NASA Earth Exchange (NEX)

A collaborative supercomputing environment for global change science

Earth Science Division/NASA Advanced Supercomputing (NAS)
Ames Research Center
http://www.nex.arc.nasa.gov
Rama Nemani (rama.nemani@nasa.gov)

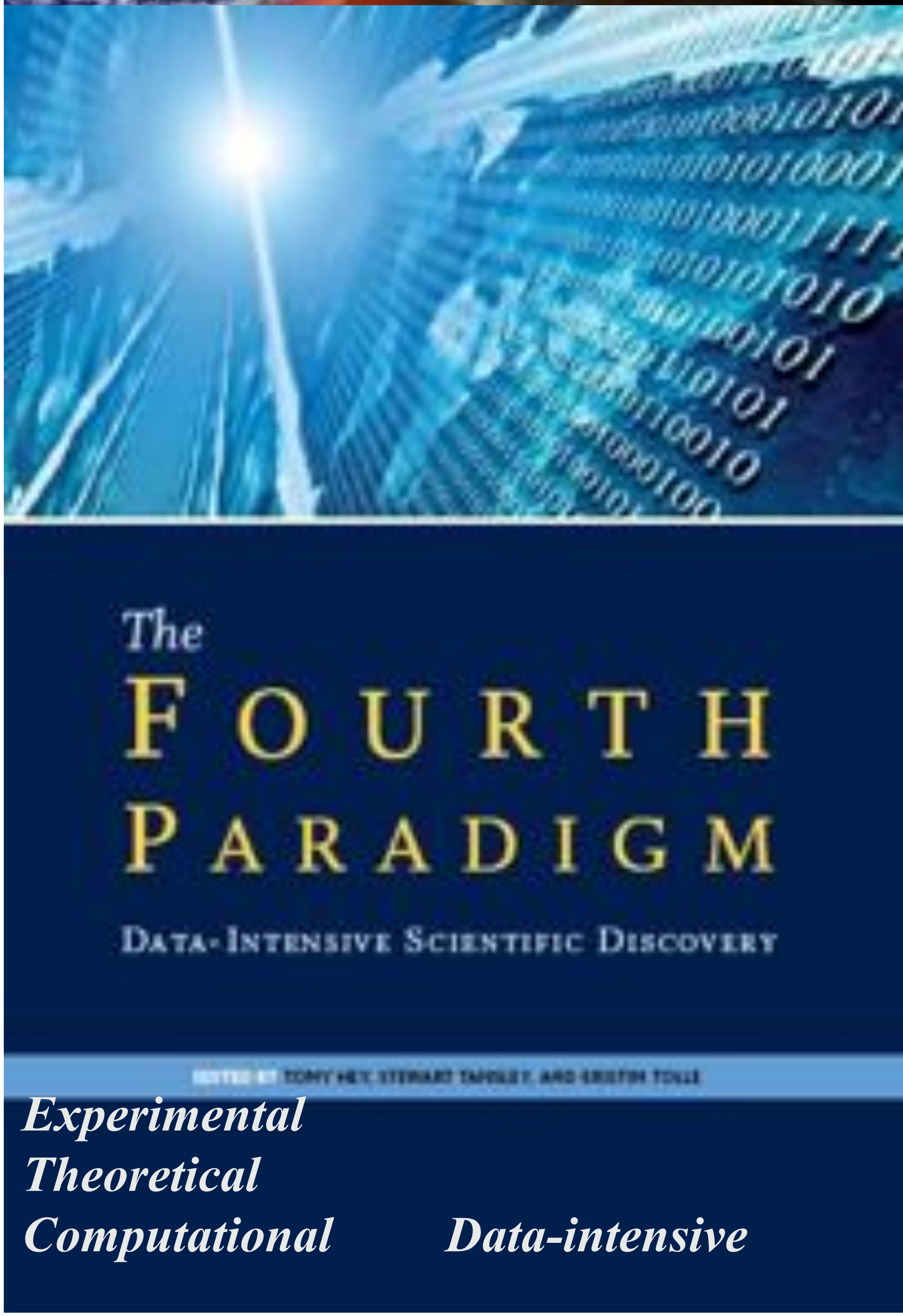
*Providing **direct** access to data, models, analysis tools, and scientific results through a supercomputing platform that fosters knowledge sharing, collaboration, and innovation.*

MODIS Science Team meeting, May 20, 2011

Outline for the presentation

- **Motivation**
- **Concept/Components**
- **Status**
- **Current projects**
  - *Modeling*
  - *Data Analysis*
  - *Applied science*
  - *National Climate Assessment*
- **Summary**

Inspiration



Experimental
Theoretical
Computational
Data-intensive

James Gray

Motivation

We spend too much time dealing with data, not enough on analysis
- 80/20 to 20/80

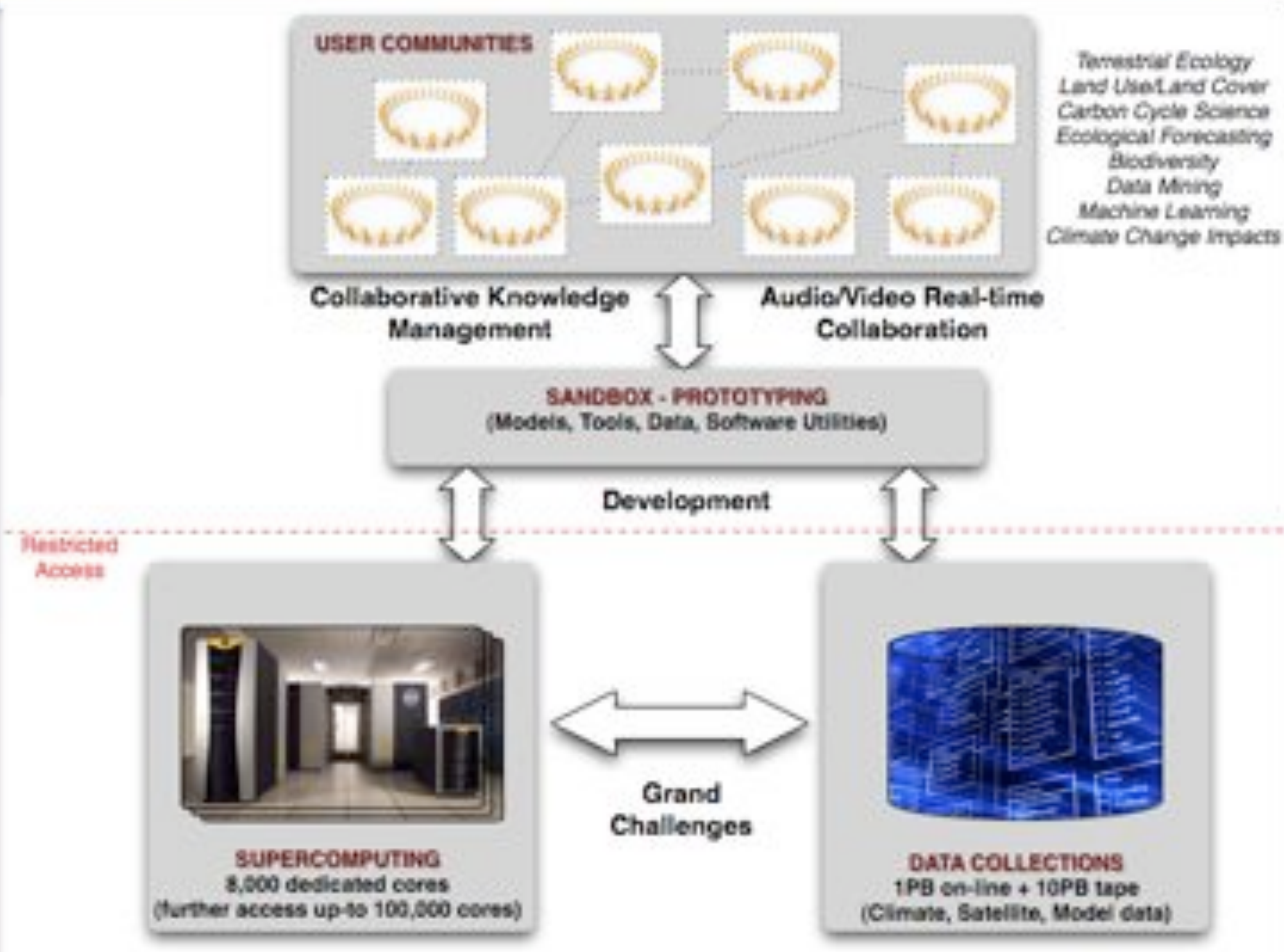
Moving large data sets through network is not easy
- Move the code to data

One scientist/one project paradigm is old for multi-disciplinary-era
- Collaborate

Repeated low-level IT efforts waste time/resources
- reuse and share software

Need mechanisms that allow transparency/repeatability
- Capture workflows

NASA Earth Exchange components



Nemani et al., 2011, Trans. of Eos

NASA Ames Supercomputing Overview

HECC Project conducts work in four major technical areas

Supercomputing Systems

Provide computational power, mass storage, and user-friendly runtime environment through continuous development and deployment of management tools, IT security, systems engineering

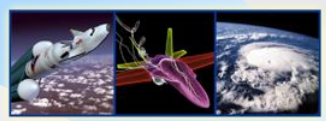
Application Performance and Productivity

Facilitate advances in science and engineering for NASA programs by enhancing user productivity and code performance of high-end computing applications of interest

HECC

Data Analysis and Visualization

Create functional data analysis and visualization software to enhance engineering decision support and scientific discovery by incorporating advanced visualization technologies and graphics displays



Networking

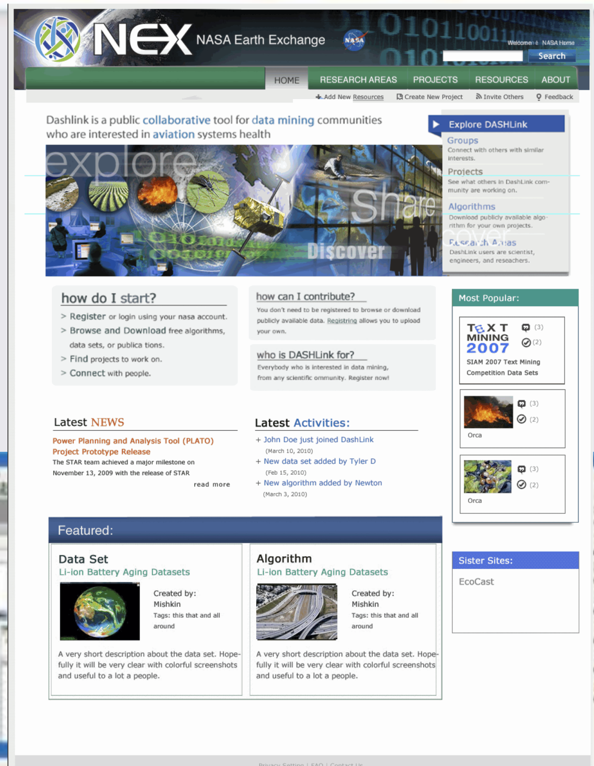
Provide end-to-end high-performance networking to meet massive modeling and simulation data distribution and access requirements of geographically dispersed users

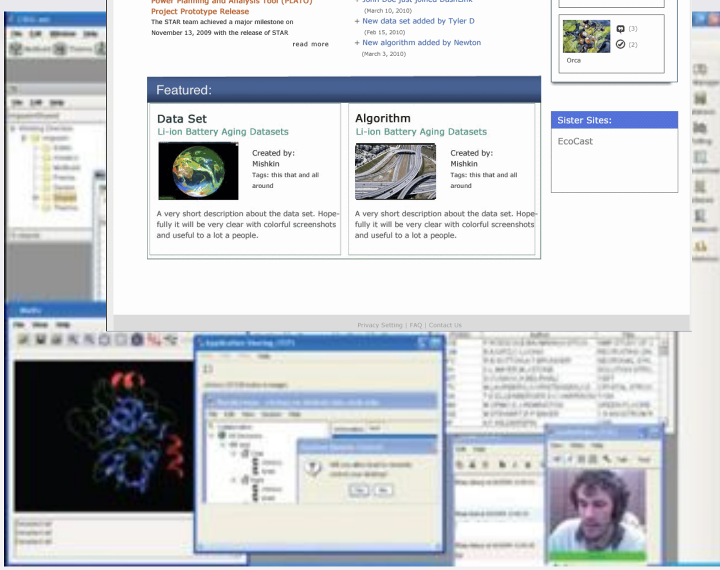
Supporting Tasks

- Facility, Plant Engineering, and Operations: Necessary engineering and facility support to ensure the safety of HECC assets and staff
- Information Technology Security: Provide management, operation, monitoring, and safeguards to protect information and IT assets
- User Services: Account management and reporting, system monitoring and operations, first-tier 24x7 support

Collaboration Tools

Share





Conference

Collaborate

Leveraging collaborative research tools at NASA

Visualize


Global TOPS 1KM Project Meeting | Connect Pro Meeting
Process Design and Knowledge Capture
Process Design
Iterative Refinement/
Distributed Collaboration
Full-scale Execution
On-Line Result Sharing and
Process Re-use
Petr Votava
My Status:Active
GLOBAL_TOPS_1KM.mmap
NASA

Example Usage

- User registers and specifies resource requirements: data, tools, models and computing resources
- A temporary environment is created for the user containing the requested resources
  - Within this environment user can:
    - Allocate CPUs and disk space
    - Run existing models
    - Bring in new data, models, and algorithms
    - Extend existing models
    - Share models and data with community
    - Provide access to the results and ***environment***
- When work is completed, the resources are recycled, but the knowledge is captured
  - The specific environment can be saved and reused

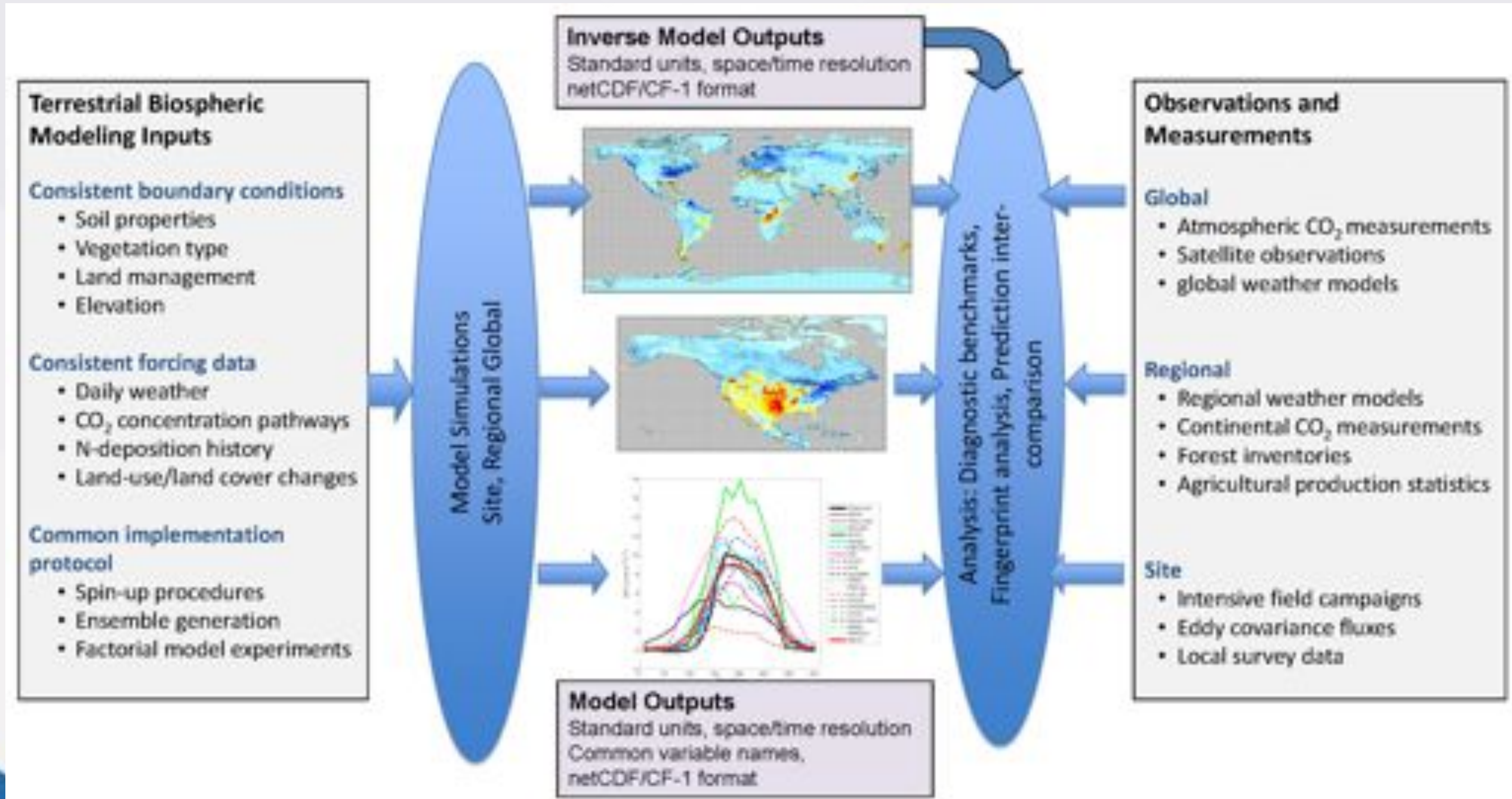
NEX Status

- **Computing**
  - Supercomputing and storage through NASA Advanced Supercomputing Division (NAS)
  - Almost 3PB storage, 80,000 CPU cores
- **Data**
  - Global MODIS, AVHRR, Landsat, GCM Scenarios, weather data, etc.
- **Models**
  - Publicly available models (ecosystem, weather, climate, hydrology)
- **Software Utilities**
  - Open source and commercial

Community (154 Members, 22 Active Research Teams)

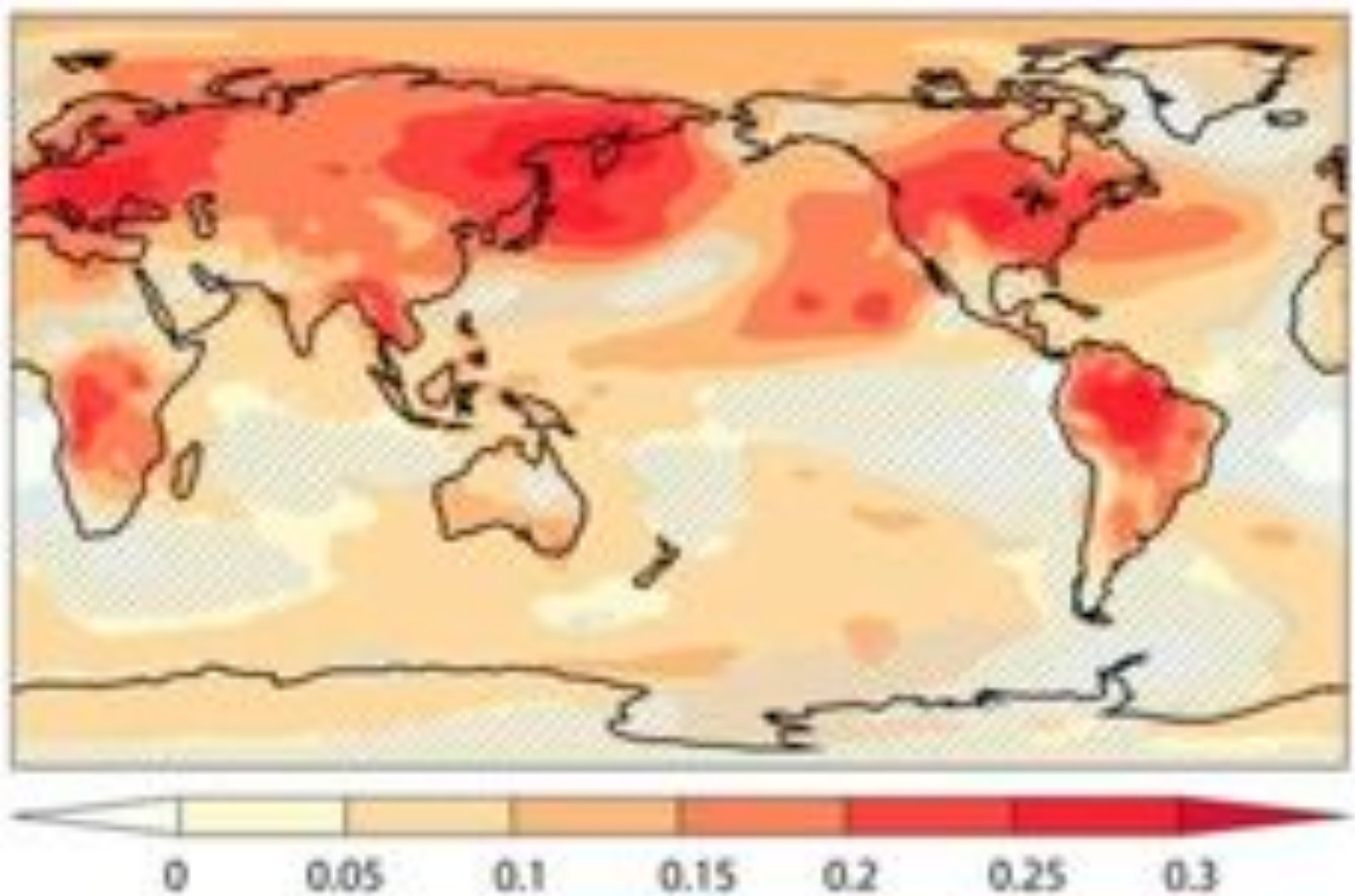
Community modeling (Carbon)

Multi-Scale Synthesis and Terrestrial Model Intercomparison (MsTMIP)



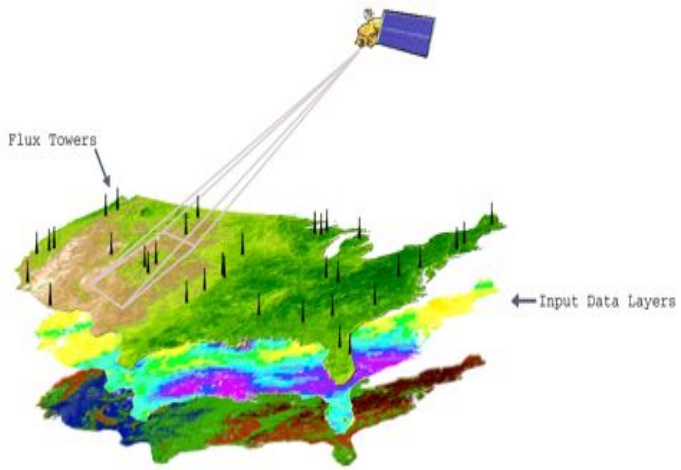
Hosting inputs, model codes, and analysis

Climate modeling: Warming from CO2-induced physiology



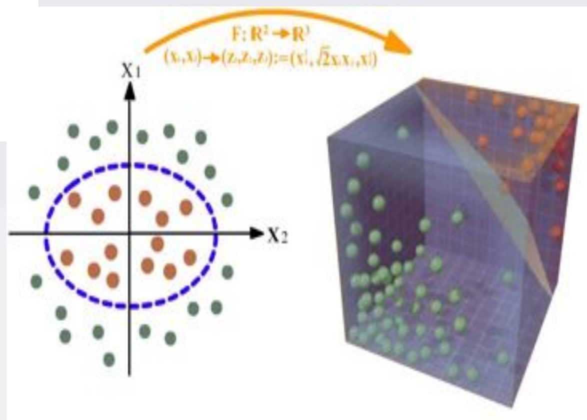
Cao et al., PNAS 2010, CCSM/CAM, workflow available on NEX

Data-driven modeling

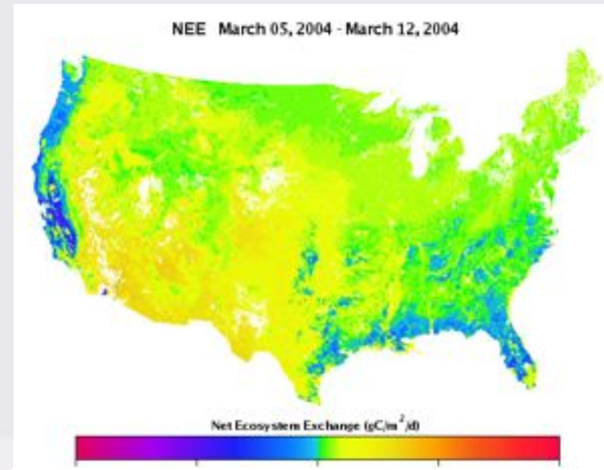


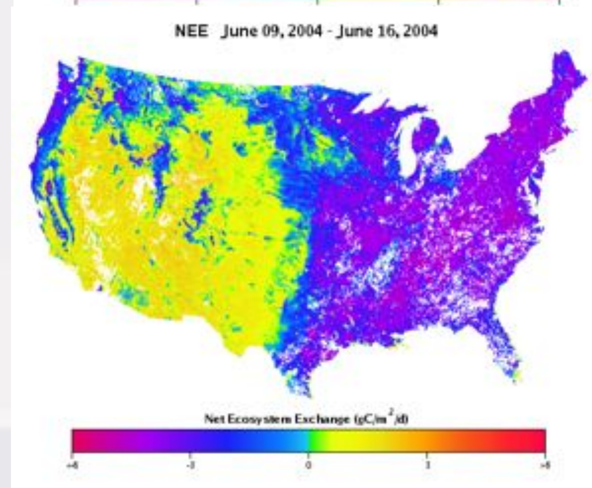
Fluxnet, MODIS, GOES

SVM/ANN



Net Ecosystem Exchange

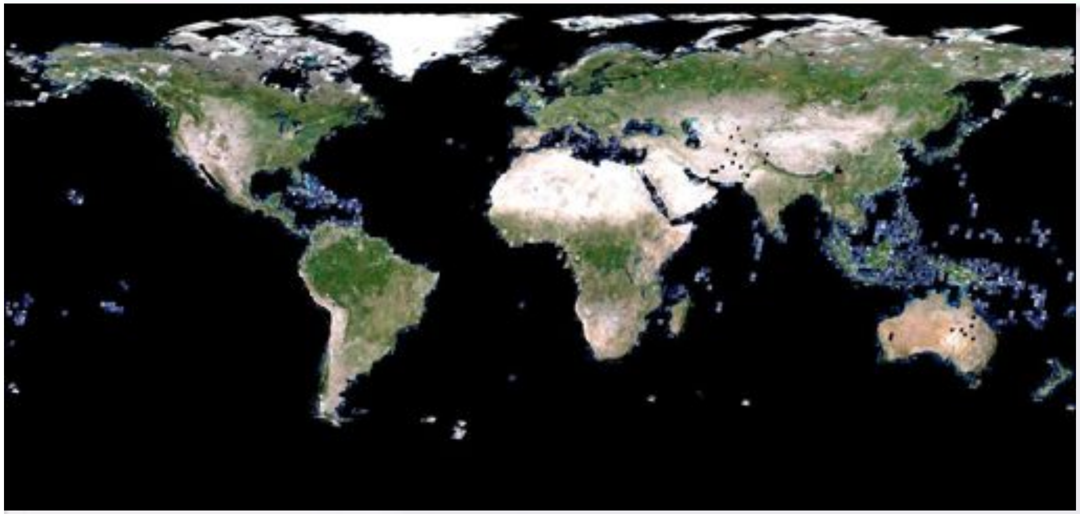




Blending time-continuous observations from Fluxnet with spatially continuous data from satellites globally.

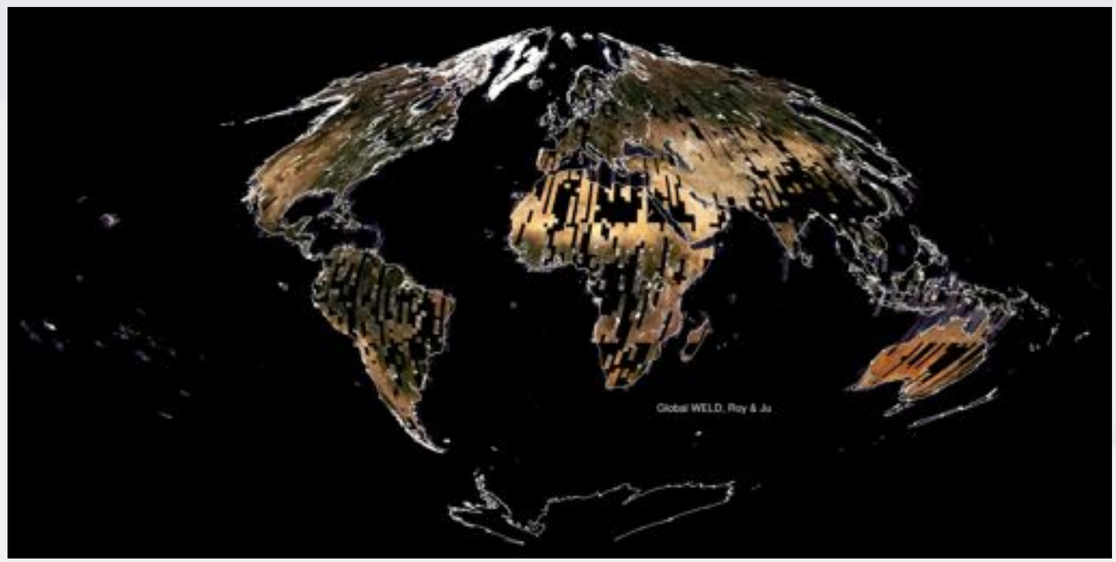
Yang et al., 2006 (RSE), 2007 (TGRS), 2008 (JGR)

Understanding global landscapes at high resolution
Modisizing the Landsat project



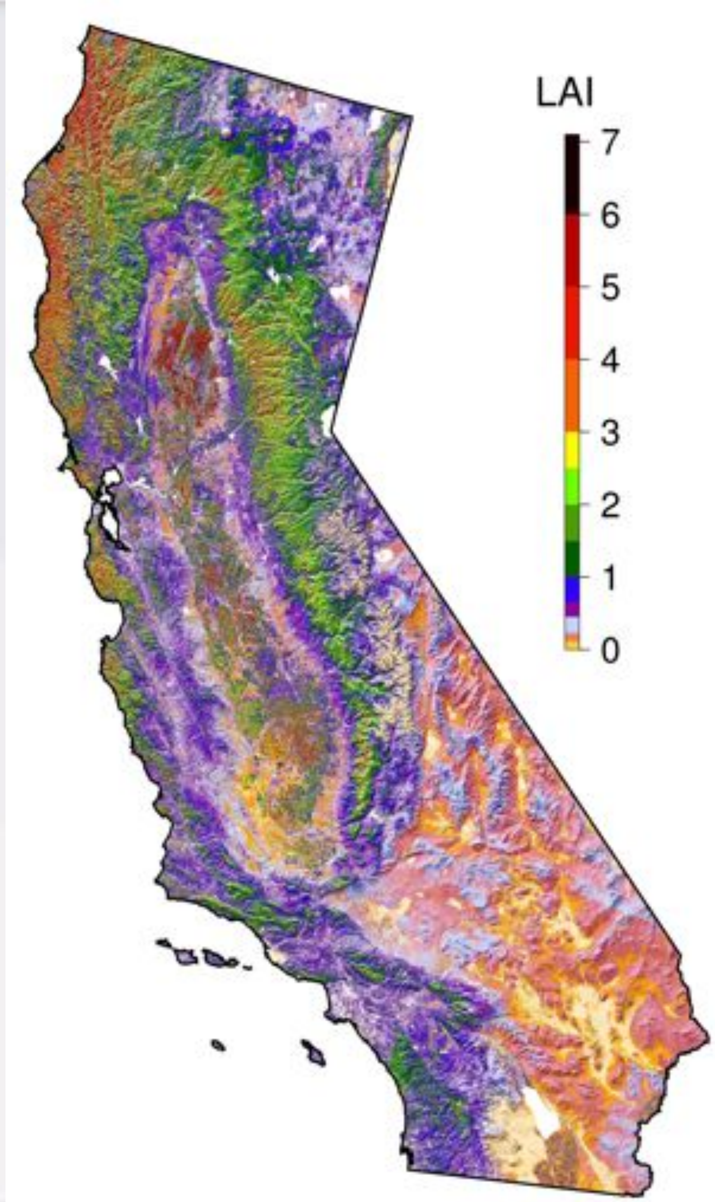
For the first time in Landsat history of nearly 30 years, we can now process and create quantitative information about changes in global landscapes, **in a matter of hours.**

Building virtual teams
- Boston U.
- Columbia U.
- U of Maryland
- South Dakota State U
- Oregon State U



For monitoring crop growth, deforestation, and impacts of natural disasters

Biophysical products from Landsat - Leaf Area Index

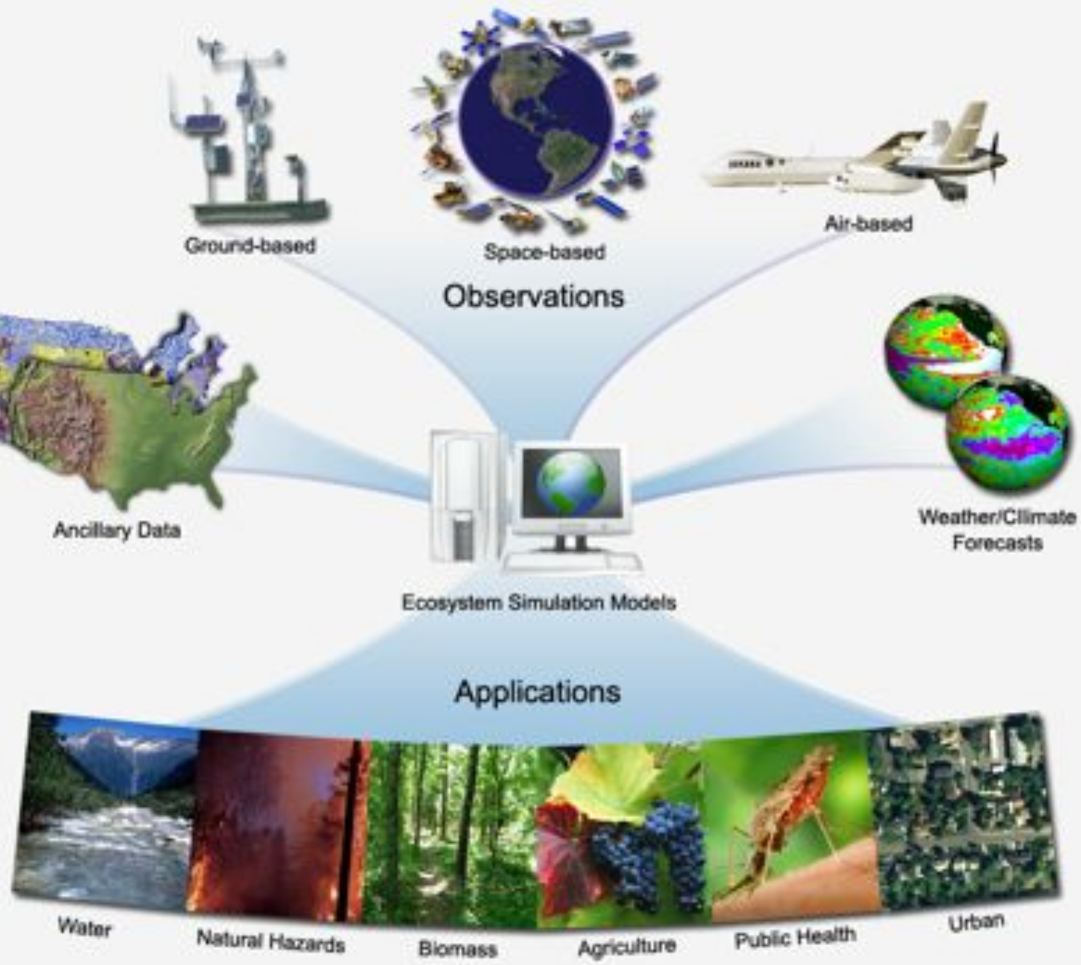


Ecological Forecasting

predicting the effects of changes in physical, chemical, and biological environment on ecosystem state and function

Terrestrial Observation and prediction system (TOPS)

A data - modeling system for integrating satellite, surface data with simulation models to produce ecological nowcasts and forecasts



<u>Key elements</u>:

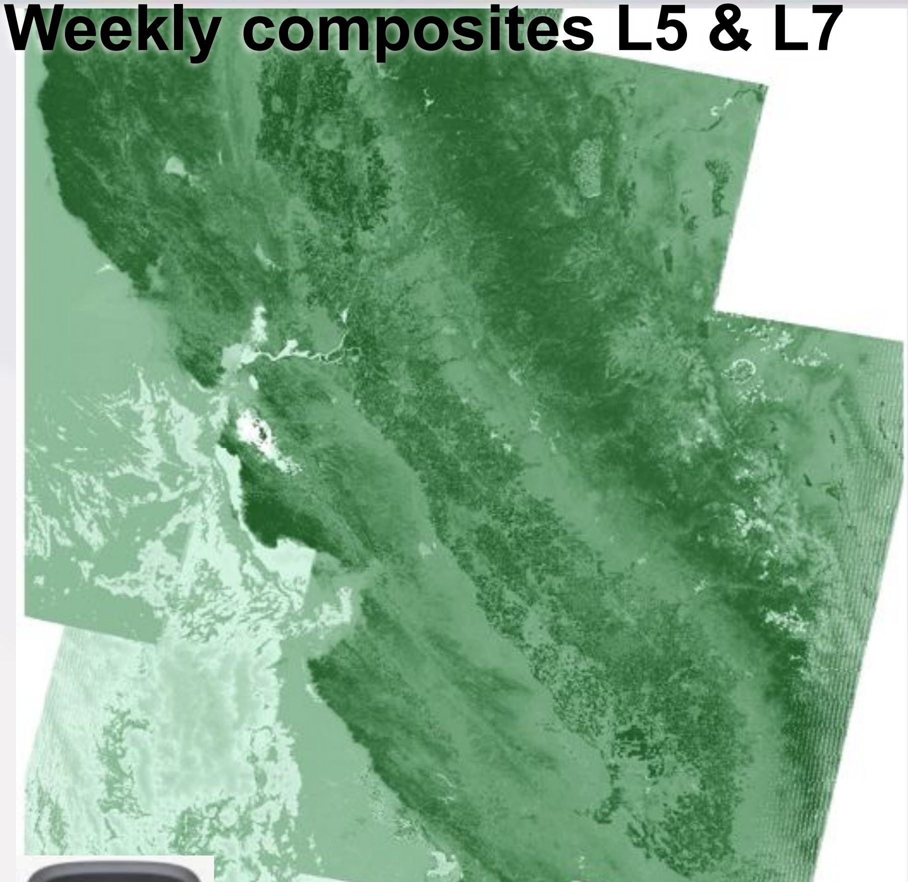
- Monitoring
- Modeling
- Forecasting
- Local to Global

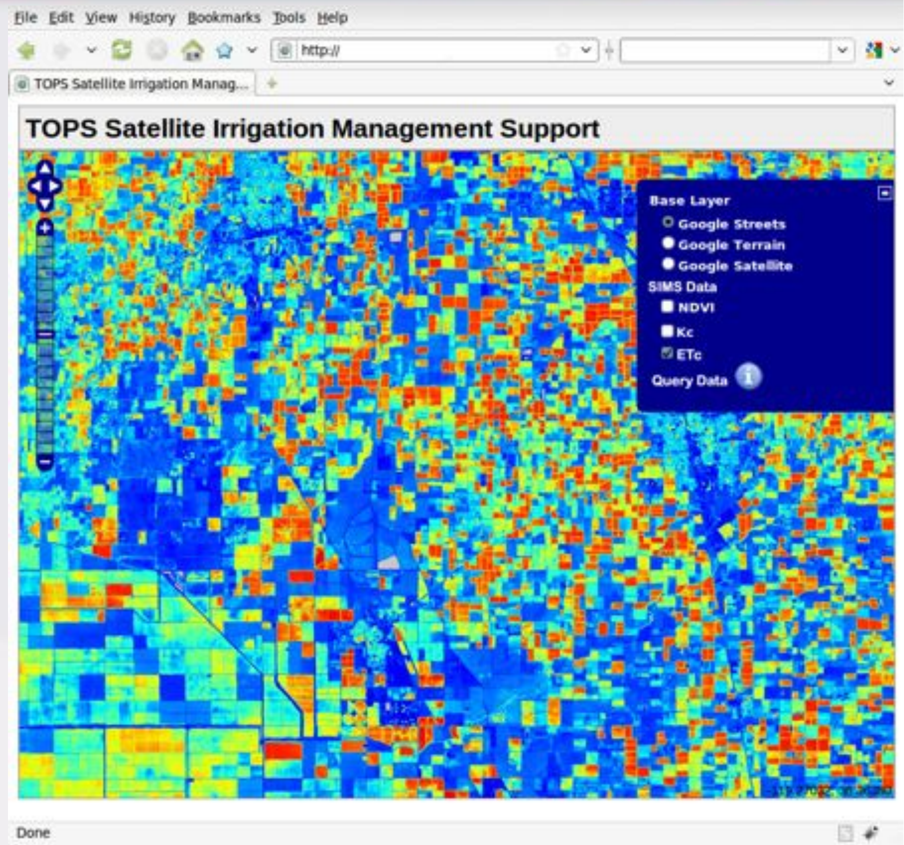
Focus on biogeo-chemical cycles

Nemani et al., 2009, RSE

Crop water management

Weekly composites L5 & L7



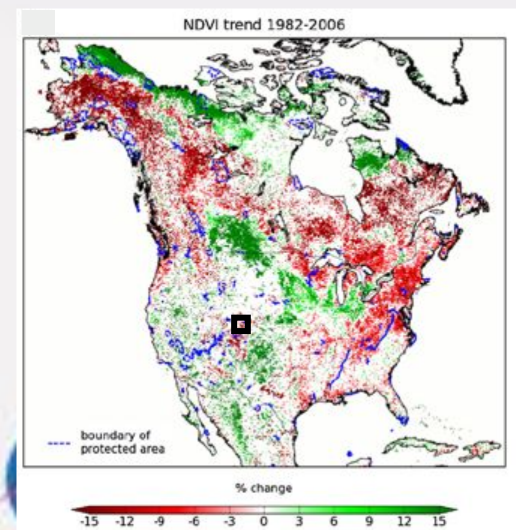


Irrigation reports

Past 3 days:								Next 3 days:			
Field	ETo	Kcb mean	Kcb max	Kcb min	ETcb	SWB	Runtime 8/8/2010	ETo	ETcb	SWB	Runtime 8/11/2010
F-01	0.83	1.05	1.09	0.95	0.87	+0.30	--	0.79	0.83	-0.53	2.5
F-02	0.83	0.97	1.04	0.91	0.81	-0.12	1	0.79	0.77	-0.77	4
F-03	0.83	1.09	1.12	1.02	0.90	+1.14	--	0.79	0.86	0.28	--
...											

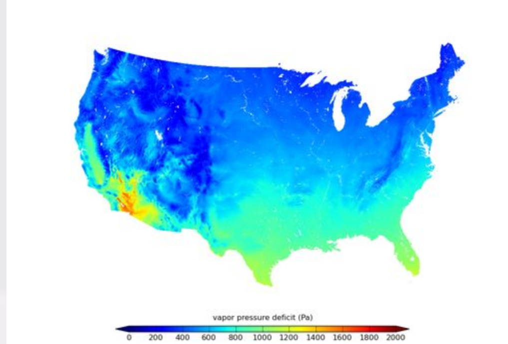
NEX in support of the National Climate Assessment

Satellite monitoring

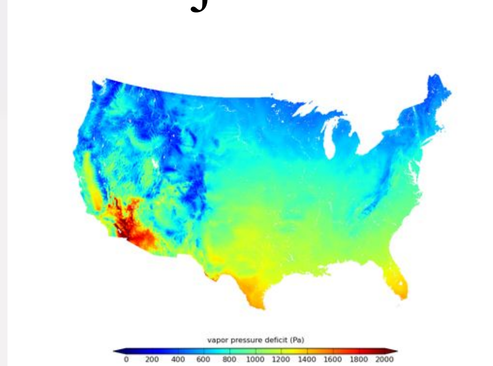


High-Resolution (1km) climate data

Historical

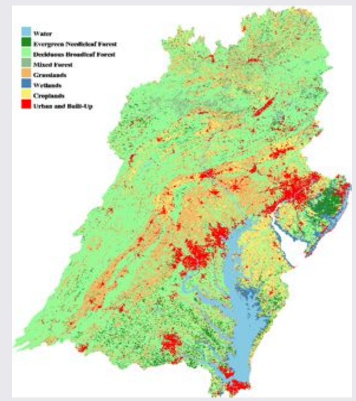


Projected

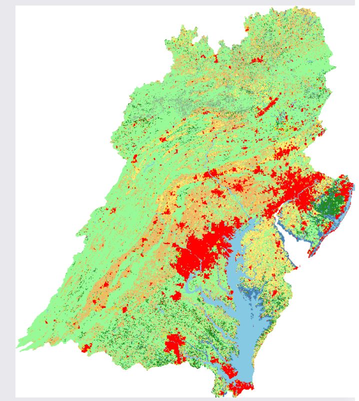


10+ teams working

Modeling Land cover changes

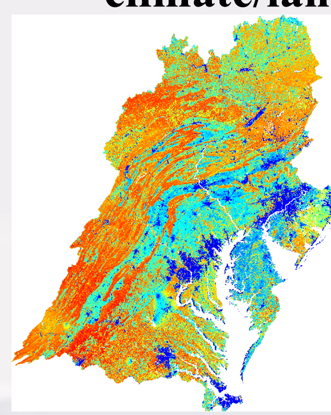


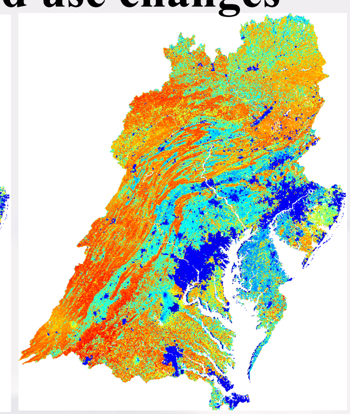
Baseline LC (2000)



Forecast LC (2030)

Vegetation responses to climate/land use changes

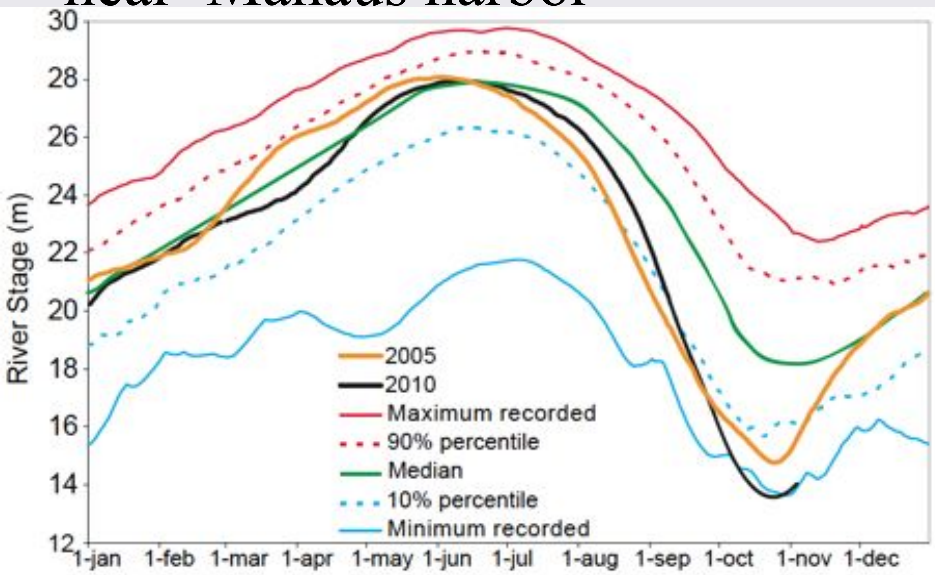




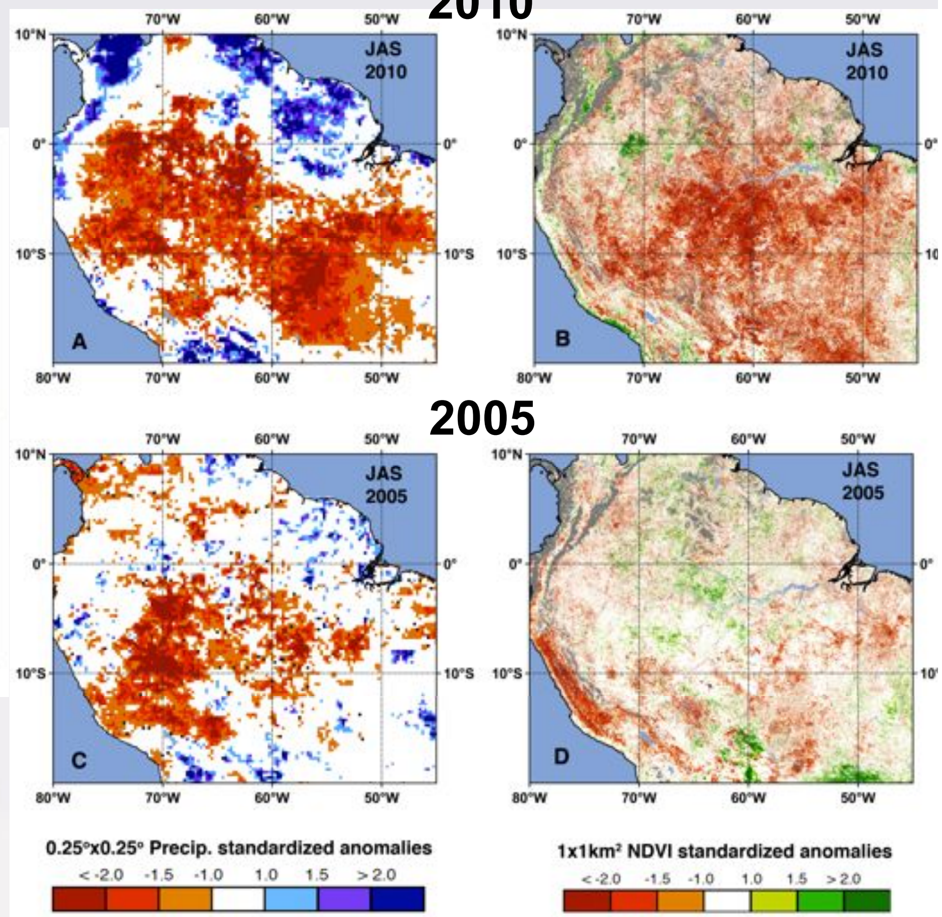
Leveraging past work: using workflows

Drought of the Century (2005, 2010)

River stages of the Rio Negro near Manaus harbor



Satellite observations



TRMM MODIS

Analyzed 2010 MODIS/TRMM data till October 15, written and submitted a publication by Jan 15, 2011, published March 29th 2011.

Samantha et al., 2010 and Xu et al., 2011, Geophys Res. Letters

Summary

- Create a knowledge network starting with all the ES funded projects.

- Engage a larger community by lowering the barrier of entry by co-locating data, model codes, and compute resources.

- Capture research through workflows and virtual machines to create the ability to build on past work, and accelerate future scientific endeavors.

- Foster inter-disciplinary efforts.

- THINK BIG!