

Global Data Bases for Integrated Assessment Modeling

Presented by Thomas Hertel
Center for Global Trade Analysis
Purdue University

Presentation for the 15th Annual Snowmass Workshop on IAMs, July 24, 2012

Outline

- **GTAP: The core data base for economy-wide analysis of global economic issues**
 - A bit of history
 - Current state of play (version 8.0)
 - Satellite data bases for analysis of land use and GHG emissions
- **GEOSHARE: Geospatial data for analysis of global agriculture and environment issues (Might GEOSHARE be GTAP **2?)**
 - Motivation for the project
 - GTAP as inspiration for GEOSHARE
 - Design: GEOSHARE as a network of research nodes
 - Current state of play: Pilot effort launched July 1

GTAP: The Global Trade Analysis Project

Global Trade Analysis before GTAP (1)

- **Agency-based *global* economic models:**
 - Dominated by agencies in OECD countries and international institutions in Paris, Rome and Washington
 - Early 1990's evaluation of the WTO-Uruguay Round's *impact on developing countries* conducted almost exclusively by economists *based in developed countries*
 - Data base treated as strategic asset, rarely shared with outsiders
 - Unfavorable findings typically resulted in project termination
 - Methodology often poorly documented; rarely peer reviewed

Global Trade Analysis before GTAP (2)

- **Agency-based *global* economic models:**
 - Dominated by developed countries
 - Data base treated as strategic asset
- **University-based *global* economics projects:**
 - Insufficient contact with policy makers (with a few exceptions)
 - Limited financial support – cross-subsidized from other grants
 - Little incentive to invest in data base; focus instead on “toy” models
 - When do invest, keep model and data until rents and pubs extracted
 - Impossible to maintain over the long haul

GTAP as Institutional Innovation

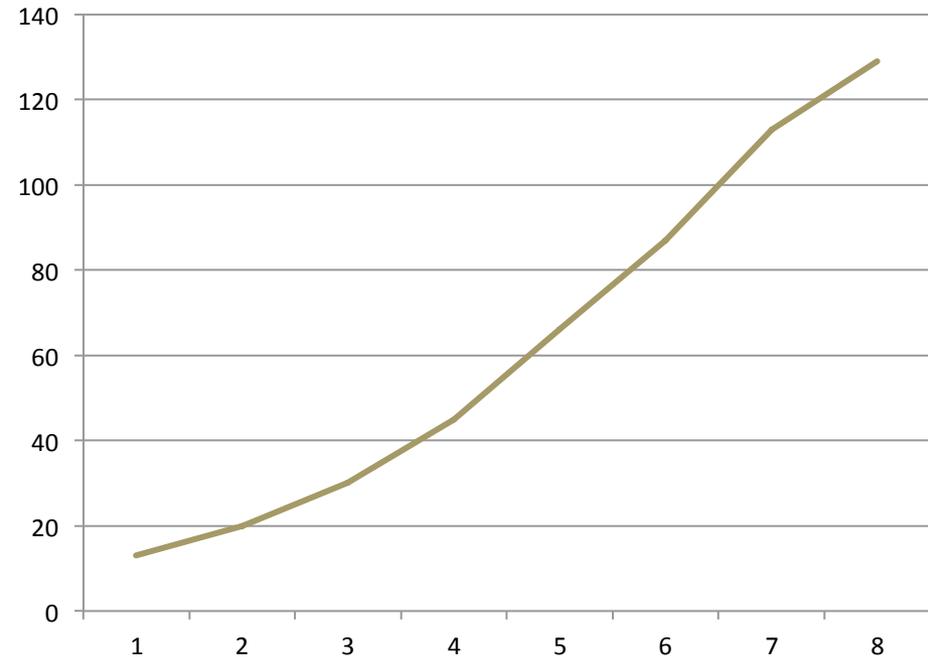
- Publicly funded project, based in academia:
 - Core support from 28 national/international agencies
 - Database is sold to public; free for contributors
 - Contributors and consortium obtain early access
- Advantages:
 - Academic base ensures continuity and encourages openness (vs. control often exerted by gov't institutions)
 - Consortium ensures policy relevance; access to current data and ongoing support
 - Individuals work to their comparative advantage; capitalize on network externalities
 - Has proven to be an effective means of spreading the use of analytical data base and modeling tools to developing countries

GTAP as a Database

V.8 includes: Updated data for 124 economies covers 99% global GDP

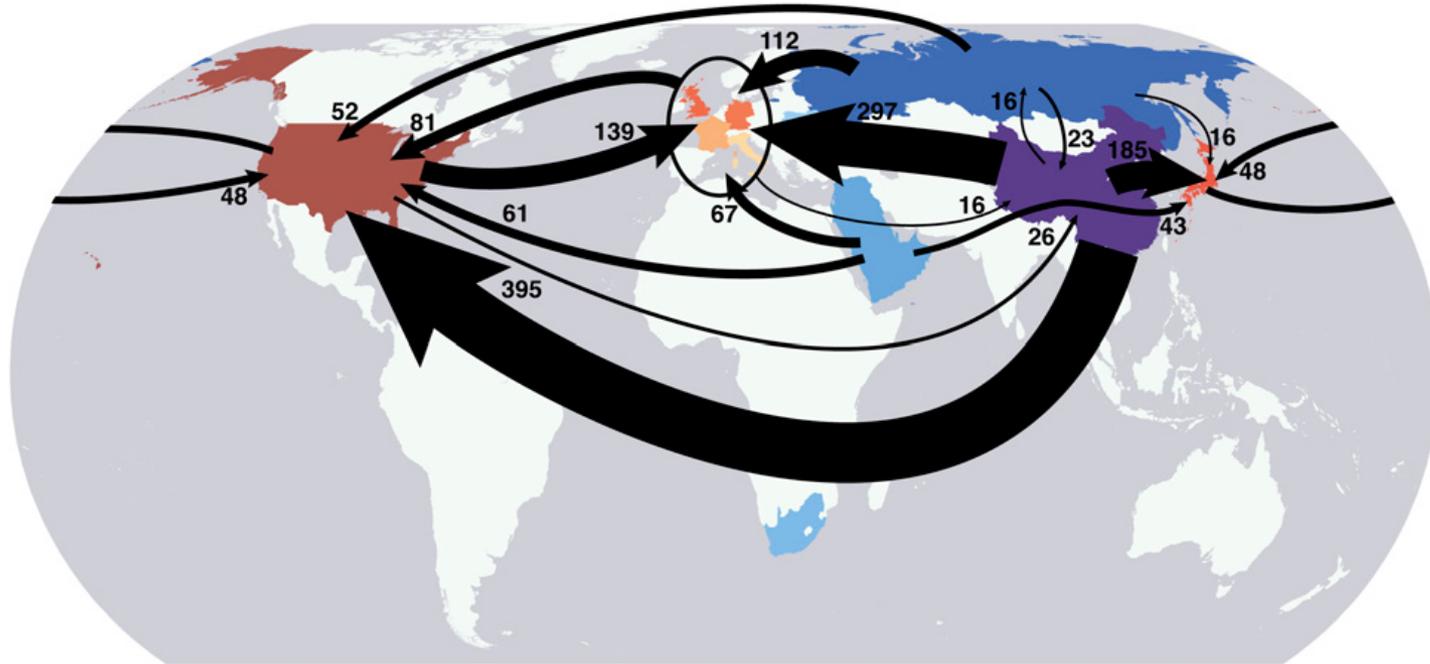
Critical role for reconciliation just one GTAP version which is widely accepted for use in policy analysis

Satellite data bases to deal with energy, land use, GHG emissions, migration, international capital ownership



Growth in number of regions:
GTAP v.1 – v.8

GTAP is increasingly linked with biophysical data bases



International fluxes of emissions embodied in trade,
Calculated from the GTAP data base (Mt CO₂ y⁻¹)
Davis and Caldeira, PNAS, 2010

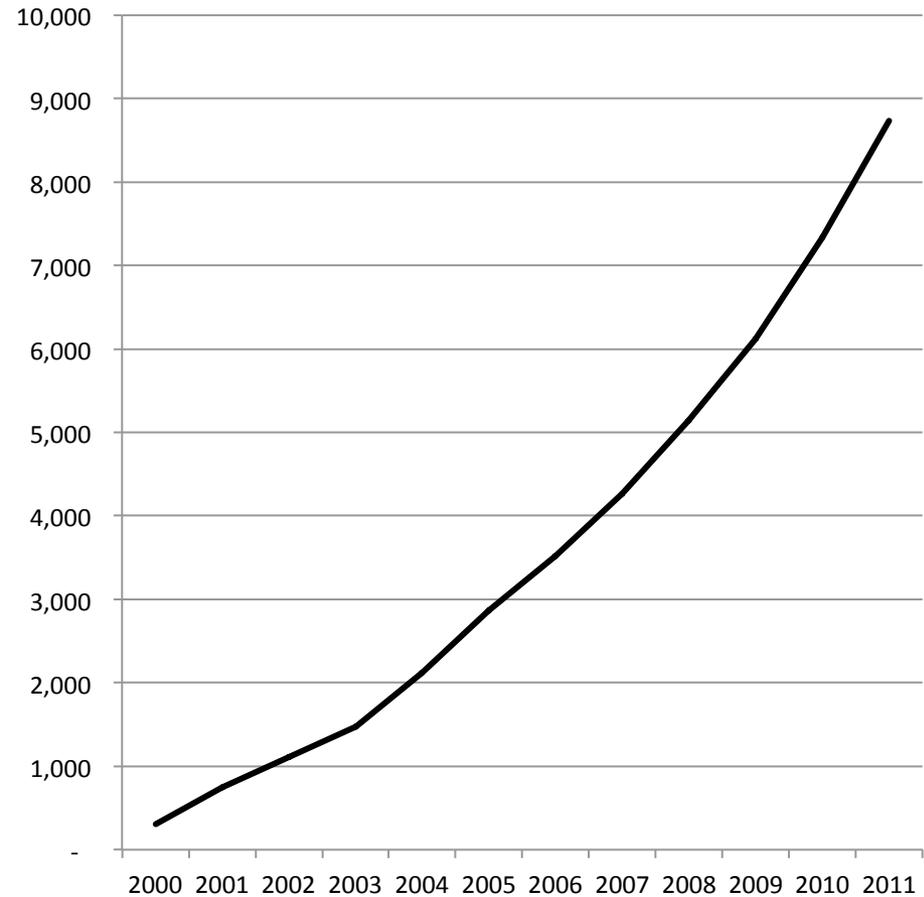
GTAP as a Network

Nearly 9,000 members from 159 countries

Demand factors: convergence of interest in a few global issues

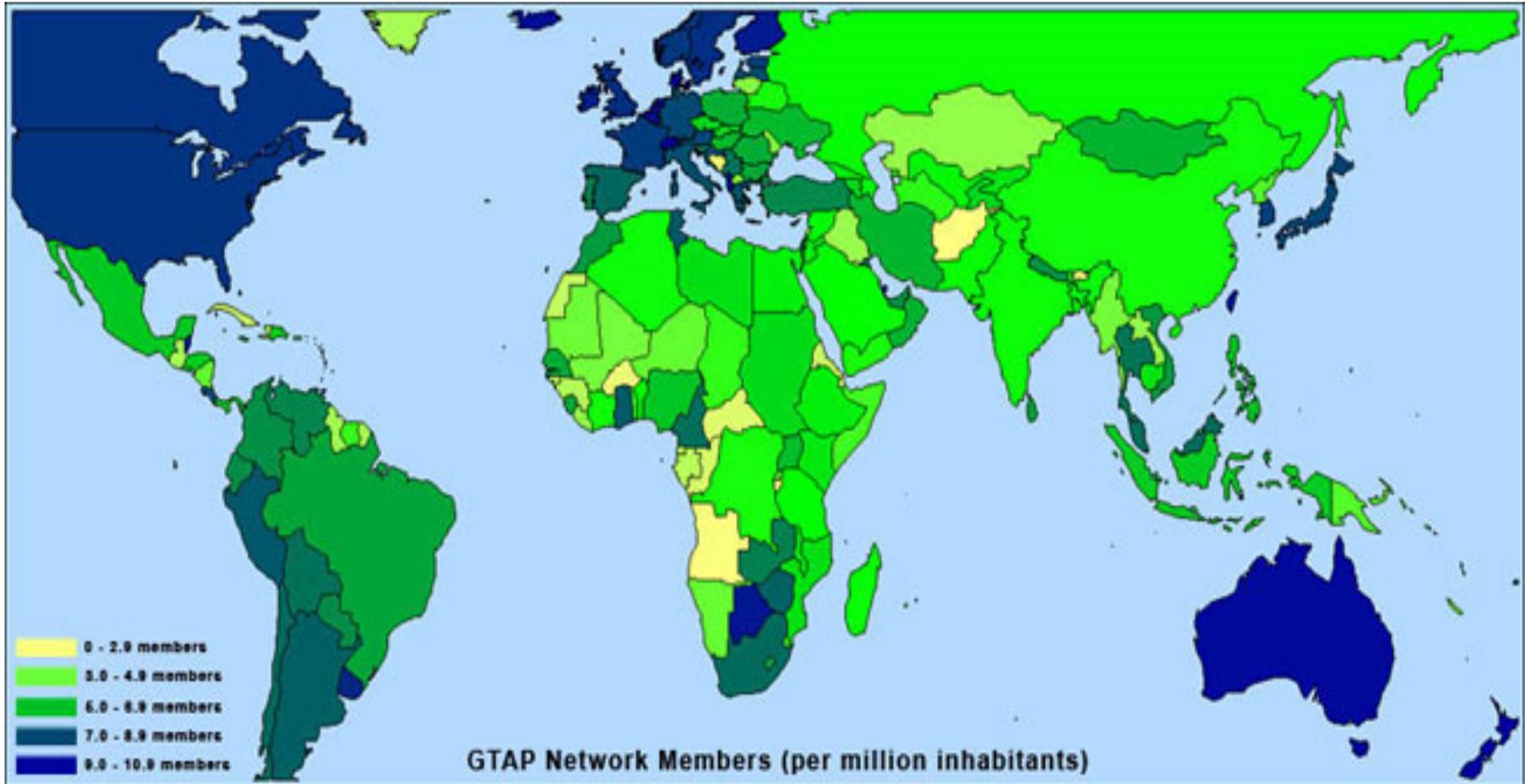
Supply factors: dozens of courses, 15 annual conferences; partnerships with UN in Africa, Asia and Latin America

Open network: half million page views/yr on web site



Growth in network membership

Graphical Distribution of GTAP Network



GEOSHARE:

Geospatial, Open-Source Hosting of
Agriculture, Resources and
Environmental Data

Q: Can we repeat the success of
GTAP
with the geospatial community?

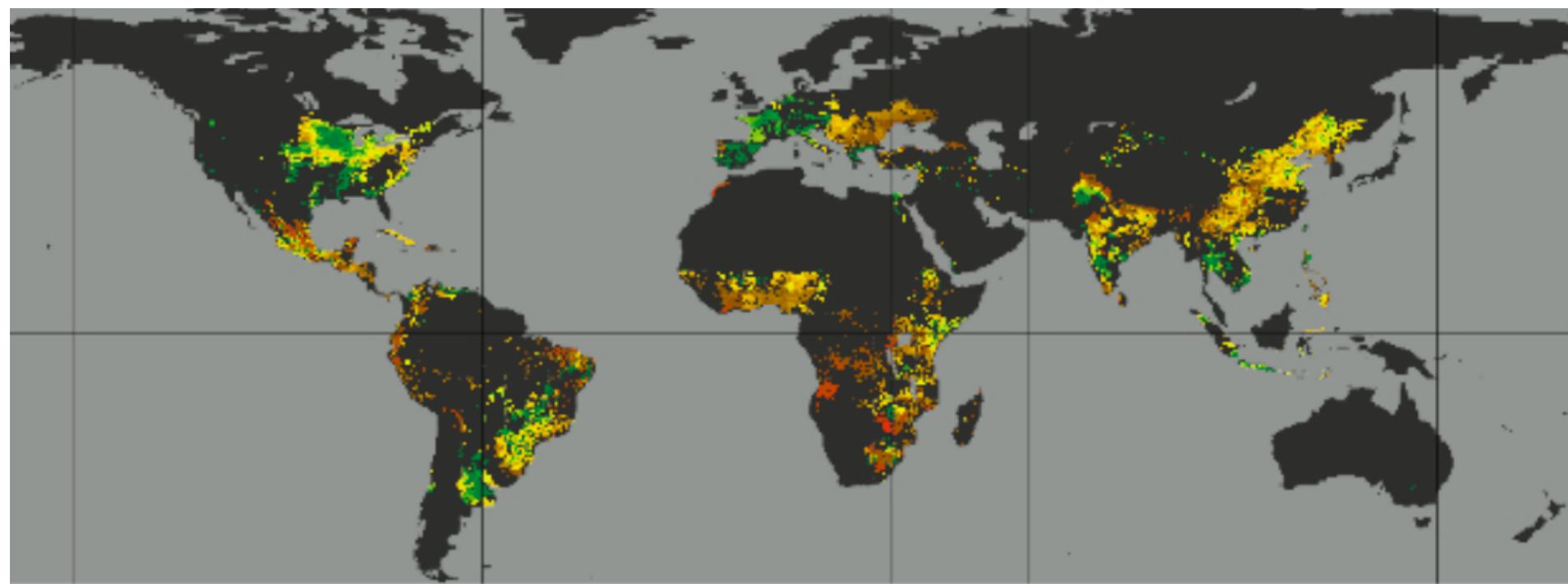
GEOSHARE Objectives

- *Provide a globally consistent, temporally opportune, and locally relevant database* for better decision making.
- *Assist decision makers, policy analysts and researchers seeking to use geospatial data and analysis tools to inform activities relating to agriculture, poverty, land use and the environment.*
- *Build capacity* throughout the world in individuals who can effectively bridge disciplines to make decisions and to identify solutions to complex resource use and development problems using geo-spatial data and analysis tools.

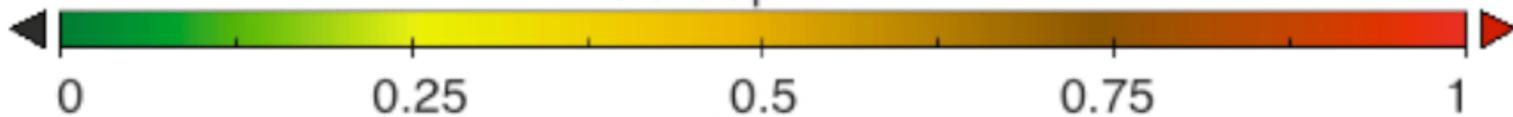
Motivation for GEOSHARE

- Feeding 9 billion people in 2050 in the face of a changing climate, while preserving the environment and eliminating extreme poverty, is one of the most important challenges facing us today.
- Agriculture is at the heart of this challenge:
 - Farming/land use change account for quarter of global GHG emissions
 - Land-based activities are very sensitive to climate change
 - Agriculture remains the predominant source of income for the world's poorest households and is therefore central to poverty reduction
- Key questions facing decision makers all require time series, spatial data:
 - Potential for boosting yields to meet projected growth in global demand
 - Optimal use of REDD+ funds to limit deforestation and sequester carbon
 - Impact of water scarcity on agricultural output
 - Impact of climate change on global agricultural productivity
 - Impacts of increasing climate extremes on vulnerability of the worlds' poor

Spatial detail is key in identifying yield gaps for crops (e.g., maize circa 2000)



Yield Gap Fraction

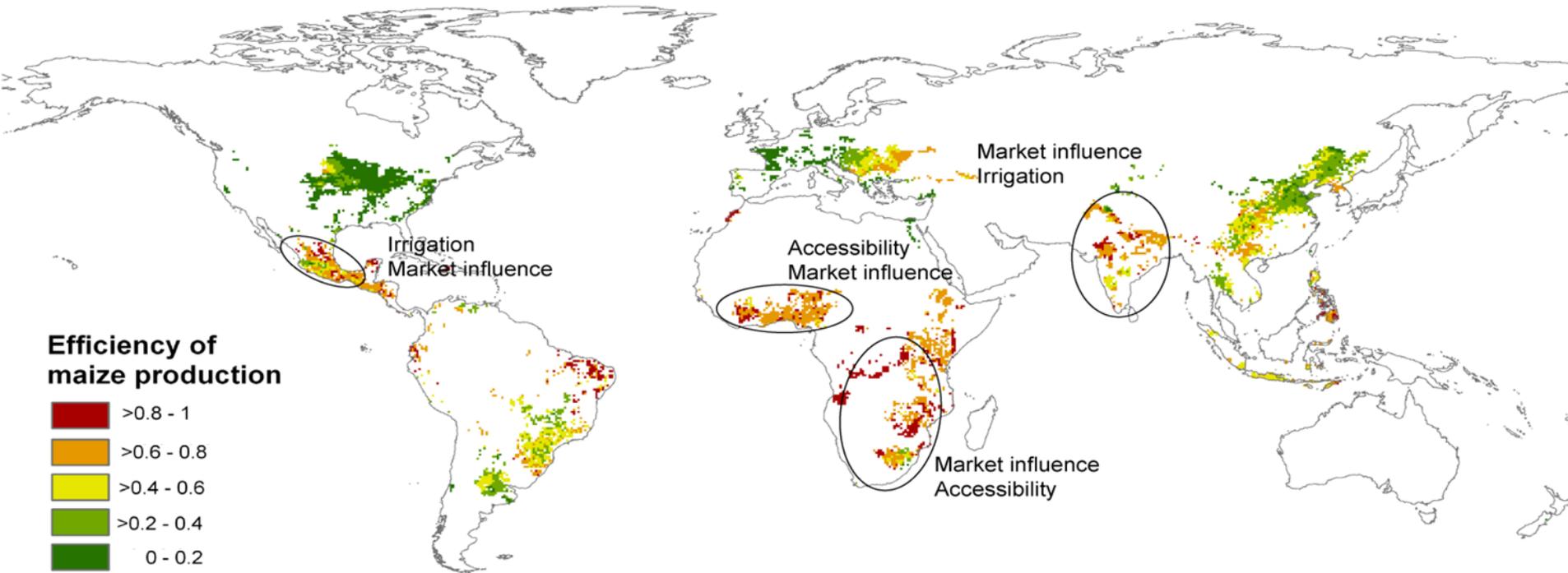


$$\text{Gap} = (1 - \text{Actual yield} / \text{Climatic potential yield})$$

So 0 = on the production frontier, 1 = no productivity

Source: Licker et al. (2010)

As well as explaining their causes: Factors affecting maize production inefficiency

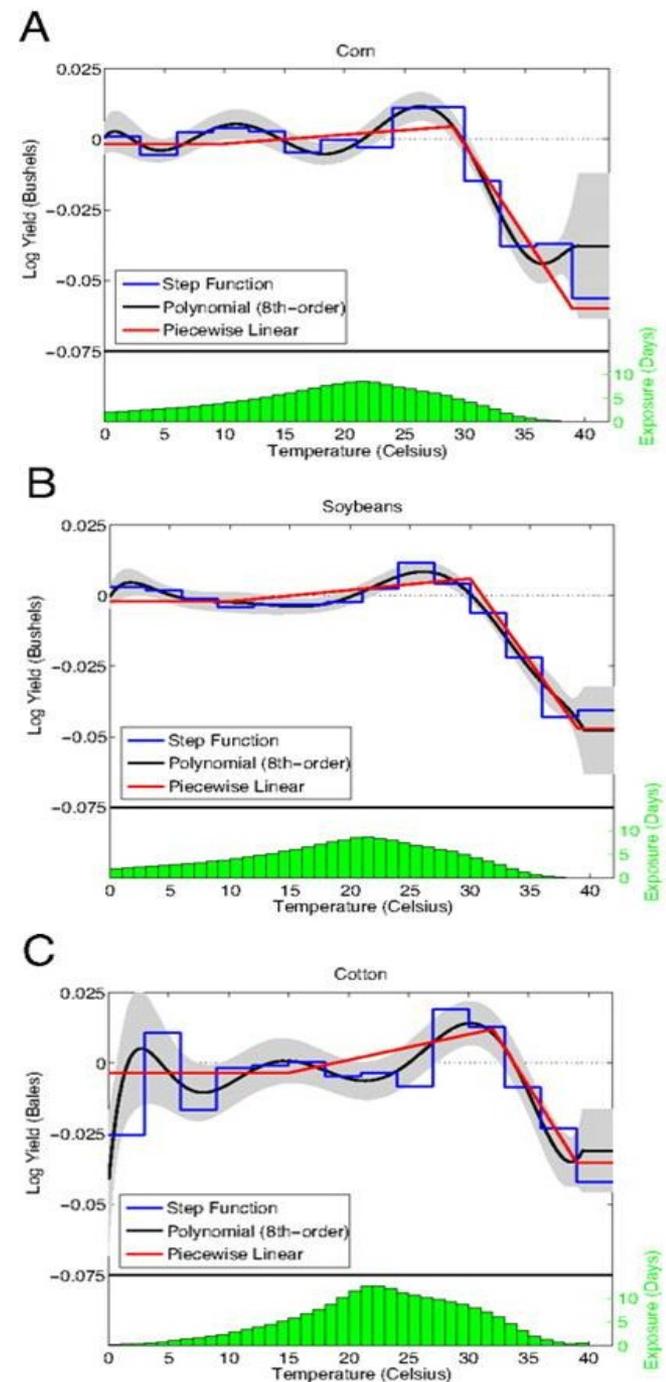


Darkened areas are more efficient – serve to “set the frontier”.

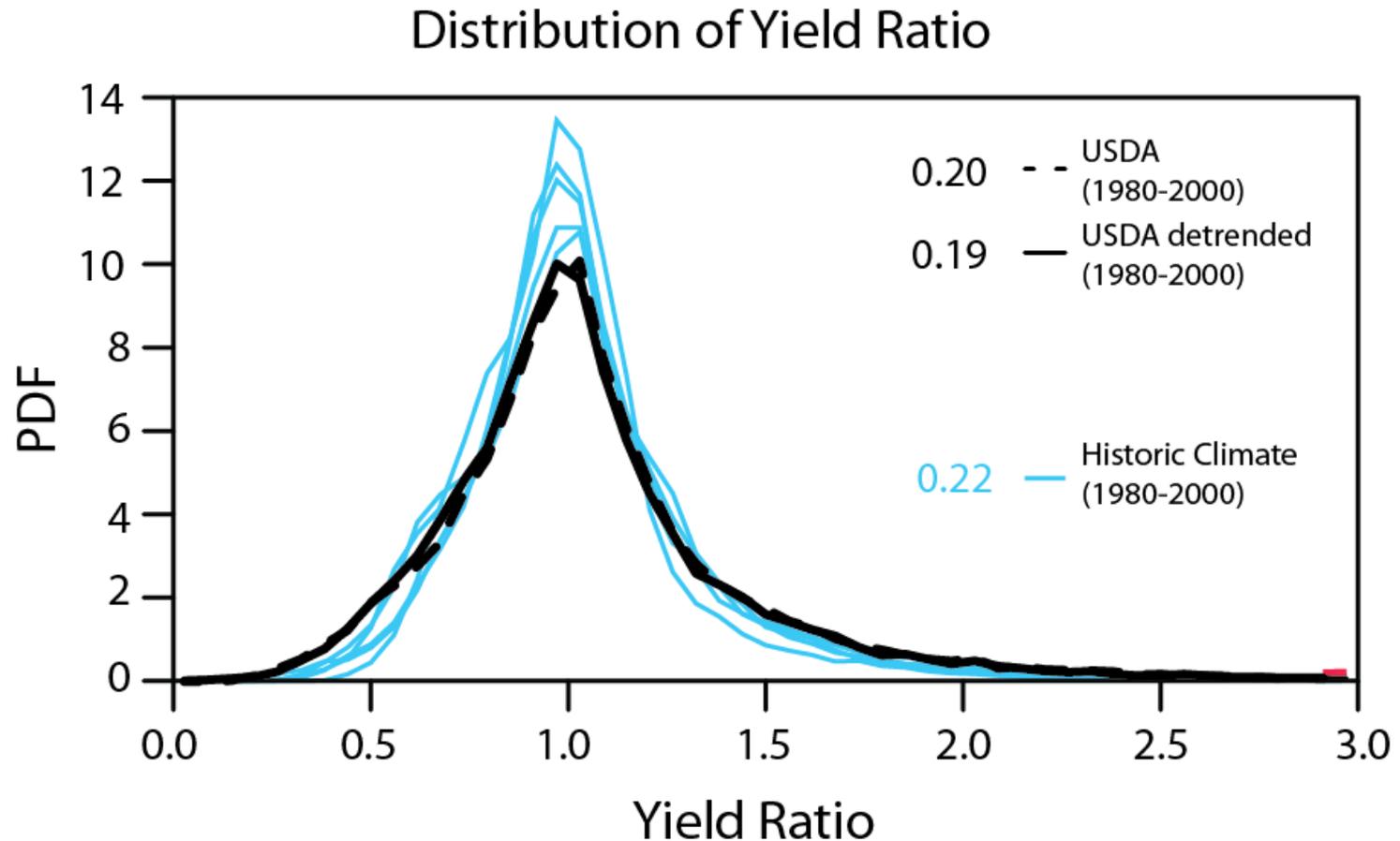
Circled areas are inefficient; primary source of production inefficiency is identified

Combining time series and cross-section data is key to identifying threshold effects of climate on yields

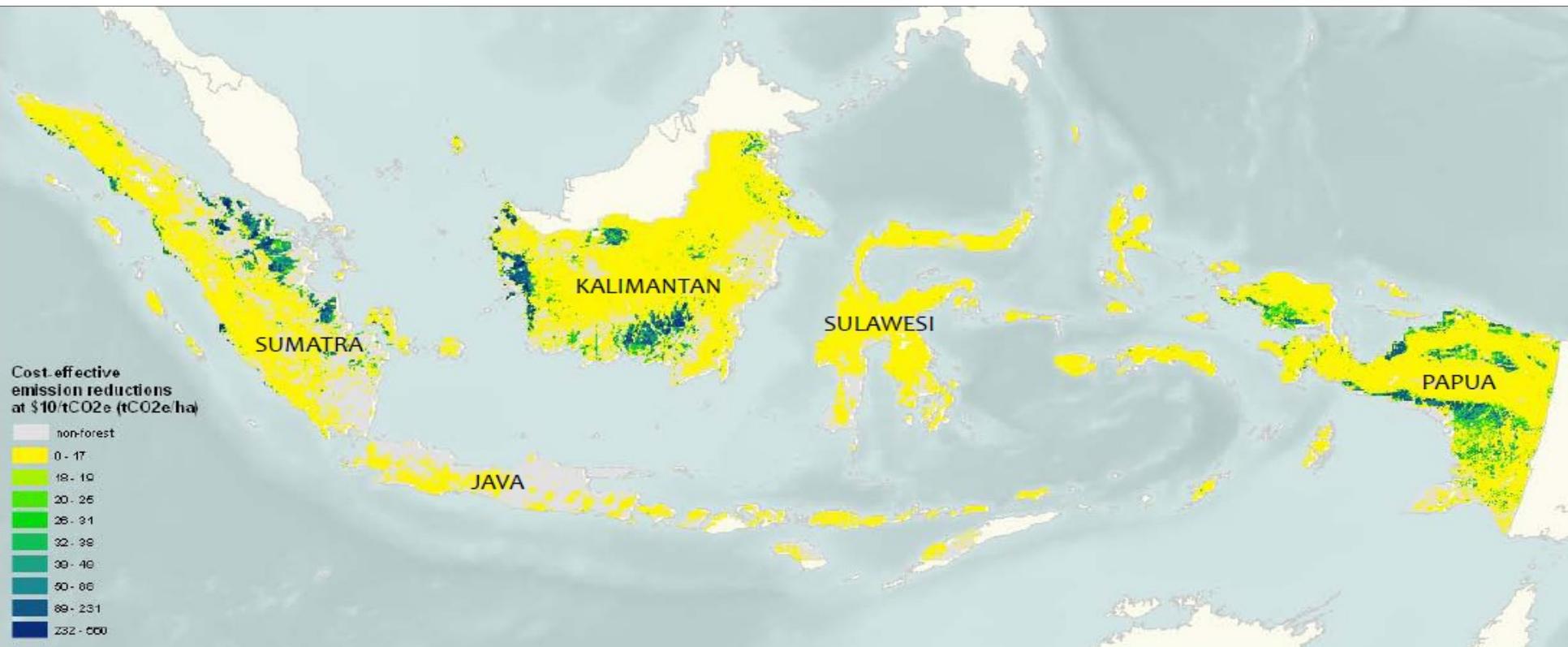
- Schlenker and Roberts (2009):
 - pair US counties' crop yields with fine-scale weather dataset
 - incorporates the distribution of temperatures within each day and across all days in growing season
- Yields increase with temperature:
 - up to 29° C for corn
 - up to 30° C for soybeans
 - up to 32° C for cotton
- Temperatures above these thresholds are very harmful to yields



The combination of high resolution climate results with the Schlenker-Roberts regression yields performs well vs. history



Time series spatial data are being used to design REDD payments in Indonesia



Land-cover response to carbon price of \$10 tCO₂e paid for emission reductions below business-as-usual levels. Darker blue represents greater voluntary abatement of emissions from deforestation in response to incentive payments.

Source: Busch et al. (2012: PNAS), *Climate and revenue benefits of economic incentives to reduce emissions from deforestation in Indonesia*

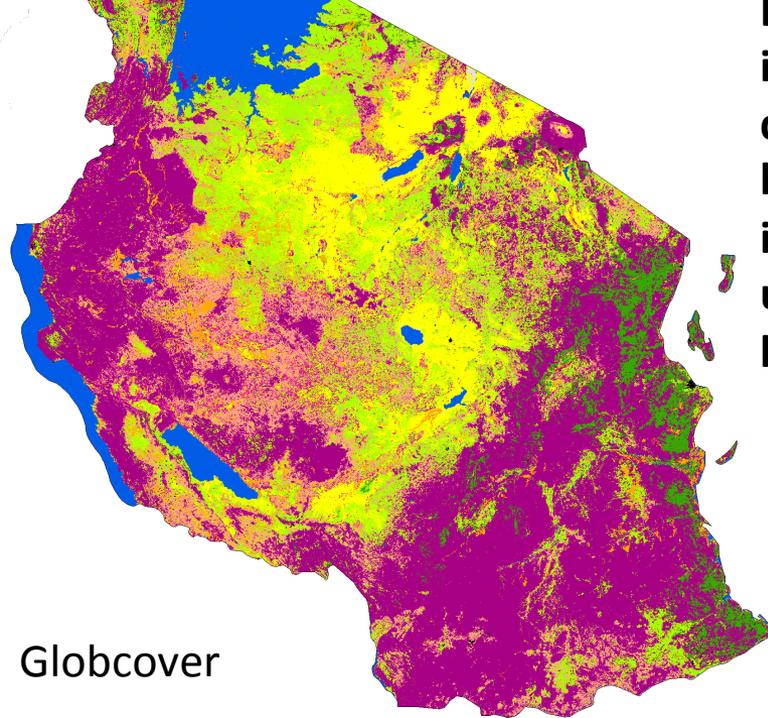
So what is the problem?

- Most spatial datasets:
 - Are regional or national, *not global*; while many of the problems are global in scope
 - Global (and many regional) data sets are *incompatible*: causes problems of inconsistency which are expensive and often impossible to resolve at the end of the data pipeline; compatibility must be designed in at outset
 - Present high barriers to entry

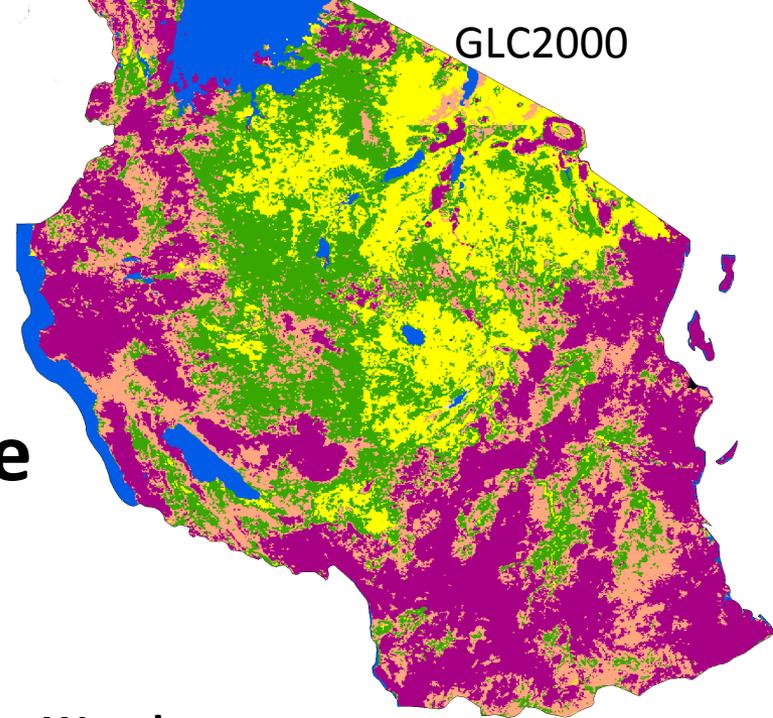
In spite of
improving satellite
data, there remain
huge discrepancies
in our
understanding of
land cover

Alternative views of Tanzania

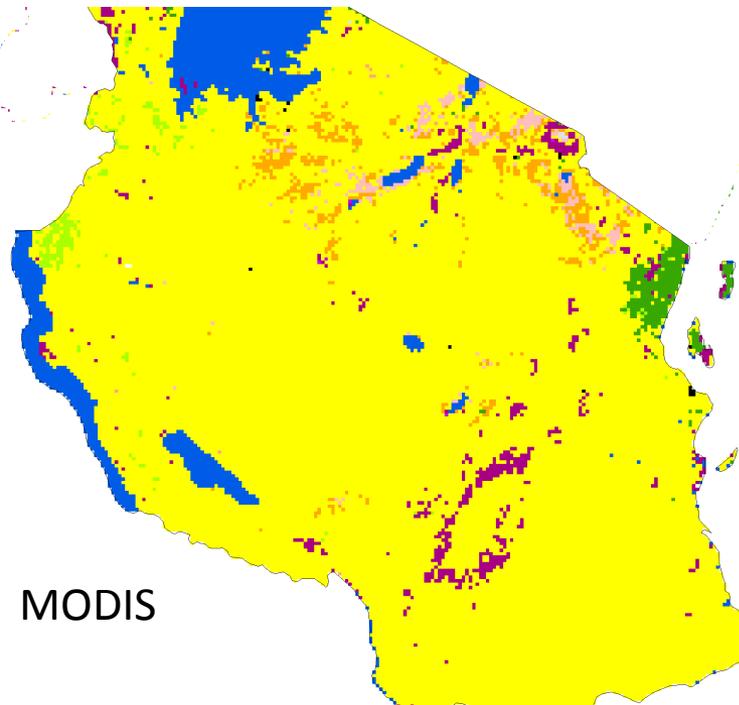
Provided by Stanley Wood



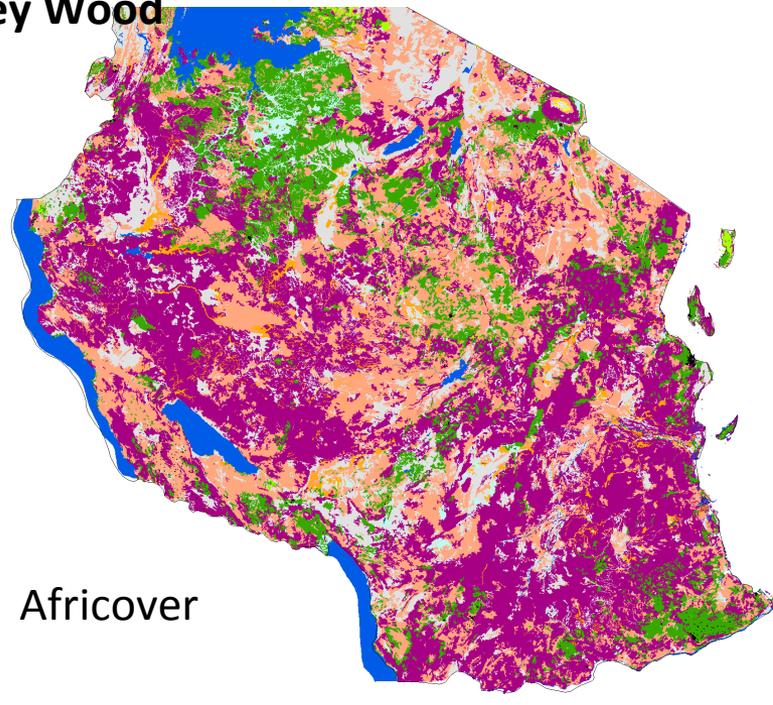
GlobeCover



GLC2000



MODIS



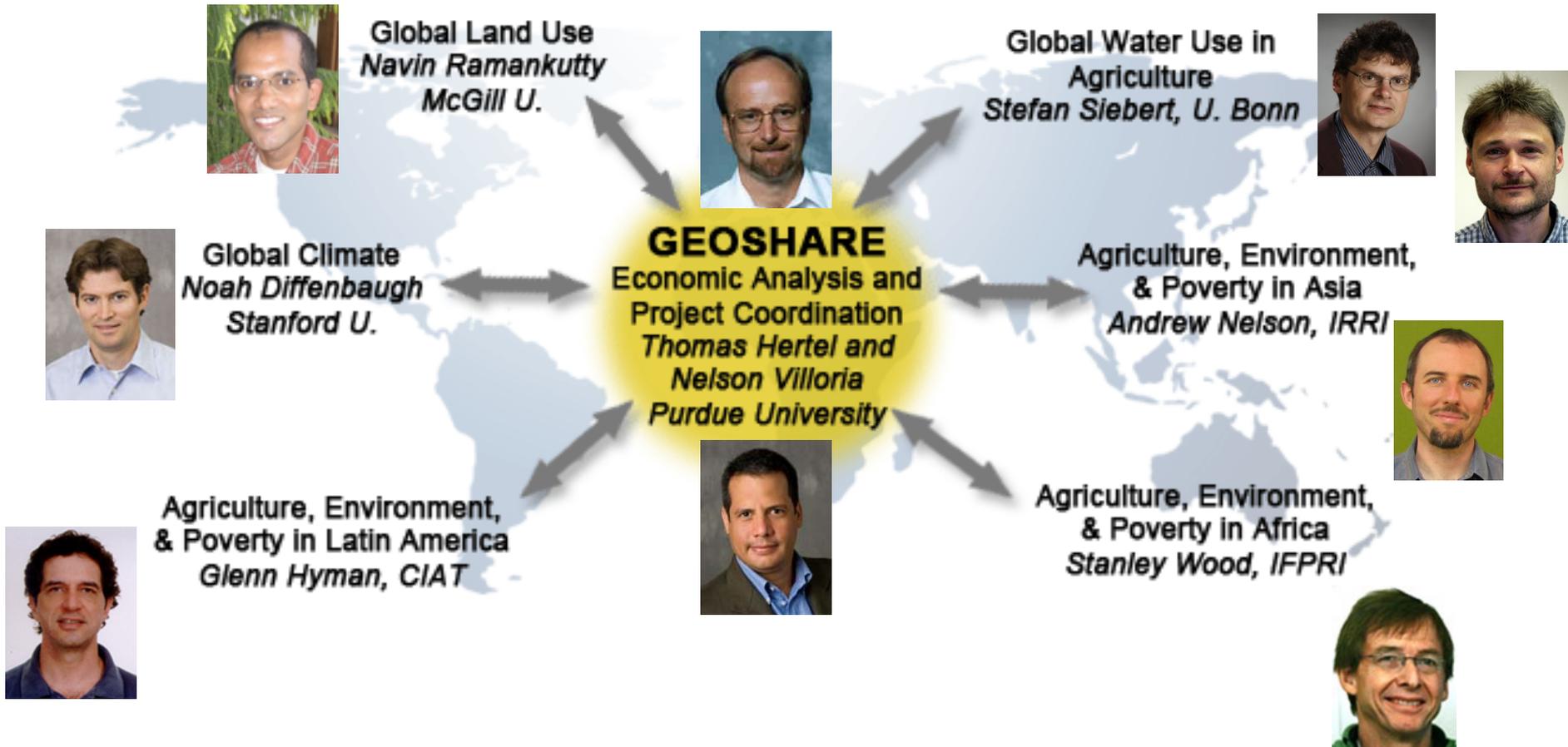
Africover

Our Diagnosis*

- The data currently available to understand how global and local phenomena affect the agriculture-environment-poverty nexus are insufficient to advance needed discovery/enable effective decision making
- This lack of information has greatly inhibited the ability of scientists, practitioners and policy makers to address the socio-economic and environmental impacts of contemporary policy issues related to poverty reduction and the long run sustainability of the world food system.

* Original proposal sponsored by the UK Foresight Programme. Available at <http://www.agecon.purdue.edu/foresight/>

GEOSHARE features a scalable structure with regional and global nodes



Additional nodes under consideration: Livestock, land tenure, population, regional node for Europe.

Data needed to understand tradeoffs between agriculture and the environment

Resources:

- Land (soil)
- Water
- Solar radiation
- Germplasm
- Energy
- Assets
- Population

Economic and institutional factors:

- Land tenure
- Market Access
- Credit Availability

Environmental and socio-economic outcomes:

- Poverty Alleviation
- Landscape protection
- Soil degradation
- Carbon fluxes
- Biodiversity



**Agriculture and
Forestry sectors**



Management:

- Planting/harvest dates
- Irrigation
- Fertilization
- Plant protection
- Labor
- Mechanization

Marketed products:

- Food
- Fiber
- Fuel
- Timber
- Carbon credits

HubZero Cyberinfrastructure

- Created by the NSF-funded Network for Computational Nanotechnology (In 2009 + 100,000 users launched 369,000 simulation runs from 160 different models)
- Users create, publish and access interactive visualization, simulation, and other analytical tools powered by cluster computing resources
- Focus on education and capacity building

Structure: Carbon Nanotube

Simulation Method: Pz orbital

Simulate

Result: Molecular structure: overall

Determine the simulation method.

Pz orbital:
The Pz Orbital model uses 1 Pz orbital/atom as the basis set. Of the two simulation methods, this has the advantage of being the faster, but the disadvantage of being the less rigorous.

Extended Huckel Theory:
The Extended Huckel Theory model uses 4 orbitals (S, Pz, Py, Pz)/atom as the basis set. Of the two simulation methods, this has the advantage of being the more rigorous, but the disadvantage of being the most time consuming.

Chirality (n,m)

n: 7

m: 5

Model parameters

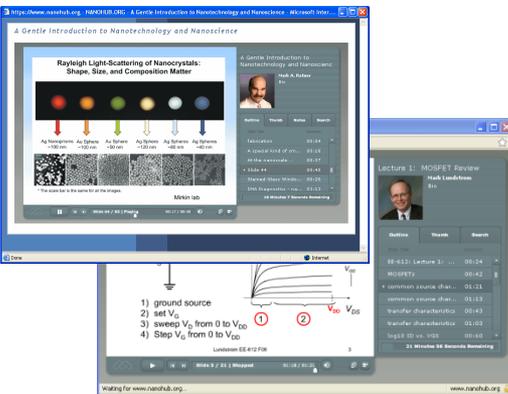
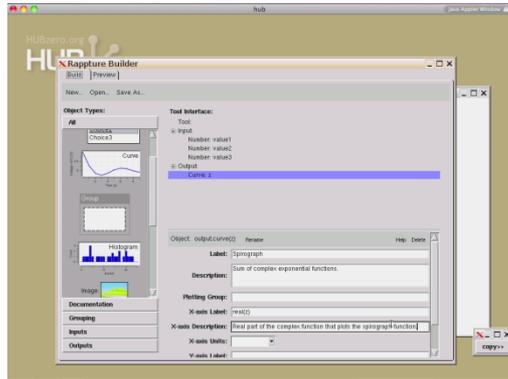
Tight Binding Energy: 3eV

Carbon-carbon spacing: 1.42A

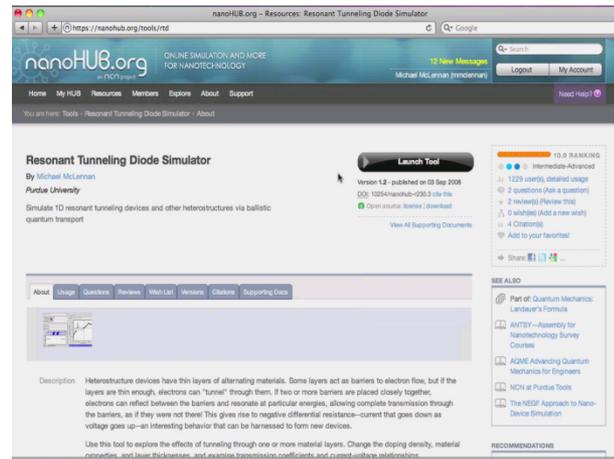
Length in 3-D view: 40

1 result Parameters... Clear

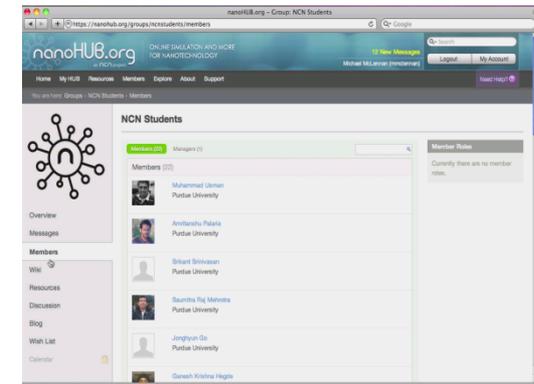
HUBzero facilitates analysis and dissemination of information



Build tools, create seminars and other resources

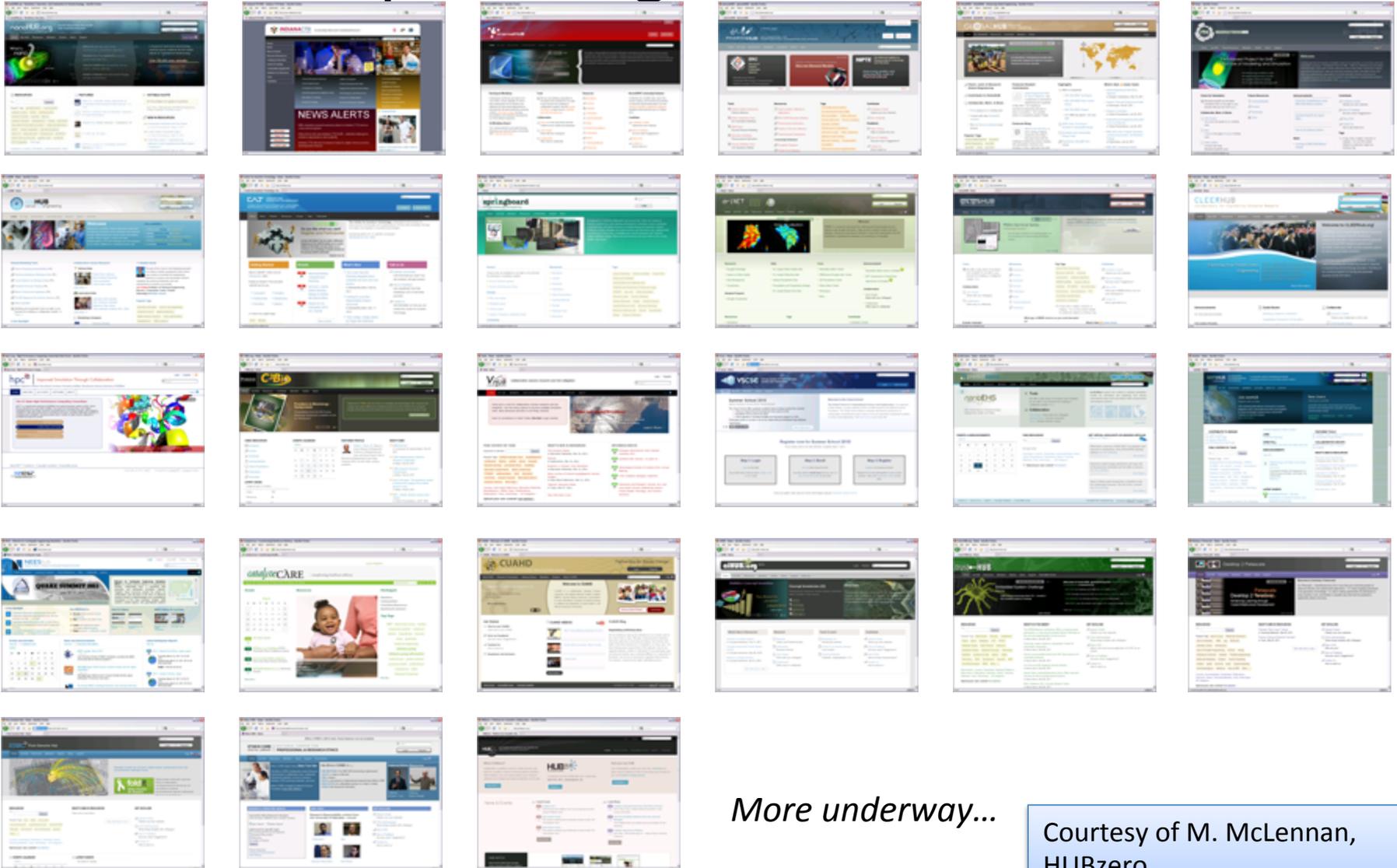


Publish online in restricted groups or for everyone



Share data, answer questions, work together in groups

Usability – HUBzero powering more than 30 hubs!



More underway...

Courtesy of M. McLennan,
HUBzero

GEOSHARE is different from existing initiatives in that....

- It emphasizes development of time series, historical data for scientific analysis:
 - Data production, consistency, validation and interoperability are front and center
- HubZero technology goes beyond spatial zooming and summary statistics, facilitating:
 - Use of analysis tools ‘in the clouds’
 - Networking and sharing of data
 - Technology transfer
 - Capacity building

GEOSHARE is complementary to existing initiatives, including:

- HarvestChoice, which has a strong regional focus, HarvestChoice Co-PI Stan Wood directs GEOSHARE's Africa region node
- Existing global data base infrastructure such as GlobalSoilMap.net and CMIP, which will be mapped to GEOSHARE
- FAOSTAT which will offer national comparisons
- GEO/GEOSS which set guidelines for geospatial data and GEO-GLAM for land use change

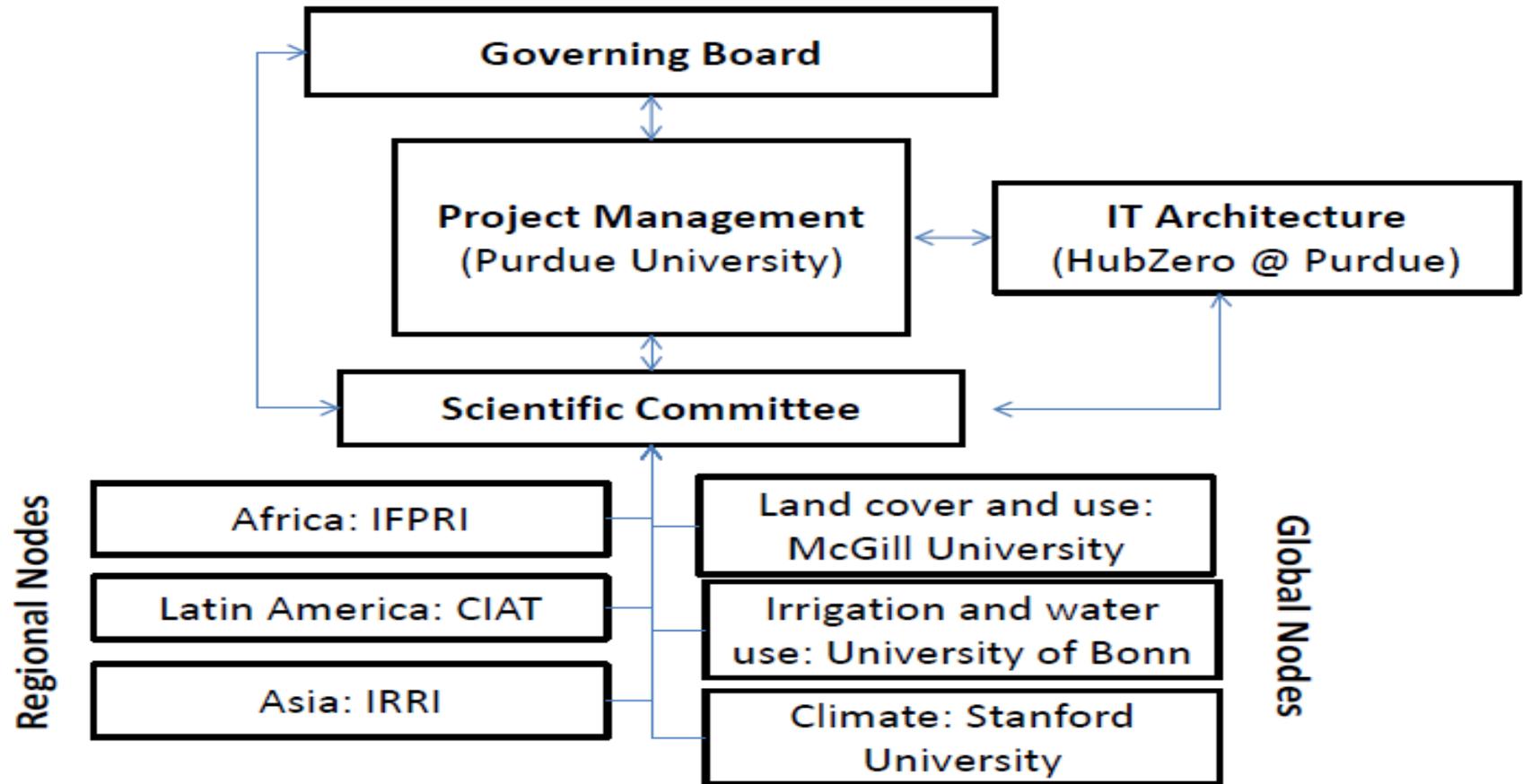
GEOSHARE is complementary to new initiatives on agricultural monitoring

- Includes new Gates-funded initiative focusing on Africa (Sachs et al., 2010)
- Three important types of complementarity:
 - Monitoring can ‘ground-truth’ GEOSHARE estimates
 - GEOSHARE can facilitate ‘scaling up’ site-specific findings, as well as choice of new monitoring sites
 - HubZero will facilitate biophysical and economic modeling to extrapolate from monitored to unobserved regions using comparable data inputs

GEOSHARE will be a source of input data for other projects

- *AgMIP*: global yield data are key for crop modeling; need rainfall/irrigation split
- Lobell et al (2011, Science) will use GEOSHARE data to update their estimates of climate impacts, disaggregating irrigated crops
- *GTAP* will use *GEOSHARE* data to disaggregate activity within countries to analyze agricultural impacts of trade and environmental policies

Proposed Governance Structure



GEOSHARE Pilot has just begun

- Modest funding from three sources:
 - UK Department For International Development: \$ 440k
 - UK Department for Environment, Food and Rural Affairs: \$ 100k
 - USDA's Economic Research Service: \$100k
 - Additional funds from HubZero (\$50k) and from CCAFS (\$30k)
- Proof of concept:
 - Two countries in South Asia, six in Africa; regional case studies demonstrating support for decision makers in Asia and Africa
 - Integrate regional and global nodes – focus on irrigation/rainfed split, overlaid by data on poverty and land tenure
 - Delivery of data and decision tools (e.g. Pegasus crop model) through HubZero infrastructure
 - Design governance; assess costs and benefits; address ethical issues
- Donor's forum in early 2014:
 - Demonstrate success of proof of concept from pilot project
 - Secure long term funding for GEOSHARE