Deb Agarwal
Berkeley Water Center - University of California and
Lawrence Berkeley National Laboratory
A major shift is happening in the way eco-science is done.

- Moving from individual studies of local processes to collaborative studies of regional and global processes. (e.g. studying the impact of climate change)

Studying global scale environmental processes requires:

- Integration of local, regional, and global spatial scales.
- Integration across disciplines, e.g., climatology, hydrology, forestry, etc., and across methodologies (field observations, remote sensing, and modeling).
Data Rich Environment

• National and International Datasets
  • USGS National Water Information System
  • NOAA National Climatic Data Center
  • FLUXNET Network
  • Satellite data (e.g. MODIS)

• Local Datasets
  • Local Agencies
  • Companies (e.g. Timber)
  • Ecology Organizations
  • Individual Researchers
Data Comes in Many Forms

- Manual Measurement
- Automated Measurement
- Sample Collection
- Typing
- Aircraft Surveys
- Model Output
- Counting
- Historical Photographs
- Relatively Ubiquitous Motes
Data Synthesis – The Challenge of Bringing it all Together

USGS stream gauges
Coho presence/absence data
MODIS evapotranspiration
Cross section locations
Samples from 1997
Stream temperature gauges
Water system
Obstructions and other
Human activities
Data Integration Challenges

- Point and spatial
- Time series and low frequency
- Vocabulary meanings (ontology)
- Sparse in time, duration, or location
• Organize data by its natural dimensions
  • Watershed
  • Stream
  • Data type
  • Time
  • etc
• Select
  • Aggregations – yearly, monthly, stream, etc
  • Filters – watershed, years,
• Search for specific characteristics
• Browse and display data in the way that makes sense to the viewer
Data Analysis Infrastructure For Time Series Data

Environmental Data Server

- Distributed Data Sets
- Catalog Database
- Archive Databases
- Data Cubes
- Models
- Reports & Visualizations
Non-Time Series Data

- One-time or infrequent event information
- Approximate measurements
- Raster images
- Photographs
- Video
- GIS information and shape files
- Model output
### Salmon Lifestages - Different Conditions

<table>
<thead>
<tr>
<th>LIFE STAGE</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult migration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spawning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Egg Incubation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergence/ Fry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juvenile rearing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emigration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A Start on this Issue

- **Mashups**
  - Browsing in a geographic view
  - Overlays of rasters
  - Popups with data
- **Geographic Information System (GIS)**
  - Connection to time series data
  - Operate in a geographic tool
• Find the data needed – Agency sites, individual researchers, networks, . . .
• Convert it to a format and units appropriate to the analysis
• Interpolate in space or time as needed
• Gap fill and quality assess
• Understand appropriate use and limitations
• Perform the analysis
• Gather all the ‘fair-use’ criteria and acknowledgments
• Publish the paper
Acknowledging Data Sources (KRXIS Data)

<table>
<thead>
<tr>
<th>Source Title</th>
<th>Source Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source 1</td>
<td>Detailed overview of market trends and statistical analysis.</td>
</tr>
<tr>
<td>Source 2</td>
<td>Comprehensive database of financial instruments.</td>
</tr>
<tr>
<td>Source 3</td>
<td>Real-time data feed for asset management.</td>
</tr>
<tr>
<td>Source 4</td>
<td>Historical data repository for research purposes.</td>
</tr>
<tr>
<td>Source 5</td>
<td>Analysis tools for risk assessment and portfolio optimization.</td>
</tr>
</tbody>
</table>

*Note: This table represents a sample of the data sources acknowledged in the document.*
Challenge - Web of Data and Papers

- Data archiving is just the beginning
- Data Publications – about the data
- Capture the contributor information, fair-use criteria, and acknowledgments
- Connect to papers that used the data to the data
- Capture analysis artifacts and connect them in
- Capture data corrections/gap-fills/quality assessment
- Capture data versions and link them
- Further increase data sharing
Acknowledgements

Berkeley Water Center, University of California, Berkeley, Lawrence Berkeley Laboratory
- Dennis Baldocchi
- Jim Hunt
- Monte Goode
- Susan Hubbard
- Keith Jackson
- Rebecca Leonardson (student)
- Carolyn Remick

Microsoft Research
- Catharine van Ingen
- Tony Hey
- Dan Fay

University of Virginia
- Marty Humphrey
- Norm Beekwilder
- Jie Li (student)

University of Indiana
- You-Wei Cheah (student)

Fluxnet Collaboration
- Dennis Baldocchi
- Youngryel Ryu (postdoc)
- Dario Papale (CarboEurope)
- Markus Reichstein (CarboEurope)
- Alan Barr (Fluxnet-Canada)
- Bob Cook
- Dorothea Frank
- Susan Holladay
- Hank Margolis (Fluxnet-Canada)
- Rodrigo Vargas (postdoc)

Ameriflux Collaboration
- Beverly Law
- Tom Boden
- Mattias Falk
- Tara Hudiburg (student)
- Hongyan Luo (postdoc)
- Gretchen Miller (student)
- Lucie Ploude (student)
- Andrew Richardson
- Andrea Scheutz (student)
- Christophe Thomas

National Marine Fisheries
- Charlotte Ambrose
- Maura Moodie
- Charleen Gavette
- William Winner

Sonoma Ecology Center
- Alex Young
- Lisa Micheli
- Deanne DiPietro

http://www.fluxdata.org