DOE Bioenergy Technologies Office (BETO)
2023 Project Peer Review

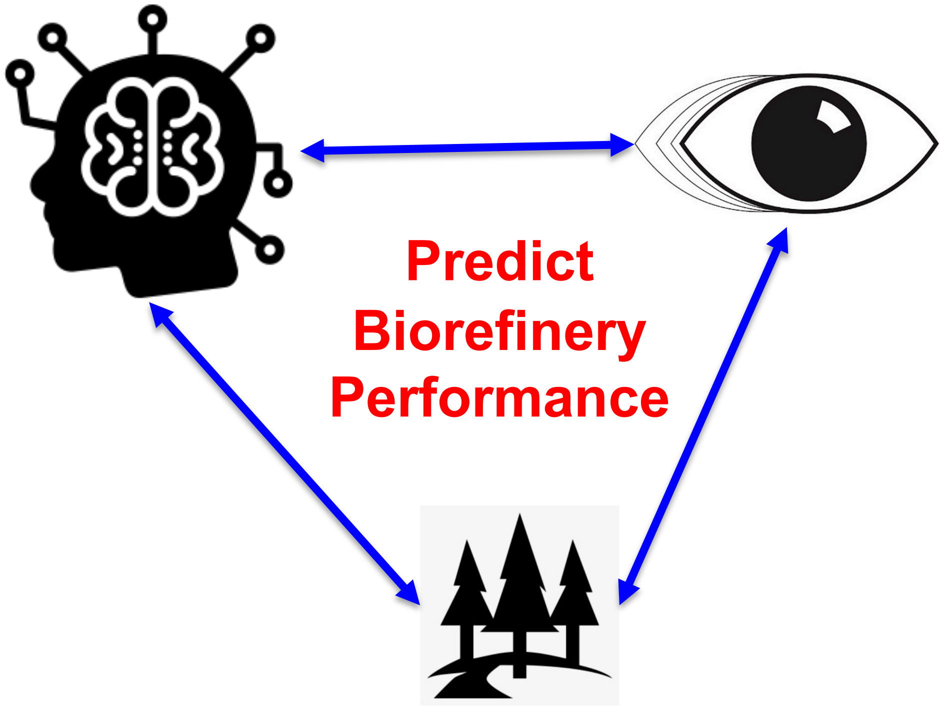
Machine learning based modeling framework to relate biomass tissue properties with handling and conversion performances (1.2.2.104)

April 05, 2023
Feedstock Technologies (FT) Session

Sudhagar Mani
University of Georgia

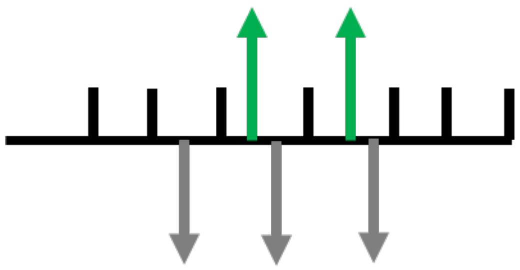
Outline

- Project Overview
- Corn Stover
  - Approach
  - Progress and Outcomes
- Southern Pine Forest Residues
  - Approach
  - Progress and Outcomes
- Impacts
- Summary



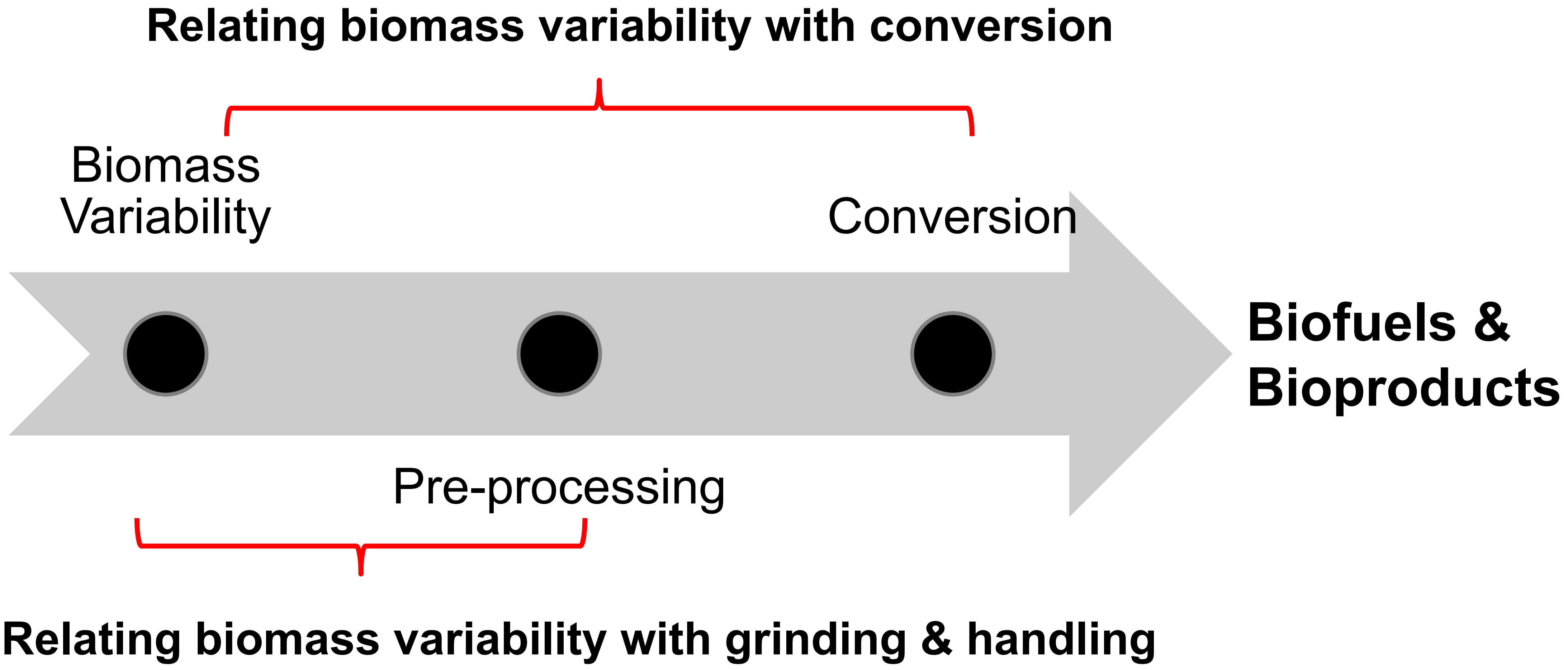
Project Overview

Biomass Variability

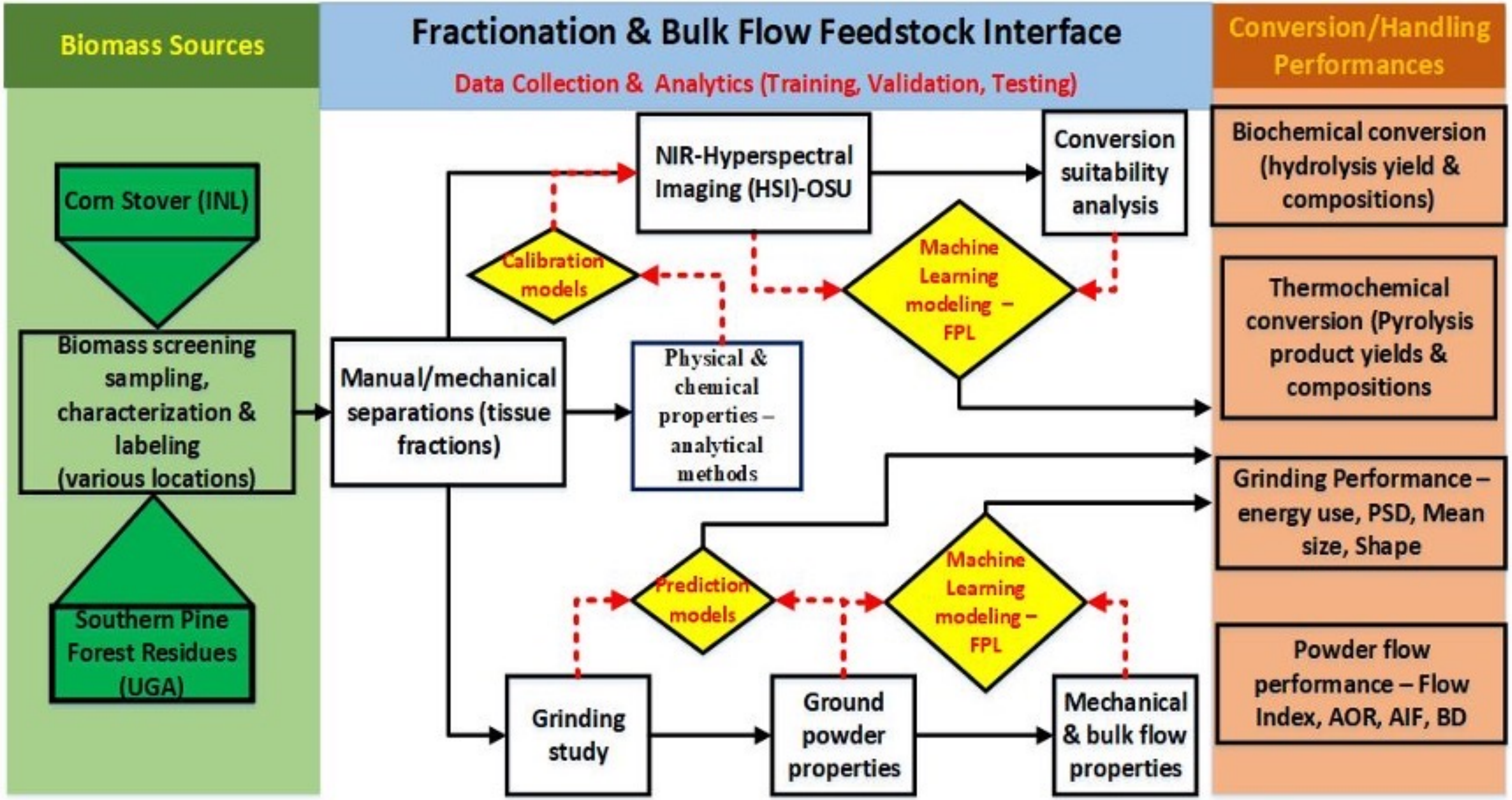


- Variability caused inconsistent product yield and operational predictability
- Moisture and ash content variations have been extensively studied towards preprocessing and conversion
- The (inherent) variability of biomass at the anatomical (tissue) component level is unknown
- Reducing the variability of lignocellulosic biomass to generate **feedstock of right quality standard** is critical for targeted conversion application

Biomass Supply to a Biorefinery



Project Overview



Project Objectives

- (1) to develop a novel tool for rapid determination of individual biomass tissue properties using NIR hyperspectral imaging (HSI) system
- (2) to apply the machine learning framework, especially Artificial Neural Network (ANN) to correlate the biomass chemical compositions with conversion performances
- (3) to investigate the grinding performances and the bulk flow properties of tissue components,
- (4) to use the machine learning modeling framework – ANN to relate physical properties of tissue powders with grinding and bulk handling

1. Approach: Corn Stover

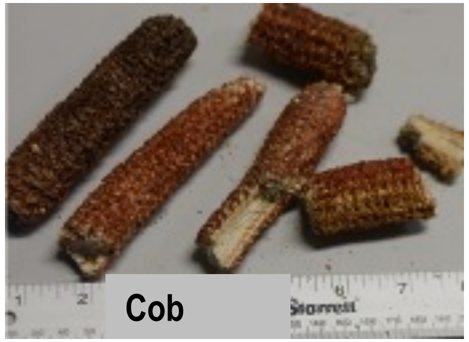
Harvest Variables Examined			
Harvest Year	2018/19	2020	
Harvest Method	Low-Cut	High-Cut	Raked
Harvest Location	Kadolph Field	Field ‘70/71’	
Total of 12 Bales			

Corn Stover Bales

Manual fractionation

12 bales x 5 fractions = 60 samples

- Whole material
- Leaf
- Husk
- Stalk
- Cob

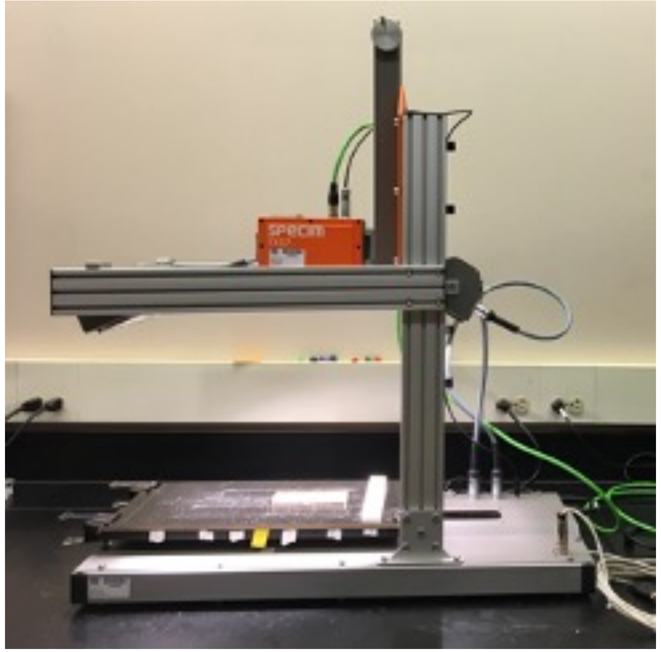


1. **Approach:** Hyperspectral Analysis

Hyperspectral Imaging (HSI) combines spectroscopy and digital imaging

Generates a hyperspectral image (a hypercube) of 3-dimensional multivariate data

Pixels in a hypercube may contain different spectral responses associated with different chemical compositions



Data collection:

Specim FX17 (931-1718 nm)

X-Axis = 321 pixels

Y-Axis = 716 pixels)

- Each pixel has spectra data from 224 wavelengths (Z-Axis)

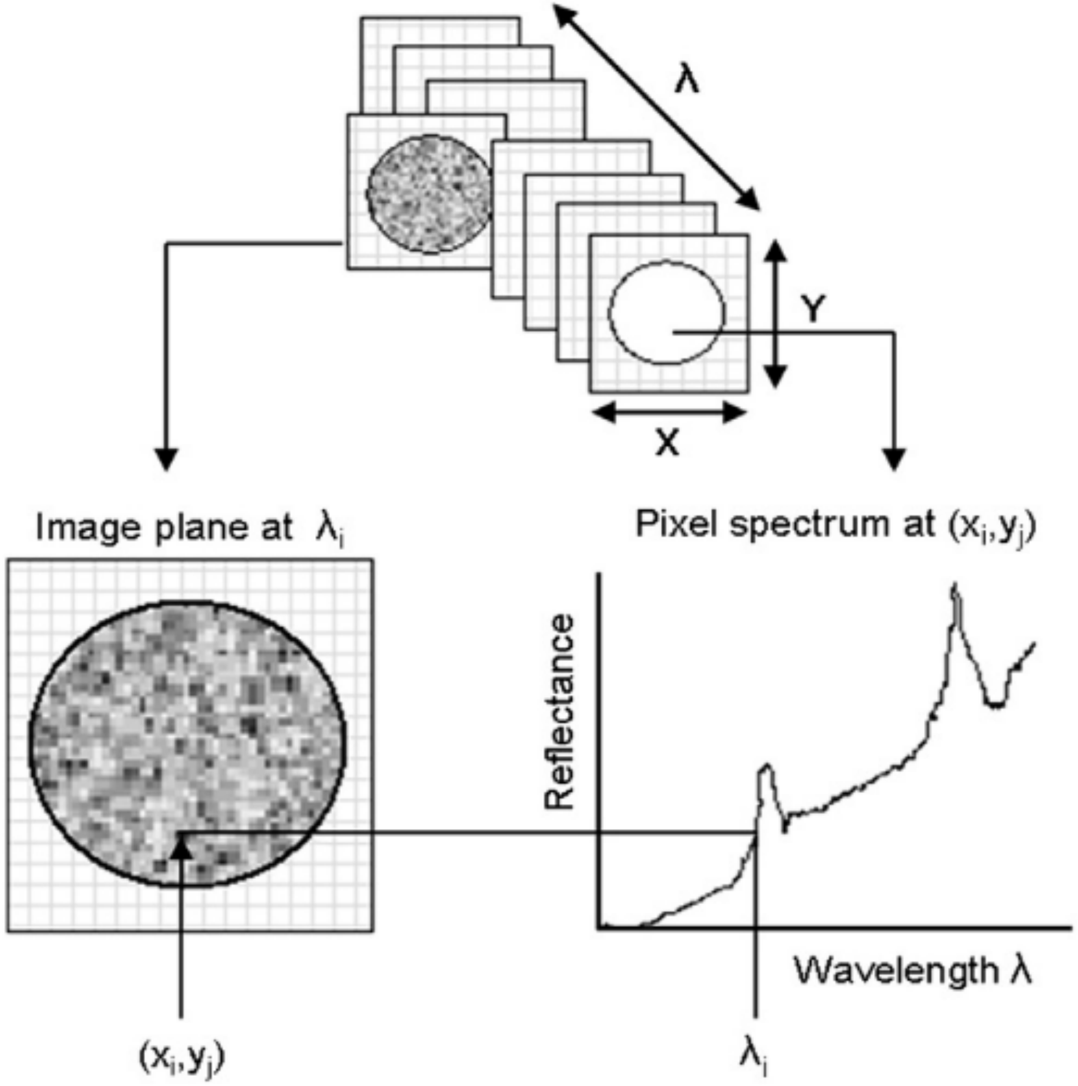
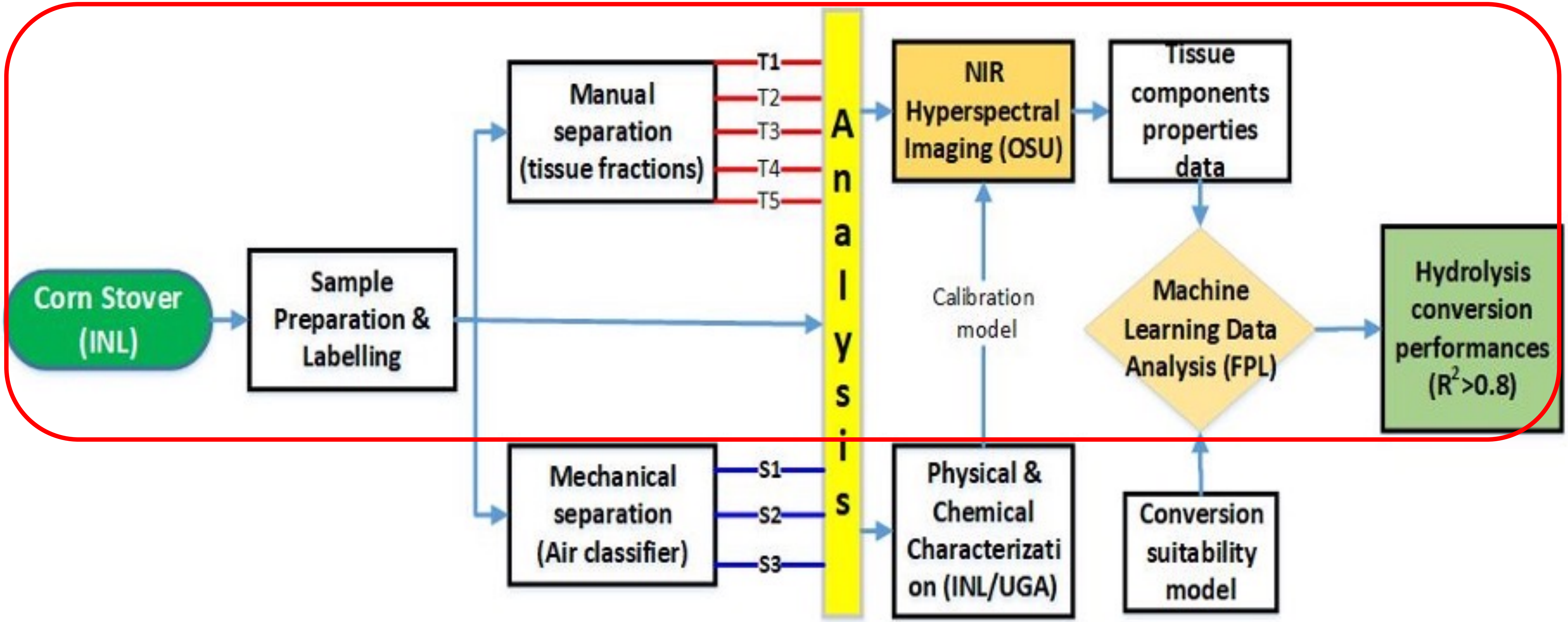


Fig. 1. Hyperspectral image cube (hypercube).

J. Burger, A. Gowen / Chemometrics and Intelligent Laboratory Systems 108 (2011) 13–22

1. Approach: Corn Stover Hydrolysis



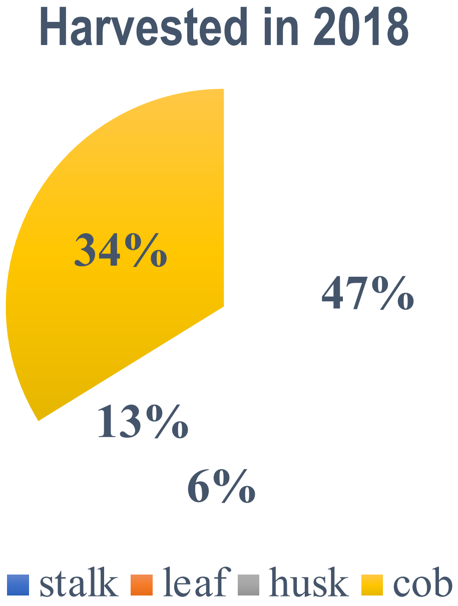
Note: T1....T5 = Tissue fractions – 1. Cobs; 2. Leaves; 3. Husk; 4. Stalk/nodes; 5.Others. S1..S3.. = Mechanically Separated fractions

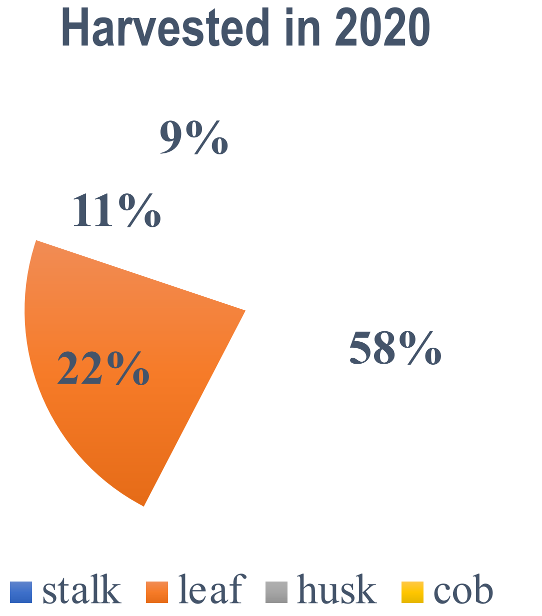
Corn stover tissue fraction conversion: Dilute acid pretreatment + Enzymatic hydrolysis

2. Progresses & Outcomes: Corn Stover

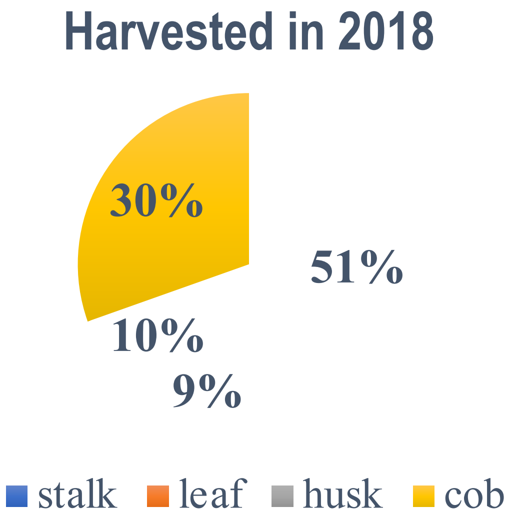
- Corn stover sourcing and collection (Completed)
- Anatomical mass balance (completed)
- Moisture content (completed)
- Particle size distribution (completed)
- Conversion study – in progress
- Chemical/elemental analyses – in progress
- Grinding studies – in progress
- Flowability tests – in progress

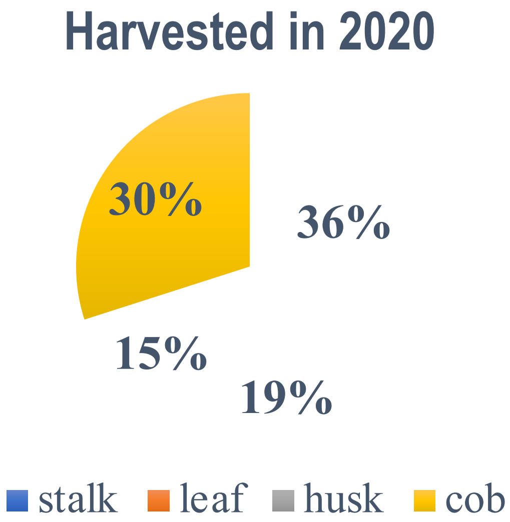
Low-cut Bale from Kadolph Field



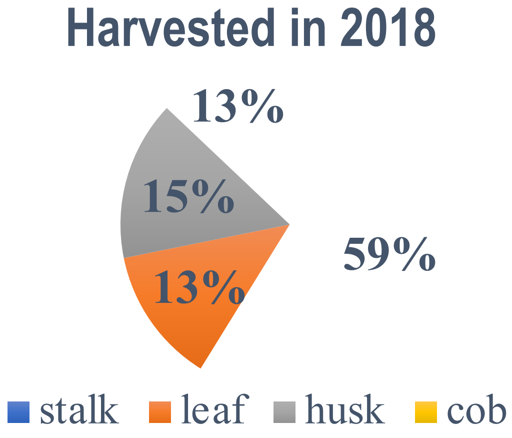


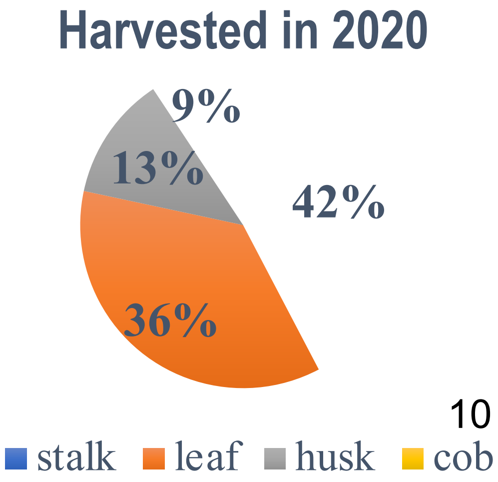
High-cut Bale from Kadolph Field





Raked Bale from Kadolph Field





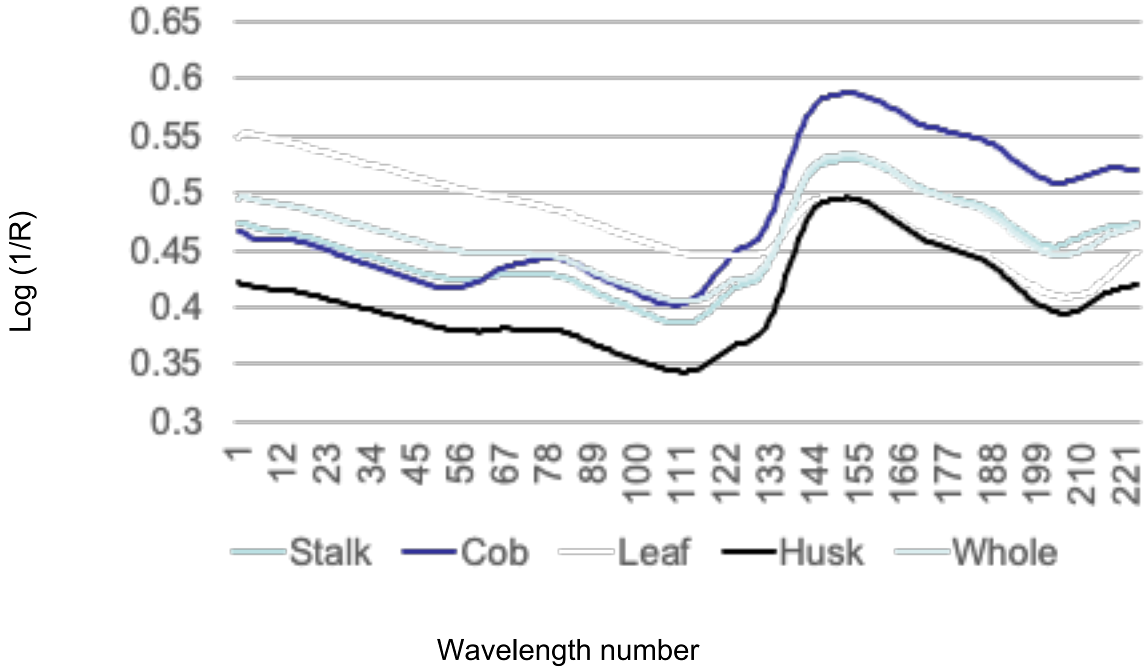
2. Progresses & Outcomes: Hyperspectral Analysis of Corn Stover

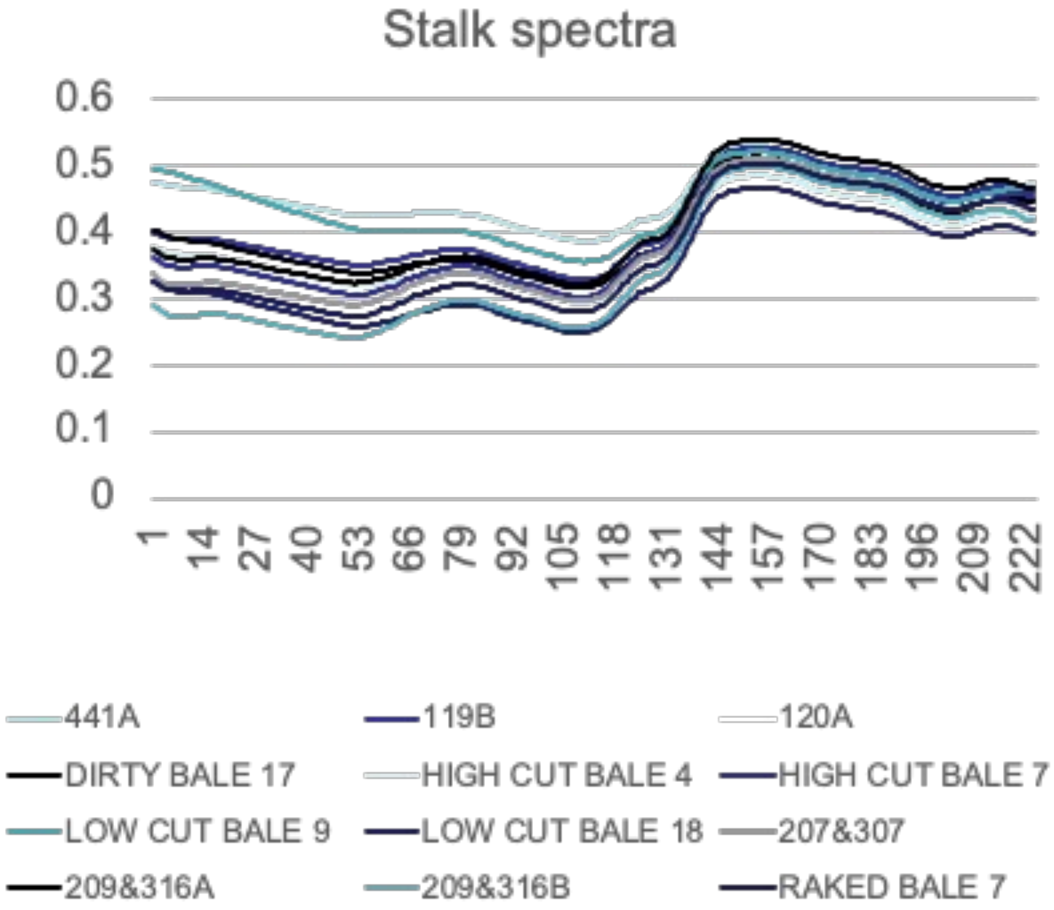
Hyperspectral Imaging (HSI) of all fractions (cobs, husk, leaf, stalks, Whole) has been completed

For each sample / fraction examined, a representative spectrum of each has been obtained from the hyperspectral image

Region of interest was identified, within the region spectral response for each pixel was averaged

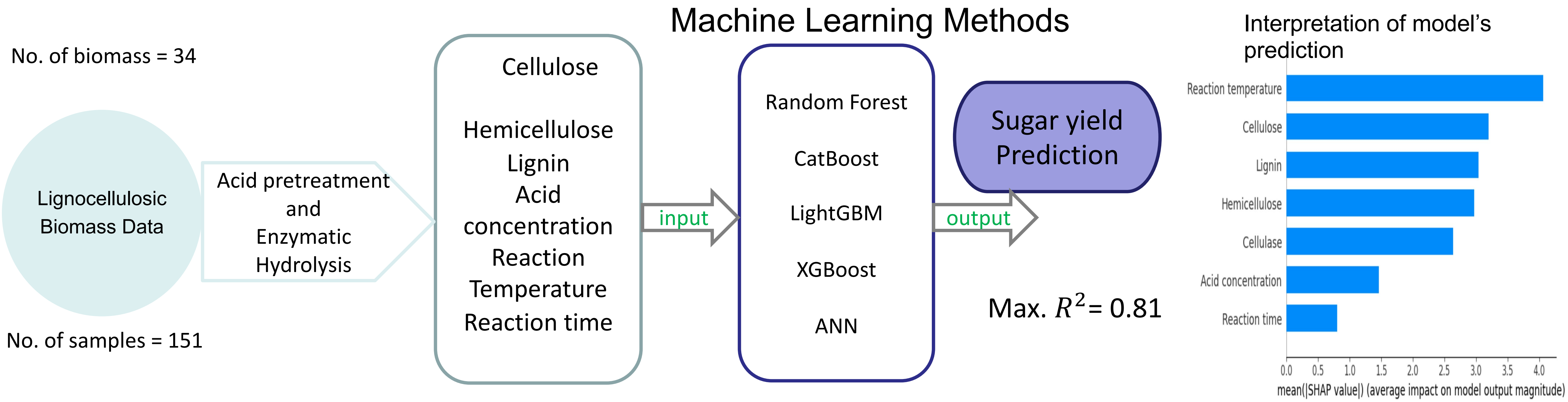
Variation in spectral response amongst different fractions observed





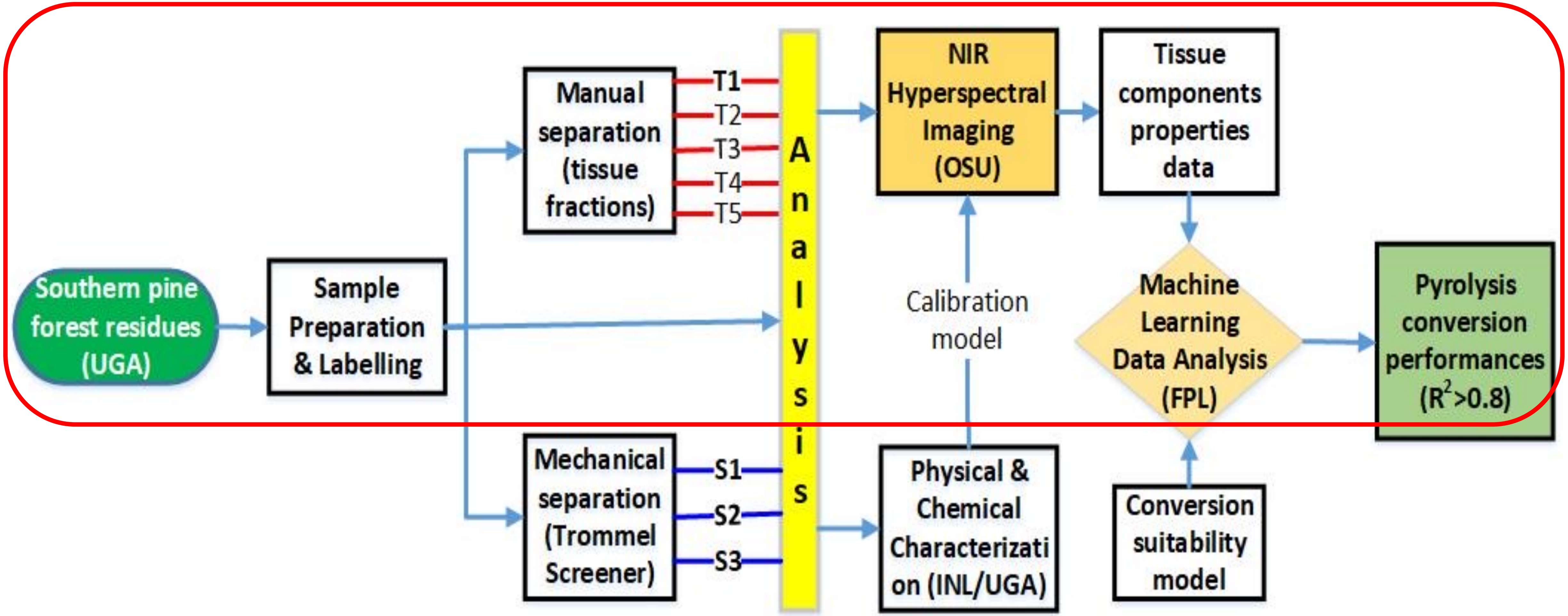
Variation in spectra will be critical for next phase of the project which will involve developing predictive models for a variety of parameters

2. Progresses & Outcomes: Hydrolysis Prediction Model



- The dataset for modeling of acid pretreatment and enzymatic hydrolysis has been collected from literature to predict the sugar yield
- The dataset consists of 151 samples, with each sample having seven features
- The dataset has been randomly divided into 80% training set and a 20% testing set
- Among the algorithms utilized, the CatBoost a Gradient-Boosting based Machine learning algorithm achieved the highest R^2 value using testing dataset
- SHapley Additive exPlanations (SHAP) method, a game theoretic approach, was used to explain the output of the machine learning model
- The most important features contributing to the model's predictions for predicting the sugar yield are ranked

1. Approach: Southern Pine Forest Residues



Note: T1....T5 = Tissue fractions – 1. Needles; 2. Bark; 3. Juvenile Wood; 4. Branches; 5.Others. S1..S3.. = Mechanically separated fractions

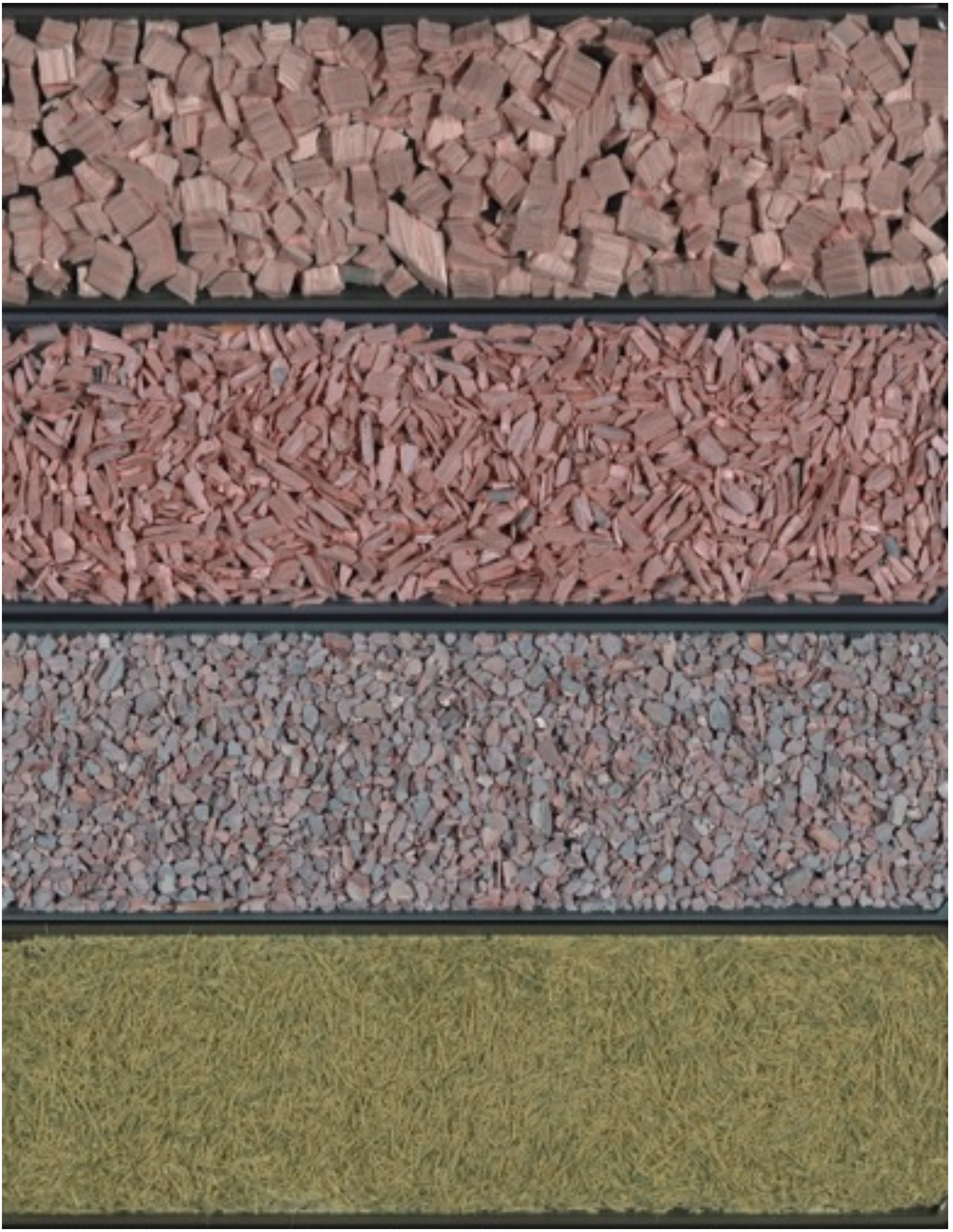
2. Progresses and Outcomes: Southern Pine Forest Residues

- Collect Southern Pine Forest Residues (SPFR) from 5 different stands
- Different regions and age classes sampled to maximize material variability
  - Lower Coastal Plain – 2 Stands (1st thin and final harvest)
  - Upper Coastal Plain – 1 Stand (final harvest)
  - Piedmont – 2 Stands (1st thin and final harvest)

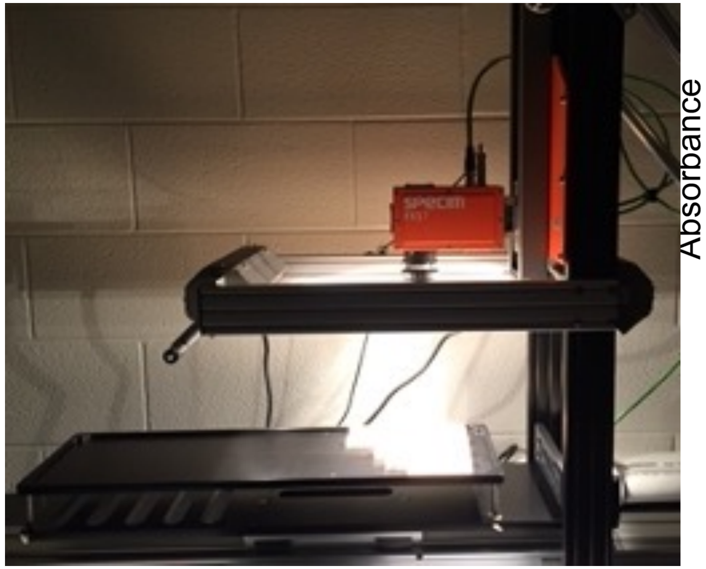
5 fractions

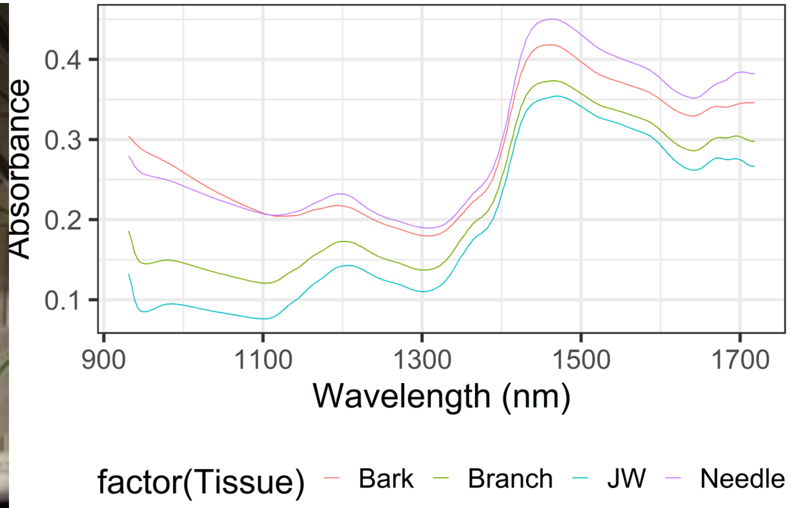
- Juvenile Wood
- Branch
- Needle
- Bark
- Bulk (mixed)

2. Progresses and Outcomes: Hyperspectral Imaging



- Near-infrared spectroscopy hyperspectral images collected from each fraction + size class
- The spectral data will be linked to measured chemical properties



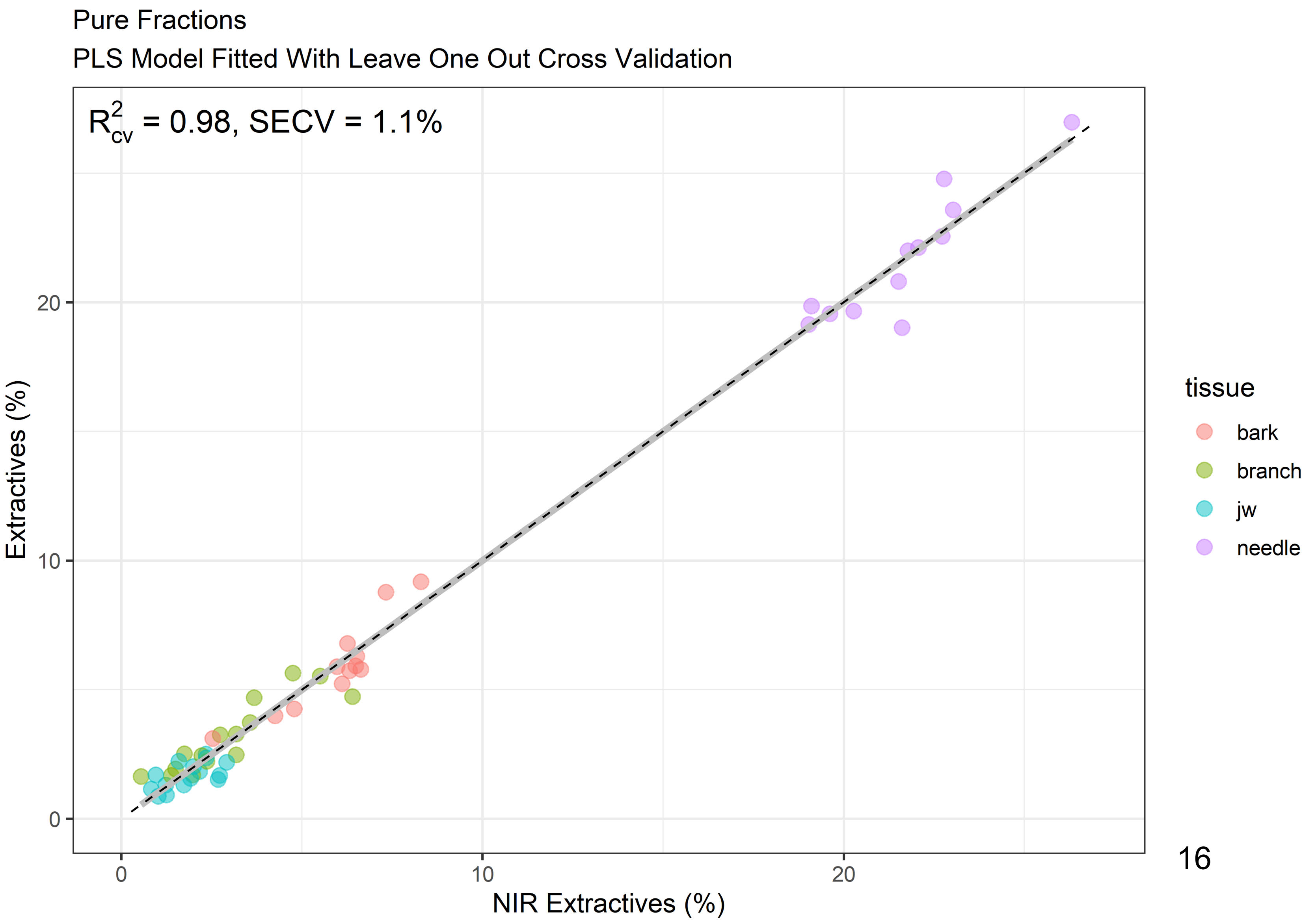


Normal Red Green Blue = 400-700 nm, 3 bands
Specim FX17 = **930 – 1718 nm**, 224 bands

2. Progresses and Outcomes: Extractives Model (in-progress)

- **Extractives**
- Lignin
- Hemicellulose
- Cellulose

Extractives from wood

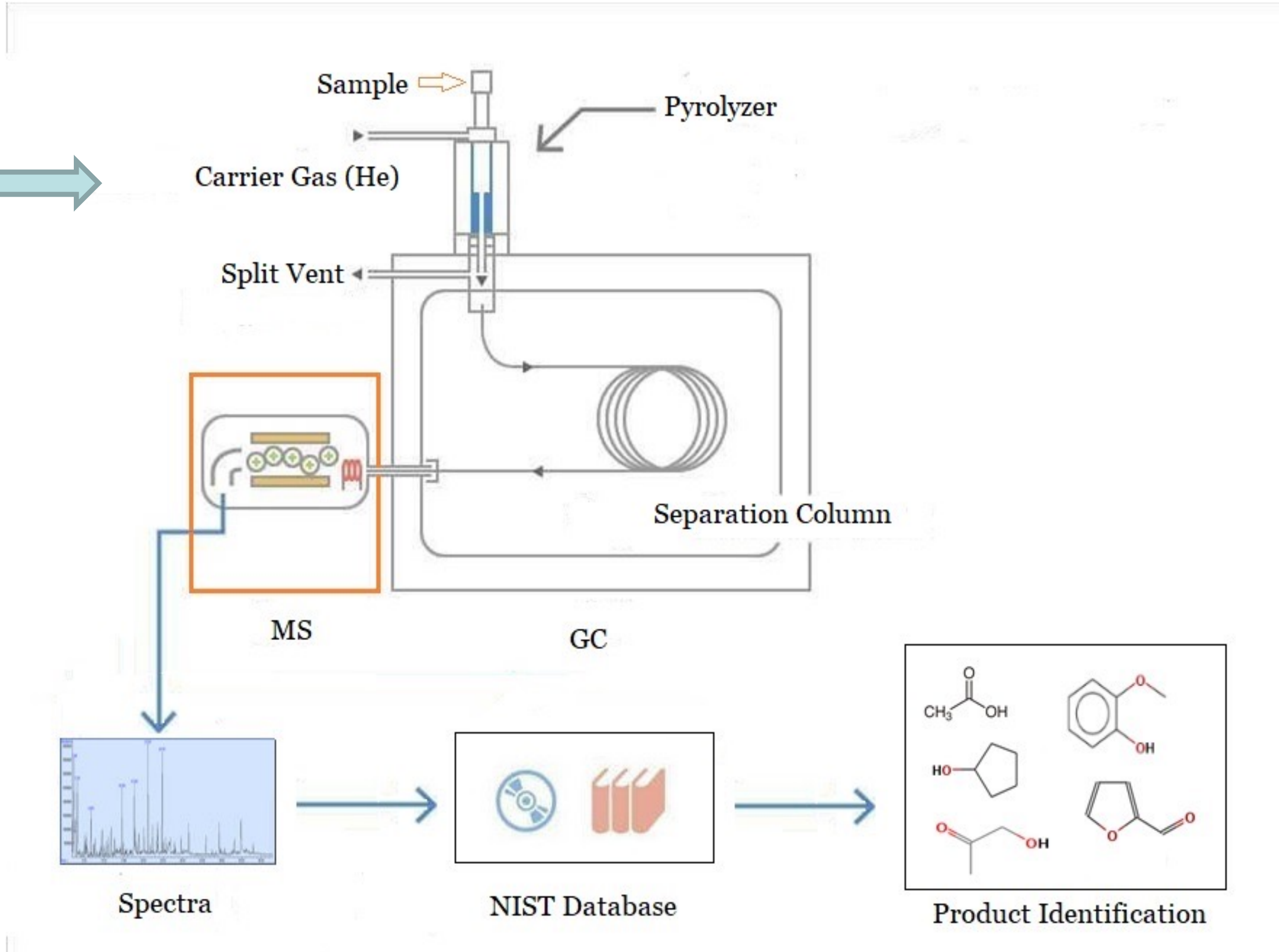


On-going: Fast Pyrolysis by Pyro/GC/MS

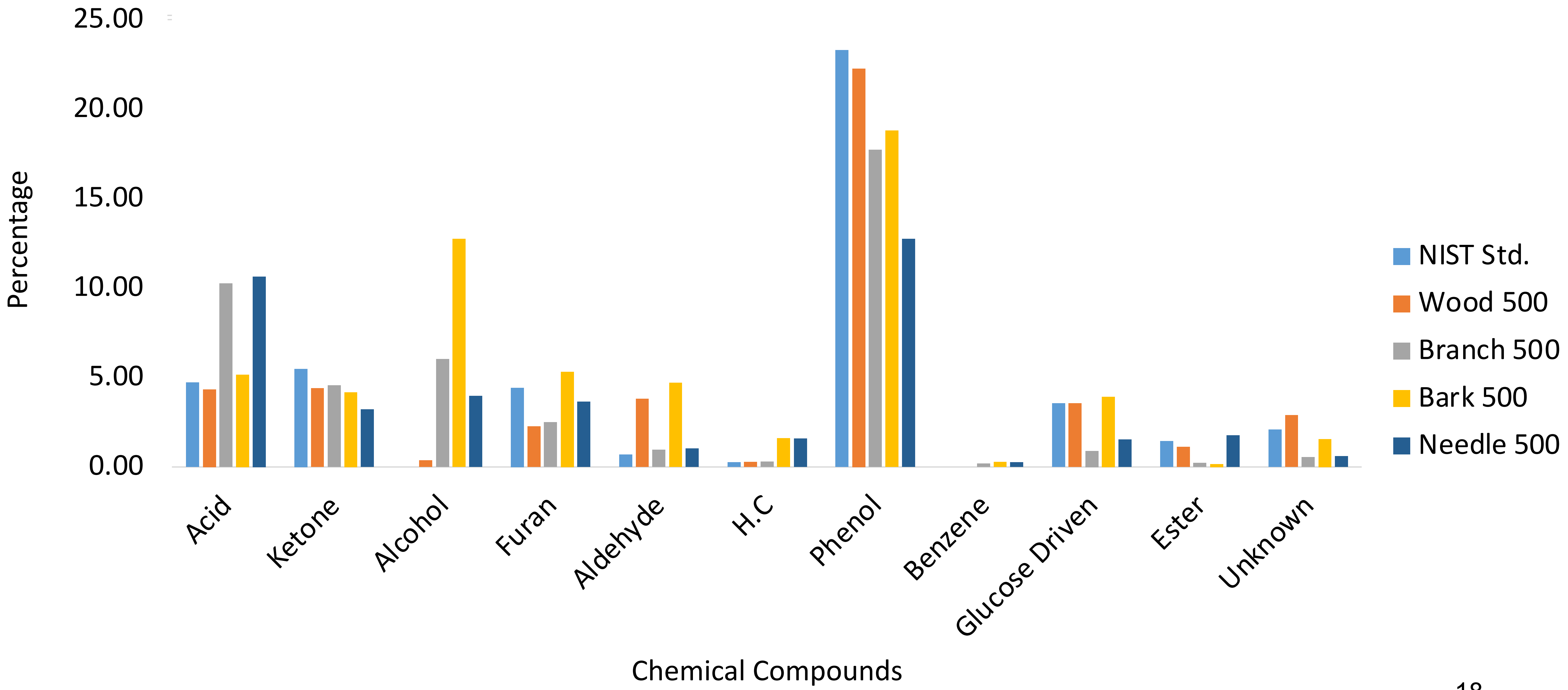
Tissue Fractions

Drying Oven T = 105 °C for 24 hr

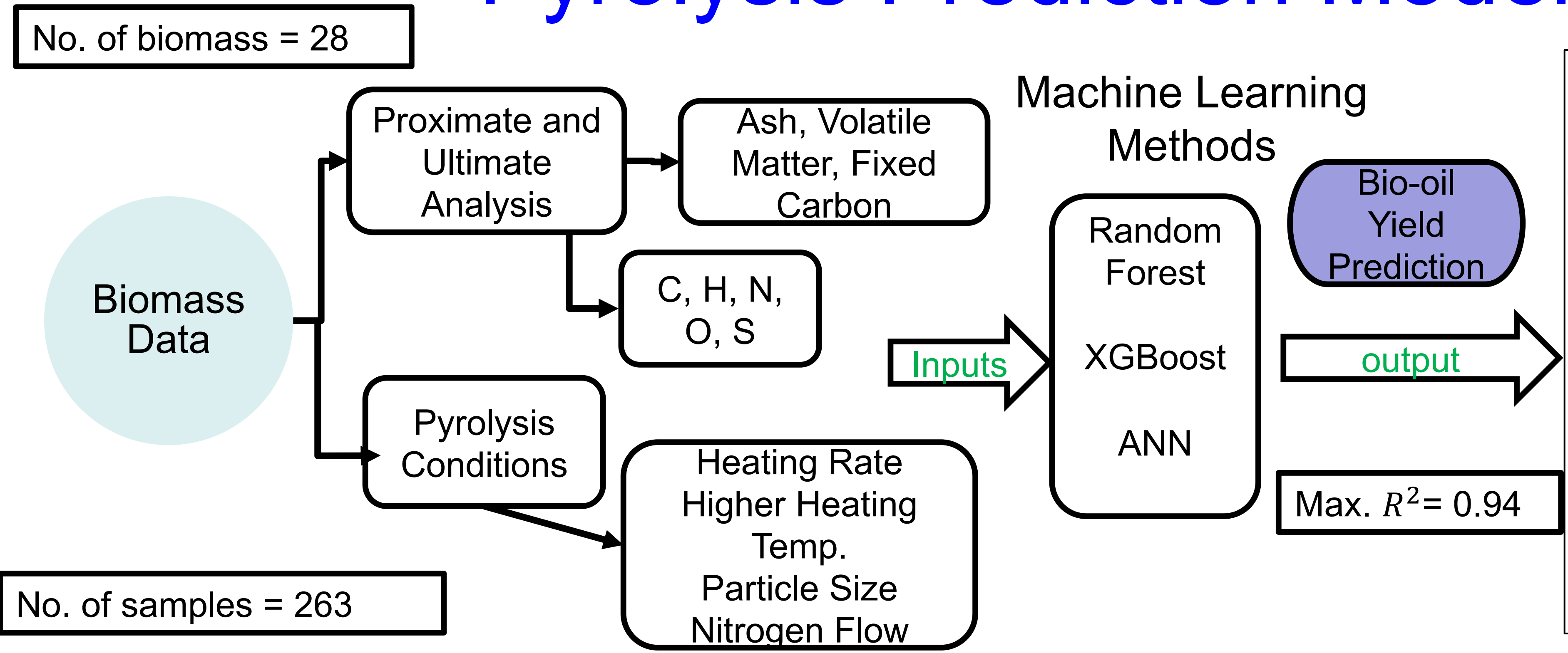
Knife Mill
P.S<0.5 mm and Between 0.5 and 1 mm



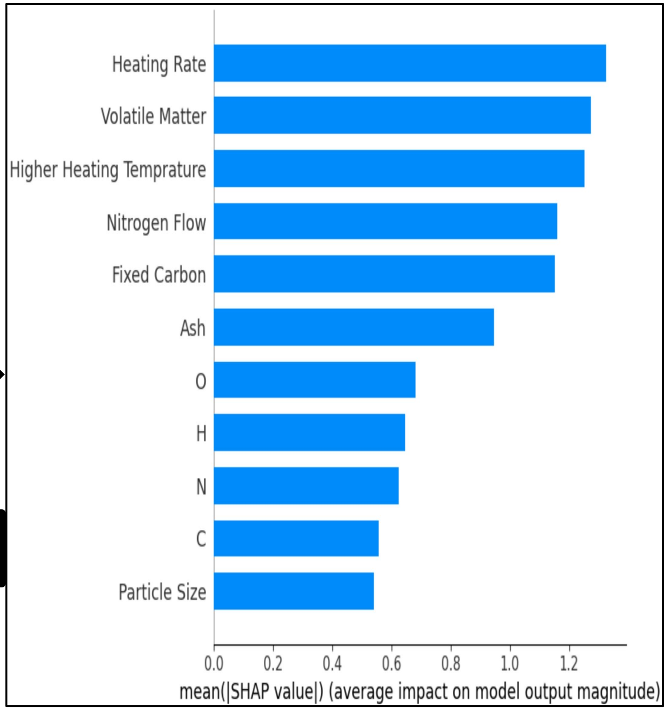
Pyro/GC/MS Preliminary Results at 500 C



Pyrolysis Prediction Model

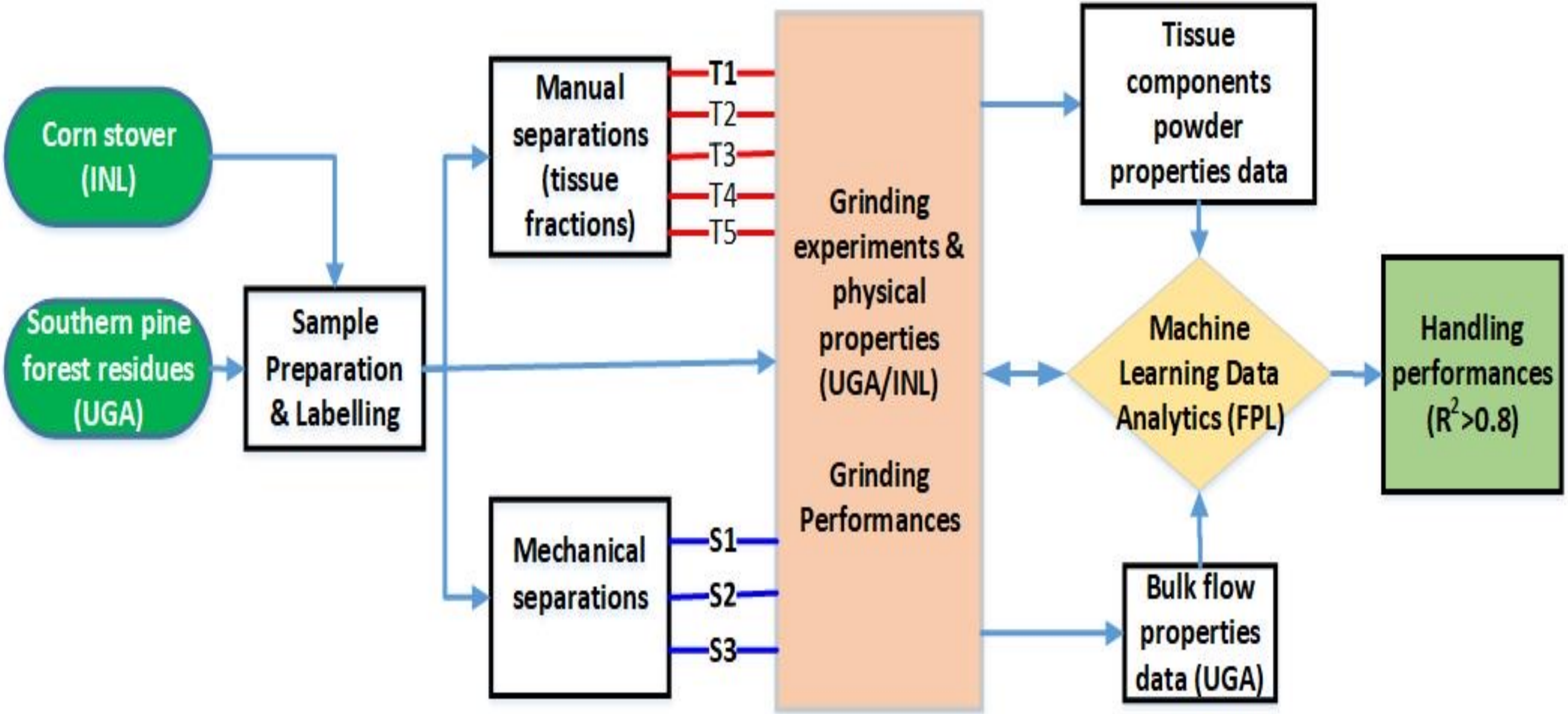


Interpretation of model's prediction



- The dataset for modeling pyrolysis has been collected from literature to predict the pyrolysis bio-oil yield
- The dataset consists of 263 samples, with each sample having seven features
- The dataset has been randomly divided into 80% training set and a 20% testing set
- Among the algorithms utilized, the XGBoost a Gradient-Boosting based machine learning algorithm achieved the highest R^2 value using testing dataset
- SHapley Additive exPlanations (SHAP) method, a game theoretic approach, was used to explain the output of the machine learning model
- The most important features contributing to the model's predictions for predicting the bio-oil yield are ranked

On going Work - Relating the tissue properties to grinding/bulk flow performances



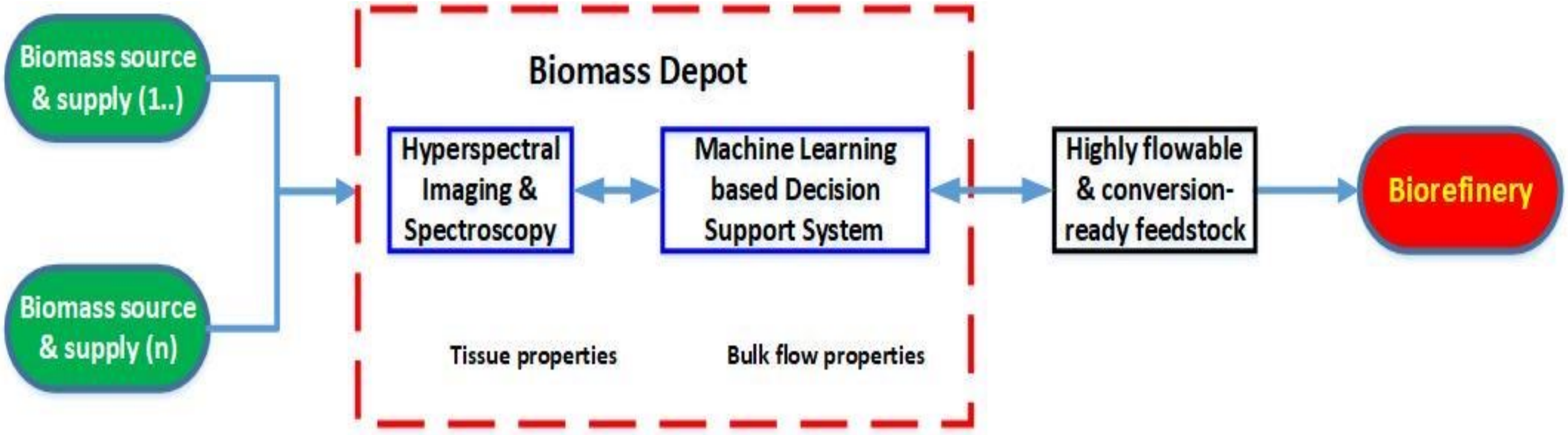
Note: T1....T5 = Tissue fractions (manually separated); S1..S3.. = Mechanically separated fractions

Project Schedule – Budget Period 2 (BP2)

Task 2	Milestone	BP2*					
		Q1	Q2	Q3	Q4	Q5	Q6
Task 2.1 - Biomass sourcing & collection	M2.1	√	√				
Task 2.2 – Tissue type separation	M2.2	√	√				
Task 2.3 – Tissue properties measurement	M2.3			√	x		
Task 2.4 – Hyperspectral Imaging	M2.4			√	x		
Task 2.5 – Conversion suitability analysis	M2.5				√	x	x
Task 2.6 – Grinding & Bulk flow properties	M2.6				√	x	x
Task 2.7 – ANN framework & data processing	M2.7				√	x	x
Task 2.8 – ANN Prediction Models	M2.8				√	x	x
Task 2.9 – External Advisory Board	M2.9		√				x
Task 2.10 – Project Management & Reporting	M2.10	√	√	√	√	x	x

*Note – Green color with √ symbol indicates that the task is completed;
Red color with x symbol indicates that the task is in progress

3. Impacts



Fundamental understanding on how physical and mechanical properties linked to complex powder flow patterns and performances

Developed machine learning based predictive models can be used as a decision support tool to design conversion-specific, highly flowable feedstock at the depot for successful operation of a biorefinery.

Developed models can also be trained and used at the biorefinery for process control, yield predictions and plant resource management and risk mitigations.

3. Impacts: Publications & Conference Presentations

1. Vasefi, S., P. Azadi, and **S. Mani**. 2022. Effects of particle size and pyrolysis temperature on fast pyrolysis of southern pine forest residue tissue fractions using Pyro-GC/MS. Presented at the International Bioenergy and Bioproducts Conference (IBBC), Oct. 30 – Nov. 02, 2022, Providence, RI.
2. Vasefi, S., and **S. Mani**. 2022. The current state of predicting the performance of biomass pyrolysis process using machine learning tools. Presented at the International Bioenergy and Bioproducts Conference (IBBC), Oct. 30 – Nov. 02, 2022, Providence, RI.
3. Vasefi, S., and **S. Mani**. 2022. Application of machine learning in predicting biomass pyrolysis performance. Presented at the Institute of Biological Engineering (IBE) Annual Meeting April 8-9, 2022, Athens, GA.
4. Osagie, Y., and **S. Mani**. 2022. Grinding and bulk flow performances of southern forest residue tissue fractions. Presented at the Institute of Biological Engineering (IBE) Annual Meeting April 8-9, 2022, Athens, GA.
5. Raut S, Dahlen J, **Mani S**, Schimleck L, Onakpoma I. 2022. Machine learning based modeling framework to relate forest residues with handling and conversion performances. Wood Quality Consortium Annual Meeting. July 12, 2022.

Summary

- *Both corn stover and southern pine forest residue samples were all sourced and manually separated into various tissue fractions*
- *The hyperspectral imaging with NIR spectra on the tissue fractions captured the variability among the chemical compositions and the prediction models are under progress*
- *The conversion performance experiments are ongoing with enzymatic hydrolysis route for corn stover tissue fractions and fast pyrolysis experiments for pine forest residues*
- *Both grinding experiments and bulk flow properties measurement studies are ongoing.*
- *The machine learning based predictive models were developed based on the existing published data in the literature and selected machine learning algorithms offer increased prediction potential for both hydrolysis and pyrolysis models*
- *The developed models can be utilized by biomass industries to predict product yield and also screen feedstocks for target conversion applications.*

Quad Chart Overview

Timeline

- *BP1 – Oct 2019 - Jul 2021*
- *BP2 - Aug 2021 – Sept 2023*
- *BP3 – Oct 2023 – Sept 2024*

	FY22 Costed (EXCLUDING FFRDC)	Total Award (DOE+FFRDC)
DOE Funding	*$279,437*	*$1,451,342*
Project Cost Share *	$199,709	$363,336

Project Partners*

- Idaho National Lab (INL)
- Oregon State University
- USDA Forest Products Lab

Project Goal

(1) To develop a novel tool for rapid determination of individual biomass tissue properties using a Near InfraRed (NIR)-Hyperspectral Imaging (HSI) system.
(2) To apply the machine learning framework to correlate the biomass chemical compositions with conversion performances.
(3) To investigate the grinding performances and the bulk flow properties of tissue components.
(4) To use the machine learning modeling framework to relate physical properties of tissue powders with grinding and bulk handling performances

End of Project Milestone

Validation of Machine learning based predictive model with R^2 > 0.8

Funding Mechanism

FOA Call on DE-FOA-002029
CFDA # 81.087

Additional Slides

Performance Metrics

(1) Relating tissue chemical properties to conversion performances
  - Corn stover tissue properties to ethanol yield
  - Forest residue tissue properties to pyrolysis yield (Bio-oil)
  - Expected predictive model R^2 >0.8

(2) Relating tissue physical properties to grinding/handling performances
  - Biomass tissue properties to grinding performances (energy use, PSD)
  - Biomass tissue properties to bulk flow properties
    - Angle of Repose (AOR) = Hopper design
    - Internal Friction & Cohesion = Hopper design, storage bin design
    - (Relative) Flow Index (FI) = material conveying system design
  - Predictive model R^2 >0.8