

# Background

## Federal (NOAA)

Eugene Burger  
Ansley Manke

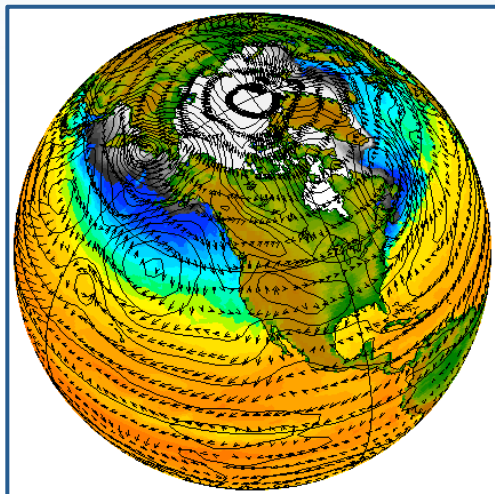
## Contractor

Roland Schweitzer  
Weathertop Consulting, LLC

Steve Hankin (emeritus)

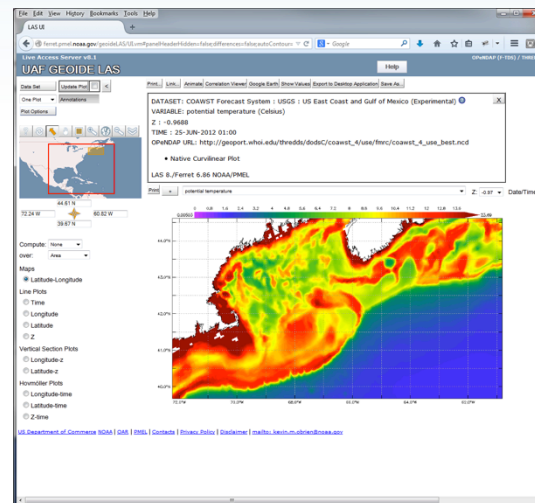
## Cooperative Insitute/JISAO

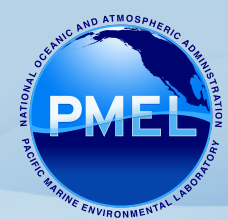
Kevin O'Brien  
Karl Smith



Ferret/PyFerret

Live Access Server





# Data Management



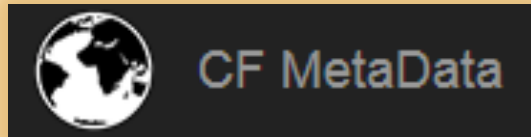
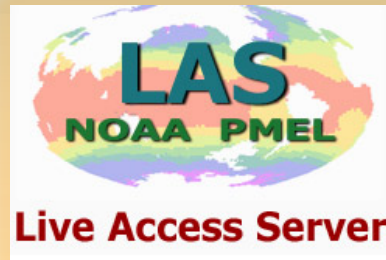
Don't solve problems, copy success

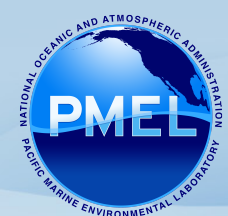
"Switch: How to Change Things When Change Is hard"  
- Chip Heath, 2010



# Data Management

Don't solve problems,  
copy success



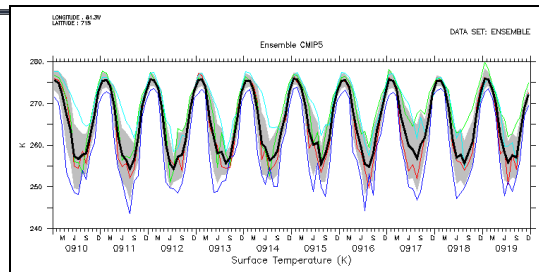
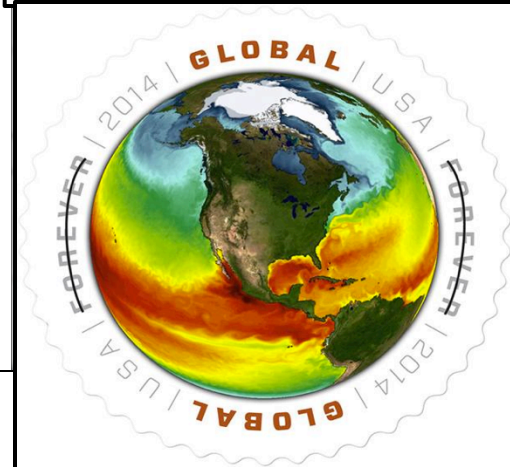
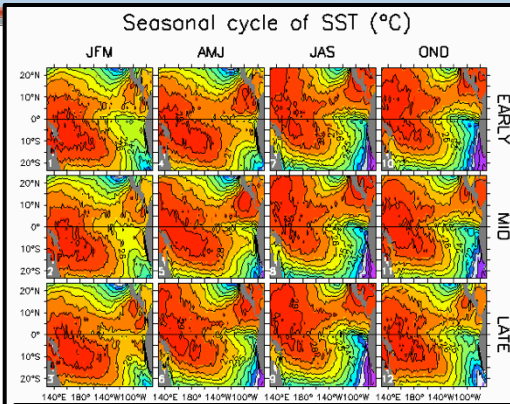
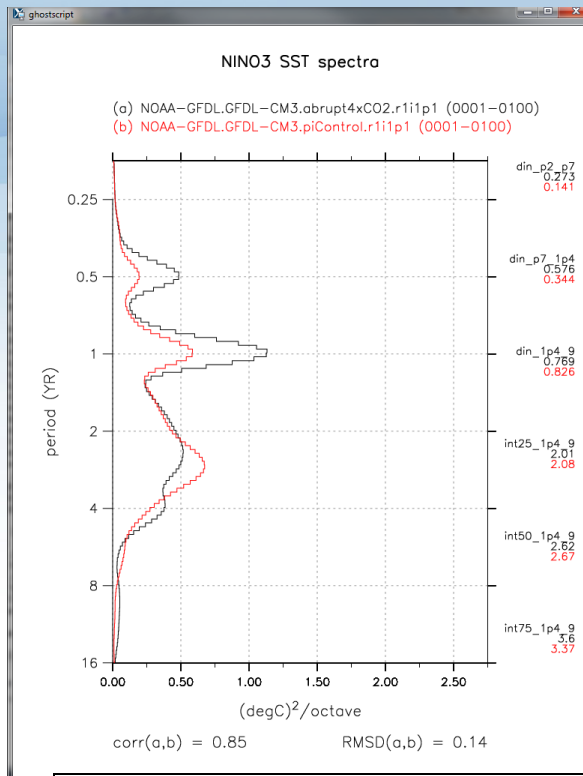


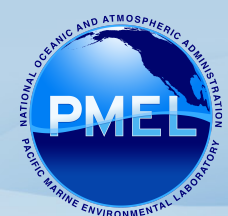
Challenge: Processing and analyzing vast amounts of model data

Integrate into model workflow  
 automated generation of graphic products and data subsets

Ferret can now do computations in 6 dimensions

Ferret graphic used for “Global Forever: Sea Temperatures” US postage stamp



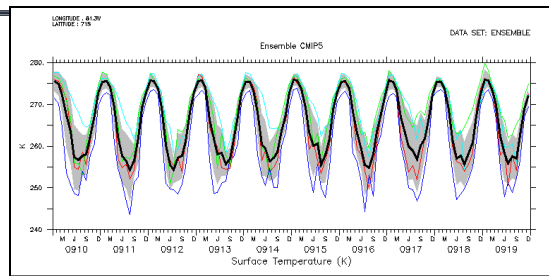
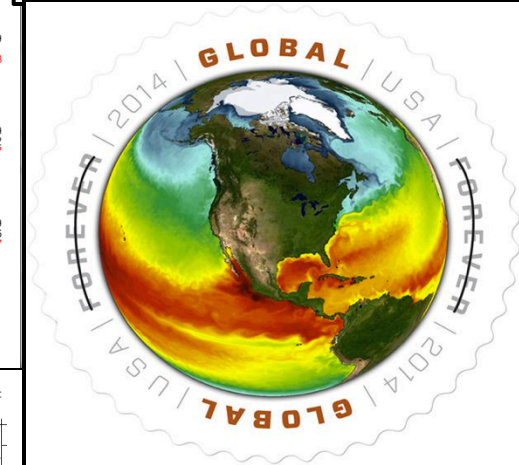
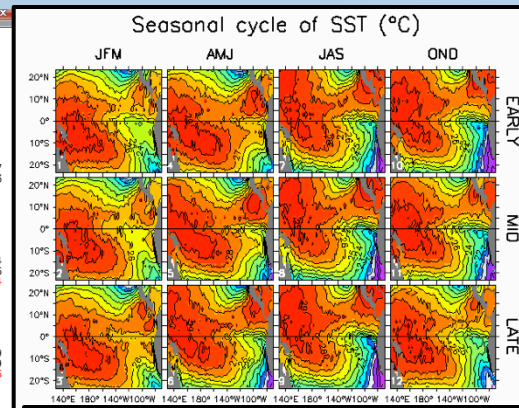
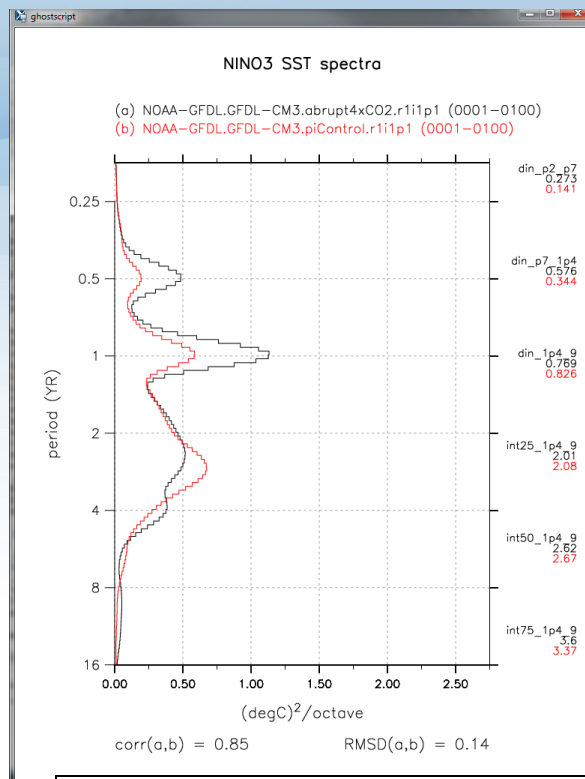


Future directions:

Recently renewed 5 year MOU

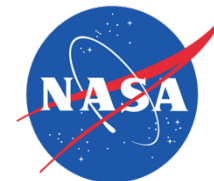
Integrate python version of Ferret into GFDL workflow

Explore and leverage new ways to present data and products



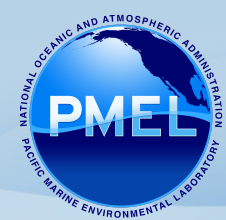


Challenge: Provide access to 5th Coupled Model Intercomparison Project (CMIP 5) data



東京大学 大気海洋研究所  
The University of Tokyo Institute of Oceanography

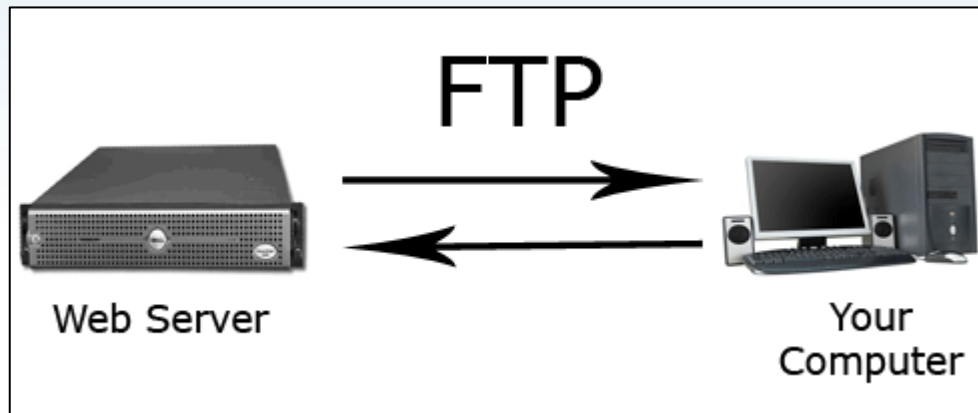


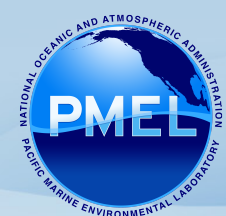


Challenge: Provide access to 5th Coupled Model Intercomparison Project (CMIP 5) data

IPCC AR4 data primarily downloaded (ftp)

This was not scalable for AR 5





# ESGF

Earth System Grid Federation

Challenge: Provide access to 5th Coupled Model Intercomparison Project (CMIP 5) data

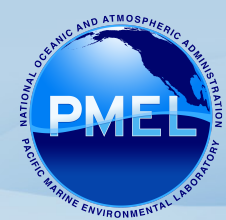
IPCC AR4 data primarily downloaded (ftp)

This was not scalable for AR 5

Live Access Server integrated into ESGF software stack to provide discovery, exploration and download of subsets

The screenshot shows the ESGF LAS web interface. At the top, there's a navigation bar with 'ESGF Portal' and 'LAS UI'. Below that, the 'Live Access Server' section is visible, with 'ESGF LAS' and a 'Help' button. The main content area displays two global maps of Sea Surface Temperature. The left map shows a color scale from blue (cold) to red (warm). The right map shows a similar view but with a different color scale. Below the maps, there are controls for 'Data Set', 'Enlarge Image', and 'Sea Surface Temperature'. A search bar is also present. At the bottom, there are date/time selectors for '1967 Jan' and '1969 Jan'. A red circle highlights the 'Visualize and Analyze' link in the 'Actions' list.





# ESGF

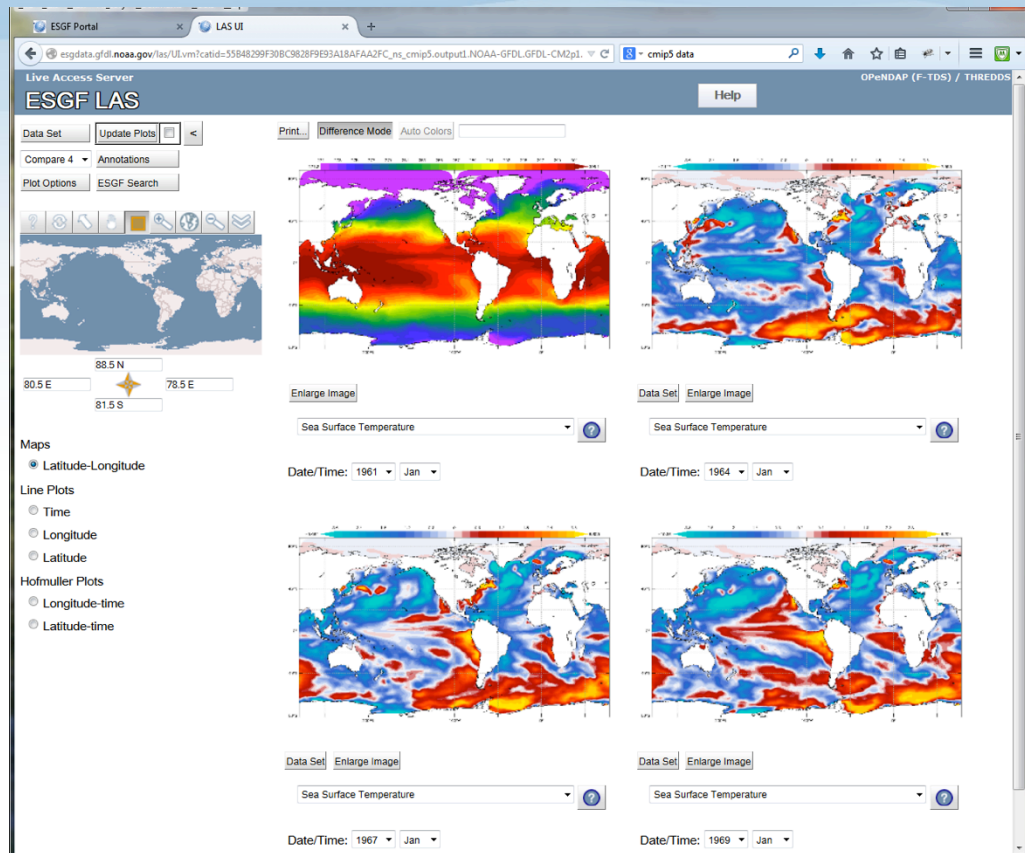
Earth System Grid Federation

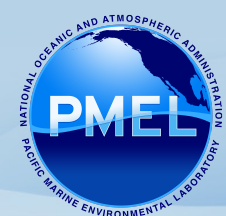


Future directions:

Represent GFDL needs in ESGF project

Move computing closer to data



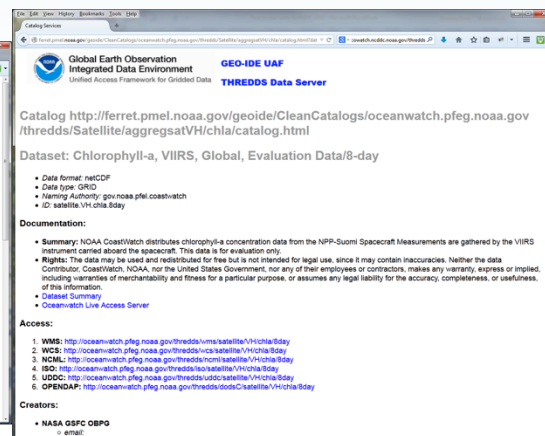
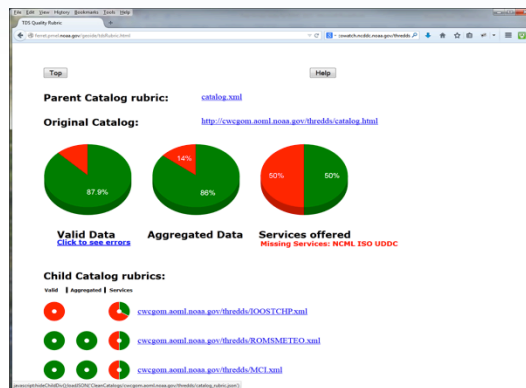
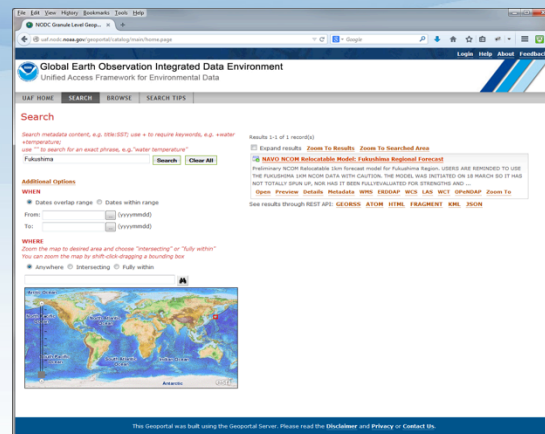
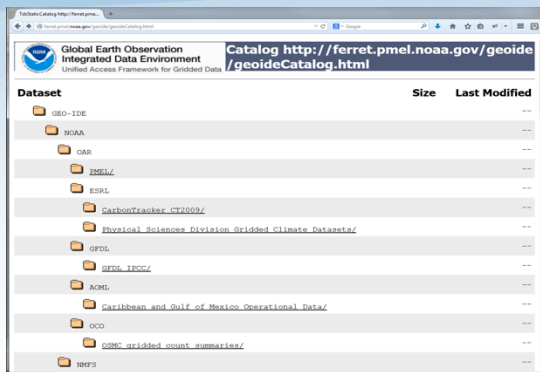


Challenge: Improve accessibility, understanding and discoverability of NOAA data

Created “Master” catalog of data with uniform set of access services

Created catalog “cleaning” tool to assist data providers

Created rubric to measure catalog quality  
 Ingest metadata into data discovery systems



# Global Earth Observation Integrated Data Environment

## Unified Access Framework

Future directions:

Growth of data catalog

Provided “unified” look at observation data

Global Earth Observation Integrated Data Environment  
Unified Access Framework for Gridded Data

Catalog <http://ferret.pmel.noaa.gov/geoid/geoidCatalog.html>

Dataset	Size	Last Modified
GE0-IDE		
NOAA		
OSR		
PMEL/		
ESRL		
CarbonTracker_CT2009/		
Physical Sciences Division Gridded Climate Datasta/		
GRSL		
GFN_TCCO/		
ACML		
Caribbean and Gulf of Mexico Operational Data/		
CCO		
CMC Gridded count summaries/		
INPS		

Global Earth Observation Integrated Data Environment  
Unified Access Framework for Environmental Data

Search

Search metadata content, e.g. 100-000 use + to require keywords, e.g. +water +temperature  
Use "!" to search for an exact phrase, e.g. "water temperature"

Results 1 of 1 words(s)

Expand results Zoom To Results Zoom To Search Area

NOV05 MCOM Subarctic Model: Feb-Mar-Apr Regional Forecast  
The NOV05 MCOM Subarctic Model is intended for planning purposes and is not intended for use in operational decision-making. The model was initialized on 18 March 2005 and has not been updated since. It is not intended for use in operational decision-making.

Additional Options

When: Dates (range) Dates within range From: to: (yy/mm/dd) (yy/mm/dd)

Where: Zoom the map to desired area and choose "intersecting" or "full within" You can zoom the map by shift-click-dragging a bounding box

Anywhere Intersecting Fully within

The Geospatial was built using the Geospatial Server. Please read the [Disclaimer](#) and [Privacy](#) or [Contact Us](#).

Parent Catalog rubric: [catalog.xml](#)

Original Catalog: <http://cvs.gom.noaa.gov/thredds/catalog.html>

Valid Data: 87.9%

Aggregated Data: 14%

Services offered: 50%

Missing Services: RCHL, ISO, UDDC

Child Catalog rubrics:

- [cvs.gom.noaa.gov/thredds/2005CTCHP.xml](http://cvs.gom.noaa.gov/thredds/2005CTCHP.xml)
- [cvs.gom.noaa.gov/thredds/ROMSMMETLO.xml](http://cvs.gom.noaa.gov/thredds/ROMSMMETLO.xml)
- [cvs.gom.noaa.gov/thredds/MCI.xml](http://cvs.gom.noaa.gov/thredds/MCI.xml)

Global Earth Observation Integrated Data Environment  
Unified Access Framework for Gridded Data

GE0-IDE UAF  
THREDDS Data Server

Catalog <http://ferret.pmel.noaa.gov/geoid/CleanCatalogs/oceanwatch.pfeg.noaa.gov/thredds/Satellite/aggregatVH/chla/catalog.html>

Dataset: Chlorophyll-a, VIIRS, Global, Evaluation Data/8-day

- Data format: netCDF
- Data type: GRID
- Staging Authority: gov.noaa.pfeg.oceanwatch
- ID: satellite.VH.chla.8day

Documentation:

- Summary: NOAA CoastWatch distributes chlorophyll-a concentration data from the NPP-Suomi Spacecraft Measurements are gathered by the VIIRS instrument carried aboard the spacecraft. This data is for evaluation only.
- Rights: The data may be used and redistributed for free but is not intended for legal use, since it may contain inaccuracies. Neither the data Contributor, CoastWatch, NOAA, nor the United States Government, nor any of their employees or contractors, makes any warranty, express or implied, including warranties of merchantability and fitness for a particular purpose, or assumes any legal liability for the accuracy, completeness, or usefulness, of this information.
- Global Summary
- OceanWatch Live Access Server

Access:

- WMS: <http://oceanwatch.pfeg.noaa.gov/thredds/lms/satellite/vh/chla/8day>
- WCS: <http://oceanwatch.pfeg.noaa.gov/thredds/lms/satellite/vh/chla/8day>
- NCML: <http://oceanwatch.pfeg.noaa.gov/thredds/lms/satellite/vh/chla/8day>
- ISO: <http://oceanwatch.pfeg.noaa.gov/thredds/lms/satellite/vh/chla/8day>
- UDDC: <http://oceanwatch.pfeg.noaa.gov/thredds/lms/satellite/vh/chla/8day>
- CHREDAF: <http://oceanwatch.pfeg.noaa.gov/thredds/lms/satellite/vh/chla/8day>

Creators:

- NASA GSFC OBPG
- emil

# Observing System Monitoring Center



Challenge: Provide interoperable access to near real-time ocean observations

Last 30 days of real time data available via ERDDAP

Profile data currently being used to force ROMS assimilation models

Access also available to a multitude of other clients

People can use the tool they prefer!

continuous satellite measurements of sea surface temperature, height, winds, ocean color, and sea ice

lat	stmp
17.4	13.2
18.3	16.3
19.3	19.2
19	12.8
17.3	17.1
16.7	14.9
-1.5	-103.5
29.5	NaN
NaN	10.7
4.9	NaN
NaN	NaN
11.7	NaN
28.7	NaN
20.2	NaN
29.5	NaN
23.8	NaN
23.2	NaN
29	NaN
NaN	NaN
28.3	NaN
27.9	NaN

```

IPython Notebook ERDDAP_GTS Last Checkpoint
File Edit View Insert Cell Kernel Help
In [4]: df = pd.read_csv(URL, index_col="time", parse_dates=True, skiprows=[1])
In [5]: df.describe()
Out[5]:
   count  latitude  longitude  stmp  stmp
count  12241  0.000000  12241.000000  5378.000000  2129.000000
mean    14.178235   -19.150626    22.332819    23.120030
std     35.825186   117.842684    8.077134    8.506198
min     -74.377000   -179.900000   -50.000000   -110.200000
25%    -24.220000   -63.870000   16.100000   18.400000
50%    29.642000   -76.007000   24.450000   23.600000
75%    41.890000   121.887200   28.800000   29.300000
max     89.416000   179.987000   58.500000   36.900000

8 rows x 4 columns

In [33]: scatter(df.longitude, df.latitude, c=df.stmp)
Out[33]: <matplotlib.collections.PathCollection at 0x2e9f66d0>

```

Suppressing ship observations for most recent 48 hours



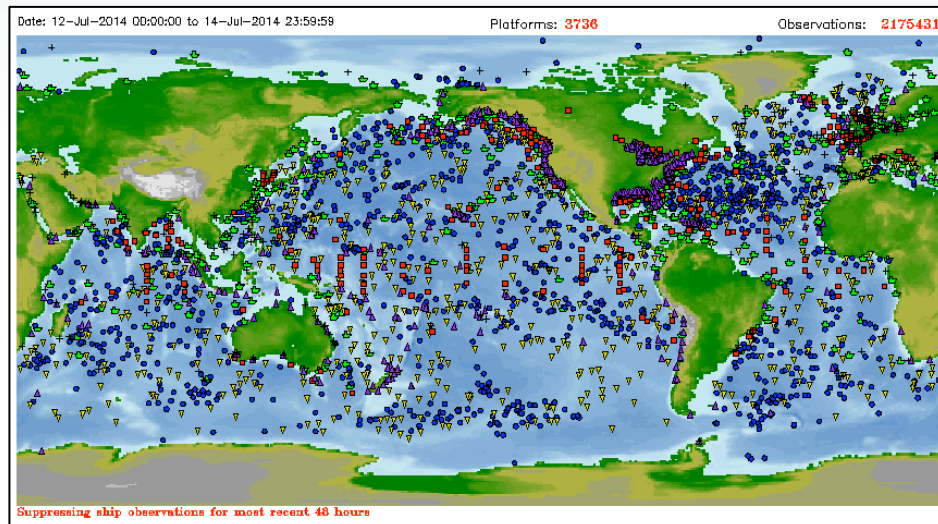
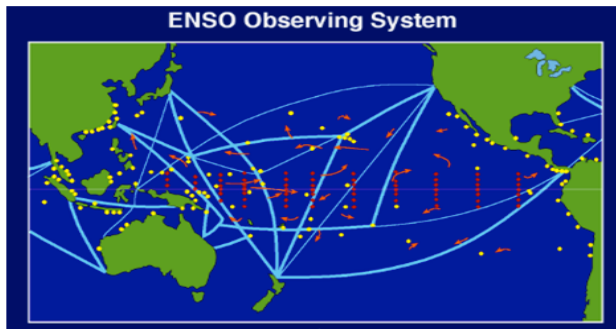
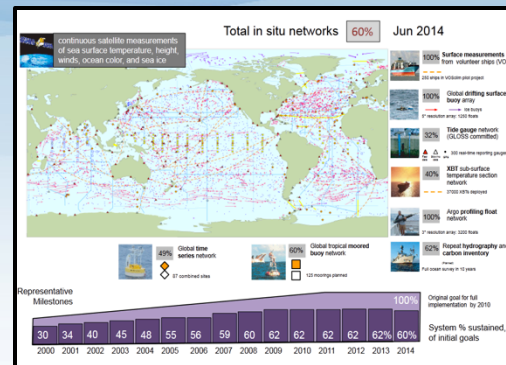
# Observing System Monitoring Center

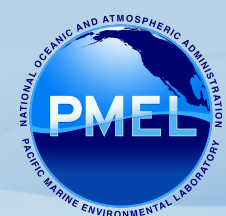


Future directions:

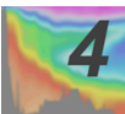
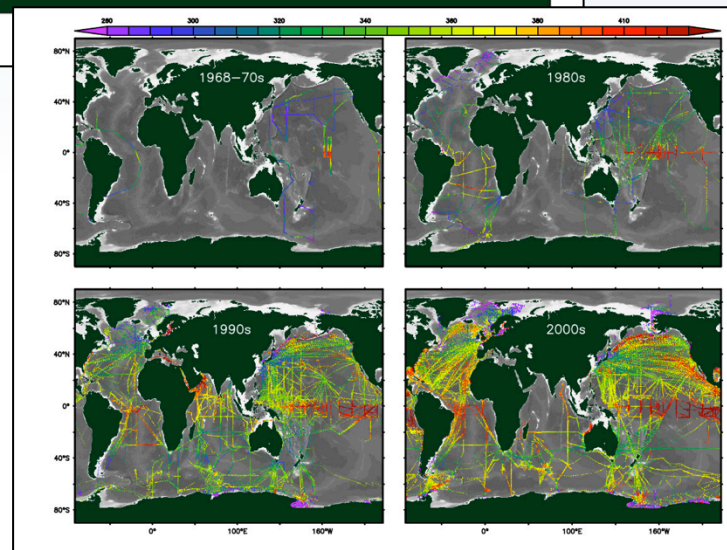
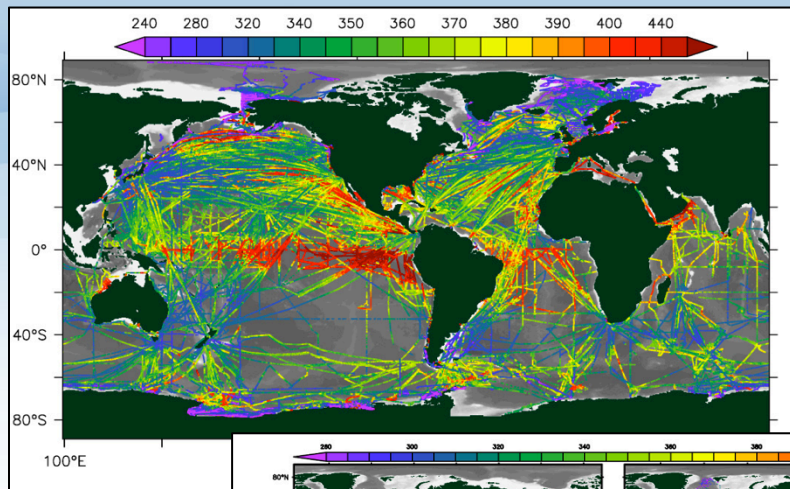
Integrate delayed mode data with real time data

Integrated ocean data prototype for potential use by TPOS





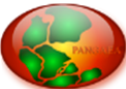
Challenge: How to create a quality controlled global surface ocean fCO<sub>2</sub> data set



IFM-GEOMAR



UNIVERSITY OF EXETER

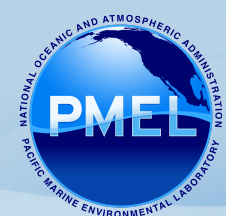


GEOCARBON



UK Ocean Acidification Research Programme



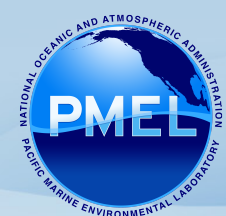


Challenge: How to create a quality controlled global surface ocean fCO2 data set

Live Access Server connected to database used for v1 and v2 QC

Users could do comparisons of data and drill down for more info

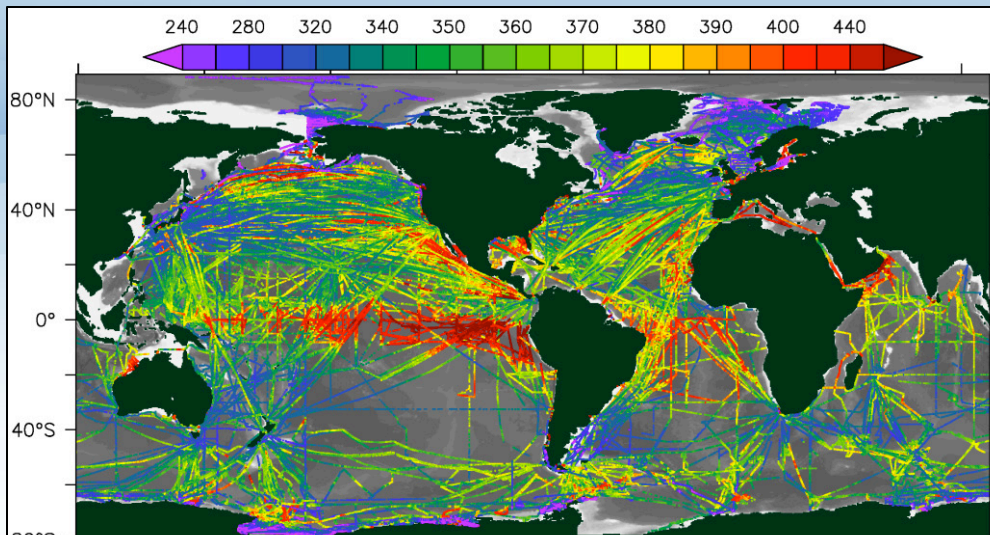
From a correlation viewer data outliers could be selected and WOCE flags applied BUT ONLY BY APPROVED USERS!



Future directions:

QC for version 3 is currently underway (~14 million obs)

Automation of data submission in development for version 4



SOCAT Upload Dashboard

My SOCAT Version 3.0 Datasets

Upload Datasets	Expcode	Upload Date	Data Status	Metadata	QC Status	Archival	Filename	Supplemental Documents	Owner
Identity	11XX00111210	2014-06-20 09:04	No warnings	(no metadata)	Private	Not specified	11XX00111210.tsv	(no documents)	a.marke
Columns	00XX00111210	2014-06-20 09:04	No warnings	(no metadata)	Private	Not specified	00XX00111210.tsv	(no documents)	a.marke
Edr	22XX00111210	2014-06-20 09:04	No warnings	(no metadata)	Private	Not specified	22XX00111210.tsv	(no documents)	a.marke
Metadata	00BB00111210	2014-06-20 08:44	No warnings	2014-06-20 08:49	Submitted	With next SOCAT	00XX00111210.tsv	00BB00111210_Cruise_report.doc (2014-06-20 08:50)	a.marke

Supplemental Documents

Submit for QC

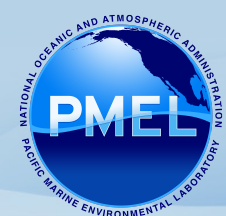
Preview Dataset

Delete Datasets





Isn't this the  
*PMEL* lab  
review?

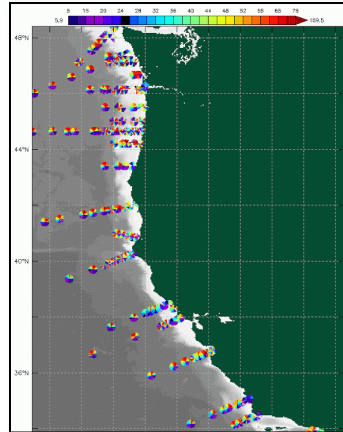
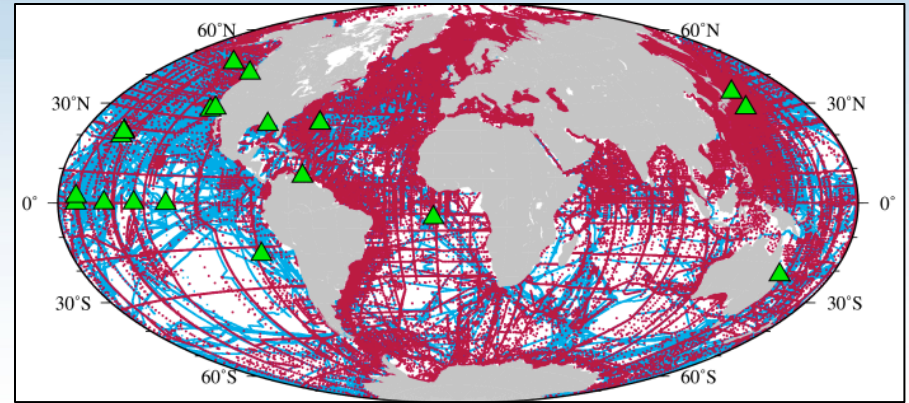


# PMEL Ocean Acidification

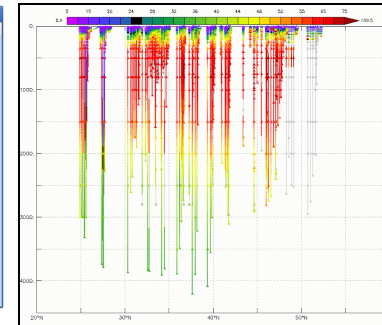
Utilize data submission dashboard

Utilize Live Access Server  
Visualize and QC system

Assist in submission to NODC and Ocean  
Acidification Scientific Data Stewardship  
(OADS)



Label	Expedise	Upload Date	Data Status	Metadata	QC Status	Archival	Filename	Supplemental	Owner
1100011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	1100011028.nc	2014-03-01	awanda
1100011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	1100011028.nc	2014-03-01	awanda
2000011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2000011028.nc	2014-03-01	awanda
2000011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2000011028.nc	2014-03-01	awanda
2000011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2000011028.nc	2014-03-01	awanda
2000011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2000011028.nc	2014-03-01	awanda
2000011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2000011028.nc	2014-03-01	awanda
2000011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2000011028.nc	2014-03-01	awanda
2000011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2000011028.nc	2014-03-01	awanda
2000011028	2014-03-01-04	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2014-03-01	2000011028.nc	2014-03-01	awanda





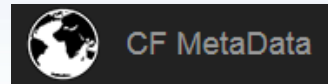
# In summary.....

Data management requirements cannot be ignored

We face limited budgets, high data volumes, “cultural” barriers



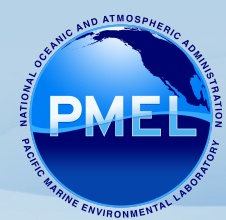
*Don't solve problems,  
copy success*



We can create data management frameworks for entire data lifecycle

Do it in a way that

- minimizes effort required by researchers
- Greatly improves access for users



# In summary.....

## Thank You!

Ferret: [ferret.pmel.noaa.gov](http://ferret.pmel.noaa.gov)

LAS: [ferret.pmel.noaa.gov/LAS/](http://ferret.pmel.noaa.gov/LAS/)

NetCDF: [www.unidata.ucar.edu/software/netcdf/](http://www.unidata.ucar.edu/software/netcdf/)

THREDDS: [www.unidata.ucar.edu/software/thredds/](http://www.unidata.ucar.edu/software/thredds/)

OPeNDAP: [opendap.org](http://opendap.org)

CF: [cfconventions.org](http://cfconventions.org)

UAF: [geo-ide.noaa.gov](http://geo-ide.noaa.gov)

OSMC: [www.osmc.noaa.gov](http://www.osmc.noaa.gov)

SOCAT: [socat.info](http://socat.info)

ESGF: [esgf.org](http://esgf.org)

GFDL: [www.gfdl.noaa.gov](http://www.gfdl.noaa.gov)

OADS: [www.nodc.noaa.gov/oceanacidification/](http://www.nodc.noaa.gov/oceanacidification/)