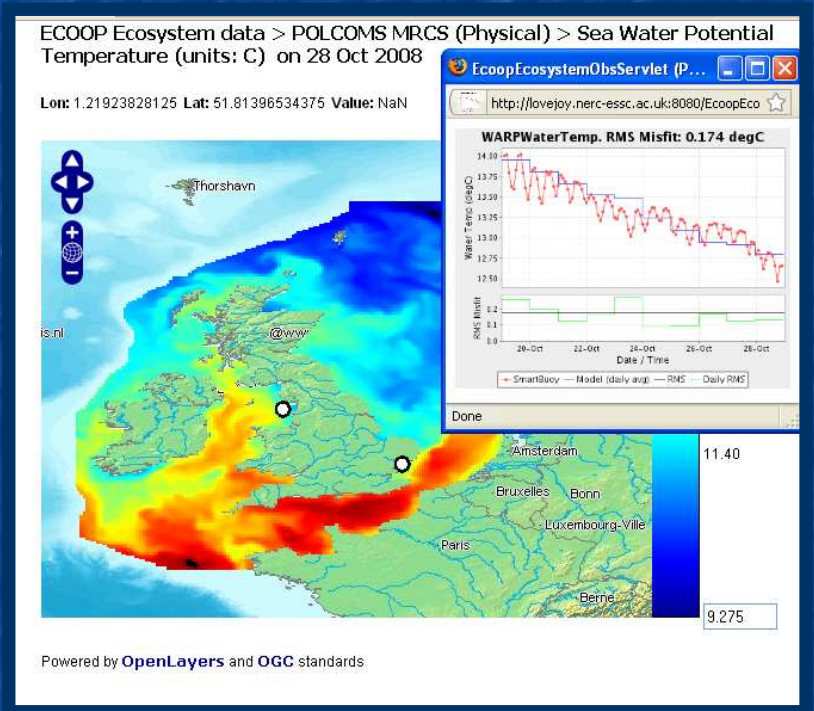
Online intercomparison of models and observations using OGC and community standards

Alastair Gemmell
* **Jon Blower**
Keith Haines
Adit Santokhee

Reading e-Science Centre,
Environmental Systems Science Centre,
University of Reading, UK

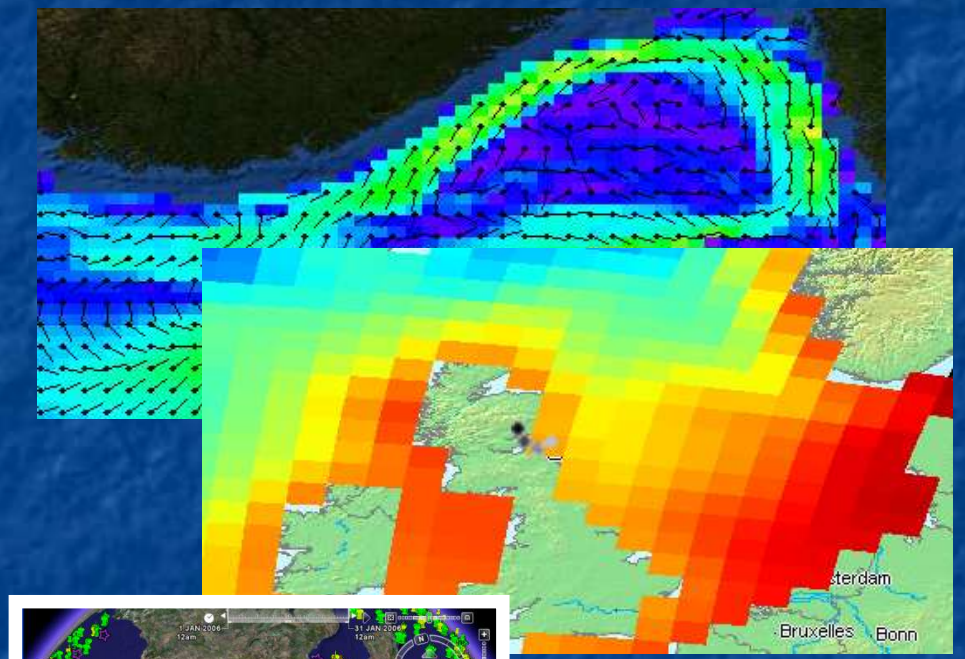
Plus lots of colleagues on ECOOP project

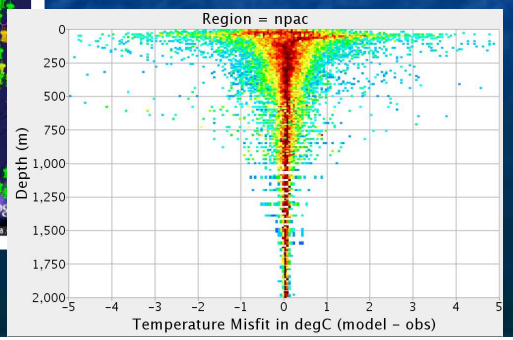


OGC/GIS/Met meeting, Toulouse, November 2009

Reading e-Science Centre

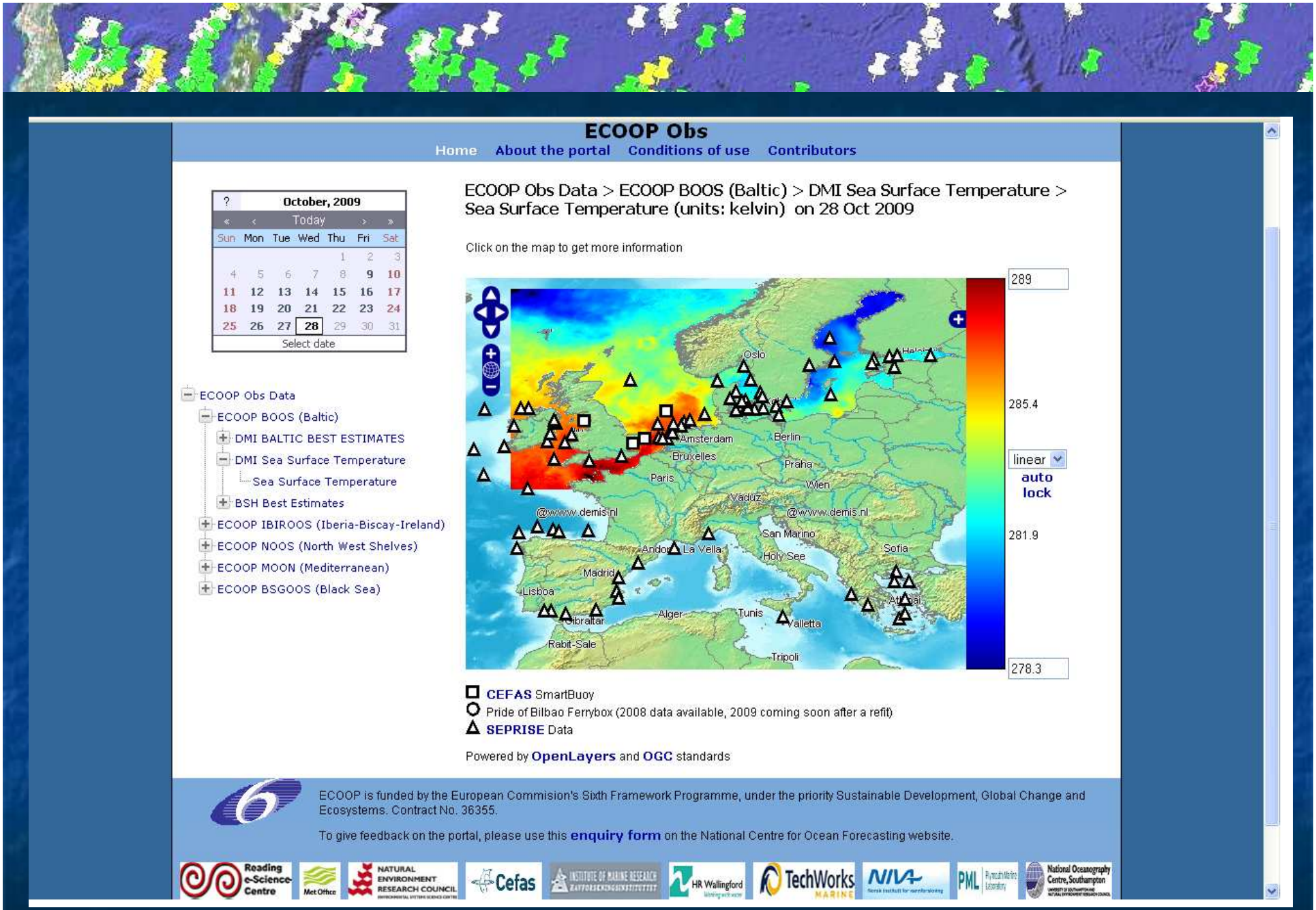
- Based in Environmental Systems Science Centre, University of Reading, UK
- Goal is to apply novel IT techniques to scientific research
- We are scientists and software engineers, not data providers
- We work with many kinds of data: model, in situ, remote sensing
- In many scientific areas (oceanography, meteorology, climate science, hydrology...)
- OpenGIS can help us to bring these datasets together



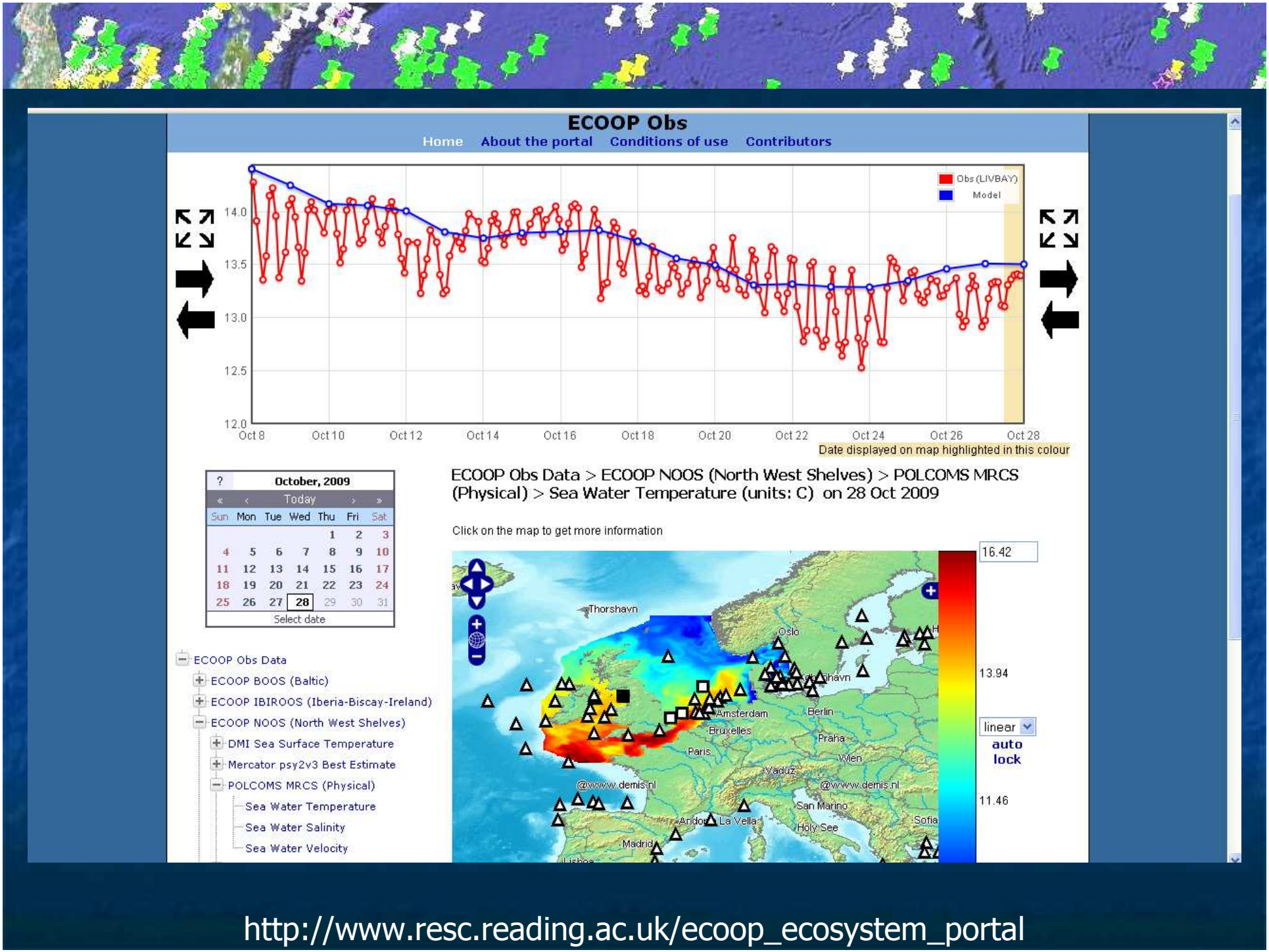


ECOOP Ecosystem Portal

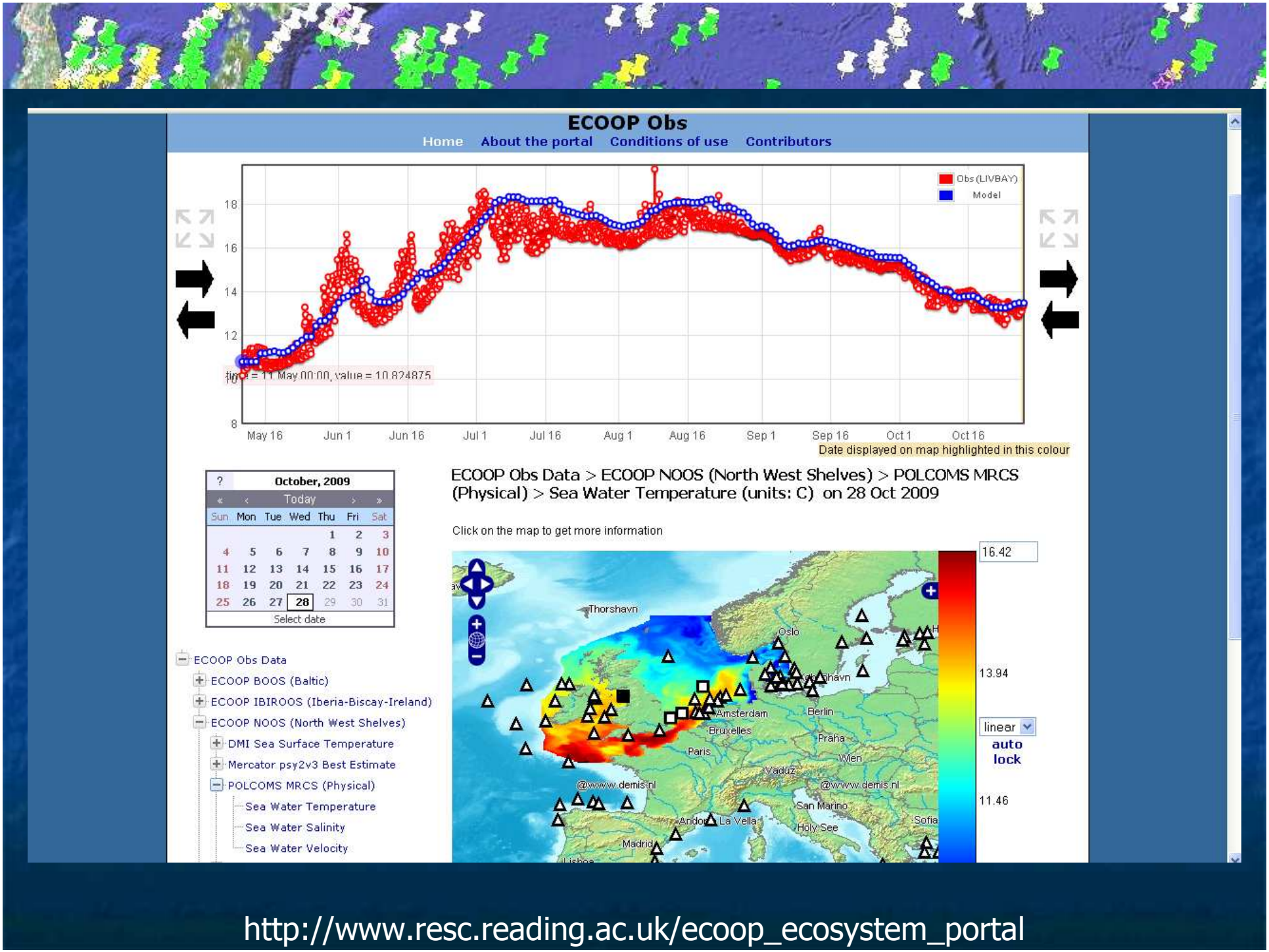
- ECOOP: European COastal Sea Operational observing and forecasting system.
- Aim is to compare **in-situ observations** and **model analyses/forecasts** as a technology demo for predicting harmful algal blooms in the North Sea.
- Intercomparison gives users more confidence in bloom predictions if past observed and model data agree relatively well
- Physical data (temperature, salinity and velocity)
- Biological data (biomass, nitrate and silicate concentrations)
- Data distributed among the partners



http://www.resc.reading.ac.uk/ecoop_ecosystem_portal

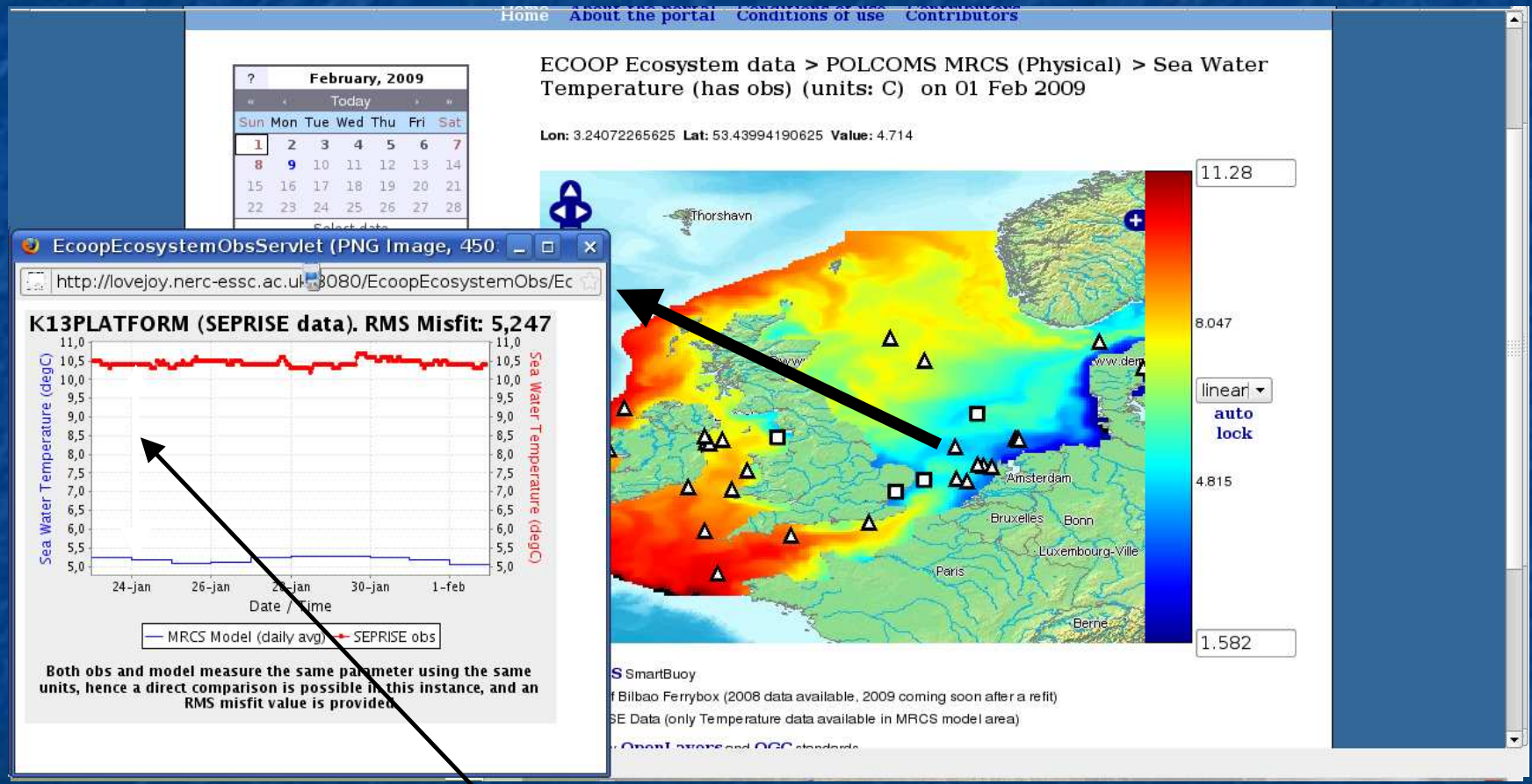


http://www.resc.reading.ac.uk/ecoop_ecosystem_portal



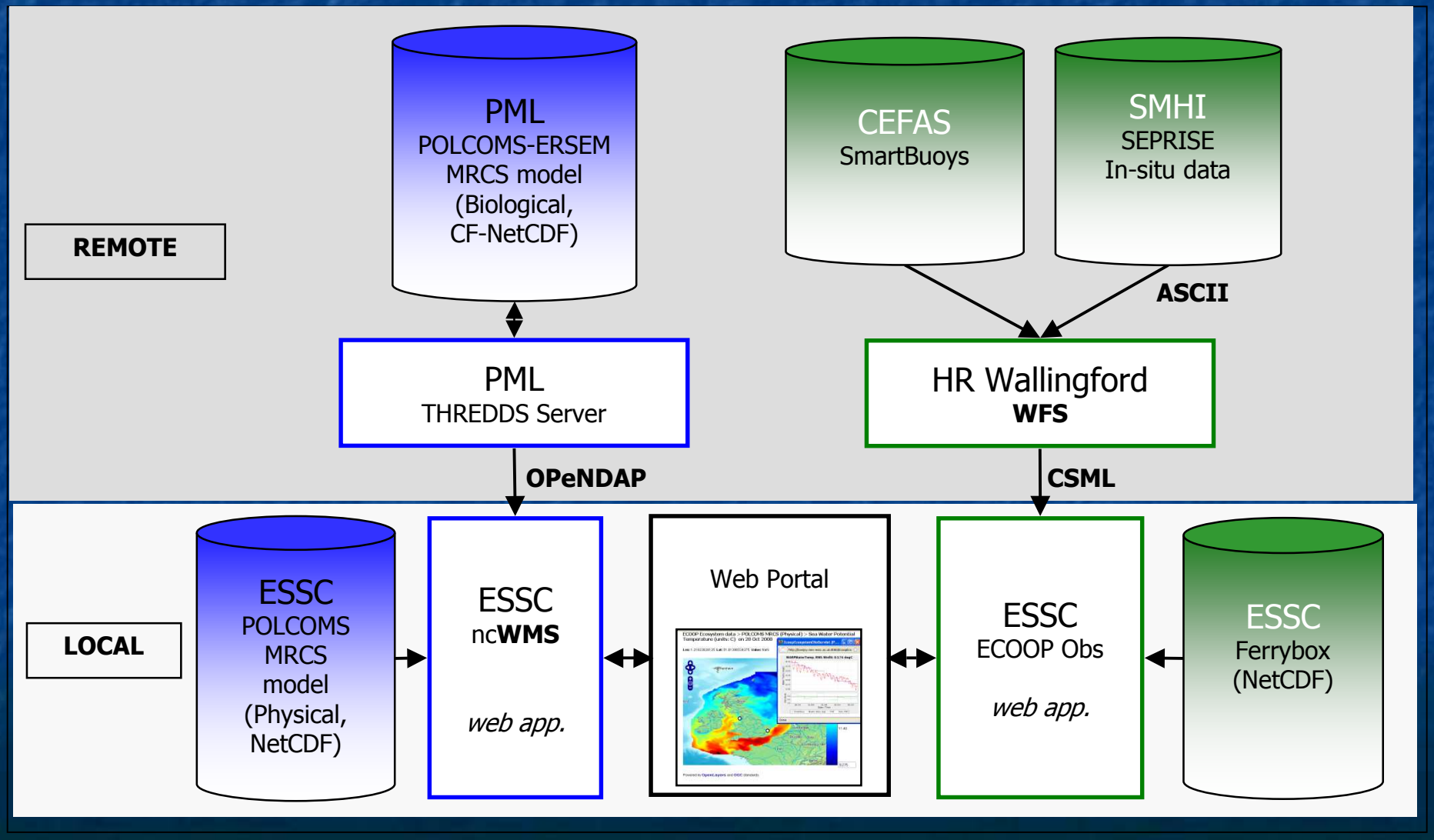
http://www.resc.reading.ac.uk/ecoop_ecosystem_portal

Comparing / co-plotting datasets can catch errors!

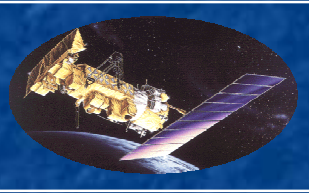
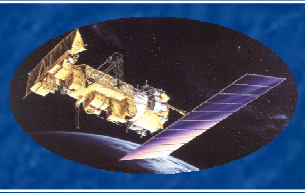


This looks like a suspiciously large and constant difference between obs and model

ECOOP Ecosystem Portal architecture



Representing data as CSML Features (selected)

PointSeriesFeature (*timeseries at a point*)	 
ProfileFeature (*vertical profile at a point*)	  
GridSeriesFeature (*series of multidimensional grids*)	 
SwathFeature (*single satellite sweep*)	
SectionFeature (*vertical section*)	 

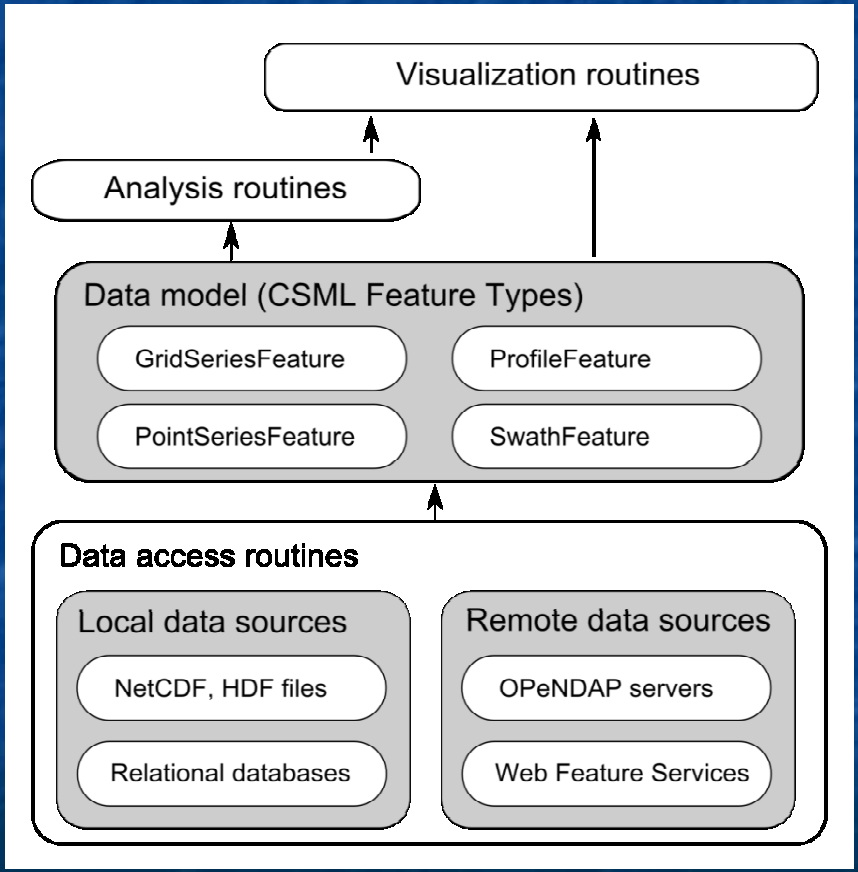
Feature Types are classified by their geometry

How it works

- Model data accessed through OPeNDAP and ncWMS
  - “pseudo-observations” extracted through WMS GetFeatureInfo
- In situ observations accessed through WFS and direct NetCDF file access
  - GeoServer, serving CSML documents
- Custom dynamic web interface based on OpenLayers

How would we like to improve this?

- Harmonize the data access layer (right)
- GetFeatureInfo should return CSML
- Bring in other data sources
- Allow data download
  - Output CSML data model in various formats, inc. NetCDF, XML, CSV
- Would SOS be more suitable than WFS?



Challenges in using CSML for data exchange

- Multiplicities
  - Inheritance from GML gives large number of encoding options
  - Many ways of encoding "soft typed" quantities like units, phenomena, coordinate systems
- So we need a "cookbook" of recommended best practice to reduce the options to something manageable
  - A "profile of a profile" essentially
  - GeoSciML did this too

What **value** do we get from OGC for this kind of app?

- The input of a much wider community
- More robust domain modelling
- Software reuse
  - Mainly for "semantically light" services
  - WMS and simple-feature WFS
- Other domains will be familiar with terms and technology
- A warm fuzzy feeling
  - Sometimes, anyway ;-)

Where is the OGC approach less helpful?

- WFS 1.0 doesn't support feature subsetting
  - We invented a syntax for timeseries subsetting
- Serving gridded (model) data
  - WCS currently adds very little value above CF-OPeNDAP in practice for this app
- Not much "real" interoperability with third parties in terms of data
  - Interop of geospatial component much better
  - All OGC services need profiling for our use
  - ... But what does that mean for other communities with different profiles?

Suggestion for maximizing value of WFS approach

- First point of entry should be a simple-feature WFS
  - Serves observation ***locations*** only
  - Plus simple text attributes (e.g. Instrument type)
  - Accessible by most OpenGIS clients and communities
  - Allows reuse of lots of existing tools
- Clicking on obs location gives fully-specified complex feature
  - In "highly precise" formats for special clients, e.g. CSML
  - In less precise formats for general audience, e.g. HTML, CSV
  - Options to subset the feature through WFS 2.0
- Requires links between services
- Looks more like the Web

Conclusions and suggestions

- Combination of met-ocean community and OGC standards used to support online data intercomparison
- CSML data model is a very practical abstraction
  - (although we still need to specialize it further to be implementable)
  - Convergence with Unidata CDM will be very welcome
  - We are producing a FOSS Java library to abstract various kinds of data
- We need to understand how best to use our robust, proven, practical (note: not "legacy" technology)
  - CF, NetCDF, OPeNDAP
- Lots of others are developing this type of app
  - How about a Best Practices document for Web-GIS?