



Don't solve problems, copy success

Leveraging standards and conventions to
improve interoperability

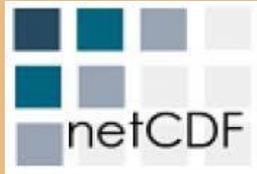
Kevin O'Brien
Science Data Integration Group
University of Washington/JISAO
NOAA/Pacific Marine Environmental Laboratory



Don't solve problems, copy success

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

Don't solve problems,
copy success



- Projects: (too many to name)

Data
formats:

netCDF

GRIB

HDF

CSV

...

Service
Stack:

NetCDF – OPeNDAP-CF-THREDDS-WMS

Applications:

Matlab

ArcGIS

Ferret

GrADS

Google Earth

IDV

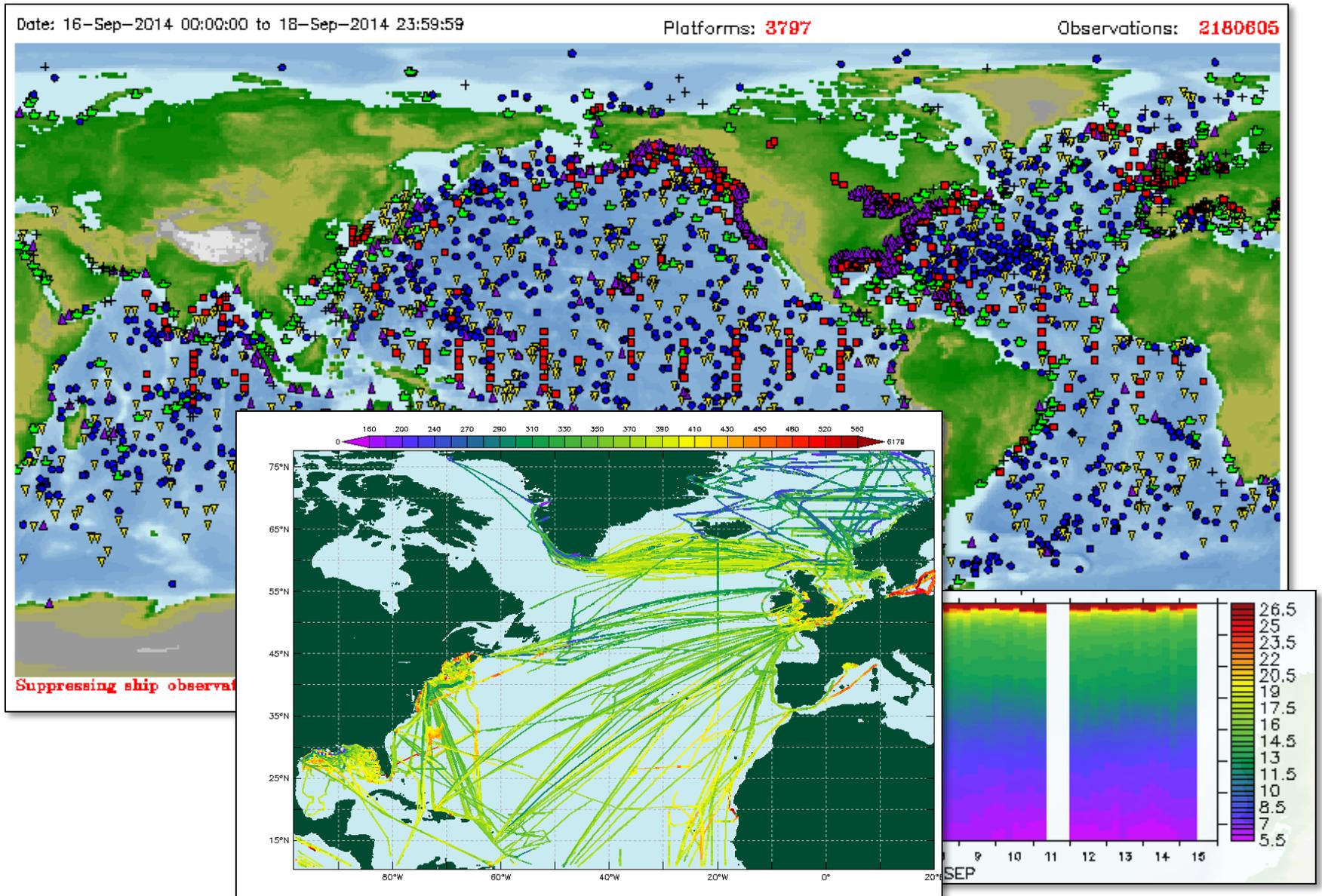
LAS

ERDDAP

...

Users: (too many to name)

An example: The real time data stream



An example: The real time data stream

How can users access real time observations?

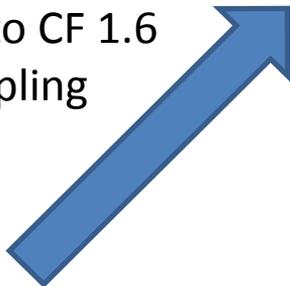
- Data has maximum societal value
 - Forecast models could utilize to improve their forecasts
- QC'd data can take months to become available
- Can be difficult to obtain (GTS)

An example: The real time data stream

Solution:

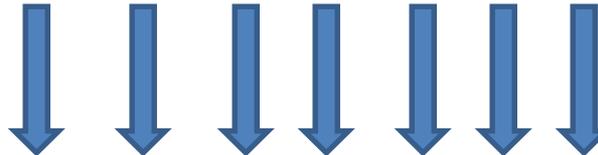
Leverage netCDF, DAP and CF conventions

Create Oracle Views
That map onto CF 1.6
Discrete Sampling
Geometries



Grid DAP	Sub-set	Table	Make	W	A	H	Title	Sum-mary	FGDC	ISO	Metadata	Back-ground	Info	RSS	E-mail	Institution	Dataset ID
set	data	graph					OSMC Argo Profile data		F	J	M	background				OSMC	OSMC_PROFILERS
set	data	graph					OSMC flattened observations from GTS		F	J	M	background				OSMC	OSMC_flattened
set	data	graph					OSMC normalized observations from GTS		F	J	M	background				OSMC	OSMC_Points
set	data	graph					OSMC Profiles		F	J	M	background				OSMC	OSMCV4_DUO_PROFILES
set	data	graph					OSMC surface trajectory data		F	J	M	background				OSMC	OSMCV4_DUO_SURFACE_TRAJECTORY
set	data	graph					OSMC TimeSeries data		F	J	M	background				OSMC	OSMCV4_DUO_TIME_SERIES

ERDDAP Server



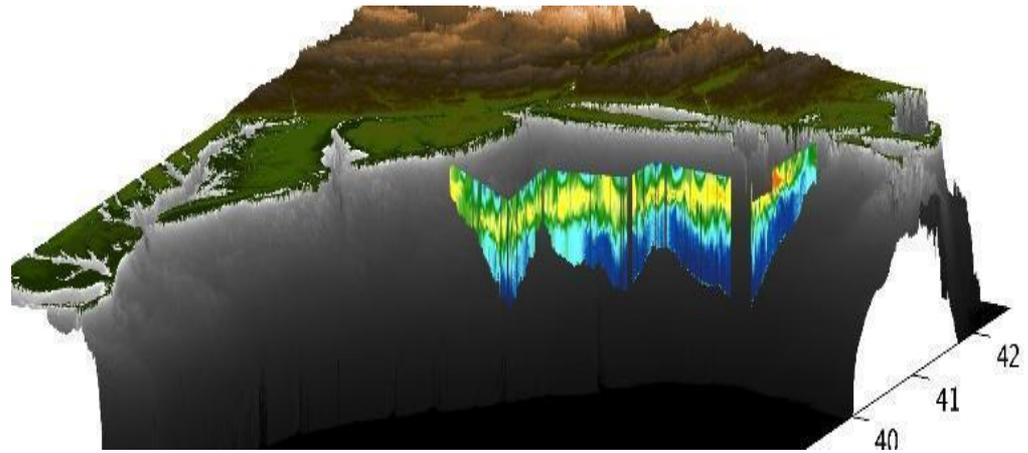
RESTful URL's
then connect a
multitude of
clients to data



An example: The real time data stream

Work flow for real-time ESPreSSO ROMS 4DVar

John Wilkin
Rutgers Ocean Modeling Group
MARACOOS



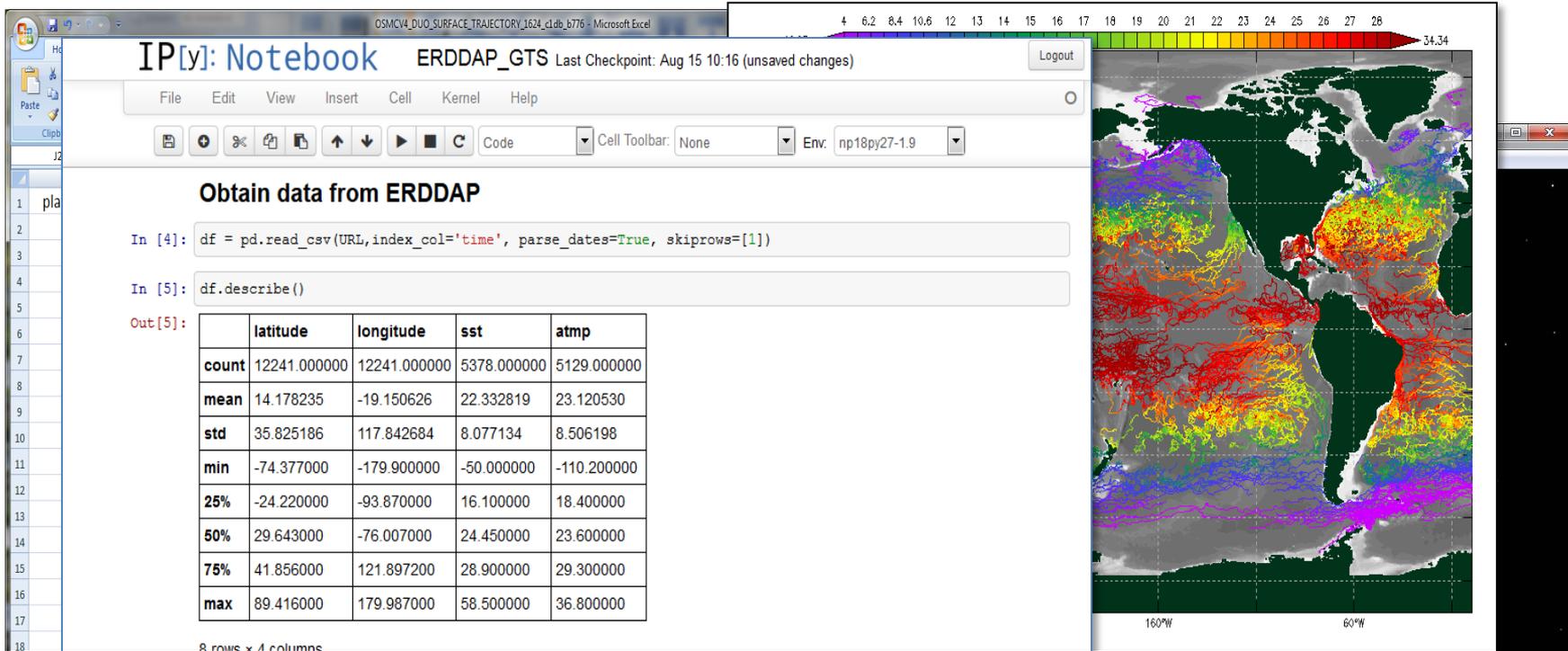
- **SOOP XBT/CTD, Argo floats, NDBC buoys on GTS** **[OSMC ERDDAP]**

http://osmc.noaa.gov/erddap/tabledap/OSMC_flattened.nc?

time,longitude,latitude,ztmp

&time>=now-3hours&time<=now

An example: The real time data stream



The screenshot shows an IPython Notebook window titled "ERDDAP_GTS" with a "Logout" button. The notebook interface includes a menu bar (File, Edit, View, Insert, Cell, Kernel, Help) and a toolbar with icons for saving, undo, redo, and running code. The notebook content is as follows:

Obtain data from ERDDAP

```
In [4]: df = pd.read_csv(URL, index_col='time', parse_dates=True, skiprows=[1])
```

```
In [5]: df.describe()
```

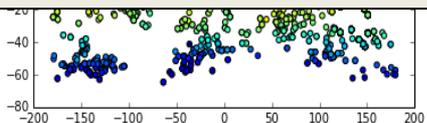
Out[5]:

	latitude	longitude	sst	atmp
count	12241.000000	12241.000000	5378.000000	5129.000000
mean	14.178235	-19.150626	22.332819	23.120530
std	35.825186	117.842684	8.077134	8.506198
min	-74.377000	-179.900000	-50.000000	-110.200000
25%	-24.220000	-93.870000	16.100000	18.400000
50%	29.643000	-76.007000	24.450000	23.600000
75%	41.856000	121.897200	28.900000	29.300000
max	89.416000	179.987000	58.500000	36.800000

8 rows x 4 columns

On the right side of the notebook, there is a global map visualization showing a data stream. A color scale at the top right of the map ranges from 4 to 34.34. The map shows a dense network of colored lines (red, orange, yellow, green, blue, purple) overlaid on a world map, representing the real-time data stream. The map includes latitude and longitude markers (160°W, 60°W).

[http://osmc.noaa.gov/erddap/tabledap/OSMC_flattened?
time,longitude,latitude,ztmp
&time>=now-3hours&time<=now](http://osmc.noaa.gov/erddap/tabledap/OSMC_flattened?time,longitude,latitude,ztmp&time>=now-3hours&time<=now)



Experimenting with GTS data from OSMC ERDDAP

Exploring use of Python to formulate ERDDAP REST requests and process the responses.

Initialize

```
In [ ]: import urllib2
import pandas as pd
import matplotlib
```

```
In [ ]: baseURL = 'http://osmc.noaa.gov/erddap/tabledap/OSMC_flattened.csv?'
params = 'time,latitude,longitude,sst,atmp'
start='now-3hours'
stop='now'
timeRange = '&time>=' + start + '&time<=' + stop
URL = baseURL + params + timeRange
```

```
In [ ]: print URL
```

Obtain data from ERDDAP

In Summary.....

Leveraging widely-used standards improves interoperability on a budget

Multiple/Overlapping standards pose compliance challenges

Uniting netCDF/GRIB would greatly ease efforts of developers and users



Thank you!

Kevin O'Brien

Kevin.M.O'Brien@noaa.gov