Understanding Global Change from Data

Vipin Kumar

University of Minnesota

kumar@cs.umn.edu
www.cs.umn.edu/~kumar
Global Change:
A Defining Issue of our Era

What is Global Change?
Global Change: A Defining Issue of our Era

Population Growth & Demographic Shifts
Global Change: A Defining Issue of our Era

Population Growth & Demographic Shifts

Industrialization & Modernization
Global Change: A Defining Issue of our Era

Population Growth & Demographic Shifts

Industrialization & Modernization

Land Use Change

Urbanization

Deforestation

Land Coversion
Global Change: A Defining Issue of our Era

Population Growth & Demographic Shifts

Industrialization & Modernization

Climate Change


1977 1989 2006

Global Surface Warming (°C)

emission scenarios

variability between models

high growth (A2)

moderate growth (A1B)

low growth (B1)

constant CO₂

ARO Workshop on Big Data
Global Change: A Defining Issue of our Era

Population Growth & Demographic Shifts

Biodiversity Loss & Natural Disasters

Climate Change

Monoculture

Ocean Acidification

Destruction of Wetlands

Industrialization & Modernization

Land Use Change

Drought

Cyclones

Fires

6/14/2012
Global Change: A Defining Issue of our Era

THIS IS GLOBAL CHANGE

Population Growth & Demographic Shifts

Industrialization & Modernization

Land Use Change

Climate Change

Biodiversity Loss

Natural Disasters

6/14/2012 ARO Workshop on Big Data
Responding to Societal Needs

- Where is population growth putting pressure on urban infrastructure and natural resources?
- What is the interplay between the global climate system, local ecosystems and natural disasters?
- How does increased biofuel production impact crop patterns and food availability?
- How do changing oceans affect the atmosphere and land climate?
- What are the major feedback mechanisms among eco-climatic processes?
Transformation: Data-Poor to Data-Rich

- Satellite Data
  - Spectral Reflectance
  - Elevation Models
  - Nighttime Lights
  - Aerosols
- Oceanographic Data
  - Temperature
  - Salinity
  - Circulation
- Climate Models
- Reanalysis Data
- River Discharge
- Agricultural Statistics
- Population Data
- Air Quality
- ...

“The future of science depends [...] on cleverness being applied to data for their own sake, complementing scientific hypotheses as a basis for exploring today’s information cornucopia.”

(Nature, September 2008)
Global Change is a Big Data Problem

- Scale and nature of the data offer numerous challenges and opportunities for research in the computational analysis of large datasets.

- Data-driven discovery methods hold great promise for advancing our understanding of the climate and ecosystem processes contributing to global change.

- Advances are of scientific importance and societal relevance.

"data-intensive science [is] so different that it is worth distinguishing [it] ... as a new, fourth paradigm for scientific exploration." – Jim Gray
Active Research Projects

• GOPHER: Global Observatory for Planetary Health and Resources
  Project Aim: Monitoring of global ecosystem for changes in land cover, land use, etc.

• NSF Expeditions: Understanding Climate Change – A Data Driven Approach
  Project Aim: Develop novel data analysis methods to help improve understanding and prediction of climate change
What is the current state of the global forest ecosystems and how are they changing as a result of logging and natural disasters?

How are the demands of a growing population affecting agriculture, e.g., creation of new farmland, changings in cropping patterns, conversion to biofuels, etc.?

How is urbanization affecting the surrounding ecosystem resources and water supply?
Traditional Approach for Change Detection

- Requires high-quality imagery
  - Available infrequently
- Requires high resolution
  - No global coverage
- Requires training data
  - Must be created manually
  - Labor-intensive, time-consuming, expensive

→ Studies are limited to small regions and unable to identify change point or rate of change
Alternate Approach: Spatio-Temporal

Multi-Spectral Data

- Provides global coverage daily
- (Relatively) coarse resolution
- Sometimes poor quality
  - Noisy
  - Missing Data

MODIS instrument on NASA Aqua/Terra Satellites

A **vegetation index** measures the surface “greenness” — proxy for total biomass

Trade-Off

lower spatial vs. higher frequency, resolution

→ opportunities and challenges for spatio-temporal data mining

This vegetation **time series** captures temporal dynamics

6/14/2012

ARO Workshop on Big Data
Time Series Change Detection

This may look easy...
Time Series Change Detection

...but there are two billion time series...and every one is different!
Novel Change Detection Techniques

Current methods are **not adequate** to address these challenges. We focus on developing algorithms that are:

- **Robust** to missing data, noise and outliers
- Able to automatically **characterize** different types of **changes**
- Capable of **incremental** update and (near) **real-time** detection
- Aware of spatial **context**
ALERTS: Automated Land change Evaluation, Reporting and Tracking System

- **Planetary Information System** for interactive investigation of ecosystem disturbances discovered by GOPHER
  - Forest Fires
  - Deforestation
  - Droughts
  - Urbanization
  - ...

- Helps quantify **carbon impact** of changes, understand the relationship between climate variability and human activity

- Provides **ubiquitous web-based access** to changes occurring across the globe, creating public awareness
Global Change Points
Northern Hemisphere Changes
Illustrative Examples

Large **forest fires in Canada** have converted the forests from a sink into source of carbon in the atmosphere.

Logging is legal in some parts of Canada, further reducing carbon sequestration.

**Brazil** Accounts for almost 50% of all humid **tropical forest clearing**, nearly 4 times that of the next highest country.

**Lake Chad** (Nigeria) shrunk by as much as 90% over the past two decades.
Illustrative Examples

Examples of afforestation can be seen in several areas around the world, including this region near Beijing (China) where new trees have been planted to prevent dust storms and erosion.

One winter the Ob River caused massive flooding due to freezing of the Bay of Ob / Kara Sea.

Hurricane Katrina caused significant damage and vegetation loss along the US Gulf Coast.

Impact on REDD+

“The [Peru] government needs to spend more than $100m a year on high-resolution satellite pictures of its billions of trees. But … a computing facility developed by the Planetary Skin Institute (PSI) … might help cut that budget.”

“ALERTS, which was launched at Cancún, uses … data-mining algorithms developed at the University of Minnesota and a lot of computing power … to spot places where land use changed.”

(The Economist 12/16/2010)
Understanding Climate Change: A Data Driven Approach

- 5-year / $10M NSF Expeditions in Computing

- Team led by UMN, consists of 15 senior personnel and ~50 students and post-docs

- Developing state of the art computational methods to address research questions in climate sciences
Understanding of Climate change is Limited

Much of what we know is derived from computer simulations of **general circulation models** (mathematical equations describing the physical processes involved in climate)

**Physics-based models are essential but not adequate**

- Relatively reliable for projections at global scale for smooth fields such as temperature, pressure
- Less reliable for variables that are crucial for impact assessment such as regional precipitation, extremes

“**The sad truth of climate science is that the most crucial information is the least reliable**” (Nature, 2010)
Expeditions Project Highlights

Hurricane Intensity Prediction and Land-Fall Modeling

Climate Extremes and Uncertainty

Teleconnections & Sparse Predictive Modeling

High-Performance Data Analytics
Thank You! Questions?

**UMN Team Members**
Vipin Kumar, Arindam Banerjee, Shyam Borah, Snigdhansu Chatterjee, Jonathan Foley, Joseph Knight, Stefan Liess, Shashi Shekhar, Peter Snyder, Michael Steinbach, Karsten Steinhaeuser

**UMN Students**
Graduate: Ivan Brugere, Yashu Chamber, Xi Chen, James Faghmous, Jaya Kawale, Arjun Kumar, Michael Lau, Varun Mithal; Undergraduate: Kelly Cutler, Ryan Haasken, Zachary O’Connor, Dominick Ormsby

**PSI Collaborators**
NASA Ames: Christopher Potter
Cal State Monterey Bay: Stephen Klooster
Michigan State: Pang-Ning Tan
PSI: Juan Carlos Castilla-Rubio

website: gopher.cs.umn.edu

**NSF Expeditions Collaborators**
NCSU: Nagiza Samatova, Fredrick Semazzi
Northeastern: Auroop Ganguly
Northwestern: Alok Choudhary, Wei-keng Liao
North Carolina A&T: Abdollah Homaifar

website: climatechange.cs.umn.edu

**Acknowledgements**

**Contributors**

Grant IIS-1029711
ARO Workshop on Big Data

6/14/2012