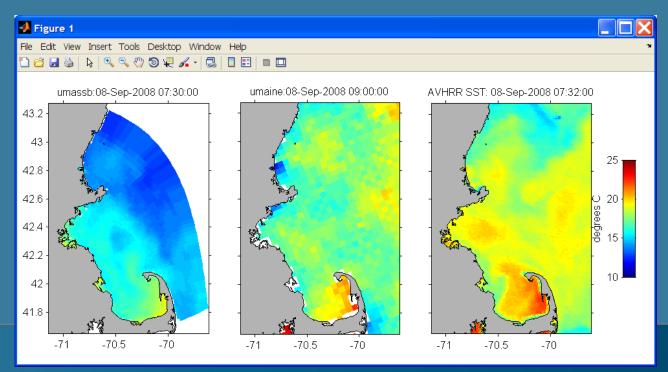
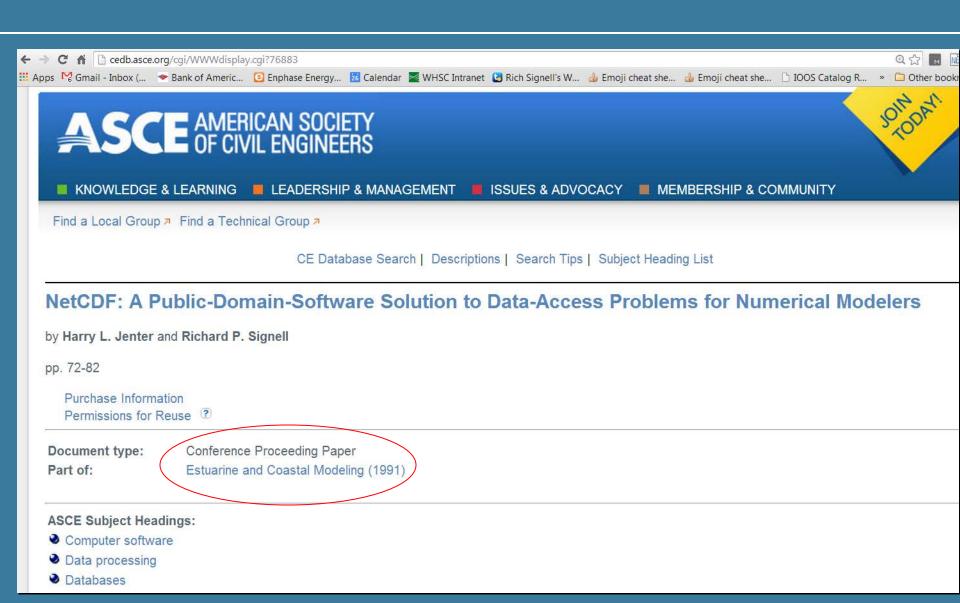
Ocean, Atmosphere & Climate Model Assessment for Everyone

Rich Signell
USGS Woods Hole, MA
Unidata 2014 DeSouza Award Presentation
Boulder, CO: Sep 15, 2014





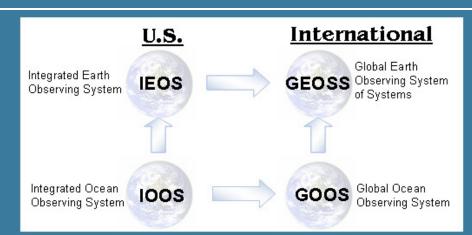


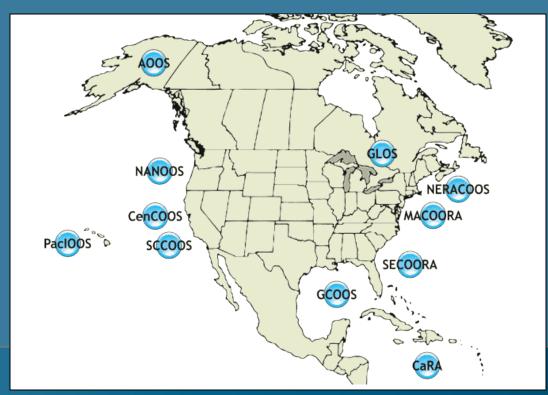


US Integrated Ocean Observing System (IOOS®)

IOOS® Plan defines:

- Global Component
- Coastal Component
 - 17 Federal Agencies
 - 11 Regional Associations







How well do models simulate events?



Hurricane Sandy, Ocean Grove Pier -New Jersey, October 29, 2012 -Photograph by Bob Bowné

Deepwater Horizon Oil Spill, Northern Gulf of Mexico, May 24, 2010. Image from MODIS on NASA's Terra Satellite.



100S Core Principles

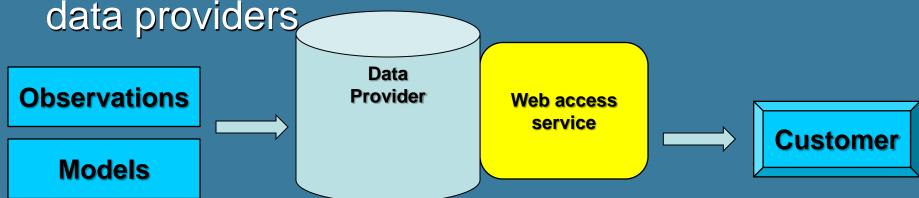
Adopt open standards & practices





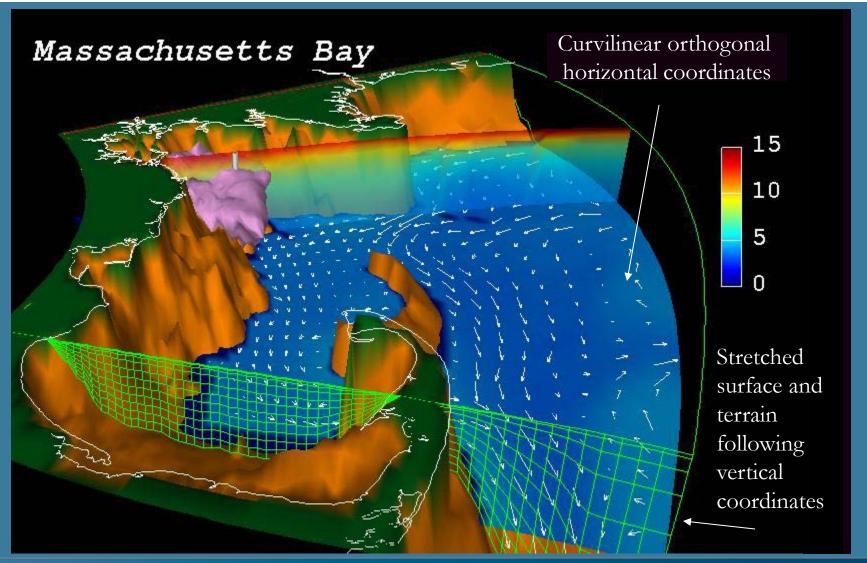


- Avoid customer-specific stovepipes
- Standardized access services implemented at data providers



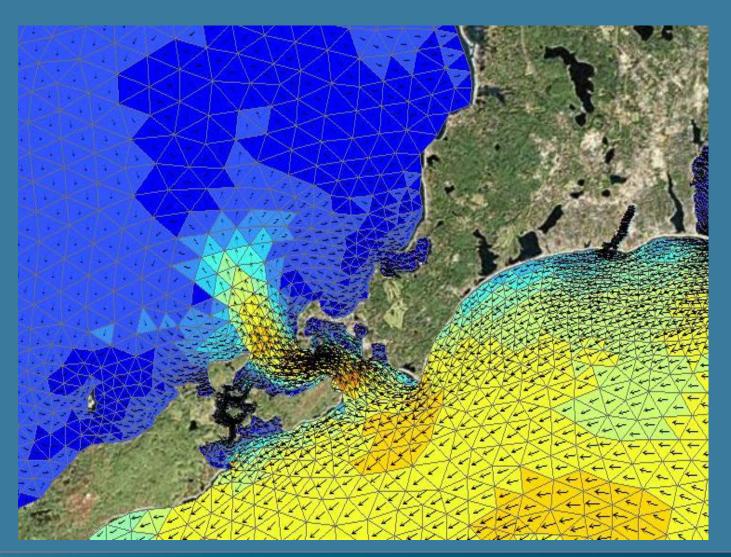


Issue: Ocean grids are not regularly spaced!





Unstructured (e.g. triangular) grid





Time Series, Trajectories



Meteorology and Wave Buoy in the Gulf of Maine. Image courtesy of NOAA.

Ocean Glider. Photo by Dave Fratantoni, Woods Hole Oceanographic Institution



NetCDF Climate and Forecast (CF) Conventions provide a solution

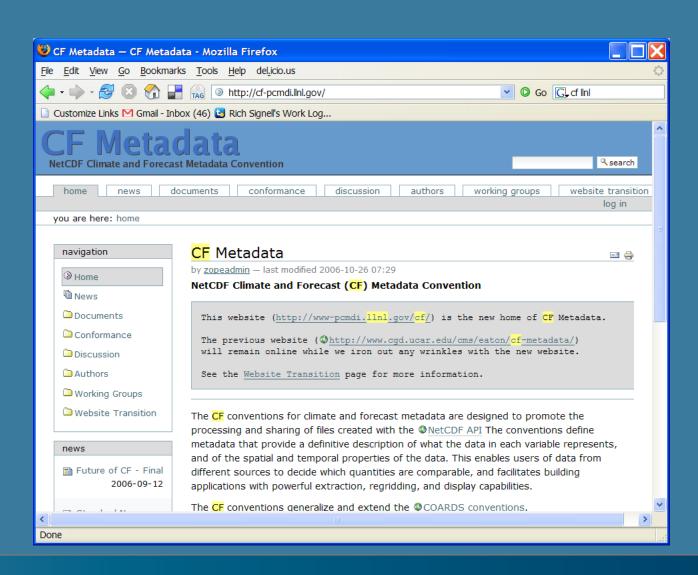
Groups using CF:

GO-ESSP: Global Organization for Earth System Science Portal

IOOS: Integrated Ocean Observing System

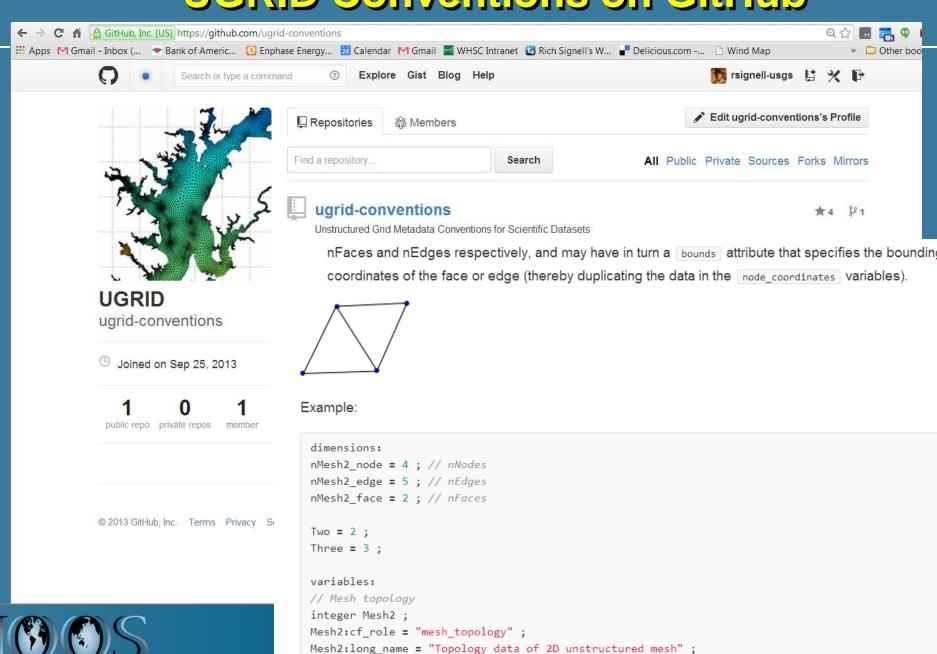
ESMF: Earth System Modeling Framework

OGC: Open Geospatial Consortium (GALEON: WCS profile)





UGRID Conventions on GitHub



Mesh2:topology dimension = 2;

Example NcML (on StackOverflow)

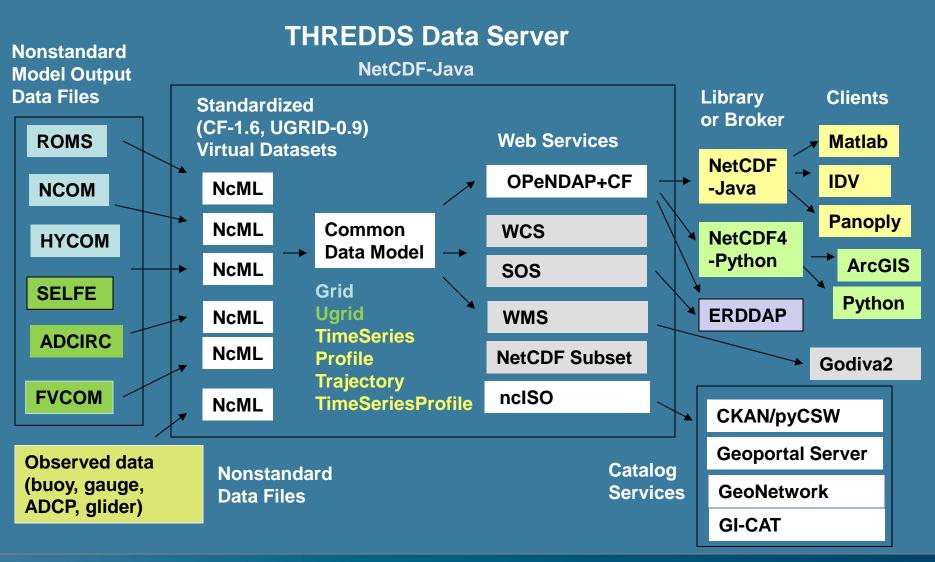


Here's an example:

```
<variable name="v" shape="time2 sigma node" type="float">
 <attribute name="standard name" value="barotropic northward sea water velocity"/>
 <attribute name="mesh" value="selfe mesh"/>
 <attribute name="location" value="node"/>
 <attribute name="coordinates" value="v x"/>
 <attribute name="units" value="m/s"/>
</variable>
<aggregation type="union">
 <netcdf xmlns="http://www.unidata.ucar.edu/namespaces/netcdf/ncml-2.2">
  <aggregation dimName="time" type="joinExisting">
   <scan
    location="/data/ftp/upload/Inundation/vims/selfe tropical/runs/Rita/3D varied row
    regExp=".*[0-9]{1}_WaterLevel\.nc$"/>
  </aggregation>
 </netcdf>
 <netcdf xmlns="http://www.unidata.ucar.edu/namespaces/netcdf/ncml-2.2">
  <aggregation dimName="time" type="joinExisting">
   <scan
    location="/data/ftp/upload/Inundation/vims/selfe_tropical/runs/Rita/3D_varied_ro
    regExp=".*[0-9]{1} SigWaveHeight\.nc$"/>
  </aggregation>
 </netcdf>
```

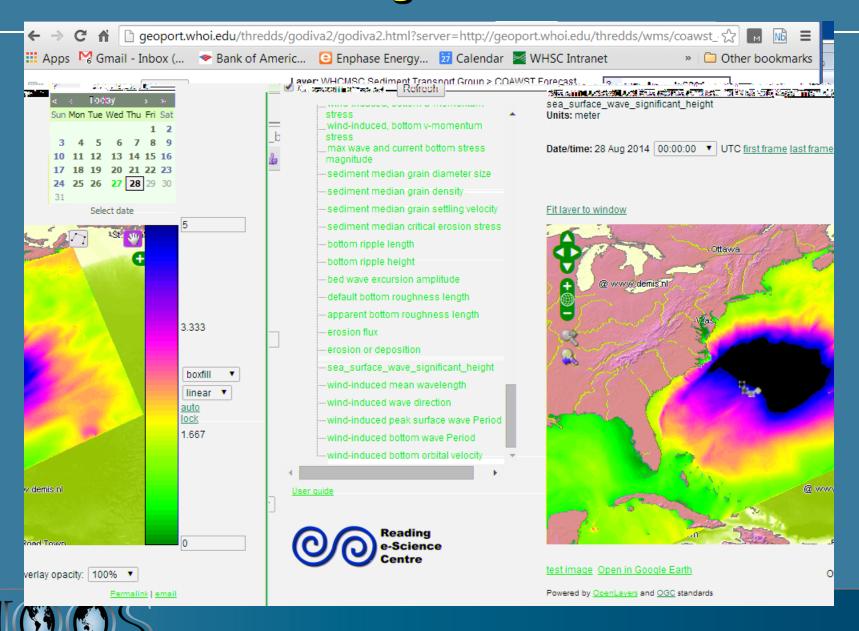


100S Model Data Interoperability Design





WMS Browsing with THREDDS/ncWMS



Matlab NCTOOLBOX https://github.com/nctoolbox

- Objective: Make it simple to access CF data
- Example function:
- [t, geo]=nj_tslice(URL,'temp',1);
- t = 22x120x180 single
- geo =
 - lat: [120x180 single]
 - lon: [120x180 single]
 - z: [22x120x180 double]
 - time: 733582 (matlab datenum)
- nj_tslice works identically for ROMS, POM, ECOM, WRF, Wavewatch3
- URL can be: local NetCDF, remote NetCDF, NcML, OpenDAP Data URL



Matlab NCTOOLBOX https://github.com/nctoolbox

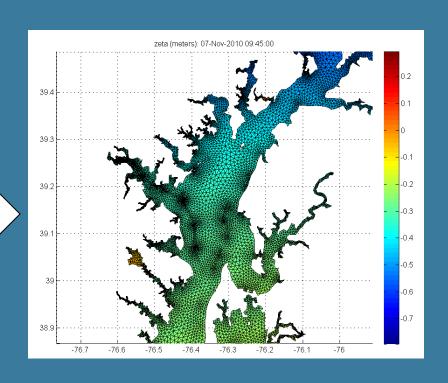
nc = ncugrid(dap_url) ncvar = nc.uvariable('zeta') z = ncvar.data(itime,:) grid = ncvar.grid(itime,:)

z = 26441x1 single grid =

lat: [26441x1 single] lon: [26441x1 single]

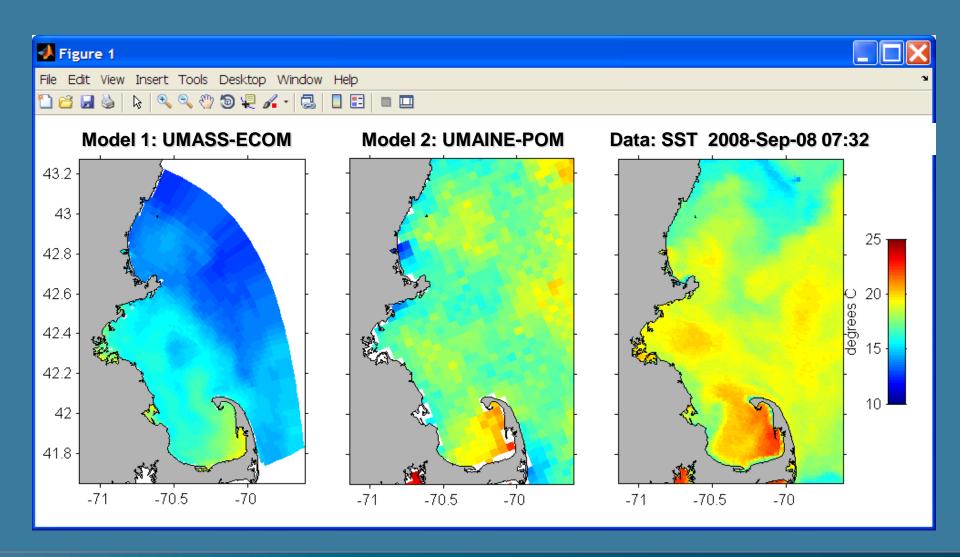
time: 730970 (matlab datenum)

connectivity: [52025x3 int32]





Comparing Models with Data in Matlab

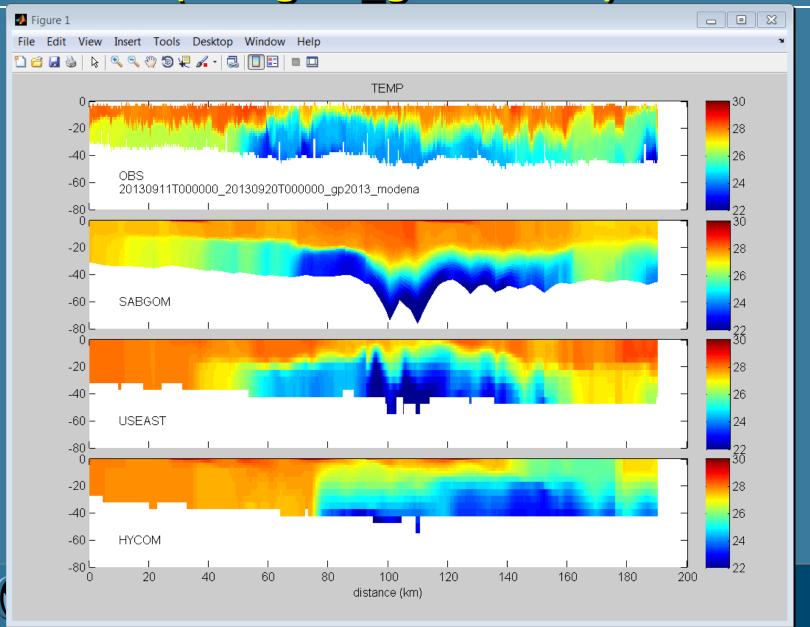




Skidaway modena glider (Sep 11-23)

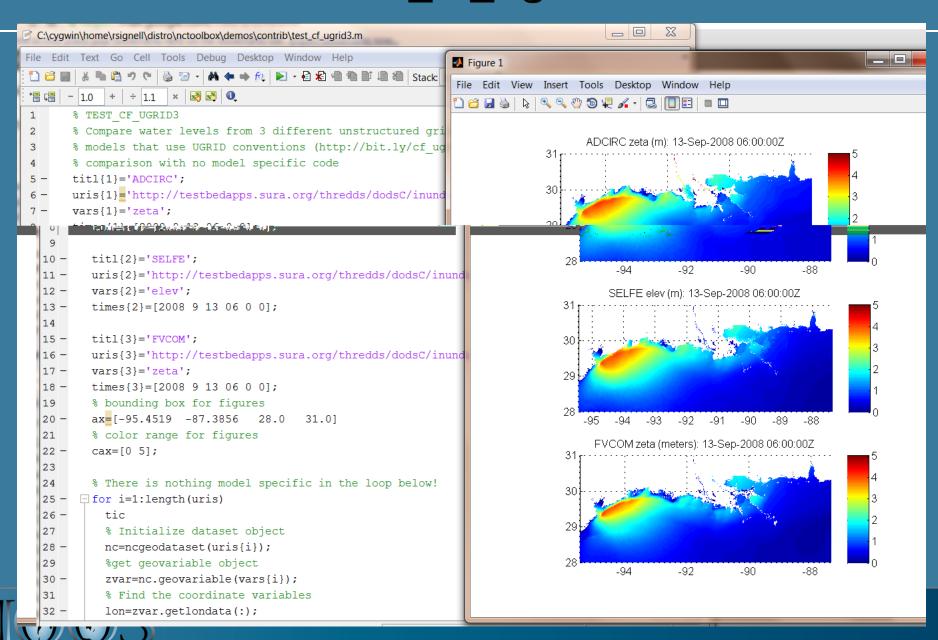


compare_secoora_model_sections.m (using nc_genslice.m)

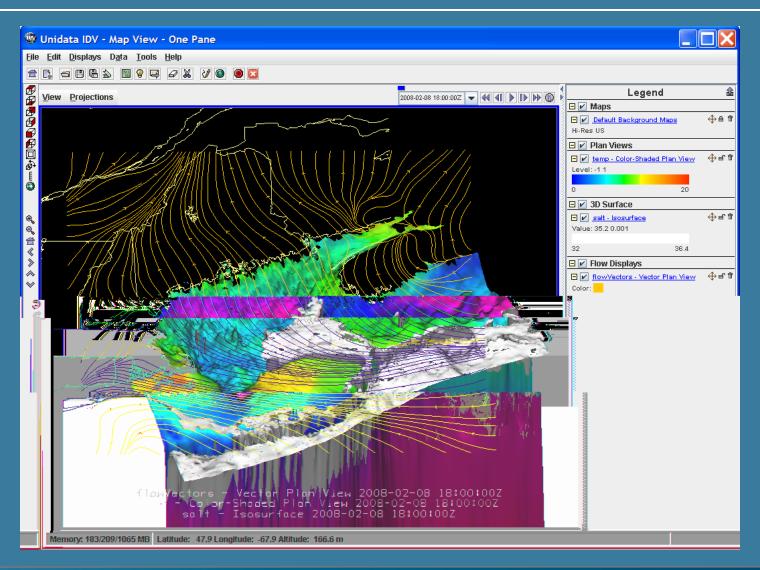




test_cf_ugrid3.m



3D visualization of data with IDV





NetCDF Subset Service to CSV



WHCMSC Sediment Transport Group

THREDDS Data Server

Catalog http://geoport.whoi.edu/thredds/catalog/coawst_4/use/fmrc/catalog.html

Dataset: coawst_4_use/Best Time Series

- Data format: netCDF
- Data type: GRID
- · Naming Authority: gov.usgs.er.whsc
- ID: coawst_4/use/fmrc/coawst_4_use_best.ncd

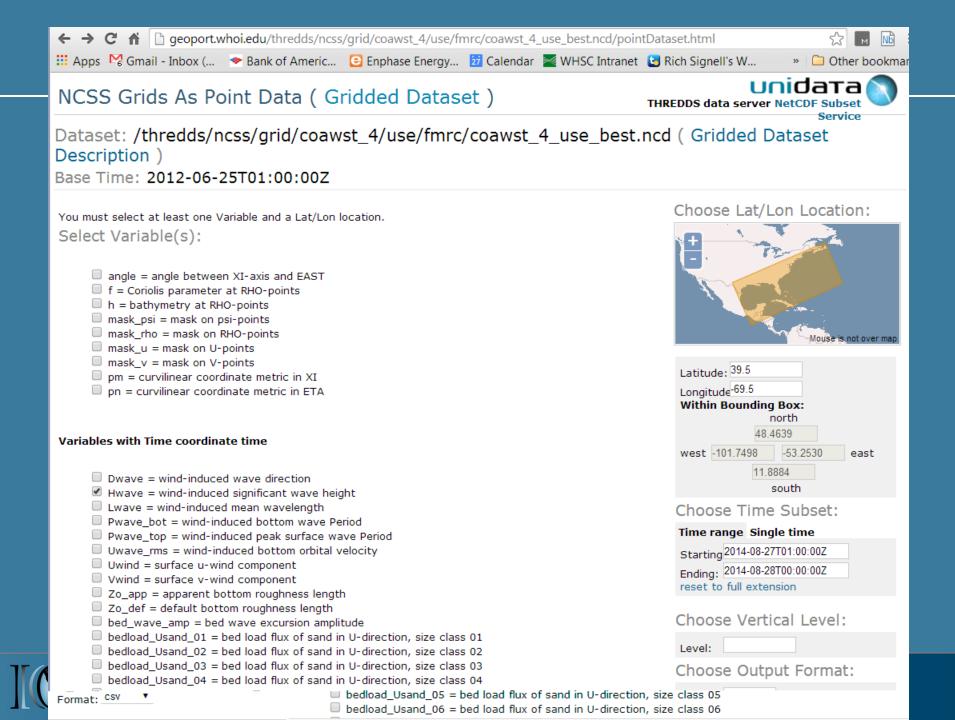
Documentation:

- summary: Best time series, taking the data from the most recent run available.
- · summary: ROMS USE Output from COAWST
- Carolinas Coastal Change Program
- ReadMe.txt

Access:

- 1. OPENDAP: https://dodsC/coawst_4/use/fmrc/coawst_4_use_best.ncd
- 2. NetcdfSubset: /thredds/ncss/grid/coawst 4/use/fmrc/coawst 4 use best.ncd
- 3. WMS: /thredds/wms/coawst 4/use/fmrc/coawst 4 use best.ncd
- 4. ISO: /thredds/iso/coawst 4/use/fmrc/coawst 4 use best.ncd
- 5. NCML: /thredds/ncml/coawst 4/use/fmrc/coawst 4 use best.ncd
- 6. UDDC: /thredds/uddc/coawst_4/use/fmrc/coawst_4_use_best.ncd





CSV Time Series from Subset Service

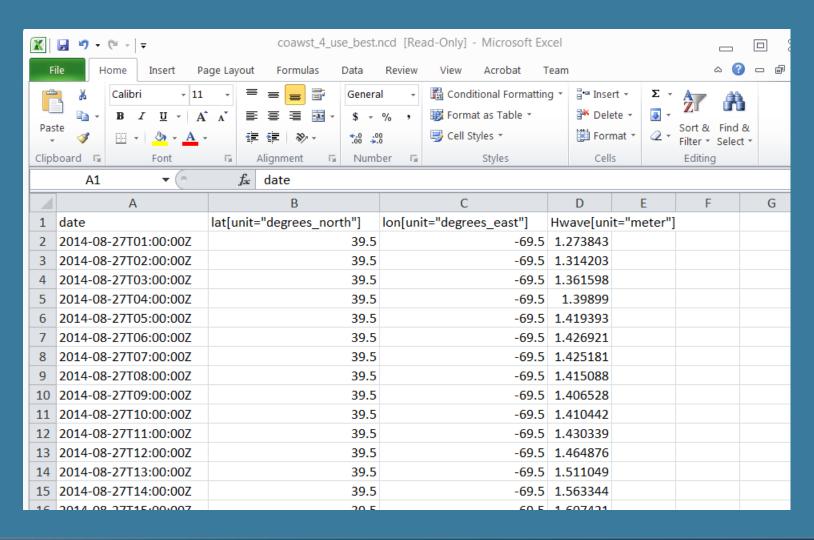
```
date, lat [unit="degrees north"], lon [unit="degrees east"], Hwave [unit="meter"]
2014-08-27T01:00:00Z,39.5,-69.5,1.273842692375183
2014-08-27T02:00:00Z,39.5,-69.5,1.3142027854919434
2014-08-27T03:00:00Z,39.5,-69.5,1.3615976572036743
2014-08-27T04:00:00Z,39.5,-69.5,1.398990273475647
2014-08-27T05:00:00Z,39.5,-69.5,1.419392704963684
2014-08-27T06:00:00Z,39.5,-69.5,1.426921010017395
2014-08-27T07:00:00Z,39.5,-69.5,1.4251813888549805
2014-08-27T08:00:00Z,39.5,-69.5,1.4150882959365845
2014-08-27T09:00:00Z,39.5,-69.5,1.406528115272522
2014-08-27T10:00:00Z,39.5,-69.5,1.4104422330856323
2014-08-27T11:00:00Z,39.5,-69.5,1.4303393363952637
2014-08-27T12:00:00Z,39.5,-69.5,1.4648756980895996
2014-08-27T13:00:00Z,39.5,-69.5,1.5110485553741455
2014-08-27T14:00:00Z,39.5,-69.5,1.5633442401885986
2014-08-27T15:00:00Z,39.5,-69.5,1.6074210405349731
2014-08-27T16:00:00Z,39.5,-69.5,1.6265686750411987
2014-08-27T17:00:00Z,39.5,-69.5,1.6184144020080566
2014-08-27T18:00:00Z,39.5,-69.5,1.5933209657669067
2014-08-27T19:00:00Z,39.5,-69.5,1.5608850717544556
2014-08-27T20:00:00Z,39.5,-69.5,1.527510643005371
2014-08-27T21:00:00Z,39.5,-69.5,1.5050616264343262
2014-08-27T22:00:00Z,39.5,-69.5,1.5094025135040283
2014-08-27T23:00:00Z,39.5,-69.5,1.5698728561401367
2014-08-28T00:00:00Z,39.5,-69.5,1.7087903022766113
```

CSV request is a "RESTful" URL:

http://geoport.whoi.edu/thredds/ncss/grid/coawst_4/use/fmrc/coawst_4_use_best.ncd?var =Hwave&latitude=39.5&longitude=-69.5&time_start=2014-08-27T01:00:00Z&time_end=2014-08-28T00:00Z&vertCoord=&accept=csv

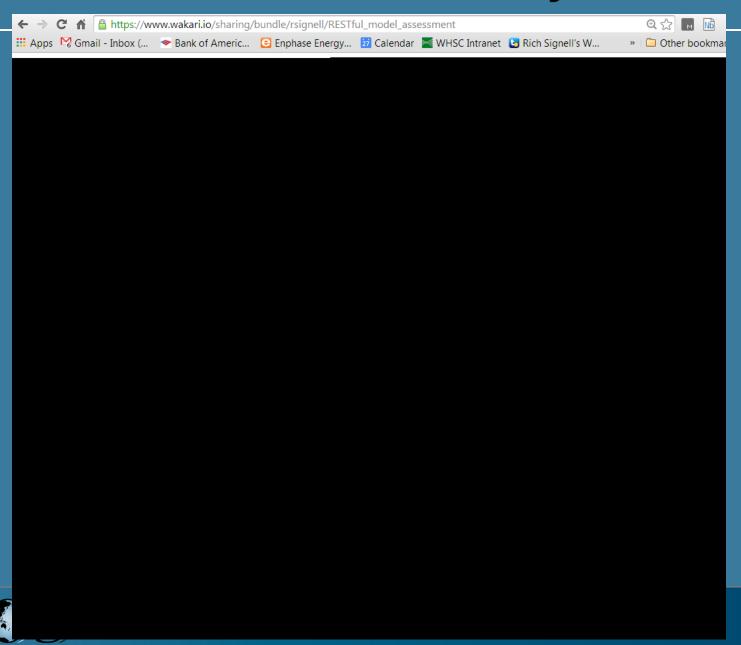


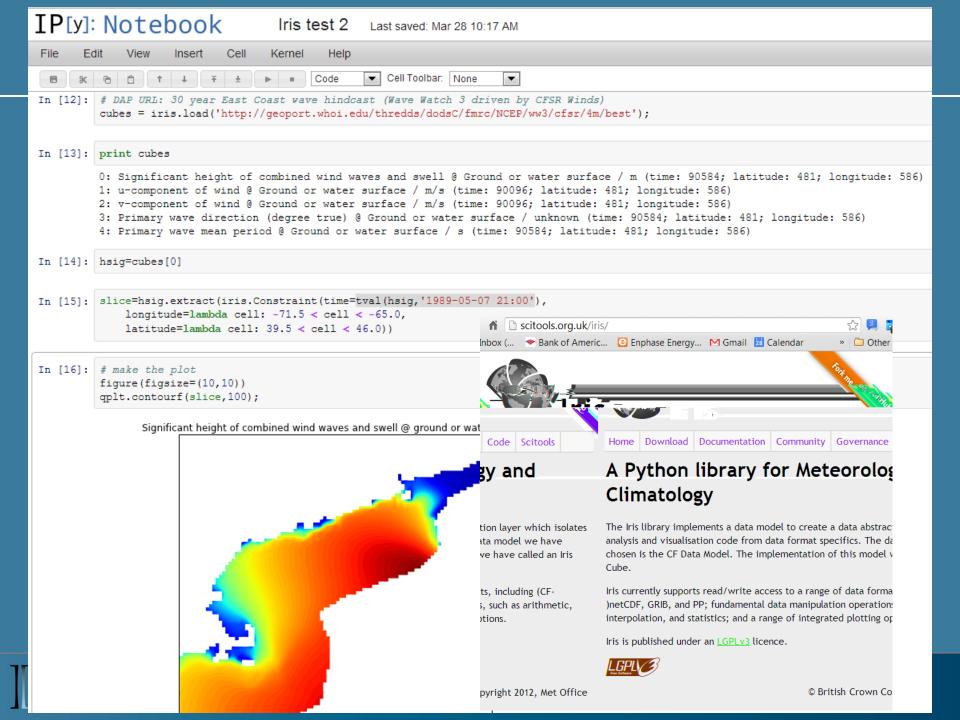
Accessing the CSV data in Excel



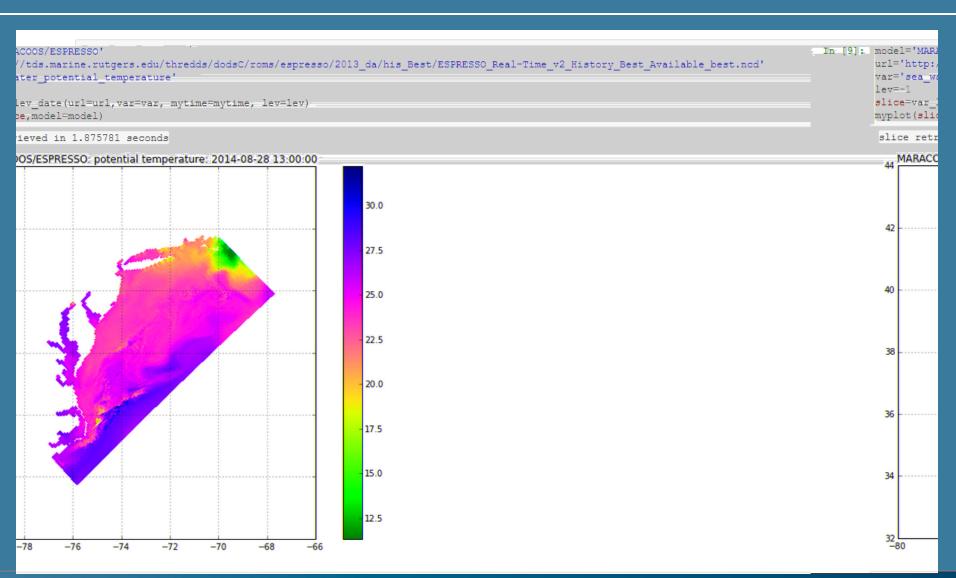


Access the CSV data in Python, R, etc





ESPresso model (Rutgers)

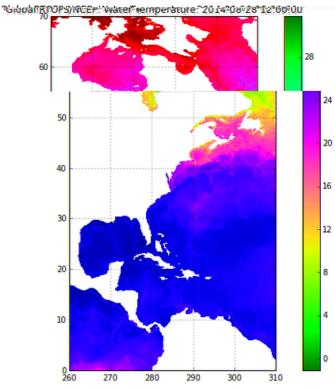




Global RTOFS (NOAA)

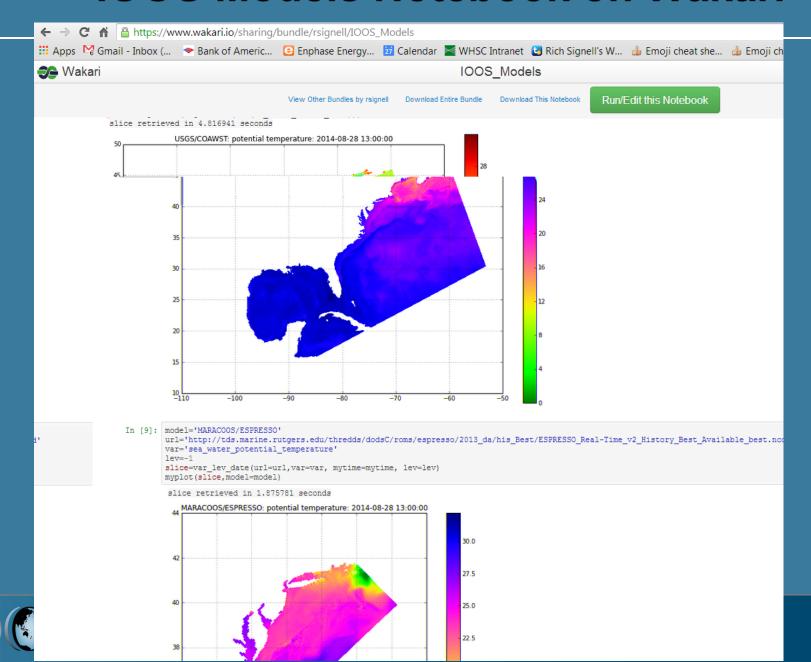
```
In [13]: model='Global RTOFS/NCEP'
url='http://ecowatch.ncddc.noaa.gov/thredds/dodsC/hycom/hycom_regl_agg/HYCOM_Region_1_Aggregation_best.ncd'
var='sea_water_temperature'
lev=1
subsample=1
slice=var_lev_date(url=url,var=var, mytime=mytime, lev=lev, subsample=subsample)
myplot(slice,model=model)

slice retrieved in 1.225454 seconds
/opt/anaconda/envs/np18py27-1.9/lib/python2.7/site-packages/iris/fileformats/_pyke_rules/compiled_krb/fc_rules_cf_fc.py:1196: UserWarning: Ignoring netCliable 'salinity' invalid units 'psu'
warnings.warn(msg.format(msg_name, msg_units))
```





IOOS Models Notebook on Wakari



60 minutes to serving your data

- 1. Install Sun/Oracle Java (10 min)
- 2. Install/configure Tomcat (15 min)
- 3. Install/configure the Thredds Data Server (15 min)
- 4. Put NetCDF (or HDF4, Grib1, Grib2, HDF5 files) in a directory (5 min)
- 5. Add NcML (XML) files for CF compliance and aggregation (15 min)



Adding data to THREDDS



alog.html Catalog http://geoport.whoi.edu/thredds/catalog/usgs/data0/bbleh/tidal/cat

Dataset: tidal/wide.ncml

• Data size: 218.0 bytes

• ID: usgs/data0/bbleh/tidal/wide.ncml

Access:

1. OPENDAP: /thredds/dodsC/usgs/data0/bbleh/tidal/wide.ncml

2. HTTPServer: /thredds/fileServer/usgs/data0/bbleh/tidal/wide.ncml

3. NetcdfSubset: /thredds/ncss/grid/usgs/data0/bbleh/tidal/wide.ncml

4. ISO: /thredds/iso/usgs/data0/bbleh/tidal/wide.ncml

5. NCML: /thredds/ncml/usgs/data0/bbleh/tidal/wide.ncml

6. UDDC: /thredds/uddc/usgs/data0/bbleh/tidal/wide.ncml

7. WMS: /thredds/wms/usgs/data0/bbleh/tidal/wide.ncml

Dates:

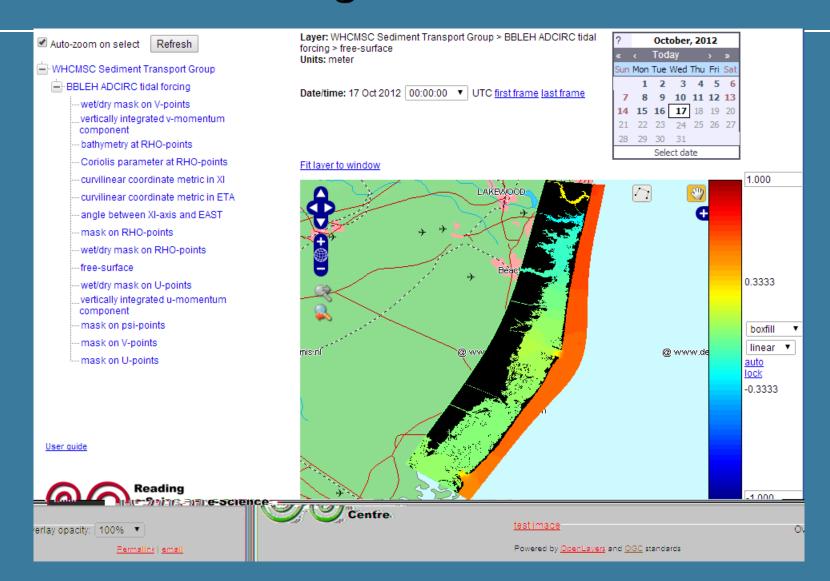
• 2014-06-03T12:44:05Z (modified)

Viewers:

- Godiva2 (browser-based)
- NetCDF-Java ToolsUI (webstart)



Browsing WMS with Godiva2

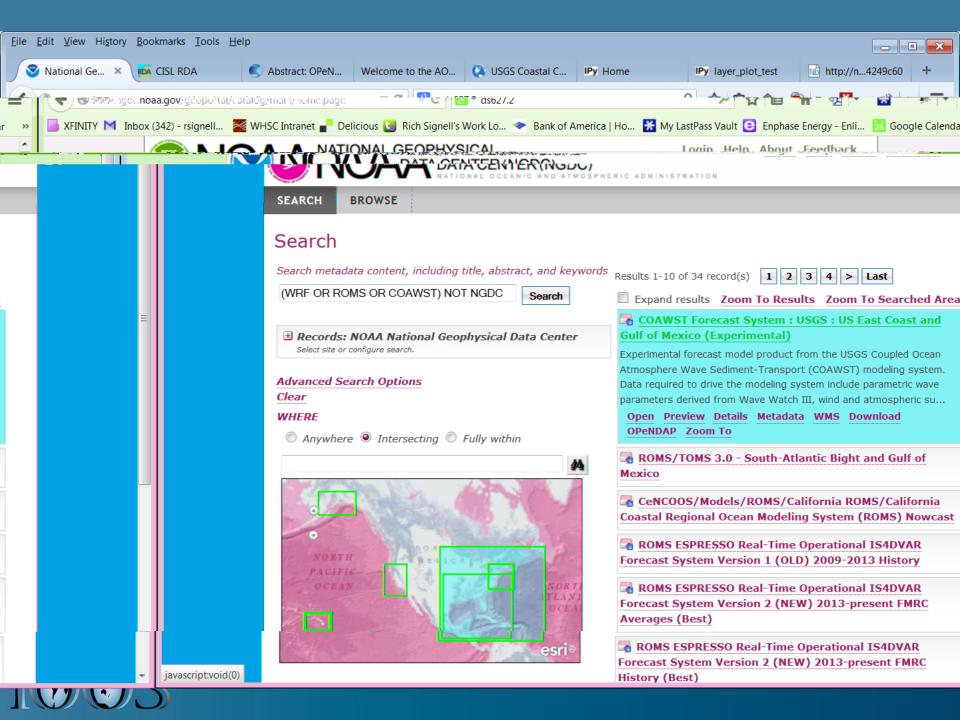




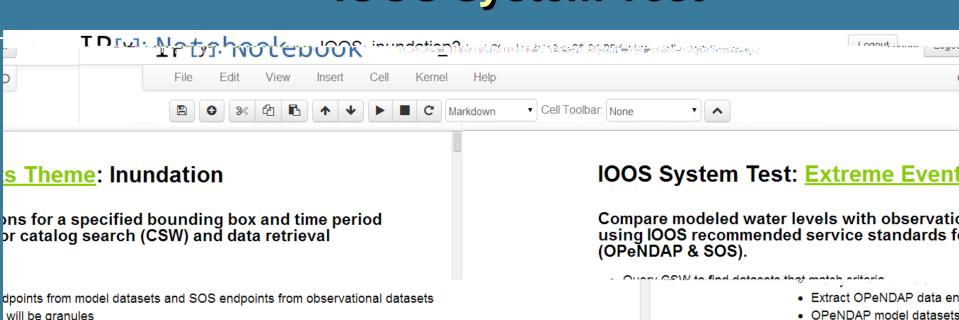
Searching for Data

GI-CAT catalog broker service architecture





100S System Test



- will be granules
- tasets (from ncSOS) or collections of datasets (from NDBC, CO-OPS SOS servers)
- in datasets
- asets
- tasets at locations of observations
- on same vertical datum

nd bounding box of interest:

- mes (UTC) ...
- (2012, 10, 26)2012,11,2)
- (2014, 2, 10)
- 2014,2,15)

· SOS endpoints may be da Filter SOS services to obta

- · Extract data from SOS data
- Extract data from model da
- Compare time series data

Specify a time range a

Project Lead: Derrick Snowden

Notebooks: Kyle Wilcox, Andy Bird Bob sandy dt.datetime

Fratantonio, Kelly Knee, Will Koeppen, dt.datetime (

Hannah Dean

2014 feb 10-15 storm https://github.com/ioos/system-testart = dt.datetime

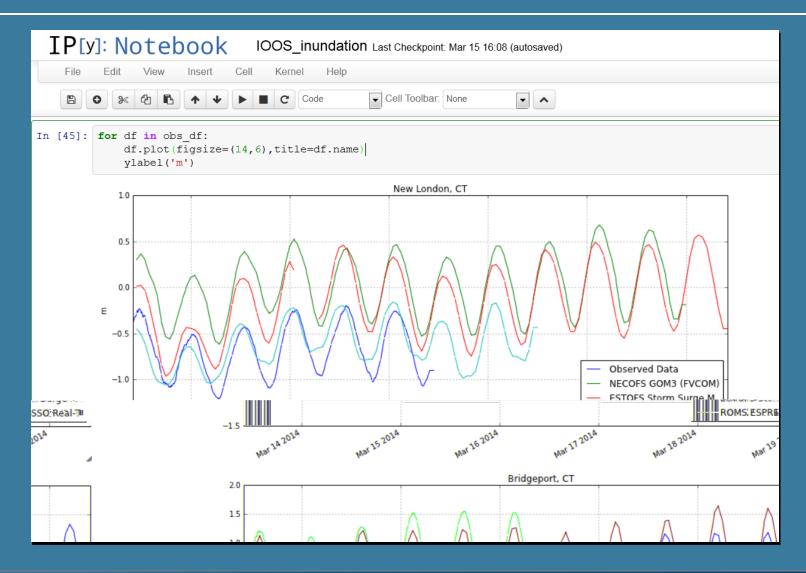
jd stop = dt.datetime(







Automated model comparison





SECOORA Model Assessment



← → C ↑ □ nbviewer.ipython.org/github/ioos/secoora/blob/master/notebooks/inundation/00-generate_page_html.ipynb 🔛 Apps 🛂 Gmail - Inbox (... 🧇 Bank of Americ... 😉 Enphase Energy... 🔞 Calendar 🌌 WHSC Intranet 🚨 Rich Signell's W... 👍 Emoji cheat she... 👍 Em with open(fname, 'w') as f: **Project Lead: Vembu Subramanian** f.writelines(table)

to html(bias.T)

Notebook: Filipe Fernandes

URL: https://github.com/ioos/secoora

Out [21]:

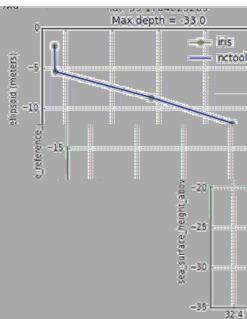
	COAWST_4	ESPRESSO	ESTOFS	HYCOM	SABGOM
Duck, NC			0.05		0.43
Oregon Inlet Marina, NC	-0.07	0.21	0.00	0.45	
Wrightsville Beach, NC	-0.03		0.07	0.32	0.45
Springmaid Pier, SC	-0.04		0.08		0.45
Oyster Landing (N Inlet Estuary), SC			0.18		
Fort Pulaski, GA			0.13		
Fernandina Beach, FL			0.04		
Mayport (Bar Pilots Dock), FL			0.01		
Trident Pier, FL			-0.13		
Lake Worth Pier, FL	-0.30		-0.11	0.07	
Virginia Key, FL			-0.06		0.25
Vaca Key, FL			-0.06		
Key West, FL	-0.28		-0.08		0.14
Naples, FL			-0.03		
Port Manatee, FL			0.02		
Clearwater Beach, FL			0.07		
Cedar Key, FL	-0.10		0.09		0.18
Apalachicola, FL	-0.04		0.15		
Panama City, FL	-0.02				
Pensacola, FL	0.04				

Python & Matlab notebook

```
In [5]: %%matlab -i url -o s -o g
        addpath(genpath('/home/filipe/IOOS/nctoolbox'));
        nc = ncgeodataset(url);
        s = nc{\text{'salt'}}(1, :, 11, 21);
        g = nc{'salt'}(1, :, 11, 21).grid;
        log4j:WARN No appenders could be found for logger (ucar.nc2.NetcdfFile).
        log4j:WARN Please initialize the log4j system properly.
        kw = dict(linewidth=2, color=(0.6, 0.1, 0.15),
                   alpha=0.75, label='nctoolbox')
        leg = dict(numpoints=1, loc='best', framealpha=0)
        c = cube[0, :, 10, 20]
        fig, ax = plot profile(c)
        ax.plot(s, g['z'], **kw)
        1 = ax.legend(**leg)
                       lon: -74.0022201538
```

lat. 20 1704005000





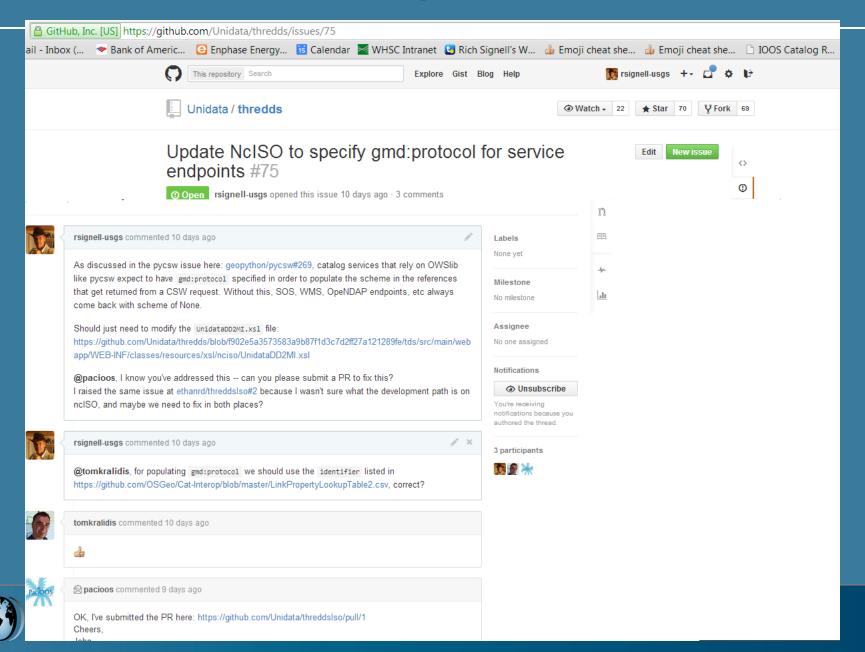


Unidata Challenges

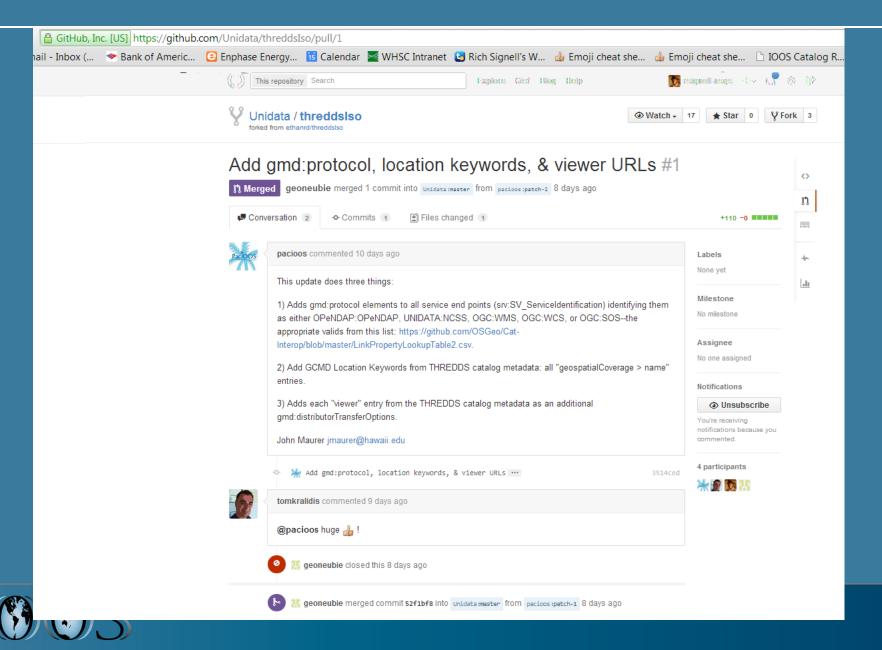
- Community THREDDS Data Servers robustness
- Support as popularity increases
- UGRID incorporated into Unidata NetCDF-Java
- Aggregation of large collections of NetCDF files
- Met/Ocean tools developed for Python on top of CF/Ugrid conventions (Iris, libCF)
- IDV-like client built on top of Python tools
- Participation in standards (e.g. OGC)
- Managing community development



Community Development



Community Development



Why do I promote Unidata technologies?

- Powerful
- Flexible
- Easy to install
- Free
- Supported
- Driven by community of users



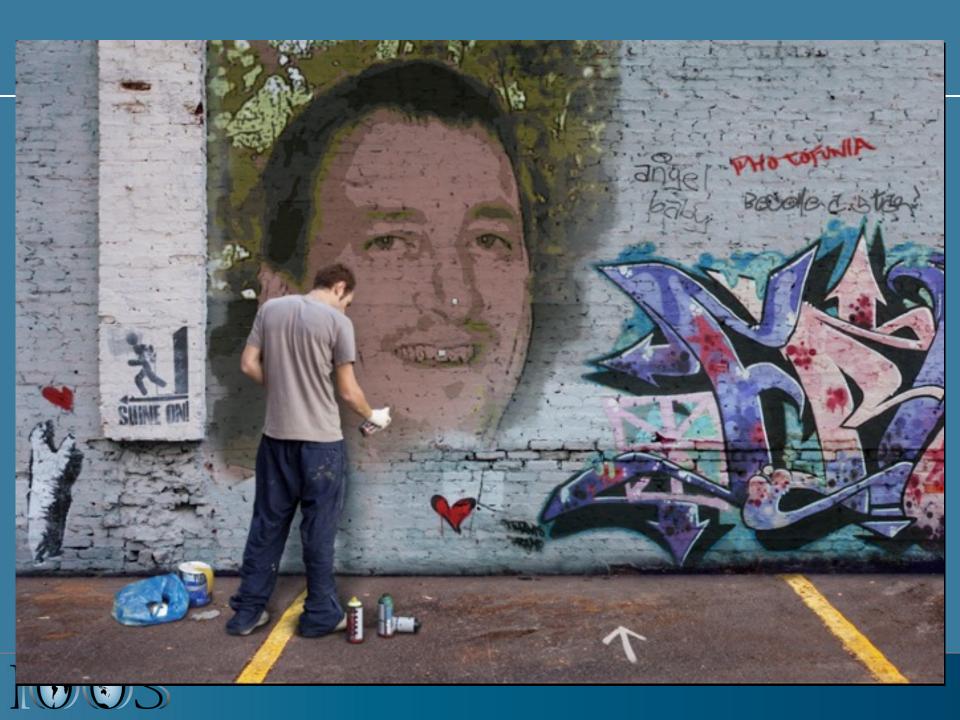




NPS160











Summary (1 of 2)

- Common data models for "feature types" (structured and unstructured grids, time series, profiles, swaths) (Unidata CDM)
- Standard web services for delivering these data and metadata (OGC, Unidata)
- Tools to access and process these services in common analysis environments: R, Matlab, Python, ArcGIS, JavaScript



Summary (2 of 2)

- It's easy (1 hour) to deploy free, supported systems that allow for standards-based delivery of aggregated data from native model grids that put little effort on the data provider
- What do you get?
 - Lots of choices for data access (Browser, Matlab, Python, Excel, IDV, R, IDL)
 - More usage of model results by more people
 - Faster feedback to modelers, leading to improved models
 - Shared code base in the community
 - Increased community support for standards-based access
 - Less time wasted messing with data, more time spent on ecosystem based management
- What should you do? Encourage providers to use these standards, and develop tools that use standardized access



More at: Github.com/rsignell-usgs

