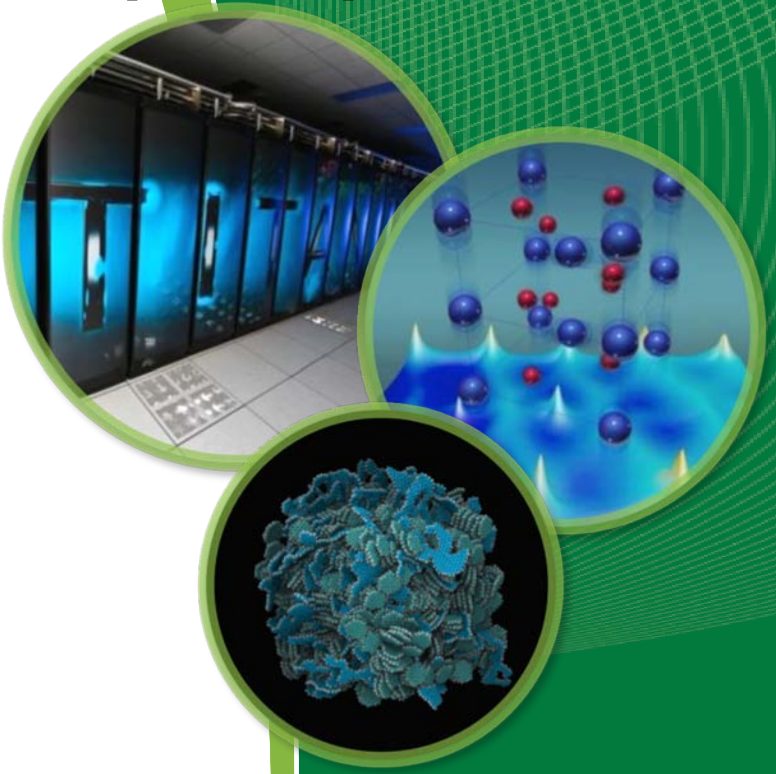
U.S. Department of Energy (DOE) Bioenergy Technologies Office (BETO) 2017 Project Peer Review

4.1.2.40 -Land-Use Change Data Analysis



03/08/2017

Analysis & Sustainability

Nagendra Singh (PI)
Oak Ridge National Laboratory

This presentation does not contain any proprietary, confidential, or otherwise restricted information.

Goal Statement

Project Goal

Design and develop tools to establish scientific basis for understanding and simulating effects of bioenergy policy on land cover and changes in land characteristics or use.

Project Outcome

An integrated platform for ingesting, visualizing and analyzing satellite data for estimating land use land cover changes, resource assessment and assessing changes in productivity by using the developed algorithms and incorporating ancillary datasets for visualization and analytics.

Project Relevance

This project addresses the fundamental challenge of capturing 'the dynamic state of land use and management' as well as establishing a framework for data analysis and visualization and provides a platform for BETO, researchers, and policy makers to understand resource dynamics at landscape scales.

Quad Chart Overview

Timeline

- Project start date :FY15
- Project end date : FY18
- Percent complete : 75 %

Budget

	Total Costs FY 12 –FY 14	FY 15 Costs	FY 16 Costs	Total Planned Funding (FY 17)
DOE Funded	NA	150K	225K	225K
Project Cost Share (Comp.)*				

Barriers

- At-A. Comparable, Transparent, and Reproducible Analyses
- St-C. Sustainability Data across the Bioenergy Supply Chain
- St-G. Land-Use and Innovative Landscape Design

Partners

Interactions/collaborations

- USGS, USDA, NCSU
- Inputs from other DOE labs

1- Project Overview

Context

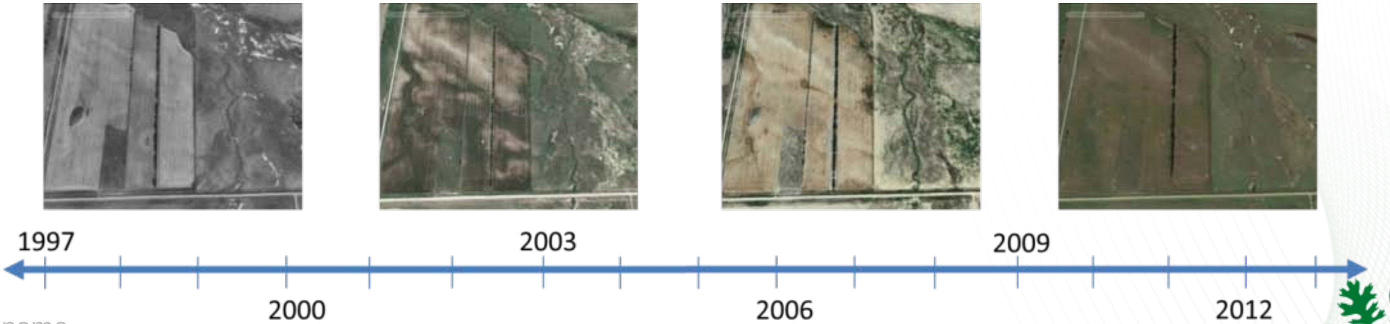
Concern: bioenergy policy leads to effects (conversion of forestland; displacement of food production) with significant social and environmental (e.g., GHG emissions) consequences.

History

- Gaps & uncertainties from using derived datasets, area/class aggregations, pixel-level calculations, and lack of required resolution.
- Lack of data, algorithms to monitor and detect dynamic change at required temporal resolution.

Objective

- Develop tools and reliable data to :
  - Accurately assess changes and trends in vegetation patterns
  - Monitor changes in vegetation (annual, range)
  - Scalable over time and space



2 – Approach (Management)

Regular project meetings with team members

- Assess project progress, discuss issues, analyze results
- Project SharePoint to share literature, codes, data
- Discuss issues (computing, data) and deviations from desired outcome
- Track quarterly deliverable as well as target overall objective

Quarterly progress reports and discussions with BETO managers; monthly written updates.

Project Structure

- Algorithm- Nagendra Singh, Dalton Lunga, Su Lim
- Computing – Dilip Patlolla
- Data/ Analysis – Nagendra Singh, Ronan Lucey, Huina Mao
- Guidance – Budhendra Bhaduri, Raju Vatsavai (NCSU), Keith Kline, Virginia Dale.

2 – Approach (Technical)

Vegetation is easily impacted by changes in climate, environment and policies and therefore it is critical that such changes are monitored and detected for planning and mitigation strategies.

Ability to analyze large volumes of satellite data **at high temporal resolution** over **long periods** provides the ability to capture the dynamic nature of land use change.

Challenges

- Conventional classification algorithms do not work on noisy time-series data.
- Lack of consistent ground truth data
- Existing techniques does not capture the temporal correlations (or phenology) between the NDVI signals
- Computational -Scaling the algorithms to process billions of pixels.

Critical success factors

- Ability to monitor and measure changes as they occur
- Scalability over time and space
- Platform to visualize and analyze results

3 –Technical Accomplishments/ Progress/Results(FY15/16)

Development of semi-supervised classification algorithm

Attains same accuracy as the supervised methods but works with a limited amount of user supplied ground-truth data.

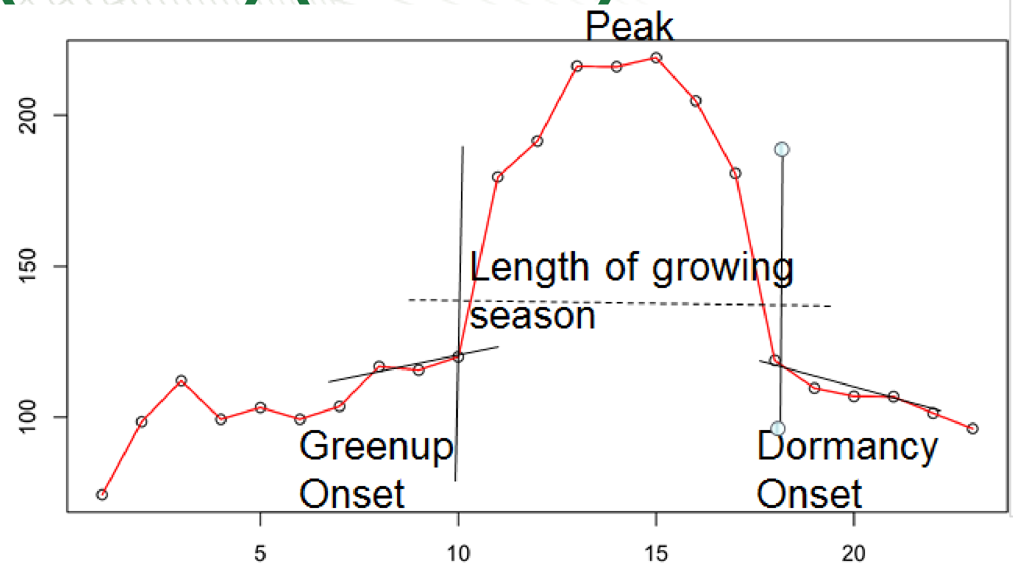
- Temporal variability of vegetation types used to differentiate between land cover types
- Considerable improvement in learning accuracy

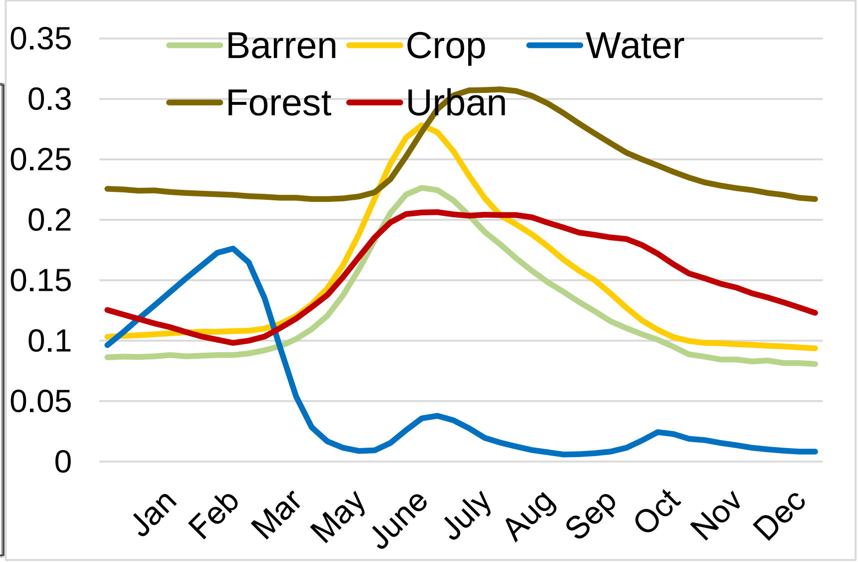
Key milestones and status

✓ Validation and testing of semi-supervised model for 2-3 areas in the USA and compare results with ground truth data (GO/NO-GO)
  - Tested over selected areas of Iowa, North Dakota, South Dakota (accuracy > 80 %)

✓ Parametrize and calibrate the semi-supervised model

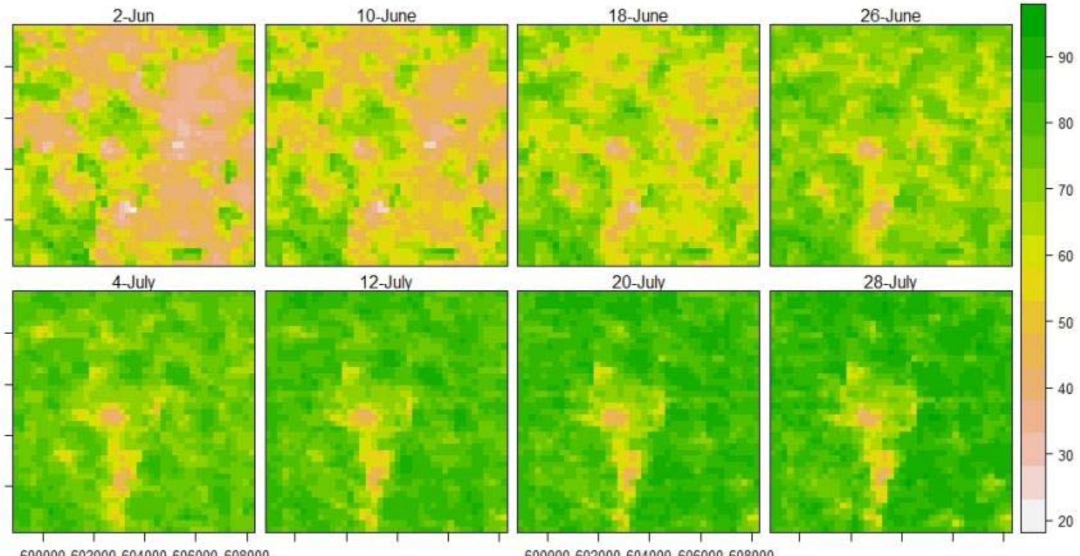
3-Technical Accomplishments/ Progress/Results (cont'd) (FY15/16)





Use Phenology of Crop Types to differentiate them from surrounding vegetation

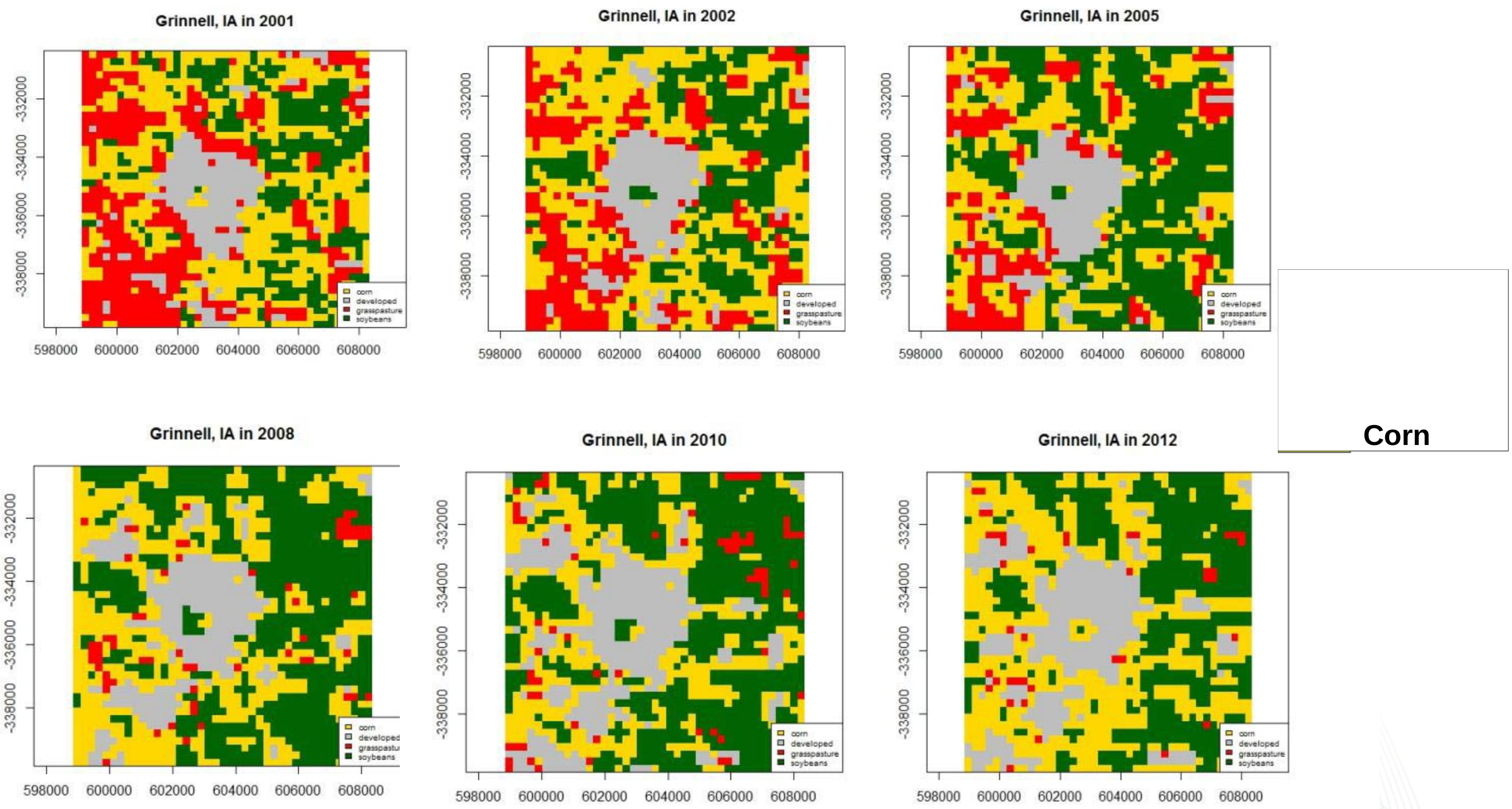
Extract time-series plots of major land cover types: cropland, urban areas, forestland, and grassland.



OAK RIDGE National Laboratory

3 – Technical Accomplishments/ Progress/Results (cont’d) (FY15/16)

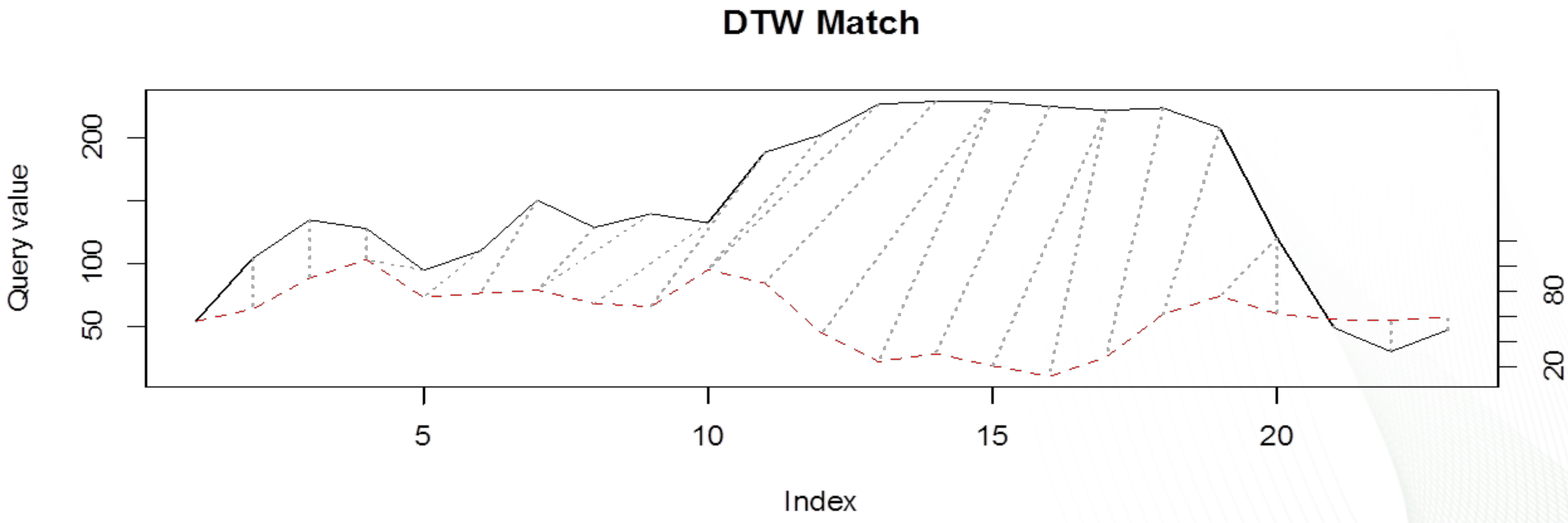
Identifying trends in Land use Land Cover



OAK RIDGE National Laboratory

3 -Technical Accomplishments/ Progress/Results (cont'd) (FY16/17)

- Development of a spatio-temporal change detection method
  - Use novel Hierarchical clustering method on stacked data
  - Pick the optimal number of clusters and cluster groupings through dendrogram exploration.
  - Conventional clustering is not effective in modeling temporal correlations and crop phenology in particular
  - Use Dynamic Time Warping (DTW) to compute the similarity matrix.



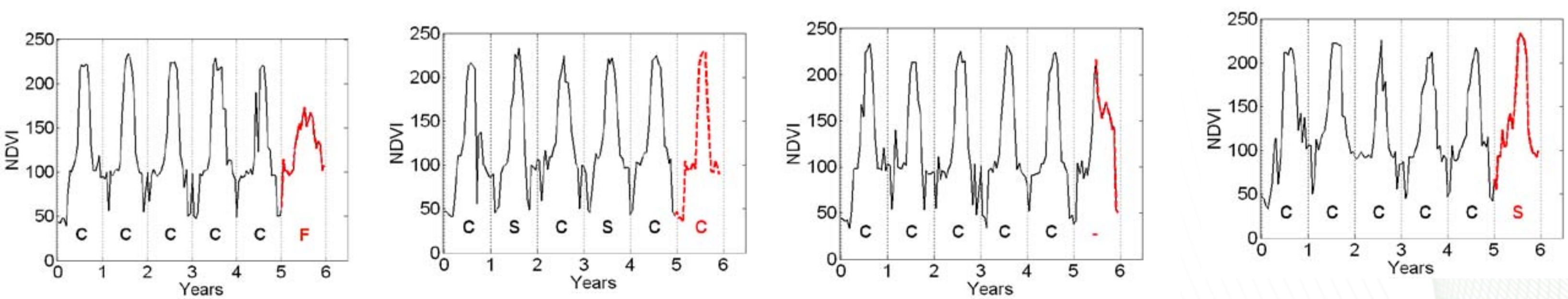
3 -Technical Accomplishments/ Progress/Results (cont'd) (FY16/17)

Gaps in Existing Research

- Not suitable for near-real time monitoring
- Make parametric assumptions that cannot be easily handled (e.g., non-linearity)

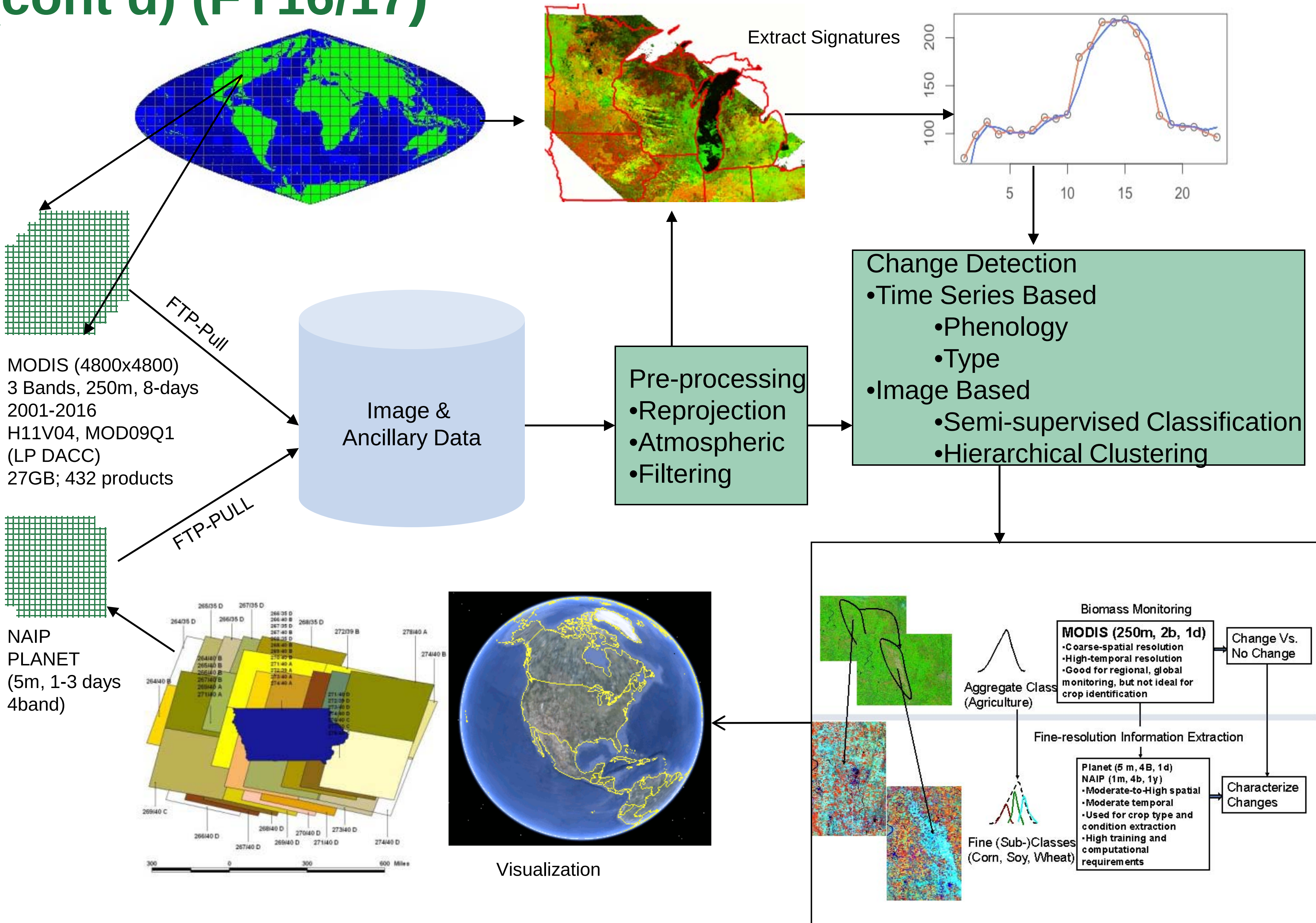
Planned Approach

- Gaussian Process (GP) based time series prediction model for periodic time series data
- A novel covariance function that captures the temporal proximity as well as seasonality effects
- Used with EWMA for near-real time monitoring and detection algorithm



Examples of changes (last year) in NDVI data (C - corn, S - soybean, F - fallow).

3 - Technical Accomplishments/ Progress/Results (cont'd) (FY16/17)



4. Relevance

Project consistent with key DOE- BETO Objectives

- ✓ The developed tools and data enable accurate estimation and management of biomass resources -a key part of the bioenergy supply chain.
- ✓ Transparent and open-source algorithms and tools based on sound scientific principle enable “consistent, reproducible, peer-reviewed analyses”
- ✓ An integrated platform for biomass monitoring and visualization.

Addresses key challenges & barriers and supports A&S platform activities

- ✓ Development of analytical tools and algorithms to understand bioenergy dynamics.
- ✓ Methodology developed ensures “comparable transparent and reproducible analysis” through use of robust algorithms and raw satellite data.
- ✓ Enables analysis at various geographical and temporal scale to “capture the dynamic state of land use and management”.
- ✓ Data-driven prioritization of sustainability efforts and establishment “of framework for data collection, integration, and visualization support analysis, research”.

OAK RIDGE National Laboratory

4 – Relevance

Relevance to stakeholder and strategic plan

- ✓Provides a platform for biomass resource assessment and visualization and analytics.
- ✓Provides tools and data to accurately assess the impact of biofuel development on environment.
- ✓Reducing gaps and uncertainty in analysis to enhance social acceptance and adoption of bioenergy.

OAK RIDGE National Laboratory

5 – Future Work

- Integration of visualization, classification, and change detection algorithm in a single platform for biomass analytics
- Incorporate a feedback loop to enable the system users to interactively label instances that can then be incorporated into training models.
- Integrate climate, yield and acreage data and other socio-economic data in the system to better understand drivers of land use change and management.
- Deployment of monitoring, visualization and prediction system allowing users to visualize, analyze and forecast changes in land use change and understand changes in trends in productivity, yield and future productivity of biomass.

Summary

Overview

- This project aims to provide a suite of tools and informative datasets to accurately assess and visualize changes in land use and land cover in near real time.

Approach

- Apply novel classification and change-detection techniques to time-series satellite data to understand and characterize changes in biomass productivity.

Technical accomplishments

- Development of semi-supervised model that works with limited ground truth data.
- Development of hierarchal clustering technique which works on noisy irregular data.
- Development of change detection platform for estimating and visualizing changes.

Relevance

- Approach removes bias and uncertainty associated with derived data products.
- Enables analysis over any geographic area over any range of time.
- Allows to ‘measure’ and ‘monitor’ trends in vegetation pattern at scale.

Future Work

- Develop a system for monitoring and visualization of changes in biomass near real time.
- A model will be created to forecast the expected land cover value at any given time based on past observations at resolution of the data (8/16 days)

(Not a template slide – for information purposes only)

- *The following slides are to be included in your submission for Peer Evaluation purposes, but will **not** be part of your oral presentation –*
- *You may refer to them during the Q&A period if they are helpful to you in explaining certain points.*

Responses to Previous Reviewers' Comments

On 'Transparency of the algorithms…. and scaling will be a challenge due to the considerable amount of data required'

All data and algorithms will be publically shared through KDF. Results have been and will be continued to be communicated through presentations and publications. Data volume is indeed a big challenge in these types of analysis where million of time series have to be analyzed and therefore there is lack of such research. However we have parallelized the algorithms so they can run on regular desktops by harnessing the power of Graphical Processing Units (GPU by dividing the data into several parts to accommodate into the global memory of GPU. After the GPU finishes operating on all sections of the data, the CPU is used to assemble the result and output for further analysis.

"Further refinement of the model and validation of the results is needed….. It would be great if the model could be expanded in the future to look at what impact climatic events, major disruptions, and population growth, etc. have on land use change.\

Over the past two years we have vigorously tested, refined, and validated the model as defined in our scope of work thereby increasing the accuracy and the speed of all tools and algorithms. We would be more than willing to integrate other variables like population growth, climate events and natural hazards (areas in which we have immense expertise based on our work for other sponsors) in the analysis if BETO wanted us to further expand this project in that direction.

OAK RIDGE National Laboratory

Publications, Patents, Presentations, Awards, and Commercialization

Singh, N., Lucey, R., and Lunga, D. Use of dynamic time warping and clustering for mapping trends and patterns in vegetation using time series MODIS data. AGU Fall meeting, Dec, 2016

Efroymson, R.A., Kline, K.L., Angelsen, A., Verburg,P.H., Dale,V.H.,Langeveld,J.W.A, & McBride, A(2016)A causal analysis framework for land-use change and the potential role of bioenergy policy, Land Use Policy,59, 516-527.

Vatsavai, R. R. (2015, July). Multitemporal data mining: From biomass monitoring to nuclear proliferation detection. In Analysis of Multitemporal Remote Sensing Images (Multi-Temp), 2015 8th International Workshop (pp. 1-4). IEEE.

Singh, N; Vatsavai, RR; Patlolla, D; Bhaduri, BL; Lim, SJ; ",Uncertainty in Estimation of Bioenergy Induced LULC Change: Development of a New Change Detection Technique. AGU Fall Meeting Abstracts,2015.

Bhaduri, B. L., Sridharan, H., Patlolla, D. R., McKee, J. J., & Vatsavai, R. (2015). Monitoring Landscape Dynamics at Global Scale: Emerging Computational Trends and Successes. International Workshop on Spatiotemporal Computing, Fairfax, VA,2015.

Vatsavai, RR; Bhaduri, BL; Singh, N; (2015) Scalable Algorithms for Global Scale Remote Sensing Applications, AGU Fall Meeting Abstracts,2015.

Publications, Patents, Presentations, Awards, and Commercialization

- Singh N. & Vatsavai R.R. (2014).Unsupervised Land Cover Change Detection Using Biweekly MODIS NDVI Data. AGU Fall Meeting 2014.
- Vatsavai R.R. & Singh N., (2014) Unsupervised Phenological Clustering Technique for Biomass Monitoring. International Workshop on Spatial and Spatiotemporal Data Mining
- Kline, K. L., Singh, N., & Dale, V. H. (2013). Cultivated hay and fallow/idle cropland confound analysis of grassland conversion in the Western Corn Belt. Proceedings of the National Academy of Sciences, 110(31), E2863-E2863.
- Singh, N. (2013, July). Spatio-temporal analysis of cropland changes in US in the last decade. In Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International (pp. 2817-2820). IEEE.
- Chandola V. Large Scale Machine Learning for Massive Remote Sensing Data — A Case Study in Biomass Monitoring. ASPRS Annual Conference, Baltimore, MD, 2013