Lindsay Barbieri, ESIP Cluster Chair
University of Vermont

Earth Science Data Analytics (ESDA)

Social-Environmental Systems

Drone Cluster

lindsay.barbieri@uvm.edu - @barbieriiv

Definition

DEFINITION: The process of examining, preparing, reducing, and analyzing large amounts of spatial (multi-dimensional), temporal, or spectral data using a variety of data types to uncover patterns, correlations and other information, to better understand our Earth. This encompasses:

---Data Preparation – Preparing heterogeneous data so that they can be jointly analyzed

---Data Reduction – Correcting, ordering and simplifying data in support of analytic objectives

---Data Analysis – Applying techniques/methods to derive results

EARTH SCIENCE DATA ANALYTICS: Goals

1. To calibrate data

2. To validate data (note it does not have to be via data intercomparison)

3. To assess data quality

4. To perform coarse data preparation (e.g., subsetting data, mining data, transforming data, recovering data)

5. To intercompare datasets (i.e., any data intercomparison; Could be used to better define validation/quality)

6. To tease out information from data

7. To glean knowledge from data and information

8. To forecast/predict/model phenomena (i.e., Special kind of conclusion)

9. To derive conclusions (i.e., that do not easily fall into another type)

10. To derive new analytics tools

EARTH SCIENCE DATA ANALYTICS: Techniques

Data Preparation	Data Reduction	Data Analysis
Aggregation	Aggregation	Aggregation
Anomaly Detection	Anomaly Detection	Anomaly Detection
Bias Correction	Bias Correction	Bias Correction
Bayesian Techniques	Bayesian Techniques	Bayesian Techniques
Bivariant Regression	Bivariant Regression	Bivariant Regression
Classification	Classification	Classification
Clustering; Heirarchical Clustering	Clustering; Heirarchical Clustering	Clustering; Heirarchical Clustering
Constrained Variational Analysis	Constrained Variational Analysis	Constrained Variational Analysis
Coordinate Transformation	Coordinate Transformation	Coordinate Transformation
Data Fusion	Data Fusion	Data Fusion

full google sheet:
bit.ly/2i9qvoG

etc..

EARTH SCIENCE DATA ANALYTICS

Many of these techniques could benefit from and/or involve/require semantics

- Data Aggregation
- Data Fusion
- Data Mining
- Graph Analytics
- Filtering
- Machine Learning
- Text Analytics
- Data Classification
- Anomaly Detection

EARTH SCIENCE DATA ANALYTICS

USE CASES

- Data intercomparisons
- Instrument diagnostics

    Requires intercomparison among data from multiple instruments
- Automated model coupling

    Find other model outputs and/or observational data that can be used by my model
- Multi-variate data analysis

    Finding patterns and correlations among heterogeneous variables
- Dataset-tool compatibility

EARTH SCIENCE DATA ANALYTICS

Programs, Courses & Skills

Program Pertaining to Data Science/Data Analytics: Course Topics Most Offered

Overall:
- Statistics, Data Mining, Database Management/Analysis

Data Science, Data Analytics, and Computer Science:
- Data Mining, Mathematics, Statistics, Machine Learning, Data Visualization

Data Science, Data Analytics, and Information Systems:
- Database Management/Analysis

Quantitative Analysis:
- Data Mining, Mathematics, Statistics

Other Relevant Courses Offered:
- Programming, Neural Networks, Data Analysis, GIS, Clustering, Pattern Recognition, Remote Sensing, Artificial Intelligence, Data Warehousing, Text Mining, Time Series, Information/Knowledge Management

But also…

Data Analytics/Data Science Techniques Practiced:

- Mathematics, • Statistics
- Signal Processing
- Machine Learning
- Data Mining, • Database
- Data Engineering
- Pattern Recognition
- Visualization
- Data Warehousing
- Data Compression
- Artificial Intelligence
- High Performance Computing
- Programming
- Classification, • Modeling
- Anomaly Detection
- Data Fusion

Programs offer most of the essential training (blue), but equally essential:
- Knowledge of the discipline that is data analytics/science is being applied to
- Knowledge of data life cycle

Must be interdisciplinary: Every Earth science program should contain training in Data analyitcs/science and Programming (Fox, others)

Social-Environmental Systems

- Challenging environmental and social issues occur at the intersection of complex natural and social systems:

Humans ⟵⟶ Earth
ex: natural disasters, climate change, deforestation
& natural resource planning and use

- These challenges need to be addressed in an integrative & interdisciplinary way to capture <u>biogeophysical AND social</u> dynamics
    - frameworks have been developed such as SESF, Ecosystem Services, etc.

Social-Environmental Systems

Table 3. Purposes of the different frameworks as stated by the authors.

Framework	Purpose	References (selection)
Driver, Pressure, State, Impact, Response (DPSIR)	Develop an improved understanding of, indicators for, and appropriate responses to impacts of human activities on the environment along the causal chain-drivers-pressure-state-impact-responses.	Eurostat 1999, Carr et al. 2007, Svarstad et al. 2008
Earth Systems Analysis (ESA)	Understand the global interactions in and dynamics of the earth system as well as its sustainable evolutions.	Schellnhuber 1998, 1999, Schellnhuber et al. 2005
Ecosystem Services (ES)	Analyze the integral, dynamic, and complex interactions of biotic and abiotic components of an ecosystem in relation to the supply of services this system provides to support life on Earth.	Costanza et al. 1997, Daily 1997, de Groot et al. 2002, Limburg et al. 2002
Human Environment Systems Framework (HES)	Provide a methodological guide or template for analyzing the structure of social-ecological systems and understanding the processes and dynamics between the social and ecological systems as well as within different scales of the social system.	Scholz and Binder 2004, Scholz et al. 2011*a*
Material and Energy Flow Analysis (MEFA)	Analyze the metabolic profiles of societies. Analyze the material and energy flows as representing the metabolism of a society, region, or nation.	Ayres 1978, Baccini and Bader 1996, Haberl et al. 2004, Brunner and Rechberger 2005
Management and Transition Framework (MTF)	Support the understanding of water systems, management regimes, and transition processes toward more adaptive management; enable comparative analyses of a wide range of diverse case studies; and facilitate the development of simulation models based on empirical evidence.	Pahl-Wostl 2009, Knieper et al. 2010, Pahl-Wostl and Kranz 2010
Social-Ecological Systems Framework (SESF)	Provide a common language for case comparison for organizing the many variables relevant in the analysis of SES into a multitier hierarchy that can be unfolded when needed, and for facilitating the selection of variables in a case study.	Ostrom 2007, 2009
Sustainable Livelihood Approach (SLA)	Analyze which combination of livelihood assets enable the following of what combination of livelihood strategies with sustainable outcomes.	Ashley and Carney 1999, Scoones 1998
The Natural Step (TNS)	Provide a framework for planning toward sustainability based on: constitutional principles (how the system is constituted); outcome (principles for sustainability); and process to reach this outcome (principles for sustainable development).	Burns and Katz 1997, Robèrt 2000, Upham 2000, Missimer et al. 2010
Vulnerability Framework (TVUL)	Analyzes who and what are vulnerable to multiple environmental and human changes, and what can be done to reduce these vulnerabilities.	Turner et al. 2003*a*,*b*

Semantics?

- Finding / Sharing Data
- Dataset & Method Integration
- Data Management
- Analysis
- Comparing Frameworks ??

Binder, C. R., J. Hinkel, P. W. G. Bots, and C. Pahl-Wostl. 2013. Comparison of frameworks for analyzing social-ecological systems. *Ecology and Society* **18**(4): 26. http://dx.doi.org/10.5751/ES-05551-180426

ESIP Drone Cluster

Drones or small Unmanned Aerial Systems (sUAS or UAS) or Unmanned Aerial Vehicles (UAV) ...or?

Rapidly becoming important tools for data capture!

- higher temporal and spatial resolutions
- instrument stacking & data linking
- less impact on the environments being monitored
- access to new locations and parameters
- lowered costs and ease of use
- increased human safety during data capture

One drone: Platform for a variety of different sensors (imagery, atmospheric, etc)
Swarm of drones: Many platforms?

Data collected for recreational or industry, could be useful for science

ESIP Drone Cluster

Challenges: relatively new technology, currently no industry-wide accepted best practices for sUAS sensor and flight data handling and management.

Data relationships, metadata, provenance & what's important

ex: Instrument "A" flies on Platform "B"

- (1) The creation of standards would lower the barrier to entry and innovation in terms of what might be monitored with sUAS, by reducing the number of unknowns a new user faces and providing working examples to serve as guides.
- (2) With no common goal standards to build to, the development of mature tools for sUAS captured data processing and fusion (with sUAS and other data sources) is currently hampered. As a consequence, each use case generally develops a unique custom pipeline that only sees one-time use.
- (3) sUAS captured data is - for the most part - not being managed according to data stewardship best practices, such as would ensure the data is FAIR, as articulated by Force11 (Findable, Accessible, Interoperable, and Re-usable).

Research Data Alliance (RDA) Interest Group: sUAS Data → http://bit.ly/2ikGMmu

Lindsay Barbieri, ESIP Cluster Chair
University of Vermont

Drone Cluster

Thursday - 2:00pm-3:30pm - Forest Glen Room

Earth Science Data Analytics

Thursday - 4:00pm-5:30pm - Salon E

Social-Environmental Systems

Friday - 11:00am-12:30pm - Linden Oak

lindsay.barbieri@uvm.edu - @barbieriiv