UAV-based Image Analysis and Machine Learning for Bioenergy Crop Modeling and Precision Agriculture

Yaping Xu, Ph.D.
Plant Sciences, University of Tennessee, Knoxville
04/21/2023

THE CENTER FOR BIOENERGY INNOVATION

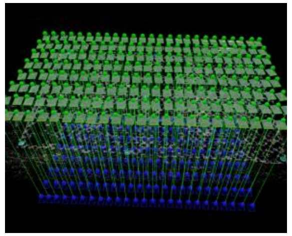
About Me

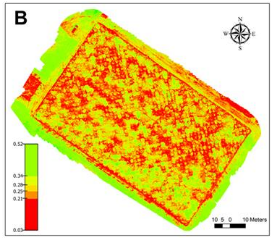
- Postdoc Research Associate, Plant Sciences, University of Tennessee, Knoxville

- Early Career Fellow, Center for BioEnergy Innovation CBI), Oak Ridge National Lab

- Assistant Professor, Environmental & Geoscience Sam Houston State University (starting from 08/14/2023)

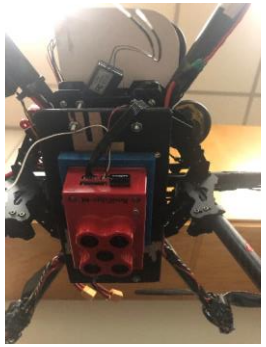
How does UAV System Work





Mission Planning → Data collection → Photogrammetry → Mapping → Data Analysis

multispectral

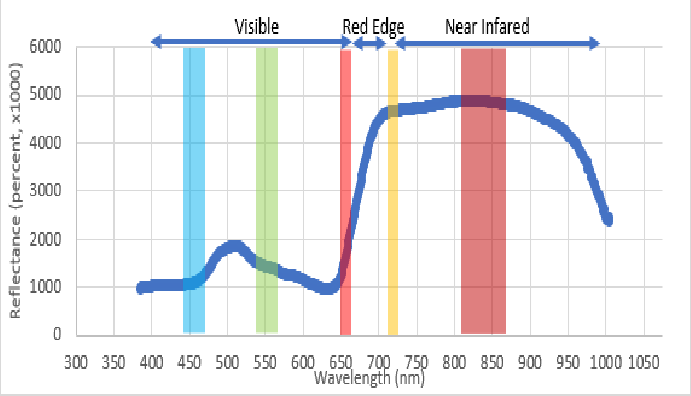


hyperspectral

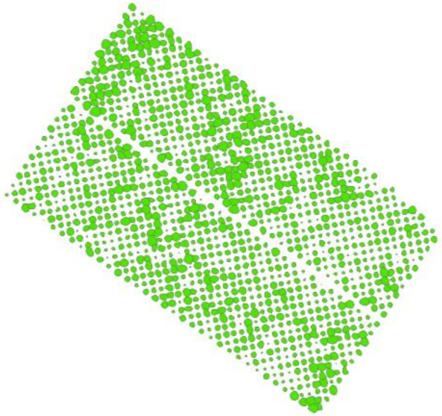
LiDAR

What information do we "mine" from the UAV data?

Spectra

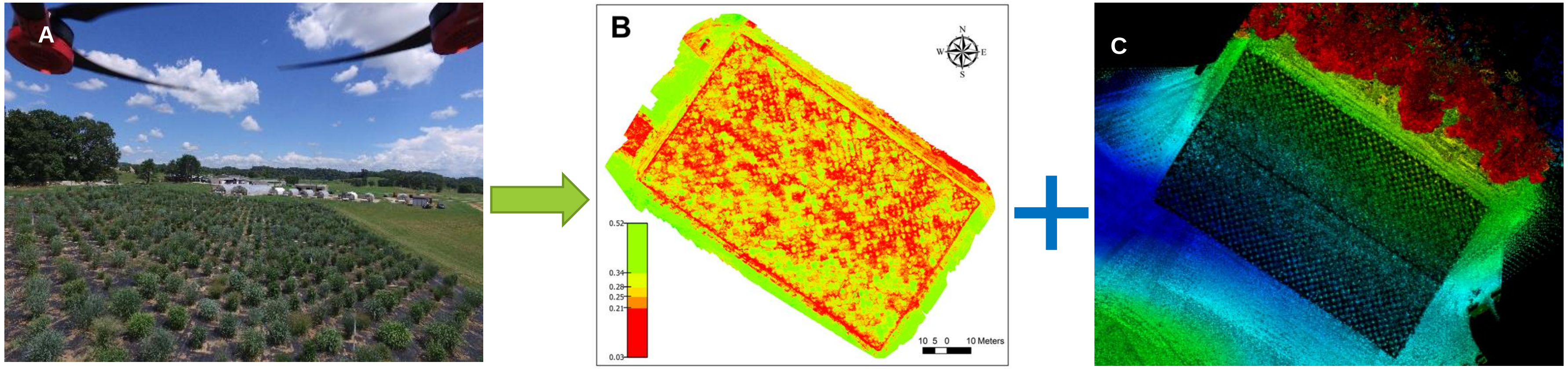


Feature



Overview: the Need of High-throughput Phenotyping and Crop Breeding for Bioenergy Feedstock

- Sustainable Aviation Fuels (bio-jet fuel): expected to reach 30 billion gallons per year by 2040.
- Switchgrass (*Panicum virgatum* L.) is a perennial grass with great potential as a bioenergy feedstock.
- Measuring the traits is labor-intensive and costly.
- UAV-based RGB and hyperspectral/multispectral imagery
- Explored the feasibility of applying UAV remote sensing to the high-throughput phenotyping for switchgrass.
- Four traits: leaf chlorophyll, nitrogen, lignin content, and rust disease (Xu et al., 2021).



UAV with multispectral camera. A) DJI Matrice 600 pro with MicaSense RedEdge-M multispectral camera. B) Vegetation index map of the field calculated from the multispectral imagery. Green color represents higher vegetation index value compared to red color; C) Digital surface model from UAV-based LiDAR sensor; D) UAV map of the field.

Research Objectives and Questions

- Objectives
- To leverage spectral information for modeling the four traits
- To provide scientific basis for a sustainable field-grown switchgrass
- To test the feasibility of automated phenotyping for precision agriculture and plant breeding
- Questions

(1) Can we model the leaf chlorophyll, nitrogen, lignin, and rust disease with UAV multi-spectral data?

(2) Which vegetation index developed from the UAV multispectral data performs best?

(3) Can we detect the rust disease from hyperspectral imagery?

Data Collected from Switchgrass GWAS 2020 in Field

- **UAV:** weekly flights
- **Chlorophyll Measurements (mg/L)**

- measured chlorophyll content of the leaves using Opti sense handheld device

- **Rust Severity Score (percentage)**

- scored the plant rust as percentage ranging from 0-100% where 0 indicates no rust and 100% indicates severe rust.

- **Nitrogen and lignin quantification (percentage)**

- 600 plant samples; 150 accessions grown under Low N (2 replicates), 150 grown under Mod N (2 replicates)

- two tillers containing both stem and leaves were collected from each plant

- samples were oven dried at 45°C for 72 hours, then milled

- nitrogen and lignin quantification were done at Noble Research Institute, Oklahoma

UAV Data Time-series:

Rust	Chlorophyll	Nitrogen and Lignin
07/29/2020 (5-day later)	08/26/2020 (1-day later)	07/29/2020 (6-day earlier)
08/12/2020* (6-day later)		11/05/2020 (5-day earlier)
08/26/2020 (5-day later)		
10/15/2020 (1-day earlier)		

Ground Data Time-series:

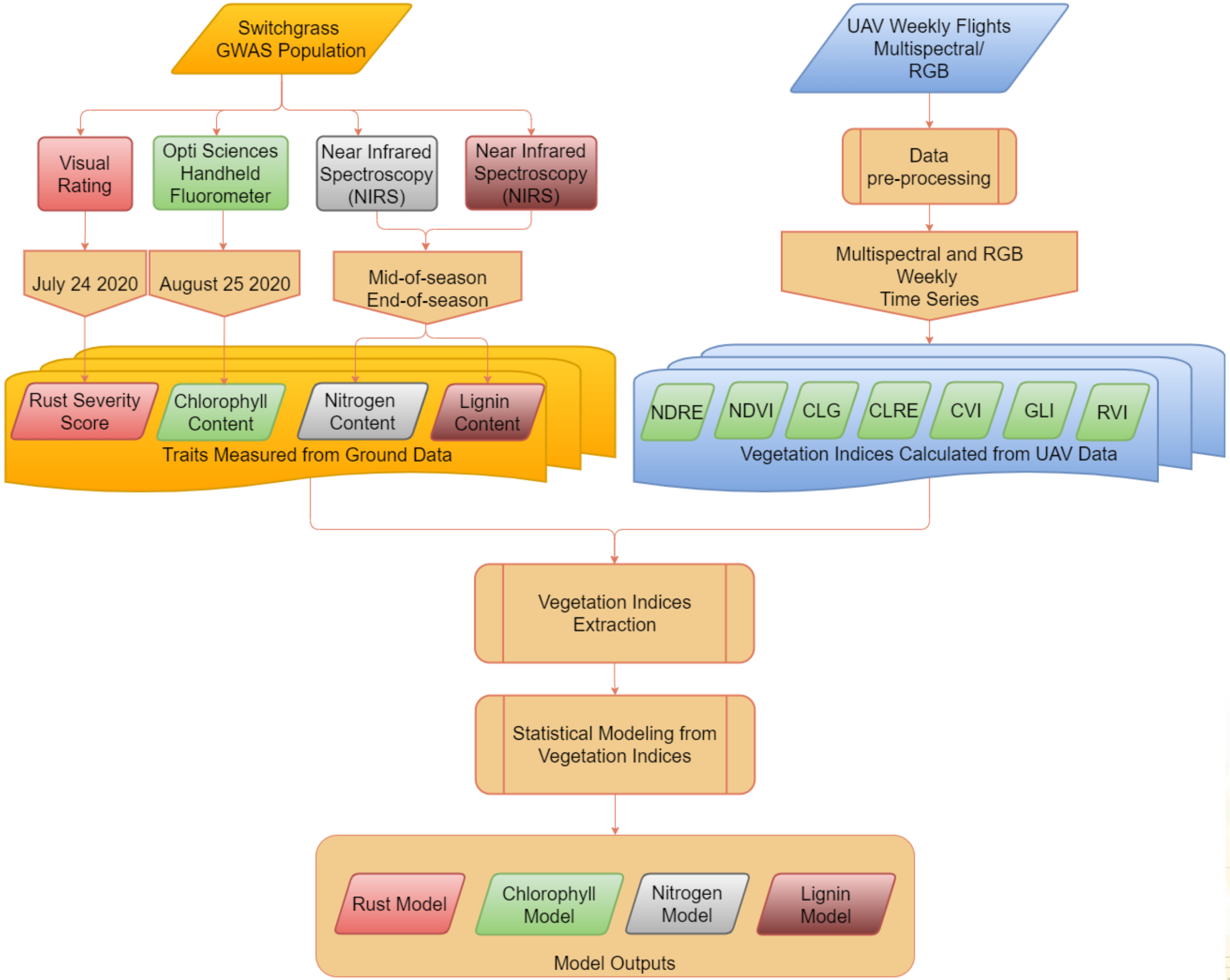
Rust	Chlorophyll	Nitrogen and Lignin
07/24/2020	08/25/2020	08/04/2020
08/06/2020*		11/10/2020
08/21/2020		
10/16/2020		

UAV-based Modeling Methods for Rust Disease, Chlorophyll, Nitrogen, and Lignin Content

Data processing and analysis pipeline for switchgrass sustainability traits modeling:

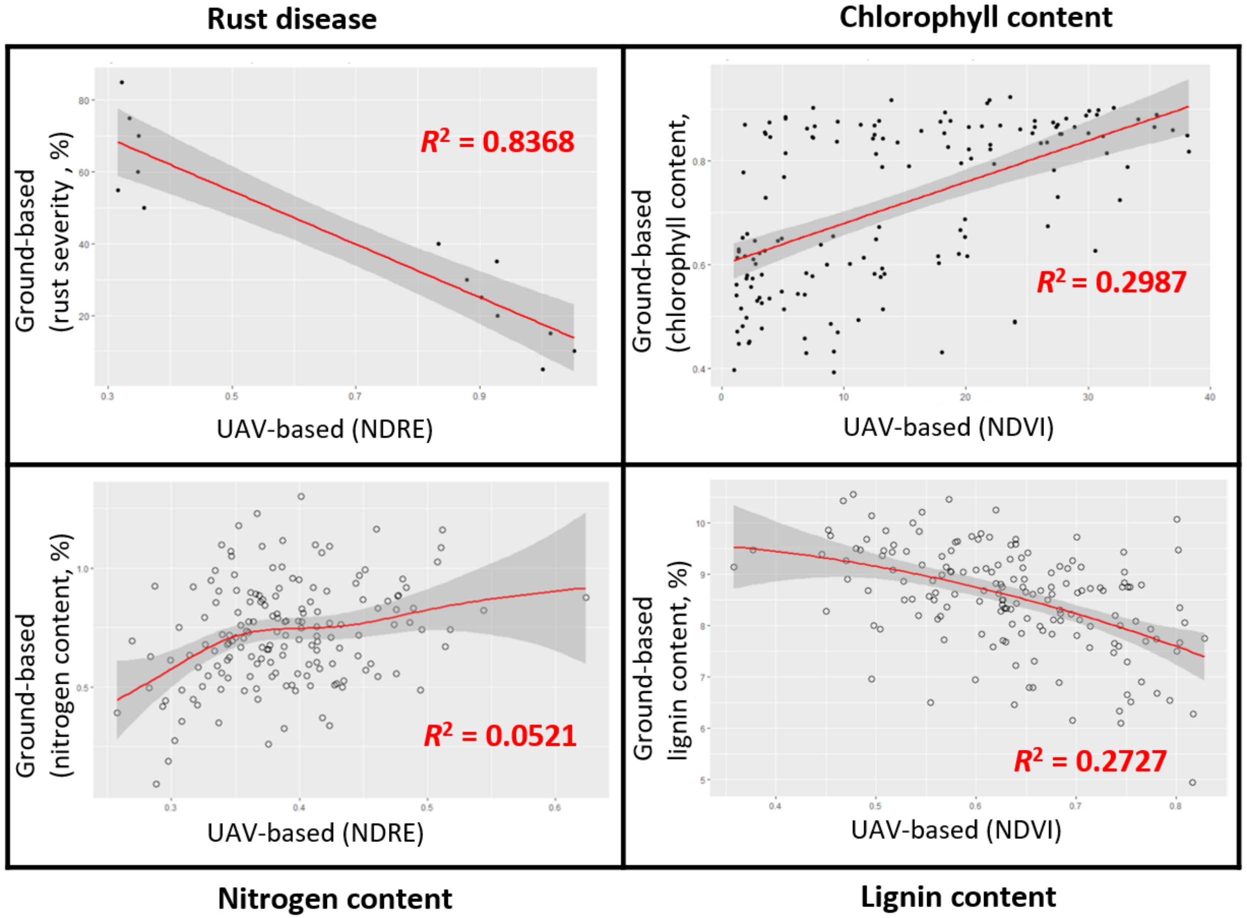
- We have ground data collected from different time points
- We have processed the UAV data from the corresponding dates. We calculated the vegetation index, including 7 vegetation indices
- We then build 4 statistical models for the four traits, one model for each trait.

NDRE: normalized difference rededge
NDVI: normalized vegetation index
CLG: chlorophyll index green
CLRE: chlorophyll index rededge
CVI: chlorophyll vegetation index
GLI: green leaf index
RVI: ratio vegetation index



Modeling Results for Rust Disease, Chlorophyll, Nitrogen, and Lignin Content

- The normalized difference red edge (NDRE) vegetation index outperforms other indices for rust and nitrogen, while NDVI performs the best for chlorophyll and lignin.
- Linear models work well for rust disease, chlorophyll and lignin.
- For nitrogen, non-linear models outperform linear models, but these regression models did not perform well (Xu et al., 2021)
- Machine learning approaches have potentials for modeling the nitrogen content



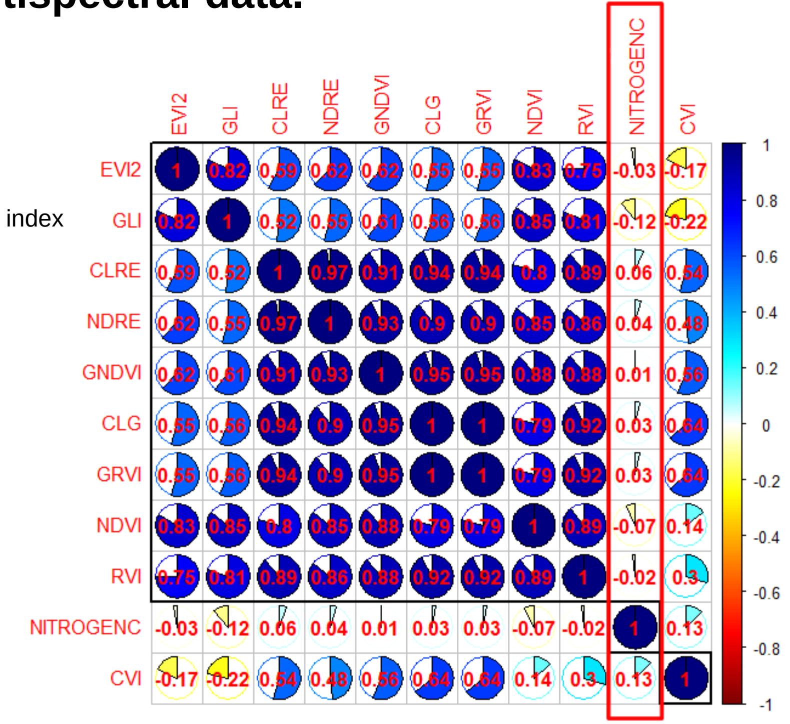
High-throughput modeling results (Xu et al., 2021)

Ongoing Work: Tiller Nitrogen Content Estimation with UAV Data – Machine Learning Methods

Advanced big data analytics with data collected from the second growing season (2020) were used to improve the nitrogen content model

- **(a) Incorporating more vegetation indices into the model. 10 Vegetation indices were calculated from multispectral data.**

GLI: green leaf index



Correlation between **nitrogen content (NITROGENC)** and the ten vegetation indices calculated from multispectral imagery.

- **(b) Machine learning approaches were used to model the nitrogen with the vegetation indices. End-of-season correlation plot shows the importance of each variable.**

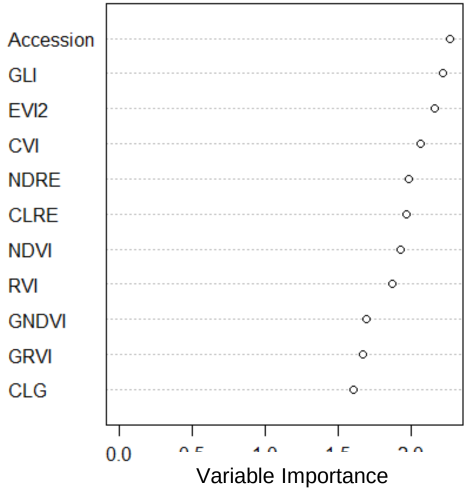
  - Random forests (RF) compared with conventional (regression-based) models is much more powerful to model the nitrogen content
  - Accession is extremely important for nitrogen modeling

Response Variables (Y) : Nitrogen content

~

Predictor Variables (X):

- Accession
- CLG
- CLRE
- CVI
- EVI
- GLI
- NDRE
- GNDVI
- GRVI
- NDVI
- RVI



Vegetation indices and their importance to machine learning algorithms. Figure shows the variable importance measured from random forest machine learning algorithm.

Clustering the Samples by Accessions

- **Cubist model: rule based modeling**

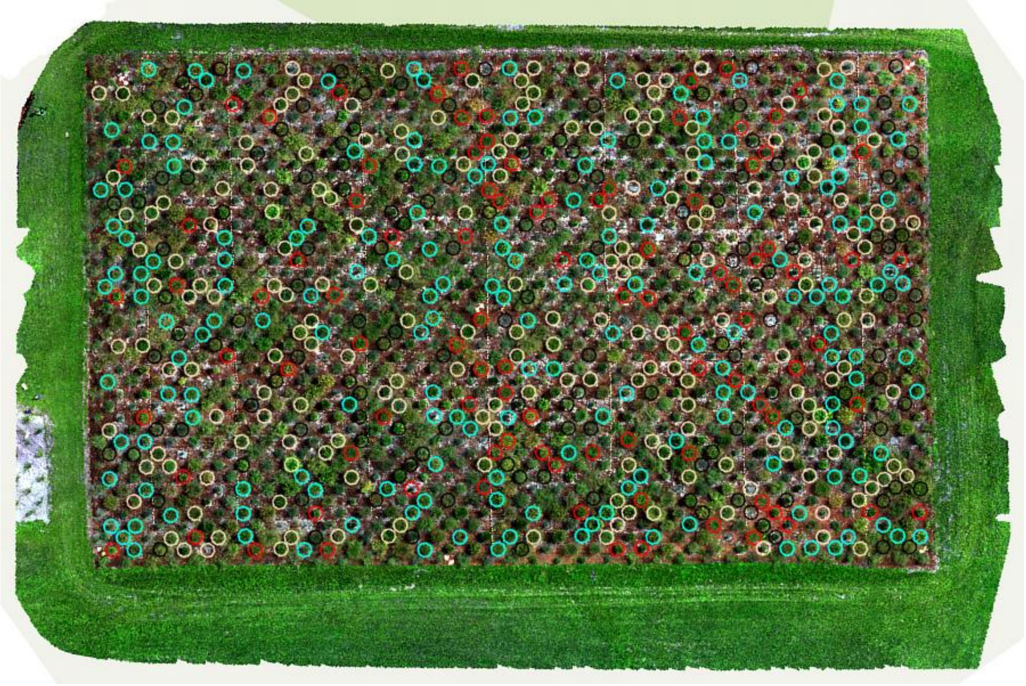
Rules applied: aggregating the accessions (e.g., J237.A) by nitrogen content

Group 1: [24 accessions, 96 plants, mean 0.86, range 0.48 to 1.16, est err 0.12]
Accessions in {J237.A, J587.B, J482.B, J275.A, J240.A, J296.A, J504.A, J280.A, J206.A, J065.A, J312.A, J483.B, J610.C, J424.A, J469.C, J339.A, J484.A, J319.A, J270.A, J008.A, J481.C, J294.A.B1, J308.A, J230.A}
Nitrogen Content = 0.85 - 0.2 **GLI** + 0.11 **EVI**

Group 2: [50 accessions, 200 plants mean 1.02, range 0.73 to 1.45, est err 0.10]
Accessions in {J324.A, J301.A, J465.A, J514.A, J337.A, J329.A, J321.A, J501.A, J331.A, J496.C, J251.C, J235.A, J018.A, J086.B, J441.A, J421.A, J028.C.B1, J020.B, J330.A, J186.A, J236.A, J001.A, J013.C, J219.A, J500.A, J216.A, J009.C, J182.A, J018.C.B1, J188.A, J313.A, J016.C, J065.B, J460.A, J231.A, J243.A, J189.A, J004.B, J041.A, J288.B, J484.C, J482.C, J320.A, J303.A, J498.B, J212.A, J458.B, J466.A, J503.C, J315.A}
Nitrogen Content = 1.01 - 0.19 **GLI** + 0.1 **EVI**

Group 3: [44 accessions, 176 plants, mean 1.15, range 0.63 to 1.52, est err 0.13]
Accessions in {J016.A, J327.A, J317.A, J013.B, J215.A.B1, J594.A, J504.C, J614.B, J022.D, J222.A, J241.A, J008.D, J022.A, J016.B, J323.A, J251.A, J502.A, J460.B, J615.A, J016.D, J464.B, J218.A, J019.A, J503.A, J305.A, J211.A.B1, J416.A, J500.B, J456.C, J311.A, J483.C, J177.A, J326.A, J497.C, J279.A, J026.B, J271.A, J023.B, J466.B, J394.C.B1, J005.B, J190.A, J268.A.B1, J226.A}
Nitrogen Content = 1.15 - 0.21 **GLI** + 0.09 **EVI**

Group 4: [28 accessions,144 plants, mean 1.29, range 0.81to 1.74, est err 0.15]
Accessions in {J272.A, AP13, J293.A, J500.C, J208.C, J073.B, J336.A, J011.B, J497.A, J274.A, J419.A, J461.C, J250.B, J003.E, J535.A, J587.A, J193.A, J229.A, J499.C, J610.A, J037.A, Performer.TCE7, J441.B, J340.A, J005.A, B6, J447.A, J030.C}
Nitrogen Content = 1.32



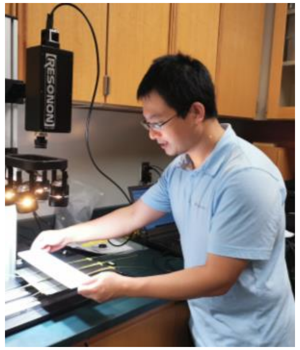
Group 1 Group 2 Group 3 Group 4

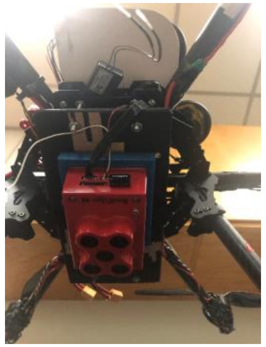
Total accessions measured: 24+50+44+28 = 146 accessions
Total plants: 96+200+176+144 = 616 plants
Evaluation on training data:
Correlation coefficient: 0.53 (vs R-squared 0.05)

Ongoing Work: Rust Disease Rating Using Hyperspectral Imagery and Artificial Intelligence

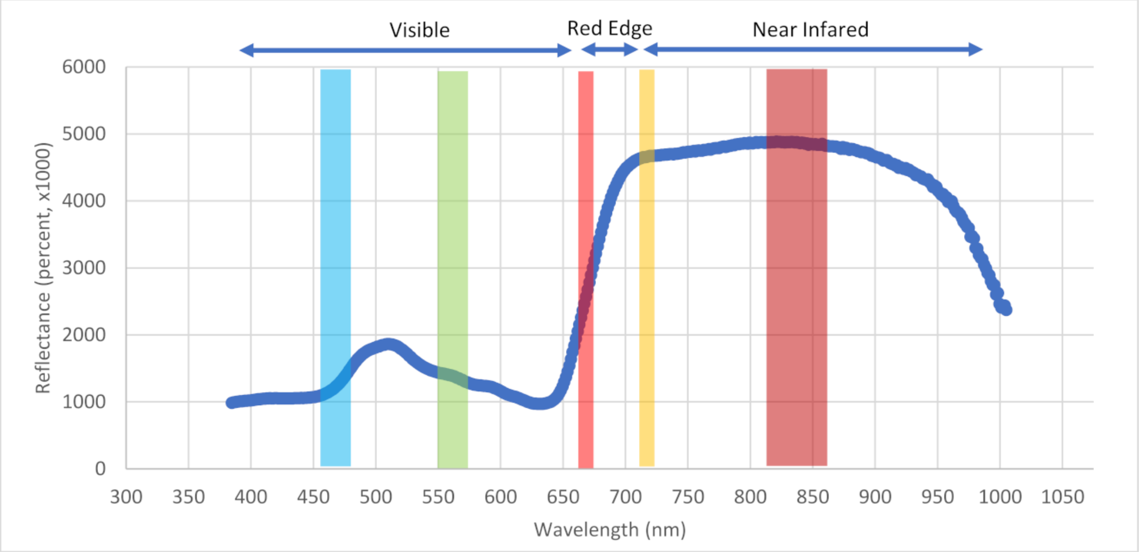
- Experimental design for switchgrass rust disease detection and rating: lab – tripod - UAV

Rust Severity Assessment → Data collection → Image Analysis → AI/ML → Rust Detection





- **Reasons for this "lab – tripod – UAV" integrated approach**:
  - downselect the spectral signature
  - less spectral variability in the lab
  - higher resolution on a tripod
  - upscaling the method with UAV once we identify the target
- Hyperspectral data will be compared with the multispectral data for rust detection.



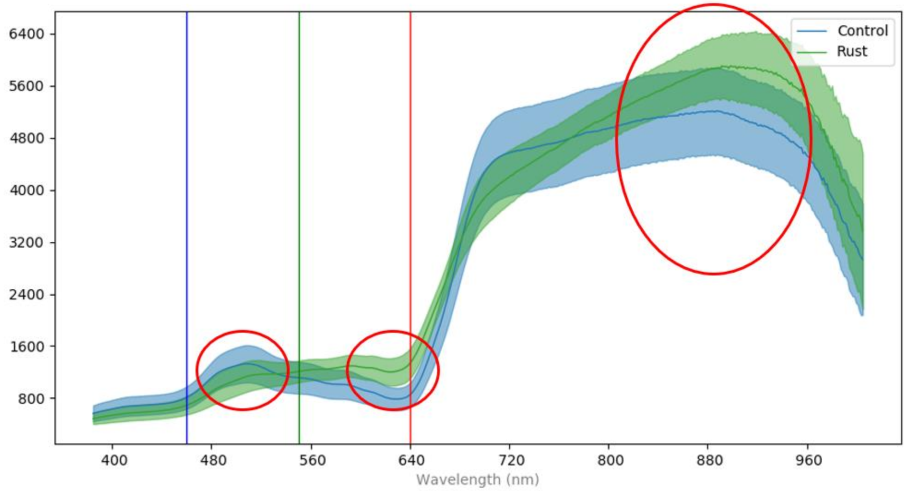
Band Number	Band Name	Wavelength (nm)
1	Blue	465–485
2	Green	550–570
3	Red	663–673
4	Near Infrared	820–860
5	Red Edge	712–722

Spectral responses for healthy switchgrass leaves (collected from the Resonon Pika XC2 hyperspectral camera). The five ribbons, from blue in the left, to red in the right, indicate the band locations of the MicaSense multispectral camera, as compared to the full spectral range of Pika XC2

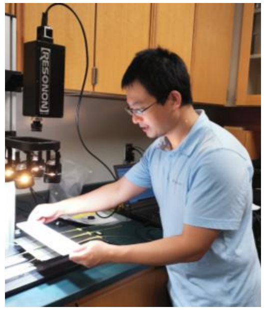
Rust Disease Rating: progress in the lab

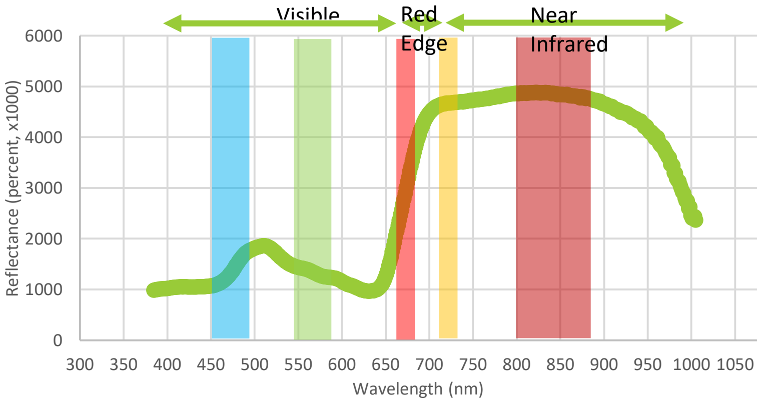
- Collected healthy and rust-infected leaf samples.
- Used Resonon Pika XC2 hyperspectral camera benchtop system.
- hyperspectral camera to scan the leaves for spectral reflectance.
- Spectral data classification analysis with spectral data processing software.
- Spectral data descriptive analysis with spectral data processing.

Benchtop hyperspectral scan, areas affected by rust disease are shown in yellow or orange

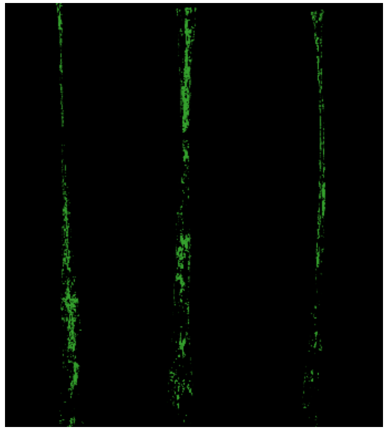


The green color and blue color lines represent the average of the spectral reflectance from the rust and healthy leaves, respectively. Imagery is pixel-based, and lines show an average of several pixels in the areas affected by rust and the healthy areas. Upper and lower boundaries of the lines represent the standard deviation of those pixels selected for the areas affected by rust or the healthy areas. Red circles show the wavelength values where differences may happen between these two groups, control (healthy) and rust.



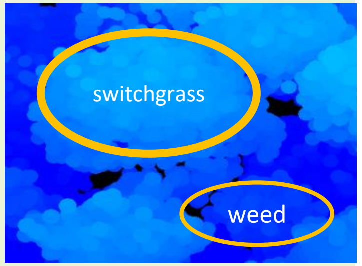
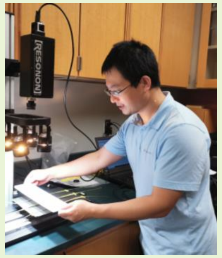
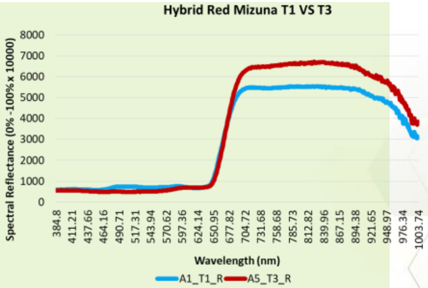


Spectral data collection in the lab and the spectral responses for healthy switchgrass leaves.
The five ribbons, from blue in the left, to red in the right, indicate the band locations of multispectral camera



Spectral Angle Mapper (SAM) for rust disease pixels

Potential Applications Beyond the Four Traits

Platform/Sensor		**Spatial Scale**		**Use**	
Satellite/UAV **multispectral** **LiDAR**		field scale		biomass/yield weed detection crop stress soil moisture precision irrigation	
Outdoor **tripod/gantry** **hyperspectral**		plant scale		canopy temperature disease/pathogen leaf area index water content	
Lab benchtop/ **hyperspectral** **camera**		leaf scale		disease/pathogen leaf expansion rate Photosynthesis rate phytonutrients	

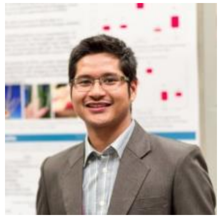
Take Home Messages

- Leaf chlorophyll, lignin content, and rust disease can be modeled with UAV multispectral data
- Nitrogen content modeling is a challenge, leveraging multiple vegetation indices and machine learning can significantly boost the model
- Hyperspectral imagery has great potential for rust disease detection
- A combination of sensors at various spatial scales can provide more opportunities for the applications

Acknowledge

- UTK Stewart Lab






Neal Stewart Reggie Millwood Mitra Mazarei Vivek Shretha Lance Hamilton Lana Howe Holly Brabzon Cassie Halvorson Britany Morgan

- UTK Plant Science and UT AgResearch

Shane Miller, Bryce Trull, Kelly Walters, Heather Kelly

- Center for BioEnergy Innovation, Oak Ridge National Lab







Jerry Tuskun, Katrien Devos, Stan Martin, Anne “Liz” Ware, John Field, Dan Jacobson, Brian Davison

UTAGRESEARCH
INSTITUTE OF AGRICULTURE
THE UNIVERSITY OF TENNESSEE

U.S. DEPARTMENT OF ENERGY | Office of Science

cbi THE CENTER FOR BIOENERGY INNOVATION

cbi

MDPI *Plants* Special Issue

- **Name of the special issue: "Modeling of Biofuel Plants Phenotyping and Biomass"**
- A special issue of ***Plants*** (ISSN 2223-7747) belongs to the section "**Plant Modeling**".
- Deadline for manuscript submissions: **31 December 2023**

Scan this QR code to access

We encourage topics from a data-driven approach, including, but not limited to:

- Perspectives of biofuel plant phenomics;
- Big data challenges for genomics and phenotyping data;
- High-throughput phenotyping: tools and techniques for assessment;
- Genomic selection in biofuel crops: Benefits of high throughput phenotyping;
- Precision agriculture association with high throughput biofuel plant phenotyping;
- Biomass quantity/quality assessment;
- Biotic/abiotic stress assessment;
- Sustainability trait assessment.

Biofuel crops ≠ switchgrass only!
corn, sugarcane, palm oil, cottonseed, sunflowers, wheat, soybean, and more…

cbi

Thanks!
Welcome to reach out to me for questions and collaborations!

Yaping Xu
yxu86@utk.edu
Website (Scan to access): https://sites.google.com/view/yapingxu

cbi