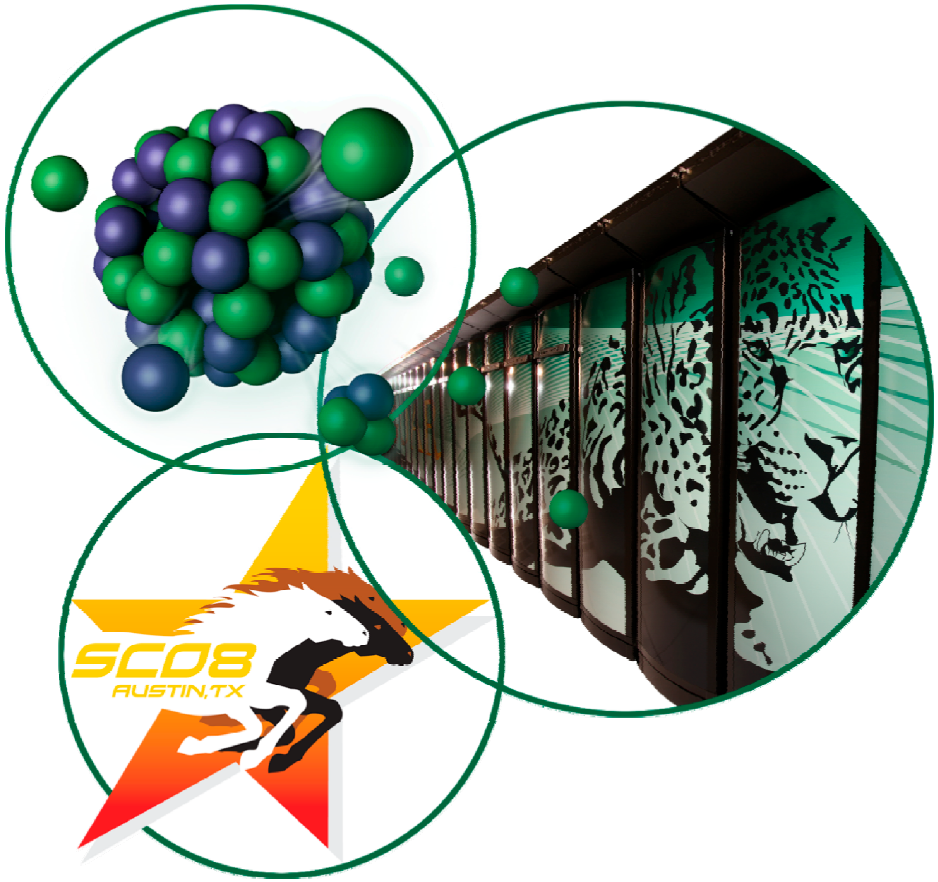
Geospatial Data Sciences

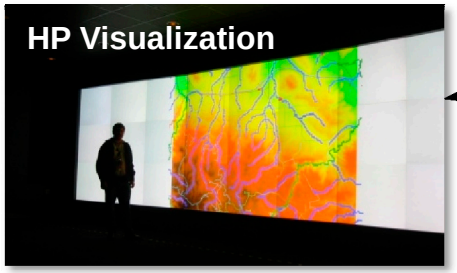


Presented by

Ranga Raju Vatsavai
Auroop Ganguly
Budhendra Bhaduri

Geographic Information Science and Technology
Computational Sciences and Engineering

Geospatial data sciences



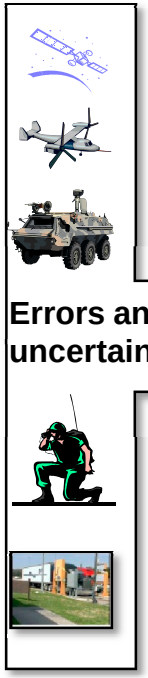
Knowledge Discovery

- Data mining
- Machine learning
- Statistics
- Spatial analysis
- Pattern recognition
- Nonlinear dynamics

Actionable insights/ knowledge

- Disaster mapping and monitoring
- Geospatial intelligence discovery
- Extremes
- Social networks
- National security

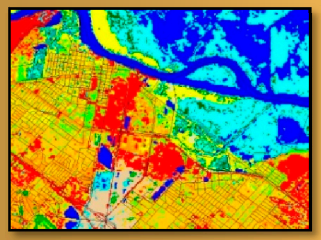
Efficient storage and retrieval methods

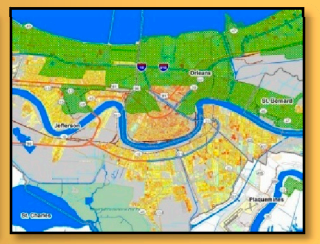


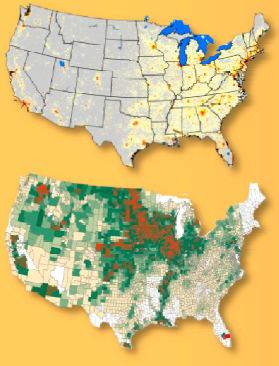
Terabytes of spatially referenced data per day

Errors and uncertainty

GeoDatabases



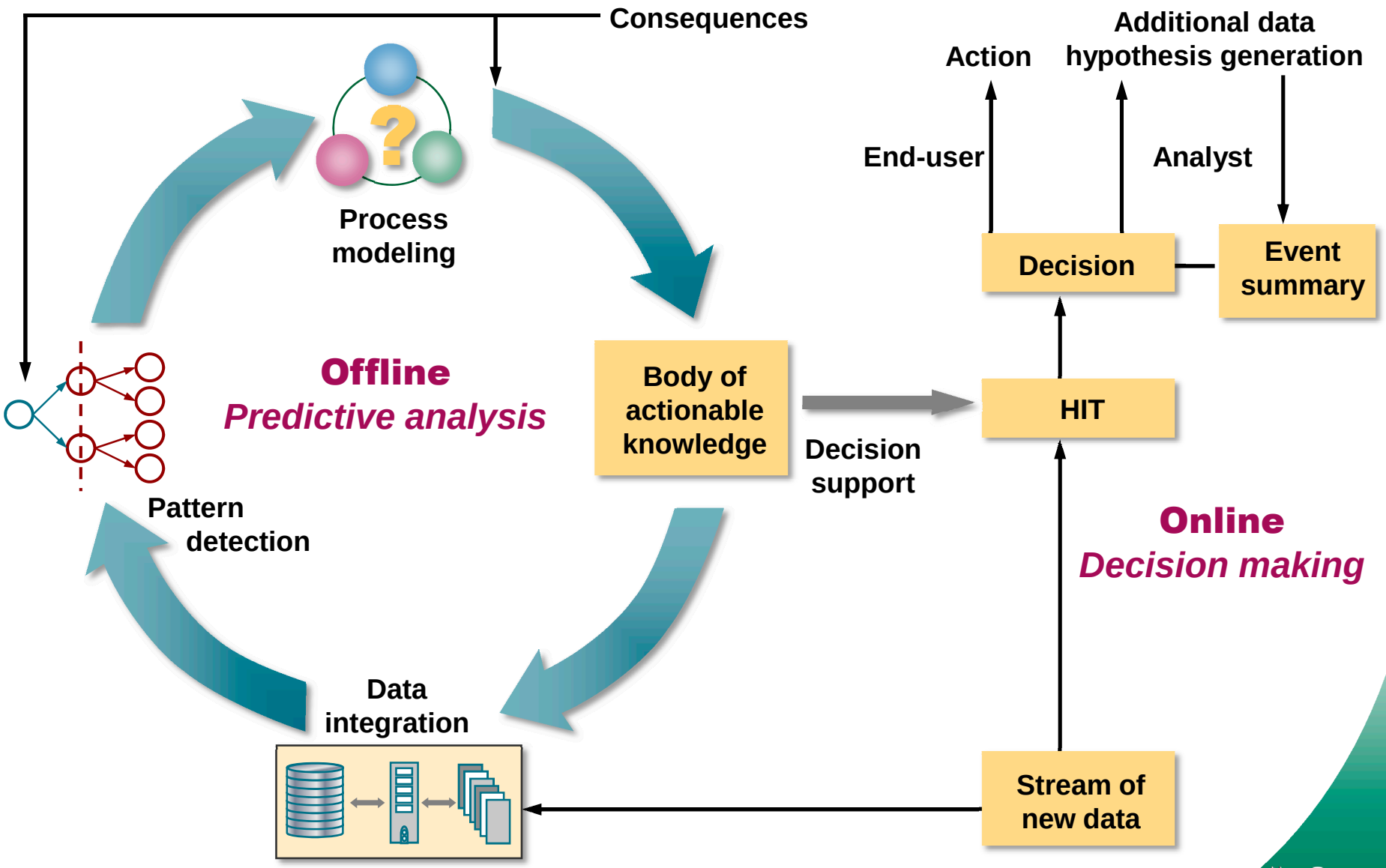




Petabytes of spatiotemporal data archives

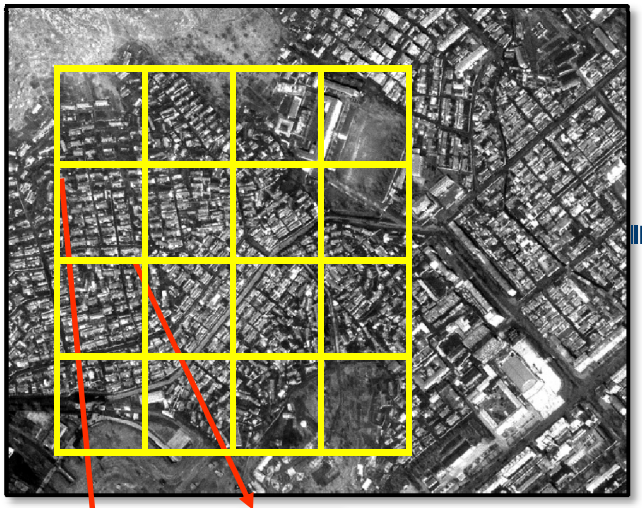
Damage	Total population	Senior
100%	791,361	90,773
75%	1,594,806	187,677
50%	1,505,196	184,372
25%	1,701,593	224,279
Total	5,592,956	687,101

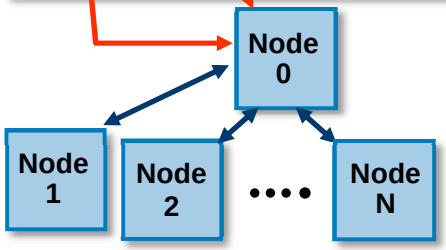
Pattern detection and decision support



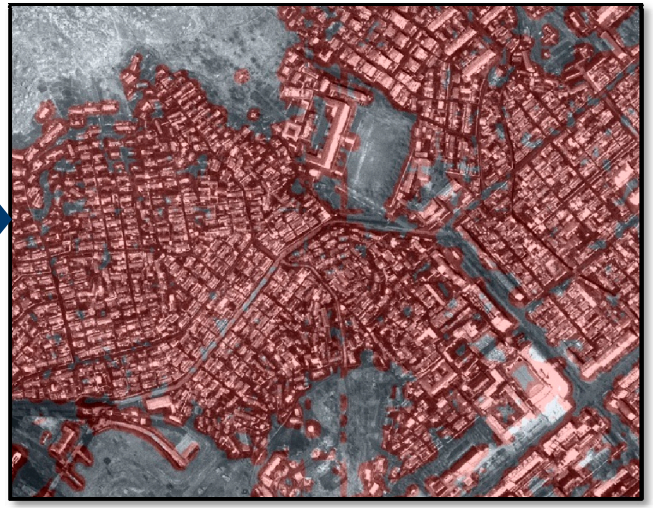
Urban mapping for global population distribution

Identify urban regions from high-resolution satellite imagery by using Gabor texture analysis





- Convert to Gabor space
- Compute energy variance inside sliding window size W
- Perform iterative data clustering

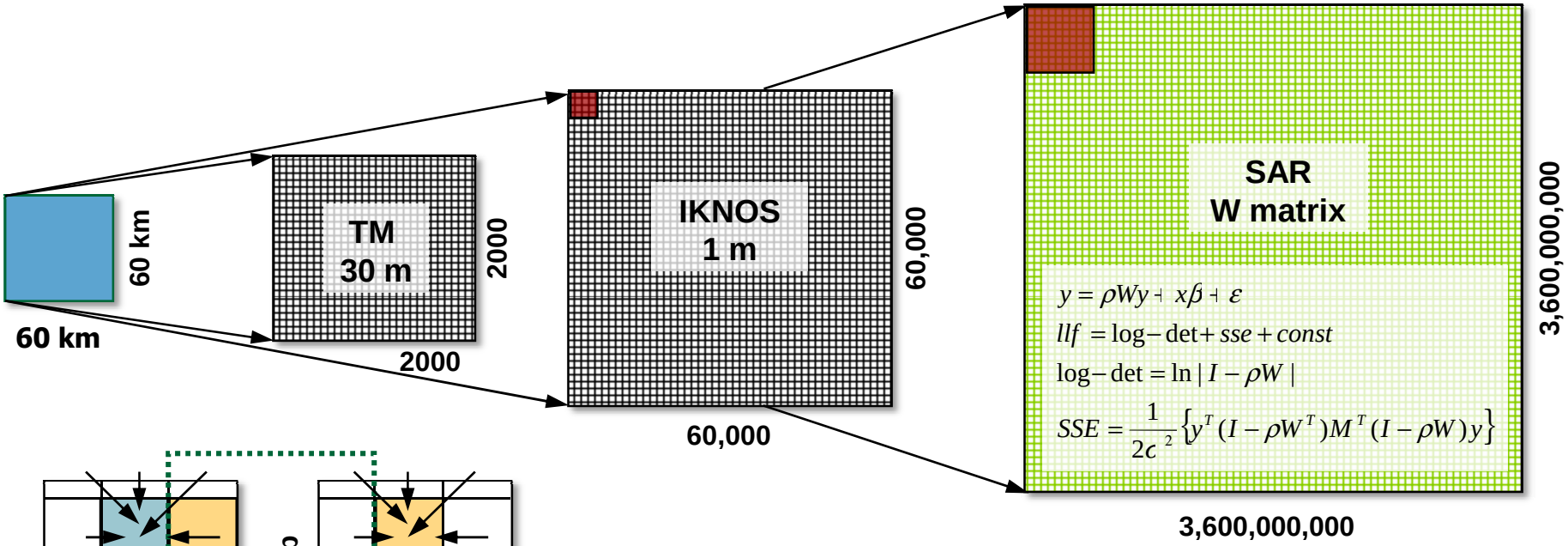


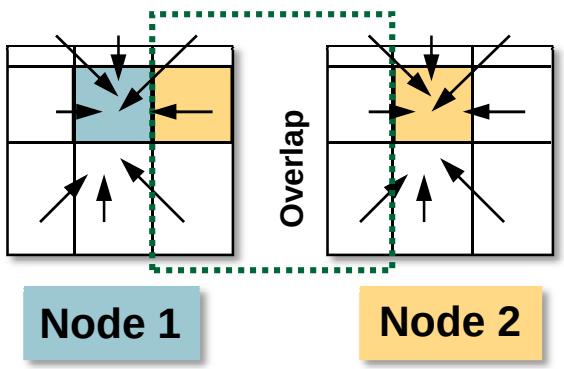
	Area (sq. km.)	Imagery data	Intermediate data created	Time 10 nodes
Sample area	114	219 MB	~3.9 GB	~30 min
Projected computing requirements				
Global populated land area	56,863,754	~103 TB	~1.8 PB	~30 years

Oak Ridge National Laboratory

Spatial classification and prediction

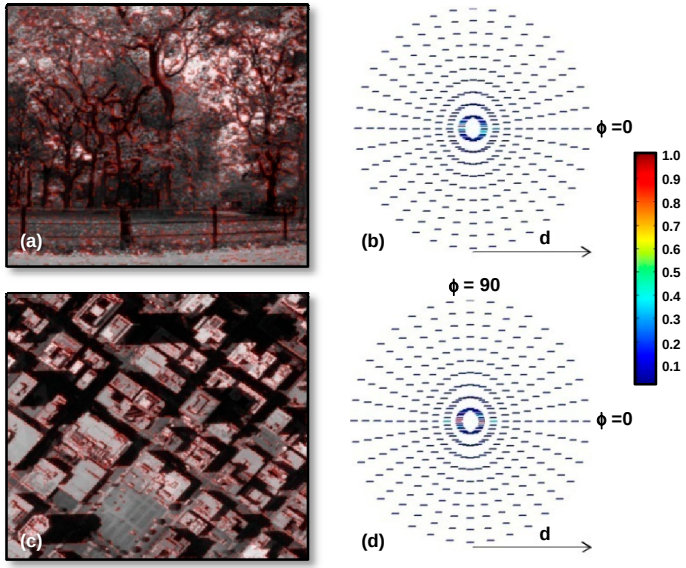
- Spatial autoregressive regression (SAR)
- Markov random fields (MRF)



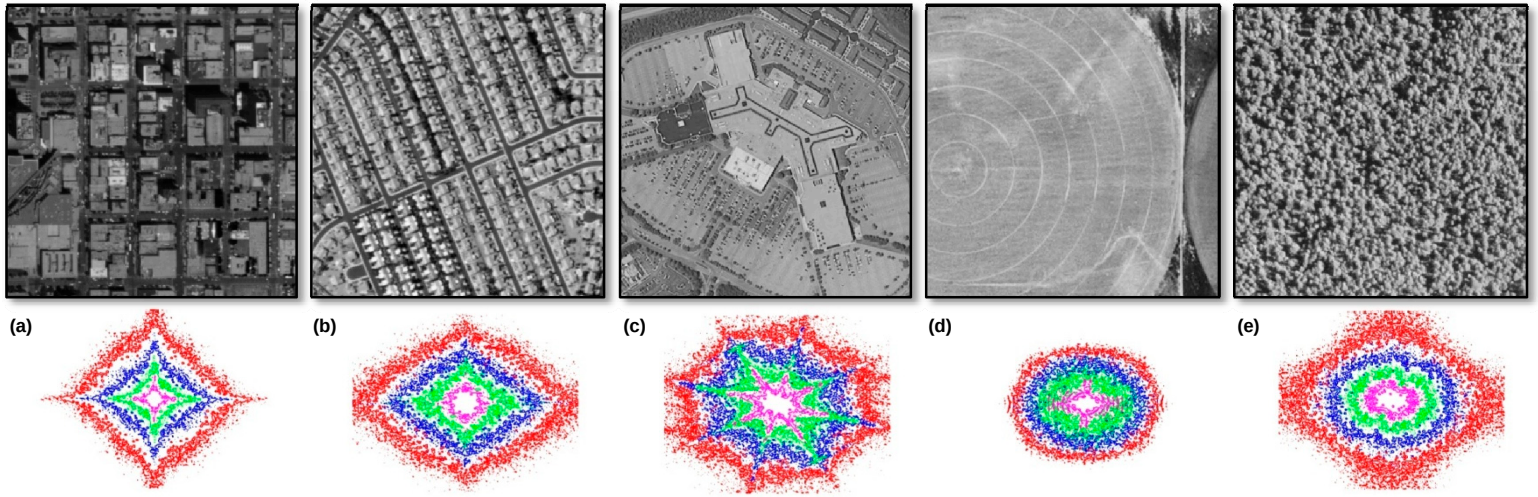


- **Challenge**
  - Large neighborhood matrices
  - Overlapping computation

Statistical analysis for satellite image characterization

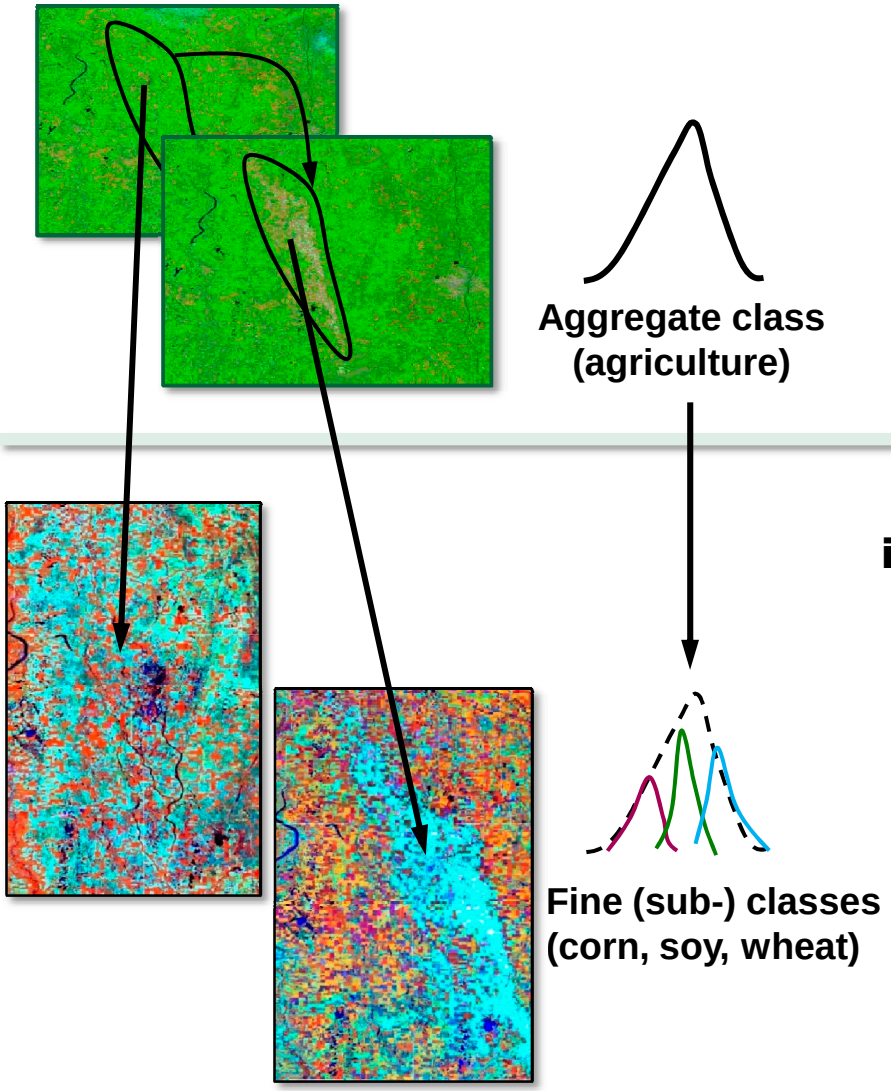


- Power spectrum analysis
- Image gradient distribution
- Wavelet analysis



- Challenge is to extend automated processing to fine-resolution images over large geographic regions

Regional and global scale biomass monitoring



Biomass monitoring

MODIS (250m, 2b, 1d)

- Coarse spatial resolution
- High temporal resolution
- Good for regional, global monitoring, but not ideal for crop identification

Change vs. no change

Fine-resolution information extraction

AWiFS (56m, 4b, 5d)
NAIP (1m, 4b, 1y)

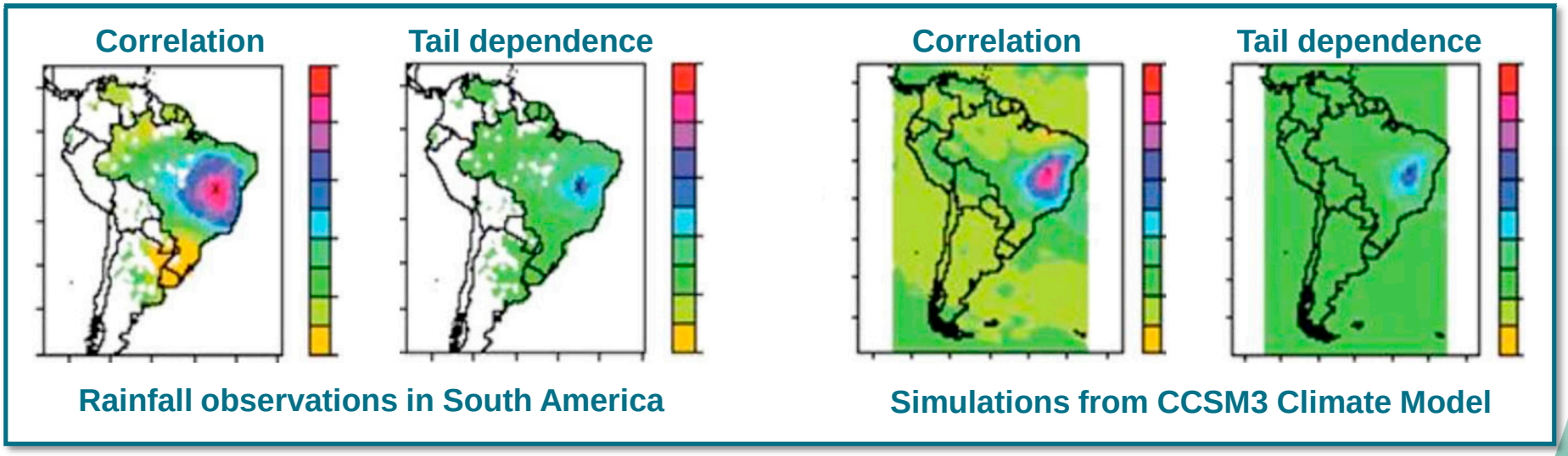
- Moderate-to-high spatial
- Moderate temporal
- Used for crop type and condition extraction
- High training and computational requirements

Characterize changes

Data complexity
(1 cycle)
MODIS~6GB
AWiFS~30GB
NAIP~38TB

Extremes

- **A new measure to describe the relationship among extremes in space and time**
  - **Simultaneous occurrence of 100-year return levels**
    - **Perfect dependence → 100 year event**
    - **No dependence → 10,000 year event**



- **Challenge**

  Spatiotemporal dependence estimation is computationally expensive (1000 grid cells results in scanning more than half a million pairs)

Contacts

Ranga Raju Vatsavai

Geographic Information Science and Technology
Computational Sciences and Engineering
(865) 576-3569
vatsavairr@ornl.gov

Auroop Ganguly

Geographic Information Science and Technology
Computational Sciences and Engineering
(865) 241-1305
gangular@ornl.gov

Budhendra Bhaduri

Geographic Information Science and Technology
Computational Sciences and Engineering
(865) 241-9272
bhaduribl@ornl.gov