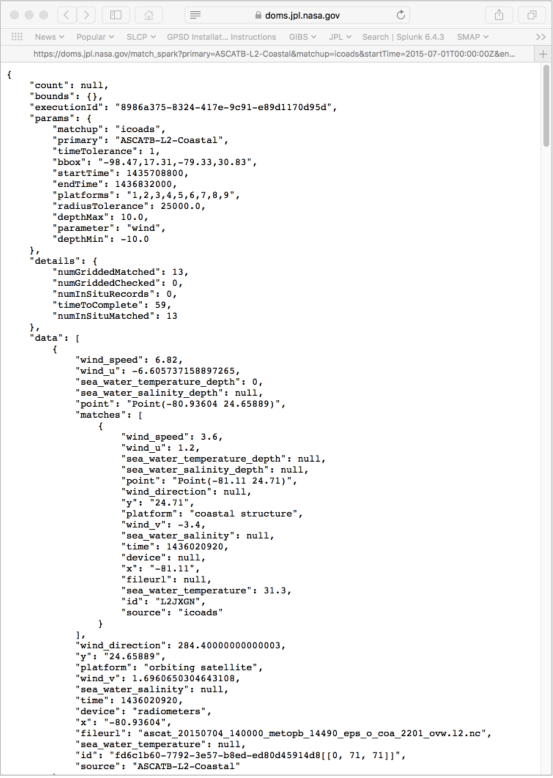


PLATFORMS

☑ SHIP
☑ MOORED SURFACE BUOY
☑ DRIFTING SURFACE FLOAT
☑ DRIFTING SUBSURFACE PROFILING FLOAT
☑ AUTONOMOUS UNDERWATER VEHICLE

OPTIONS

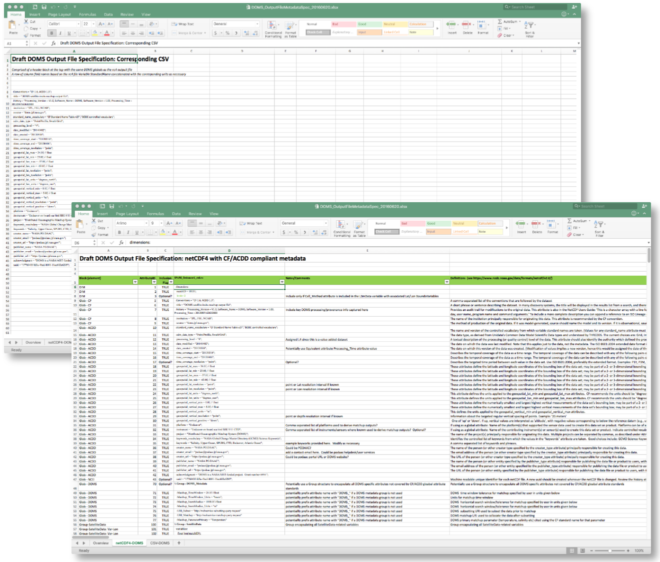
☑ NEAREST MATCHES ONLY



https://doms.jpl.nasa.gov/match_spark?primary=ASCATB-L2-Coastal&matchup=icoads&startTime=2015-07-01T00:00:00Z&en...

```
{
    "count": null,
    "bounds": {},
    "executionId": "8986a375-8324-417e-9c91-e89d1170d95d",
    "params": {
        "matchup": "icoads",
        "primary": "ASCATB-L2-Coastal",
        "timeTolerance": 1,
        "bbox": "-98.47,17.31,-79.33,30.83",
        "startTime": 1435708800,
        "endTime": 1436832000,
        "platforms": "1,2,3,4,5,6,7,8,9",
        "radiusTolerance": 25000.0,
        "depthMax": 10.0,
        "parameter": "wind",
        "depthMin": -10.0
    },
    "details": {
        "numGriddedMatched": 13,
        "numGriddedChecked": 0,
        "numInSituRecords": 0,
        "timeToComplete": 59,
        "numInSituMatched": 13
    },
    "data": [
        {
            "wind_speed": 6.82,
            "wind_u": -6.605737158897265,
            "sea_water_temperature_depth": 0,
            "sea_water_salinity_depth": null,
            "point": "Point(-80.93604 24.65889)",
            "matches": [
                {
                    "wind_speed": 3.6,
                    "wind_u": 1.2,
                    "sea_water_temperature_depth": null,
                    "sea_water_salinity_depth": null,
                    "point": "Point(-81.11 24.71)",
                    "wind_direction": null,
                    "y": "24.71",
                    "platform": "coastal structure",
                    "wind_v": -3.4,
                    "sea_water_salinity": null,
                    "time": 1436020920,
                    "device": null,
                    "x": "-81.11",
                    "fileurl": null,
                    "sea_water_temperature": 31.3,
                    "id": "L2JXGN",
                    "source": "icoads"
                }
            ],
            "wind_direction": 284.40000000000003,
            "y": "24.65889",
            "platform": "orbiting satellite",
            "wind_v": 1.6960650304643108,
            "sea_water_salinity": null,
            "time": 1436020920,
            "device": "radiometers",
            "x": "-80.93604",
            "fileurl": "ascat_20150704_140000_metopb_14490_eps_o_coa_2201_ovw.l2.nc",
            "sea_water_temperature": null,
            "id": "fd6c1b60-7792-3e57-b8ed-ed80d45914d8[[0, 71, 71]]",
            "source": "ASCATB-L2-Coastal"
```

- **Workflow driven**
- **Fast on-the-fly subseting satellite and in-situ**
- **Open webservice API**
- **ON-the-fly data analysis using NEXUS**
- **Cloud ready**



Draft DOMS Output File Specification: Corresponding CSV

Draft DOMS Output File Specification: netCDF4 with CF/ACDD compliant metadata

PO.DAAC Satellite Data

- AVHRR_OI_L4_GHRSST_NCEI
- JPL-L4-GHRSST-SSTfnd-MUR-GLOB-v02.0fv04.1
- SMAP_L2B_SSS
- ASCATB-L2-Costal

In-situ Data

- JPL SPURS 1 and 2
- FSU SAMOS
- NCAR ICOADS

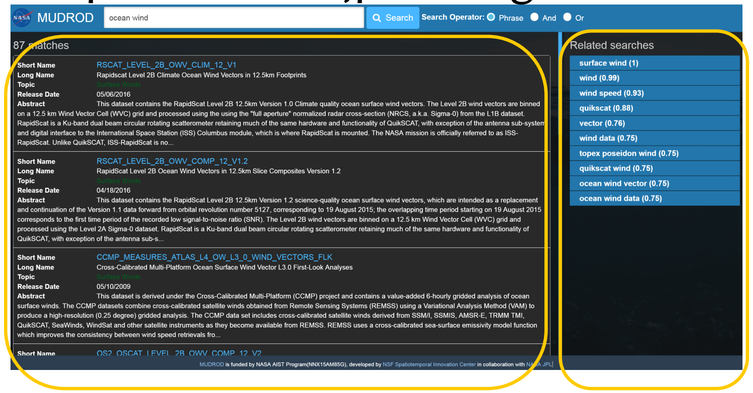
Drafted and implemented matchup specification

- Promote CF and ACDD standards
- Self-contained
- Simple to use by other tools and services

MUDROD:
Mining and Utilizing Data Relevancy from Oceanographic Dataset

- Mining and Utilizing Dataset Relevancy from Oceanographic Dataset
- **Search** – look for something you expect to exist
  - Information tagging
  - Indexed search technologies like Apache Solr or ElasticSearch
  - The solution is pretty straightforward
- **Discovery** – find something new, or in a new way
  - This is non-trivial
  - Traditional ontological method doesn't quite add up
  - The strength of semantic web is in inference
  - What happen when we have a lot of `subClassOf`, `equivalentClassOf`, `sameAs`?
  - How wide and deep should we go?
- **Relevancy**
  - It is domain-specific
  - It is personal
  - It is temporal
  - It is dynamic
- MUDROD analyzes web logs to discover user knowledge (the connections between datasets and keyword)
- Construct knowledge base by combining semantics and profile analyzer
- Improve data discovery by better ranked results

https://mudrod.jpl.nasa.gov



MUDROD ocean wind Search Search Operator: ● Phrase ○ And ○ Or

87 matches

Short Name RSCAT_LEVEL_2B_OWV_CLIM_12_V1
Long Name Rapidscat Level 2B Climate Ocean Wind Vectors in 12.5km Footprints
Topic Surface Winds
Release Date 05/06/2016
Abstract This dataset contains the RapidScat Level 2B 12.5km Version 1.0 Climate quality ocean surface wind vectors. The Level 2B wind vectors are binned on a 12.5 km Wind Vector Cell (WVC) grid and processed using the using the "full aperture" normalized radar cross-section (NRCS, a.k.a. Sigma-0) from the L1B dataset. RapidScat is a Ku-band dual beam circular rotating scatterometer retaining much of the same hardware and functionality of QuikSCAT, with exception of the antenna sub-system and digital interface to the International Space Station (ISS) Columbus module, which is where RapidScat is mounted. The NASA mission is officially referred to as ISS-RapidScat. Unlike QuikSCAT, ISS-RapidScat is no...

Short Name RSCAT_LEVEL_2B_OWV_COMP_12_V1.2
Long Name RapidScat Level 2B Ocean Wind Vectors in 12.5km Slice Composites Version 1.2
Topic Surface Winds
Release Date 04/18/2016
Abstract This dataset contains the RapidScat Level 2B 12.5km Version 1.2 science-quality ocean surface wind vectors, which are intended as a replacement and continuation of the Version 1.1 data forward from orbital revolution number 5127, corresponding to 19 August 2015; the overlapping time period starting on 19 August 2015 corresponds to the first time period of the recorded low signal-to-noise ratio (SNR). The Level 2B wind vectors are binned on a 12.5 km Wind Vector Cell (WVC) grid and processed using the Level 2A Sigma-0 dataset. RapidScat is a Ku-band dual beam circular rotating scatterometer retaining much of the same hardware and functionality of QuikSCAT, with exception of the antenna sub-s...

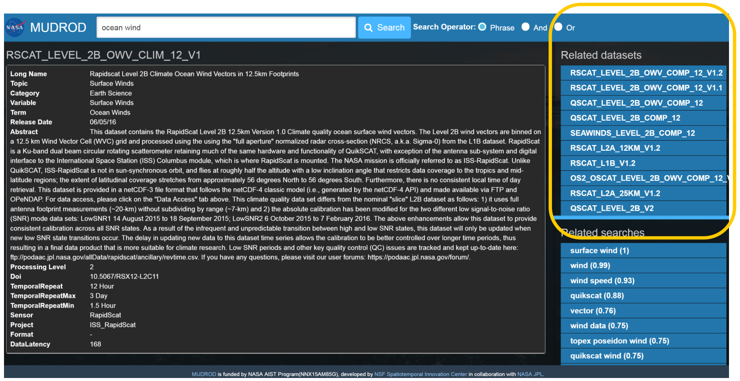
Short Name CCMP_MEASURES_ATLAS_L4_OW_L3_0_WIND_VECTORS_FLK
Long Name Cross-Calibrated Multi-Platform Ocean Surface Wind Vector L3.0 First-Look Analyses
Topic Surface Winds
Release Date 05/10/2009
Abstract This dataset is derived under the Cross-Calibrated Multi-Platform (CCMP) project and contains a value-added 6-hourly gridded analysis of ocean surface winds. The CCMP datasets combine cross-calibrated satellite winds obtained from Remote Sensing Systems (REMSS) using a Variational Analysis Method (VAM) to produce a high-resolution (0.25 degree) gridded analysis. The CCMP data set includes cross-calibrated satellite winds derived from SSM/I, SSMIS, AMSR-E, TRMM TMI, QuikSCAT, SeaWinds, WindSat and other satellite instruments as they become available from REMSS. REMSS uses a cross-calibrated sea-surface emissivity model function which improves the consistency between wind speed retrievals fro...

Short Name OS2_OSCAT_LEVEL_2B_OWV_COMP_12_V2

Related searches
- surface wind (1)
- wind (0.99)
- wind speed (0.93)
- quikscat (0.88)
- vector (0.76)
- wind data (0.75)
- topex poseidon wind (0.75)
- quikscat wind (0.75)
- ocean wind vector (0.75)
- ocean wind data (0.75)

MUDROD is funded by NASA AIST Program(NNX15AM85G), developed by NSF Spatiotemporal Innovation Center in collaboration with NASA JPL.

Search Ranking
Based on a machine learning model (RankSVM) which takes a number of features, such as vector space model, version, processing level, release date, all-time popularity, monthly-popularity, and user popularity.



MUDROD ocean wind Search Search Operator: ● Phrase ○ And ○ Or

RSCAT_LEVEL_2B_OWV_CLIM_12_V1

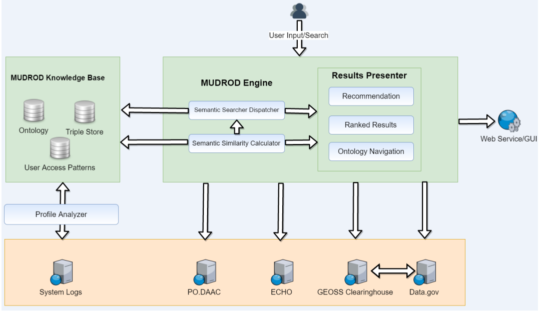
Long Name Rapidscat Level 2B Climate Ocean Wind Vectors in 12.5km Footprints
Topic Surface Winds
Category Earth Science
Variable Surface Winds
Term Ocean Winds
Release Date 06/05/16
Abstract This dataset contains the RapidScat Level 2B 12.5km Version 1.0 Climate quality ocean surface wind vectors. The Level 2B wind vectors are binned on a 12.5 km Wind Vector Cell (WVC) grid and processed using the using the "full aperture" normalized radar cross-section (NRCS, a.k.a. Sigma-0) from the L1B dataset. RapidScat is a Ku-band dual beam circular rotating scatterometer retaining much of the same hardware and functionality of QuikSCAT, with exception of the antenna sub-system and digital interface to the International Space Station (ISS) Columbus module, which is where RapidScat is mounted. The NASA mission is officially referred to as ISS-RapidScat. Unlike QuikSCAT, ISS-RapidScat is not in sun-synchronous orbit, and flies at roughly half the altitude with a low inclination angle that restricts data coverage to the tropics and mid-latitude regions; the extent of latitudinal coverage stretches from approximately 56 degrees North to 56 degrees South. Furthermore, there is no consistent local time of day retrieval. This dataset is provided in a netCDF-3 file format that follows the netCDF-4 classic model (i.e., generated by the netCDF-4 API) and made available via FTP and OPeNDAP. For data access, please click on the "Data Access" tab above. This climate quality data set differs from the nominal "slice" L2B dataset as follows: 1) it uses full antenna footprint measurements (~20-km) without subdividing by range (~7-km) and 2) the absolute calibration has been modified for the two different low signal-to-noise ratio (SNR) mode data sets: LowSNR1 14 August 2015 to 18 September 2015; LowSNR2 6 October 2015 to 7 February 2016. The above enhancements allow this dataset to provide consistent calibration across all SNR states. As a result of the infrequent and unpredictable transition between high and low SNR states, this dataset will only be updated when new low SNR state transitions occur. The delay in updating new data to this dataset time series allows the calibration to be better controlled over longer time periods, thus resulting in a final data product that is more suitable for climate research. Low SNR periods and other key quality control (QC) issues are tracked and kept up-to-date here: ftp://podaac.jpl.nasa.gov/allData/rapidscat/ancillary/revtime.csv. If you have any questions, please visit our user forums: https://podaac.jpl.nasa.gov/forum/.
Processing Level 2
Doi 10.5067/RSX12-L2C11
TemporalRepeat 12 Hour
TemporalRepeatMax 3 Day
TemporalRepeatMin 1.5 Hour
Sensor RapidScat
Project ISS_RapidScat
Format -
DataLatency 168

Related datasets
- RSCAT_LEVEL_2B_OWV_COMP_12_V1.2
- RSCAT_LEVEL_2B_OWV_COMP_12_V1.1
- QSCAT_LEVEL_2B_OWV_COMP_12
- QSCAT_LEVEL_2B_COMP_12
- SEAWINDS_LEVEL_2B_COMP_12
- RSCAT_L2A_12KM_V1.2
- RSCAT_L1B_V1.2
- OS2_OSCAT_LEVEL_2B_OWV_COMP_12_V
- RSCAT_L2A_25KM_V1.2
- QSCAT_LEVEL_2B_V2

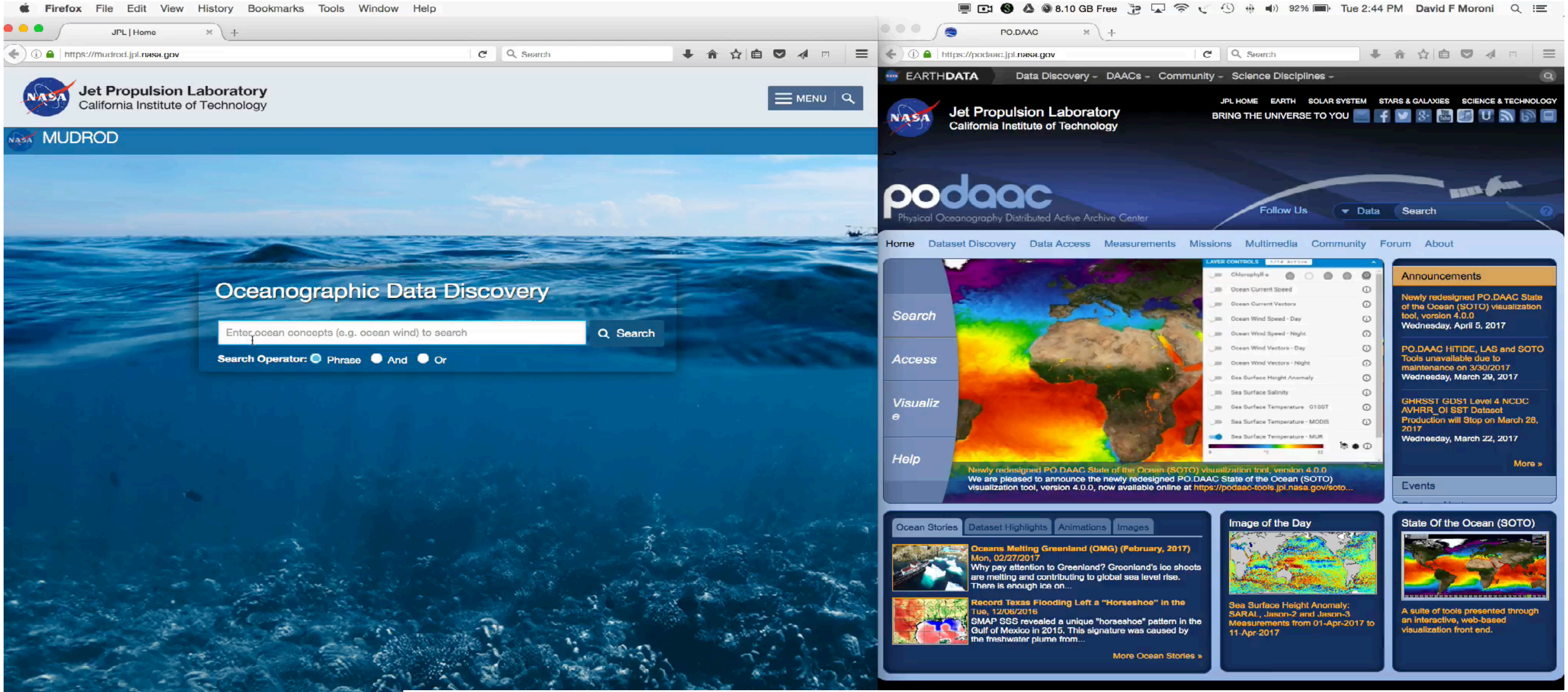
Related searches
- surface wind (1)
- wind (0.99)
- wind speed (0.93)
- quikscat (0.88)
- vector (0.76)
- wind data (0.75)
- topex poseidon wind (0.75)
- quikscat wind (0.75)

MUDROD is funded by NASA AIST Program(NNX15AM85G), developed by NSF Spatiotemporal Innovation Center in collaboration with NASA JPL.

Search Recommendation
Based on dataset metadata content and web session co-occurrence



Speed . Relevant . Discovery



Oceanographic Data Discovery

Enter ocean concepts (e.g. ocean wind) to search — Search

Search Operator: ○ Phrase ○ And ○ Or

Newly redesigned PO.DAAC State of the Ocean (SOTO) visualization tool, version 4.0.0
We are pleased to announce the newly redesigned PO.DAAC State of the Ocean (SOTO) visualization tool, version 4.0.0, now available online at https://podaac-tools.jpl.nasa.gov/soto...

Newly redesigned PO.DAAC State of the Ocean (SOTO) visualization tool, version 4.0.0
Wednesday, April 5, 2017

PO.DAAC HiTIDE, LAS and SOTO Tools unavailable due to maintenance on 3/30/2017
Wednesday, March 29, 2017

GHRSST GDS1 Level 4 NCDC AVHRR_OI SST Dataset Production will Stop on March 28, 2017
Wednesday, March 22, 2017

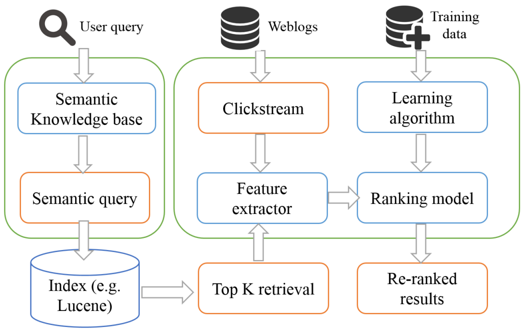
Oceans Melting Greenland (OMG) (February, 2017)
Mon, 02/27/2017
Why pay attention to Greenland? Greenland's ice sheets are melting and contributing to global sea level rise. There is enough ice on...

Record Texas Flooding Left a "Horseshoe" in the
Tue, 12/06/2016
SMAP SSS revealed a unique "horseshoe" pattern in the Gulf of Mexico in 2015. This signature was caused by the freshwater plume from...

Sea Surface Height Anomaly: SARAL, Jason-2 and Jason-3 Measurements from 01-Apr-2017 to 11-Apr-2017

A suite of tools presented through an interactive, web-based visualization front end.

MUDROD is funded by NASA AIST Program(NNX15AM85G), d

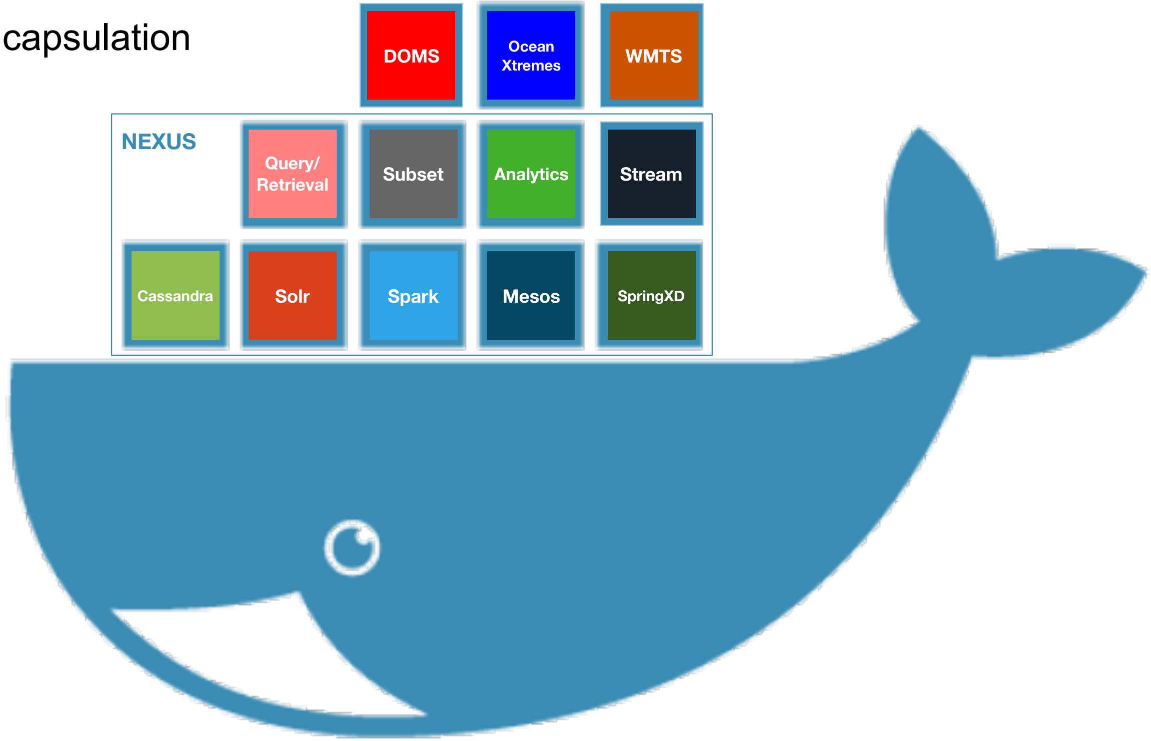


Features

- Fast web log ingestion and processing using Apache Spark, in-memory MapReduce
- Session reconstruction
- Vocabulary semantic relationship extraction
- Machine Learning Search ranking
- Integration with SWEET Ontologies for semantic-driven search and recommendations
- Recommendation

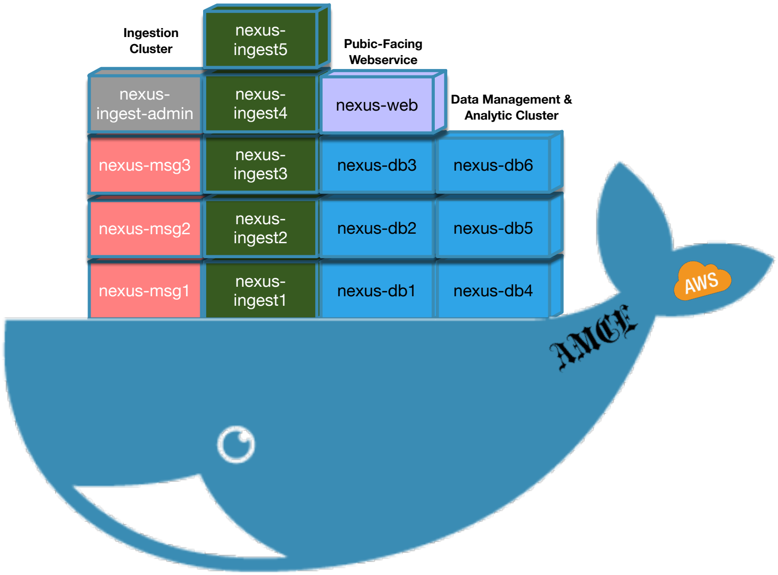
Containerization

- Dockerizing all services
- **Why?**
  - Rapid application deployment
  - Portability across machines
  - Application-centric vs machine/server-centric
  - Version control and component reuse
  - Secure due to isolation and encapsulation
  - Sharing
  - Lightweight footprint
  - Minimal overhead
  - Simplified maintenance

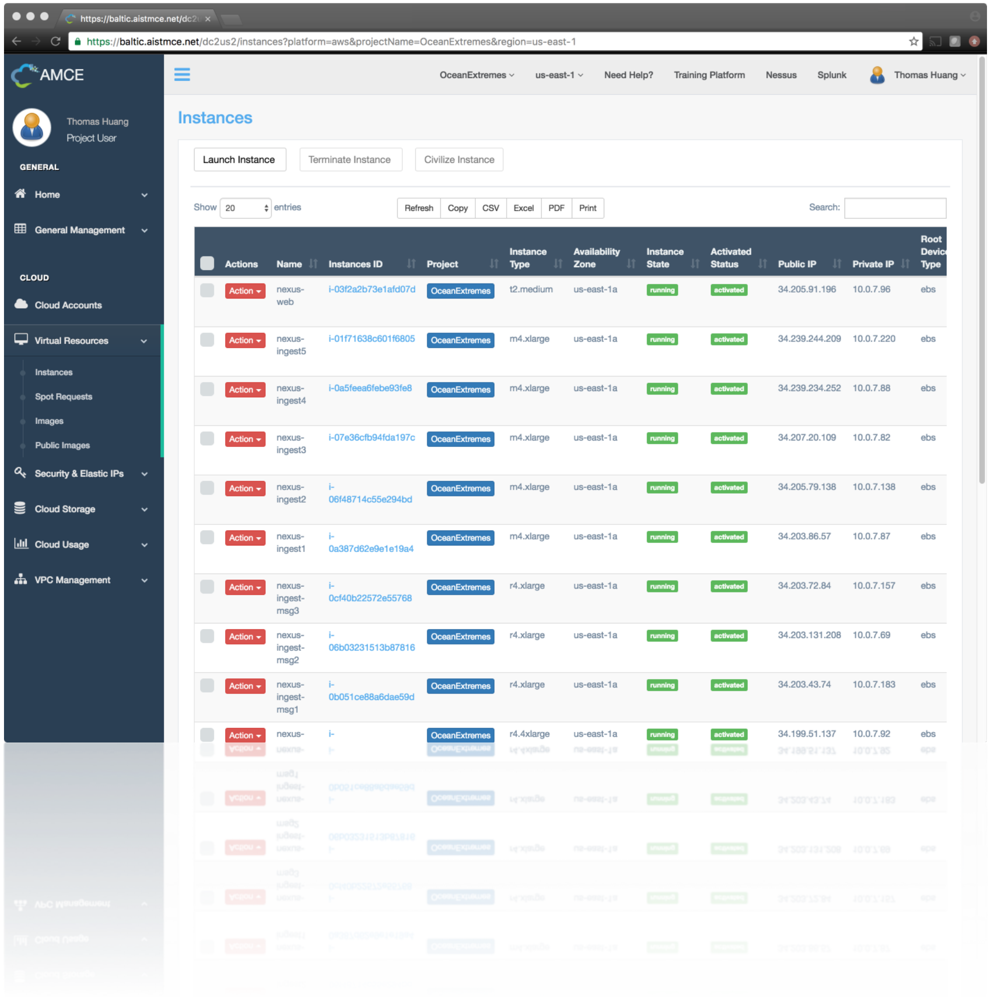


AMCE Onboarding

- OceanWorks will leverage the AIST Managed Cloud Environment (AMCE) for development – the AIST-provisioned Amazon Cloud environment
- OceanXtremes and NEXUS deployment
    - Multi-container, multi-cluster deployment
    - Leverage public DockerHub
    - Working toward dev-test pipeline automation
    - Deployed under 16 Amazon instances. Needs testing



OceanXtremes AMCE Deployment

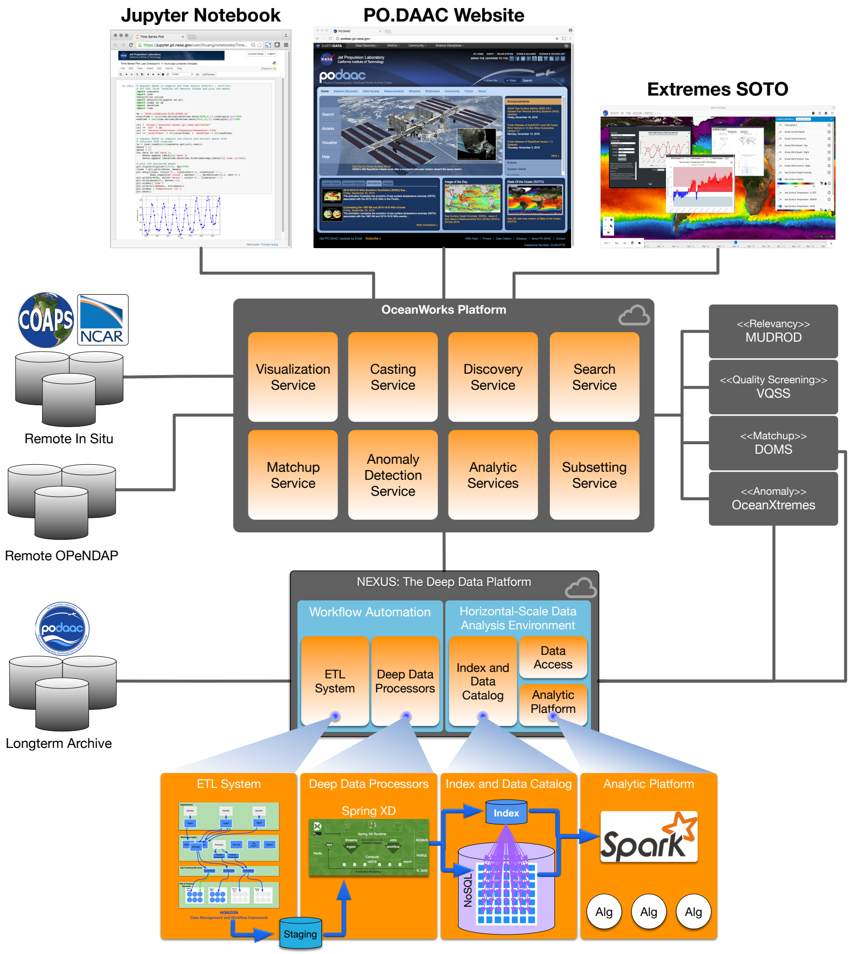


Instances

☐	Actions	Name	Instances ID	Project	Instance Type	Availability Zone	Instance State	Activated Status	Public IP	Private IP	Root Devic Type
☐	Action ▾	nexus-web	i-03f2a2b73e1afd07d	OceanExtremes	t2.medium	us-east-1a	running	activated	34.205.91.196	10.0.7.96	ebs
☐	Action ▾	nexus-ingest5	i-01f71638c601f6805	OceanExtremes	m4.xlarge	us-east-1a	running	activated	34.239.244.209	10.0.7.220	ebs
☐	Action ▾	nexus-ingest4	i-0a5feea6febe93fe8	OceanExtremes	m4.xlarge	us-east-1a	running	activated	34.239.234.252	10.0.7.88	ebs
☐	Action ▾	nexus-ingest3	i-07e36cfb94fda197c	OceanExtremes	m4.xlarge	us-east-1a	running	activated	34.207.20.109	10.0.7.82	ebs
☐	Action ▾	nexus-ingest2	i-06f48714c55e294bd	OceanExtremes	m4.xlarge	us-east-1a	running	activated	34.205.79.138	10.0.7.138	ebs
☐	Action ▾	nexus-ingest1	i-0a387d62e9e1e19a4	OceanExtremes	m4.xlarge	us-east-1a	running	activated	34.203.86.57	10.0.7.87	ebs
☐	Action ▾	nexus-ingest-msg3	i-0cf40b22572e55768	OceanExtremes	r4.xlarge	us-east-1a	running	activated	34.203.72.84	10.0.7.157	ebs
☐	Action ▾	nexus-ingest-msg2	i-06b03231513b87816	OceanExtremes	r4.xlarge	us-east-1a	running	activated	34.203.131.208	10.0.7.69	ebs
☐	Action ▾	nexus-ingest-msg1	i-0b051ce88a6dae59d	OceanExtremes	r4.xlarge	us-east-1a	running	activated	34.203.43.74	10.0.7.183	ebs
☐	Action ▾	nexus-	i-	OceanExtremes	r4.4xlarge	us-east-1a	running	activated	34.199.51.137	10.0.7.92	ebs

The Game Changer in NASA Ocean Science

- Lives on the Cloud
- Auto-Scaling
- One-The-Fly multi-parameter data analysis
- Access and matchup with in-situ measurements
- Smart subsetting
- Anomaly detection and registration
- Sharing of analytic results
- Lightning speed searches
- Discover relevant data, services, news and publications like never before
- Fully Open Source



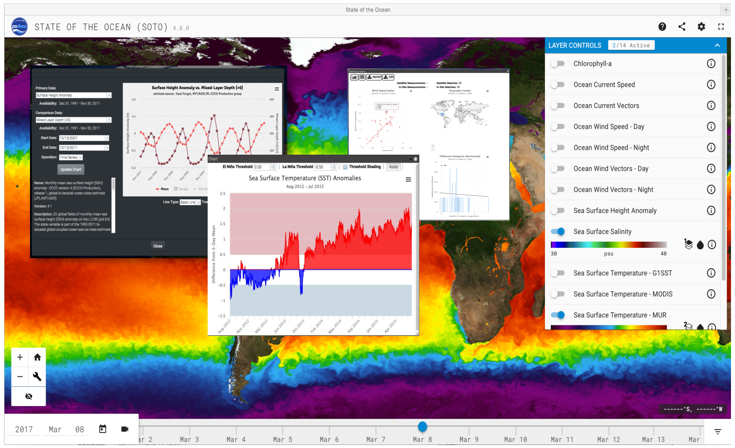
OceanWorks: Ocean Science Data Platform

PI: Thomas Huang, JPL

Objective

Develop an integrated ocean science platform to provide a virtual environment for the PO.DAAC community and applications by leveraging NEXUS, OceanXtremes, DOMS, and MUDRODS to satisfy the following goals:

- Improve data service discovery
- Subset and distribute data
- Identify and catalog ocean phenomenon
- Coordination between satellite and in-situ observations
- Analyze satellite observations
- Visualize and analyze satellite observation on the web
- Enable webservice API integration



OceanWorks State of the Ocean (SOTO) Visualization and analysis

Approach

- Define integrated system architecture and information model
- Prototype integration with PO.DAAC's existing visualization solution, SOTO, add develop web-based analytic capabilities
- Review by PO.DAAC UWG
- Improve data and service search and discovery
- Develop new data subsetting capability
- Demonstrate Jupyter notebook integration
- Facilitate validation by 4 organization and improve prformance
- PO.DAAC User Acceptance Testing environment (UAT) deployment

Co-Is/Partners: E. Armstrong, J. Jacob, N. Quach, V. Tsontos, B. Wilson, JPL; S. Smith, M. Bourassa, FSU; S. Worley, NCAR; C. Yang, Y. Jiang, Y. Li, GMU

Key Milestones

• Complete system design and dataset selection	09/17
• Perform OceanWorks system-level testing	12/17
• PO.DAAC User Working Group (UWG) CDR	04/18
• Complete discovery analysis services and performance optimization	09/18
• Data subsetting capability	09/18
• Integrate Jupyter notebook capability	12/18
• PO.DAAC UWG acceptance review and UAT deployment	04/19

TRL_{in} = 4 $TRL_{current}$ = 4

Earth Science Technology Office

Credits

NASA
Marjorie Cole
Brandi Quam
Mike Little
Chris Lynnes
Kevin Murphy

FSU COAPS
Mark Bourassa
Jocelyn Elya
Shawn Smith
Adam Stallard

NCAR
Steve Worley
Ji Zaihua

GMU
Yongyao Jiang
Yun Li
Chaowei (Phil) Yang

JPL NEXUS Engineers
Stewart (Parker) Abercrombie
Kevin Gill
Frank Greguska
Joseph Jacob
Nga Quach
Brian Wilson

JPL Science Contributors
Ed Armstrong
Andrew Bingham
Carmen Boening
Mike Chin
Michelle Gierach
Ben Holt
Tony Lee
David Moroni
Rob Toaz
Vardis Tsontos
Victor Zlotnicki

Questions, and more information

Thomas.Huang@jpl.nasa.gov