

Sun Grant Data Management and Analysis at ORNL

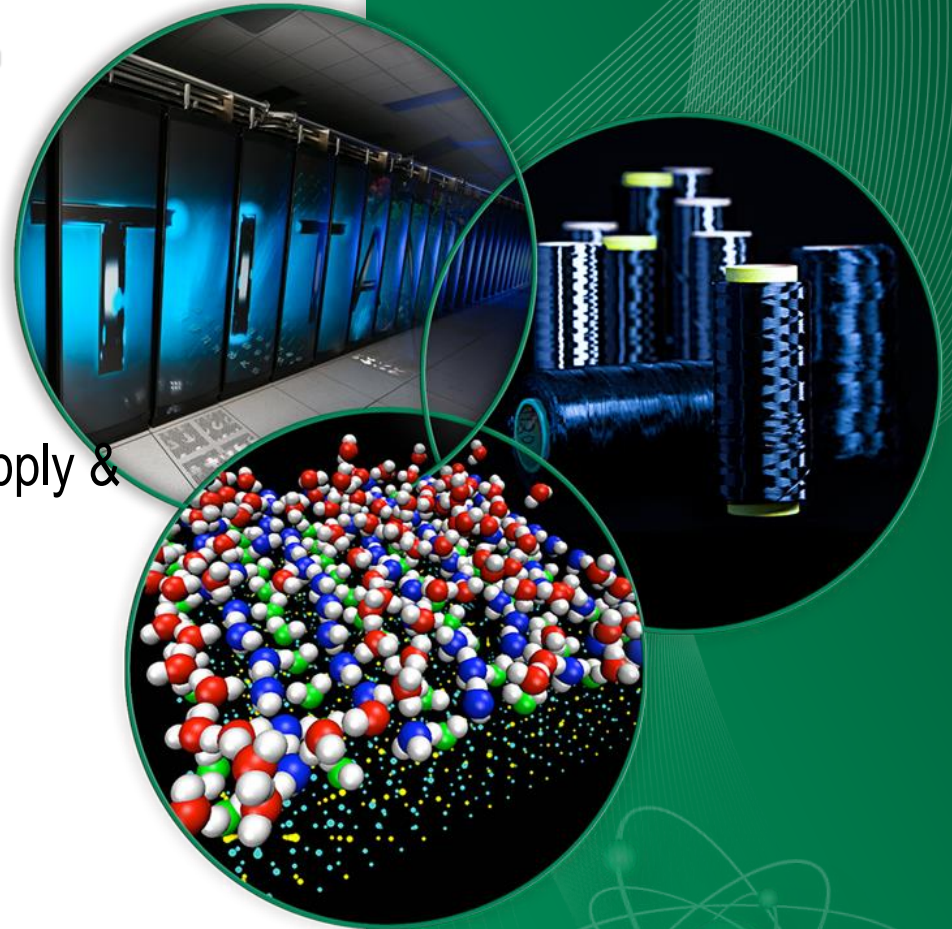
Date: Tuesday, May 21, 11:00-11:30 AM

Technology Area Review: Feedstock Supply & Logistics

Principal Investigator: Laurence Eaton

Team: Aaron Myers, Phil Nugent, Chris Abernathy, Robin Graham

Organization: Oak Ridge National Laboratory



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

Goal Statement

- Provide centralized collection, organization and standardization of data
 - 40+ Institutions/Partners, 5 regional areas, 8 crop species
 - Literature based resource assessment and modeling
- Facilitate data upload, verification, and archiving into the KDF for visualization and analysis
 - Providing consistent, comparable data across species and regions is necessary for a national assessment.
 - Maintain data for assessment across multiple user groups
- Facilitate interaction between field trial partners and resource assessment teams

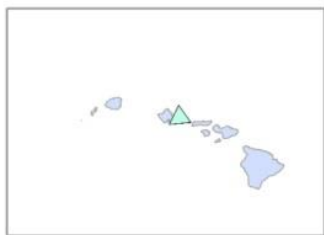
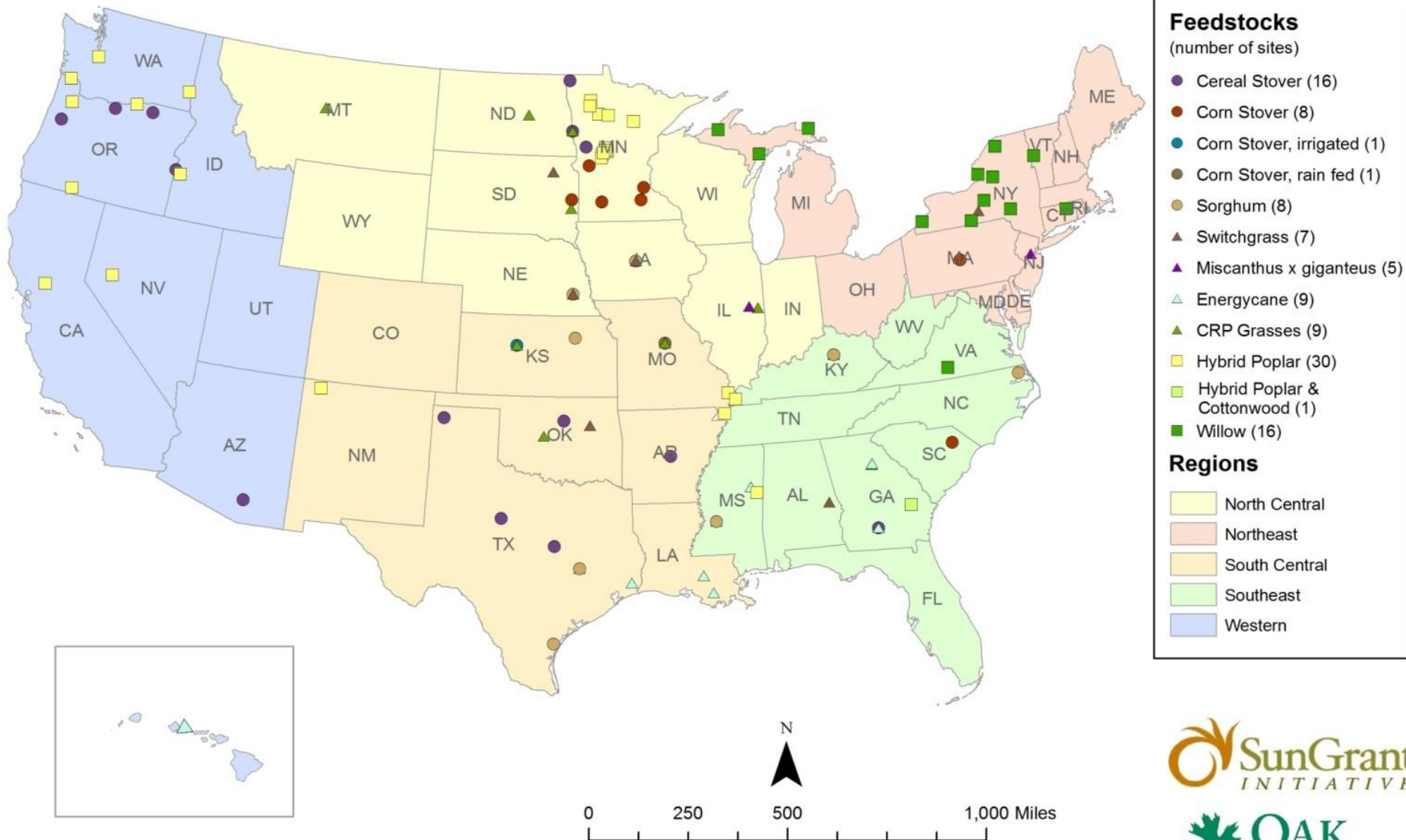
Are Billion Ton Update Report (BT2) crop yield (and corresponding supply) projections validated on the ground?

Quad Chart Overview

Timeline	Barriers
<ul style="list-style-type: none"> • Project start date: FY08 • Project end date: FY14 • Percent complete: 67% 	<ul style="list-style-type: none"> • Feedstock Availability & Cost, Ft-A • Sustainable Production, Ft-B • Feedstock Genetics & Development, Ft-C
Budget	Partners
<p>Total project funding: \$300,000</p> <ul style="list-style-type: none"> – DOE share 100% <p>Funds Received</p> <ul style="list-style-type: none"> – FY11: \$300,000 – FY12: \$0 – FY13: \$0 (included in Resource Analysis 1.6.1.3, \$150,000) 	<p>Collaborators: Sun Grant Initiative, Land Grant Resource Assessment Centers (Oregon State Univ., UofTennessee, South Dakota State Univ., Cornell, Oklahoma State Univ., USDA-ARS, INL ORNL provides all project management and oversight</p>

Sun Grant Initiative Regional Feedstock Partnership Field Site Locations

Current as of October 25, 2010



SunGrant
INITIATIVE

OAK
RIDGE
National Laboratory

Disclaimer: This map is intended for visual representation only. Many field trials occur within the same research location and may not be indicated on the map. Users of this information should contact the Department of Energy Golden Field Office for additional data information.

Project Overview & history

- Initiated with SGI Regional Field Trial Partnership to centralize collection of yield data and make public data and maps
- Facilitate activities between Field Trail Partners and Regional Resource Assessment Teams
 - Field Trials: 40+ Institutions/Partners, 5 regional areas, 8 crop species
 - Resource Assessment Teams (formerly GIS Teams): 5 Institutions, Multidisciplinary (ecologists, soil scientists, agricultural economists, among others)
 - South Dakota State University, University of Tennessee, Oklahoma State University, Oregon State University, Cornell University

Project Overview & history

- **FY11-12 Developed online tools to provide verification and group distribution of data**
 - **Populated Yield Database within KDF**
- **Created historical database of yield trials**
 - **Collected from Regional Resource Assessment Teams**
 - **Synthesized by ORNL and OSU in predictive yield model**
- **Integrate SGI field trial data into predictive model to create yield projections**

Approach

- Develop PI-friendly standard Excel templates for data collection for each of the broad feedstock types
- Develop method to ensure data security and satisfy PI needs while enabling data capture and access
- With KDF computer scientists develop data model for queryable database for Regional Partner field data (based on data entered into templates).
- Through regular calls/e-mails gather data from regional partners using templates. Attendance and presentation at all annual Sun Grant meetings
- QA gathered data
 - In some cases reentered data into templates
 - Confirm/correct data units
 - Identify missing data
- Annual Report to HQ on success of data collection – Quality, quantity, challenges
- Developed Data Synthesis Plan for field trial data using an environmental model

Technical Accomplishments/ Progress/Results (1 of 4)

- Received and processed substantial amount of field trial data
- From 2011 Peer Review, report of **specie x location x year** contributions to SGI Field Trial Database
 - 120% Increase in number of field trials reporting

Feedstocks	Pre-2008	2008	2009	2010	2011	2012	Grand Total
Cereal Stover	0	18					18
Corn Stover	5	10	12	11			38
CRP	0	6	6	6	6	6	30
Energycane	0		7	7	8		22
MXG	0	4	4	5	5	5	23
Poplar	50	11	15	15	17	27	135
Sorghum	0	6	8	7	7		28
Switchgrass	0	7	8	8	5		28
Willow	7	3	8	6	5	1	30
Grand Total	62	65	68	65	53	39	352

* Data is stored and verified in KDF

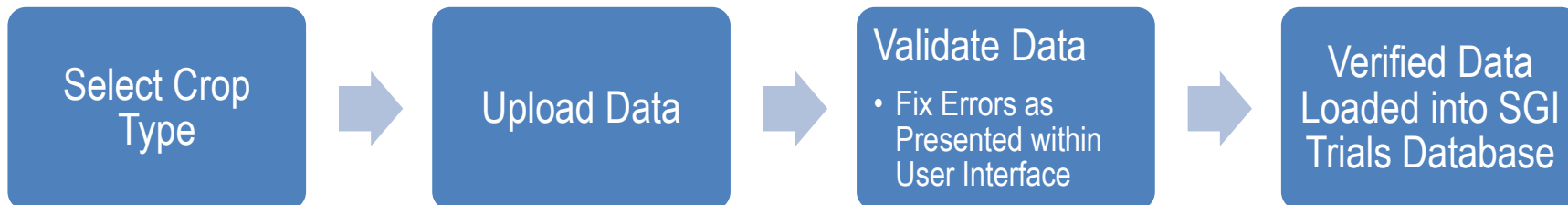
Technical Accomplishments/ Progress/Results (2 of 4)

- Developed and released **KDF Data Validation Tool** to ingest raw data and screen for data type and calculation errors
- Identified ~3500 usable mature yield observations from submissions (as of March 2012)
- **Collaboration Community** For Researchers

The screenshot shows the Sun Grant Initiative website interface. The top navigation bar includes 'Overview', 'Tools & Apps', 'Map', and 'Bioenergy Library'. The main content area features a 'Sun Grant Initiative' header with a 'Welcome' message and a 'Sun Grant Initiative' logo. Below this, there is a 'Sun Grant Initiative' sidebar with links like 'Contribute Field Trial data', 'Download Templates', '72 members', 'Manager: admin', 'My membership', and 'SGI Home'. The main content area displays 'Agronomic Data' with a table of data. The table has columns for 'Moisture Content', 'Dry Weight', 'Brix', 'Grain Yield*', 'Plant height', 'Days to Flowering', 'Lodging', 'Disease Rating*', and 'Im Rat'. The data is organized into rows, with some rows highlighted in green and others in red.

	G	H	I	J	K	L	M	N	O
	Moisture Content %	Dry Weight kg/ha	Brix %	Grain Yield* kg/ha	Plant height cm	Days to Flowering days	Lodging %	Disease Rating*	Im Rat
7	9236.4787124168			0.0					
8	17310.188531703			0.0					
9	11719.817915364			212.4					
10	56.9	18886.448691125		0.0					
11	64.0	14197.410053169		0.0					
12	68.9	18124.367318972		3780.2					
13	55.0	24562.074414337		0.0					
14	57.9	9307.3158849585		0.0					
15	70.1	14761.218028113		2455.2					
16	69.7	17076.837054311		3507.8					
17	72.0	13943.787596838		0.0					
18	63.4	14639.813774924		0.0					
19	60.7	10137.675606317		0.0					
20	61.4	24226.09731028		0.0					
21	65.6	17225.569168441		0.0					
22	70.7	10823.394796429		0.0					
23	66.8	10000.965927704		1888.7					
24	67.4	17625.166632746		2898.5					

Online Tool Workflow

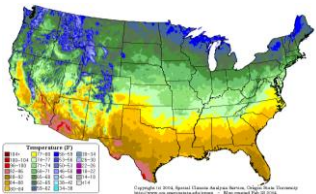


Technical Accomplishments/ Progress/Results (3 of 4)

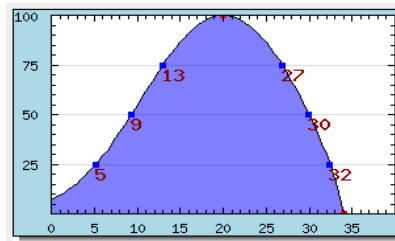
- Generated Draft Maps (July 2011) of Energycane, Switchgrass, Miscanthus, Sorghum (Oregon State University)
- Based on 30-year climate, predicted at 4km resolution



SSURGO Soil Maps

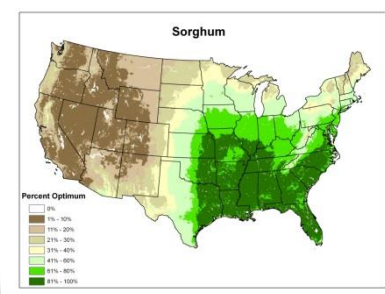
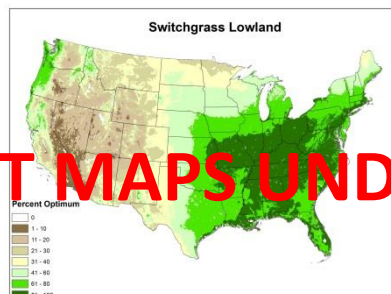
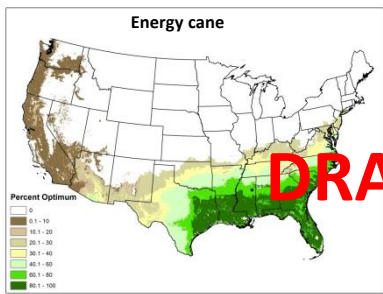
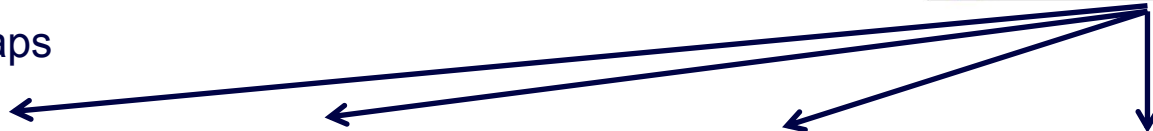
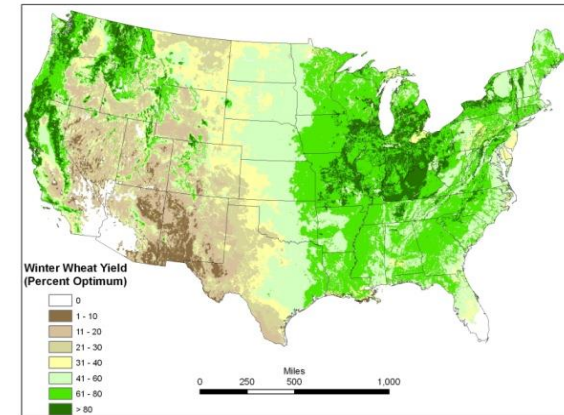


PRISM Climate Maps



Crop
Growth/Environmental
Model

Percent of Maximum Yield

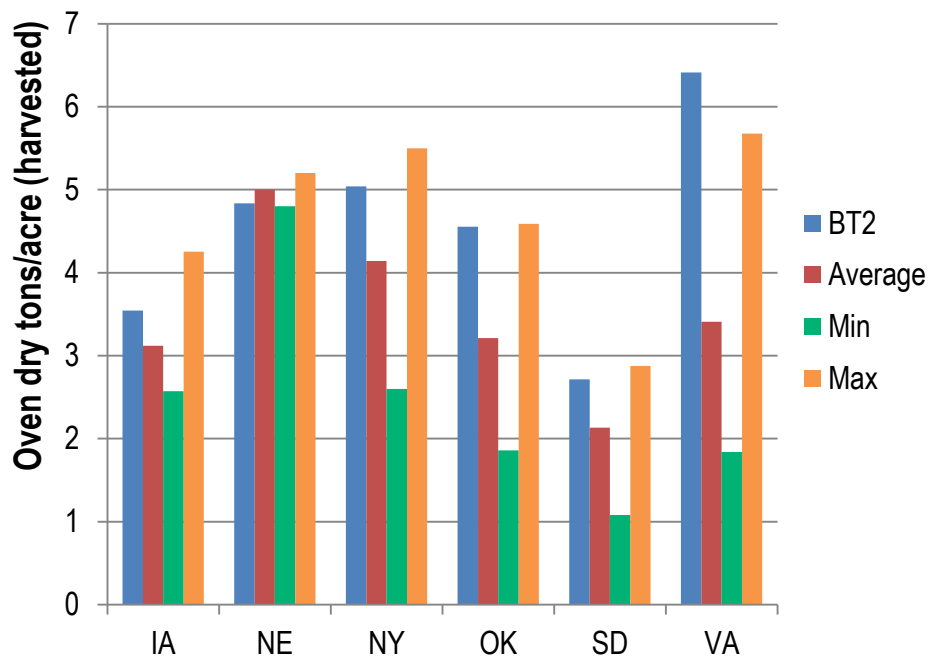


DRAFT MAPS UNDER REVIEW

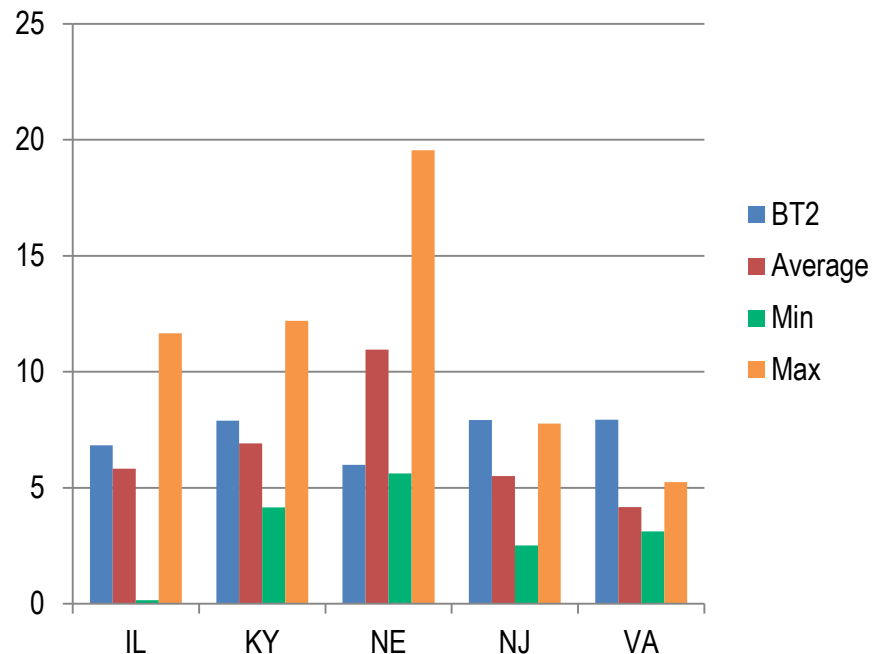
Technical Accomplishments/ Progress/Results (4 of 4)

- Comparison of SGI (Field Plot) and BT2 (County-level) Yields
- BT2= Composite Perennial Grass Category

Switchgrass*



Miscanthus**



* Average across 2009-2011 Production, all N treatments

** Average across 2010-2012 Production, all N treatments (includes non-mature years)

Relevance

Supports accurate characterization and archiving of regional feedstock production potential and variation (spatial and temporal) across the US

Supports cross-comparison of feedstock potential within a region.

Enables creation and validation of feedstock production models necessary for estimating future resources and guiding environmental sustainability

Access to credible data on biomass production and environmental characteristics reduces investor risks and promotes biorefinery establishment

Critical Success Factors

- Data Management
 - Verification of data consistency and uniformity
 - Communication of data received, expected dates, delays, errors
 - Willingness to address PI concerns for data security
- Data Distribution- Ongoing engagement with complimentary research platforms (e.g. Sustainability, Resource Assessment, Logistics, etc)
- Project Challenges

Challenges	Solutions
Data security	One POC; submission protocols
Release of data by PI	PI-relevant templates; regular communication with team leads
QA of data	Communication with PIs; data upload controls in the KDF
Relevance to region and national feedstock analysis	Engage Field Trial Partners in data model calibration and yield mapping scenarios

Future Work

- Expert work groups to bring **Case Study Yield Maps** from Draft to Production Level (to be completed by end of CY13)
 - Will occur with each team in face-to-face meetings
 - Participants to identify key management assumptions (region-specific Best Management Practices) to identify potential yields of field trial sites and accurate spatial distributions
 - Includes all Energy Crop teams
- Collect final Yield and Sustainability Data from Field Trials (measured at conclusion of field trials)
- Write-up yield map process, management assumptions
- Lead Write-up of Final Report incorporate national assessment into Resource Assessment work (1.6.1.3)

Summary

- This project ensures that the data generated by each regional partner are captured, documented and made available for regional and national-level syntheses.
- Has delivered data to KDF to accurately document DOE investment in Field Trial Partnership
- This project supports the EERE's goals of sustainable competitive biofuels, provides validation of biomass yield models, and will create a geospatial database to contribute to the national biomass assessment.
- The future will focus on analysis of four years of field trials and a literature compendium spanning 10 years of research.
- Revise Yield and Cost estimates for

Response to Previous Reviewers' Comments

- Data standards and minimum data sets lacking, Needs library sciences connection
 - KDF Data Tool is developed for data curation and documentation. The data received is much larger than minimum data sets initially established, and all records are retained that have been provided from field trial partners.
 - The KDF database is connected to the INL Biomass Library for physical trait characteristics. Additionally, the KDF templates are being used for the development of the USDA-ARS REAPnet database. Current efforts are exploring linkages between the two.
- Are geospatial tools enabling “visualizations” accurate/representative of what can be achieved on ground?
 - The Case Studies activities in FY13 will address these concerns through rigorous discussions with specie experts through Yield Mapping workshops.
- Analysis of potential fuel volumes is ignorant of sustainability metrics
 - These indicators have been recently developed (FY11-12) and, after collected, will be available to the DOE Sustainability Research Platform. To the extent that funding allows, more data may be developed and integrated in the Resource Analysis activities.

Presentations and Publications

- 2011 and 2012 Data Management Update at Annual Working Group Meetings
- Eaton et al. Updates from the SGI Resource Assessment Teams. 2012 SGI National Conference.
- Eaton, LM and Halbleib, M. Historical Yield Database from SGI Resource Assessment Teams, 2012 ORNL Technical Manuscript.
- Abernathy, CR and LM Eaton. SGI Data Synthesis Plan. 2011 ORNL Technical Manuscript.

Evaluation Criteria

- Relevance
 - Supports data collection and distribution for sustainable bioenergy production
- Technical accomplishments
 - Developed data collection model
 - Yield synthesis model to be refined in current FY with entire SGI participation
- Relevance
 - Field Trials needed to increase information on sustainable production and bioenergy resource assessment
- Critical Success Factors
 - Data uniformity and security
- Future Work
 - Yield Mapping Case Studies to synthesize field trial results for national assessment
- Technology Transfer
 - Integration of data within KDF allows for archival and dissemination of data when appropriate.

Acronyms Used

- ORNL – Oak Ridge National Laboratory
- BT2 – Billion Ton Update Report (DOE, 2011)
- INL – Idaho National Laboratory
- SGI – Sun Grant Initiative
- KDF – Knowledge Discovery Framework
- ARS – Agricultural Research Service
- USFS – U.S. Forest Service
- EERE – U.S. DOE Energy Efficiency & Renewable Energy