PRESENTER: Roland Schweitzer

OUR GOAL

- Publish the largest possible collection of high-quality OPeNDAP accessible data sets.

METHODS

1. Create a catalog of THREDDS Data Servers from NOAA and selected partners.
2. Crawl the entire collection, writing a new catalog hierarchy eliminating any candidate catalogs containing un-aggregated time series or non-OPeNDAP data sources.
3. Verify that each of the resulting data sources contains a CF-Compliant netCDF data set.
4. Collect Discrete Sampling Geometry datasets into a central ERDDAP server.
5. Publish the results in a THREDDS Data Server, a Live Access Server and an ERDDAP server (with both the grid and discrete sampling data sets).
6. Collect and publish ISO-19115 metadata for each data set into noaa.data.gov, data.gov and NOAA One-Stop.

RESULTS

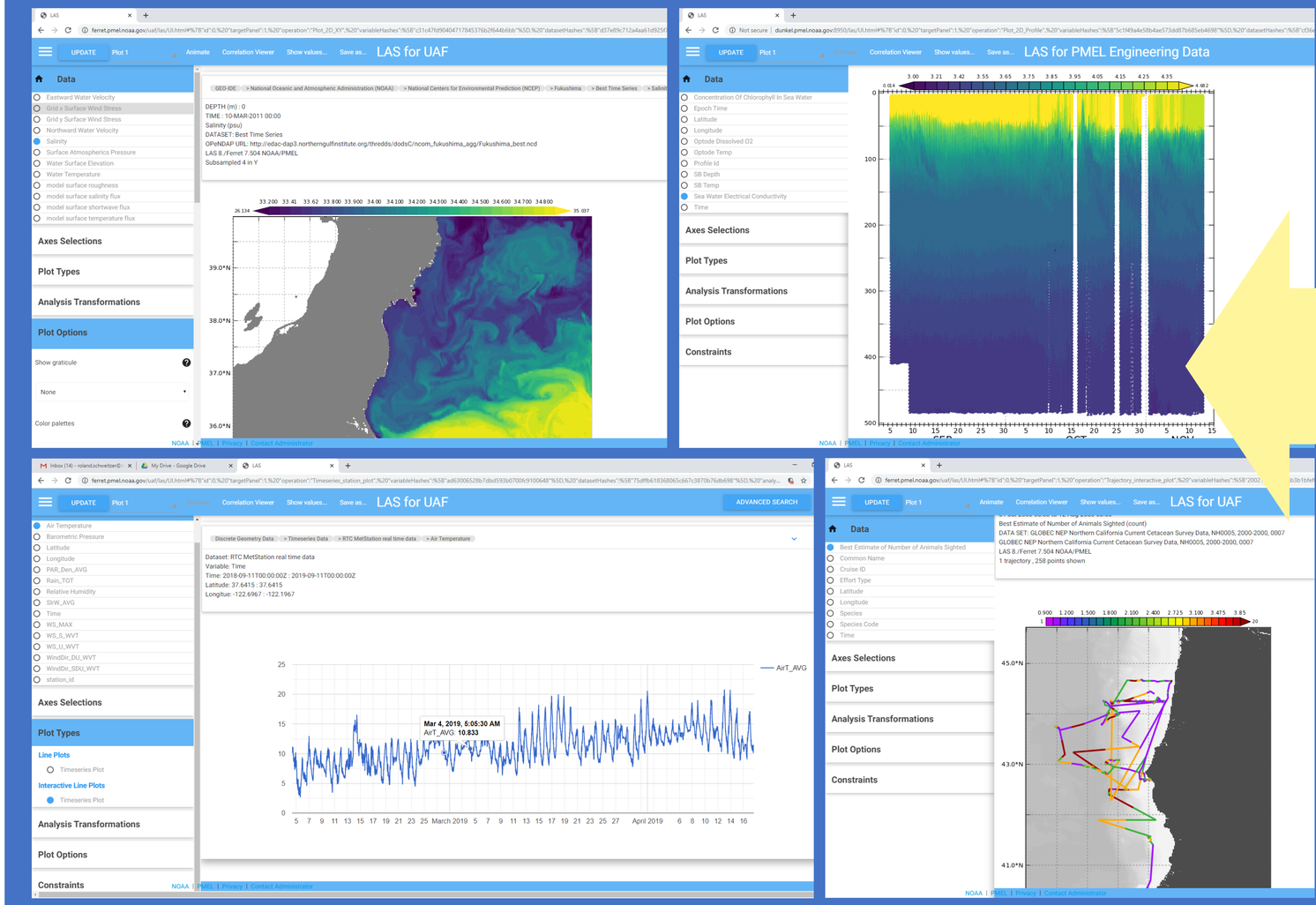
- 12420 data sets available from your favorite server using your favorite scientific analysis and visualization software.
- Searchable and discoverable via the standard NOAA and federal data repositories.

DISCUSSION

- With the right approach (copying existing successful standards and data services) and the right software (big data servers like THREDDS OPeNDAP, ERDDAP and data catalogs) and a little software glue (the UAF Catalog Cleaner and ncISO) a few people can produce and maintain a large high-quality data service filled with useful and discoverable data sets.

About LAS

Find out how a few nerds manage a data collection of 12,000+ data sets with open-source software, ingenuity and a bit of custom software glue.



AGU100 ADVANCING EARTH AND SPACE SCIENCE

Managing a Community Data Collection with Open Source Software
IN33B-0826

Roland Schweitzer, Ethan Davis, Sean Cody Arms, Robert Simons, Kevin O'Brien, David Neufeld

THE FOSS and the GLUE

FOSS:

THREDDS Data Server
https://github.com/Unidata/thredds

ERDDAP
https://github.com/BobSimons/erddap

Live Access Server (v8)
https://github.com/NOAA-PMEL/LAS

Live Access Server (v9)
https://github.com/NOAA-PMEL/las9
(coming soon)

GLUE:

UAF Catalog Cleaner
https://github.com/NOAA-PMEL/FastClean

ncISO
https://github.com/NOAA-PMEL/uafnciso

Try LAS 9
(not optimized for mobile)

NOAA National Oceanic and Atmospheric Administration U.S. Department of Commerce

PMEL Pacific Marine Environmental Laboratory