INL/MIS-24-81991-Revision-0

Optimizing Biofuel Production: Integrating NIR Imaging and Machine Learning for Corn Stover Characterization

November 2024

Changing the World's Energy Future

Sambandh Bhusan Dhal

Idaho National Laboratory

INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

DISCLAIMER

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

INL/MIS-24-81991-Revision-0

Optimizing Biofuel Production: Integrating NIR Imaging and Machine Learning for Corn Stover Characterization

Sambandh Bhusan Dhal

November 2024

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

11/13/2024

Optimizing Biofuel Production: Integrating NIR Imaging and Machine Learning for Corn Stover Characterization

- Sambandh Bhusan Dhal

Post-Doctoral Associate, Department of Analytical Chemistry, Idaho National Laboratory

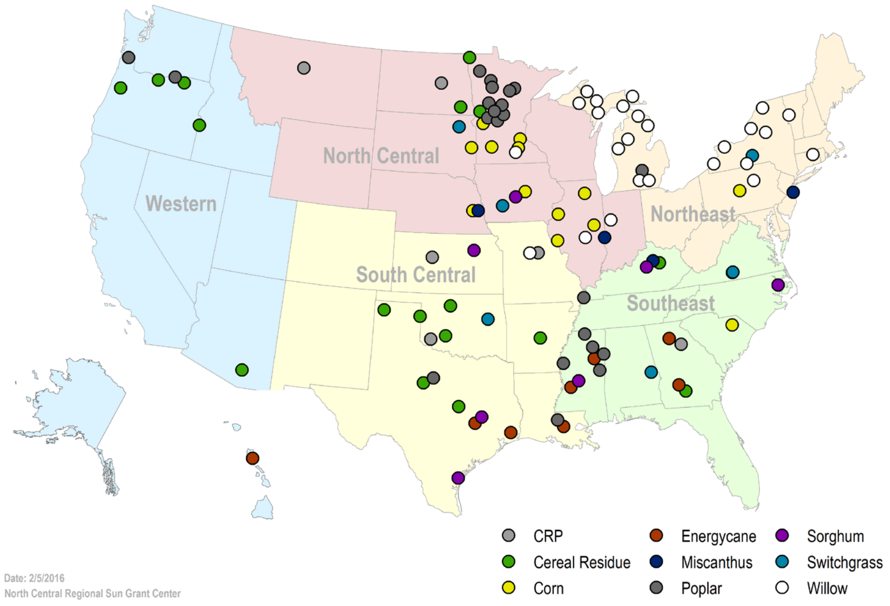
Post-Doctoral advisor and Project PI: Dr. Lorenzo Vega-Montoto

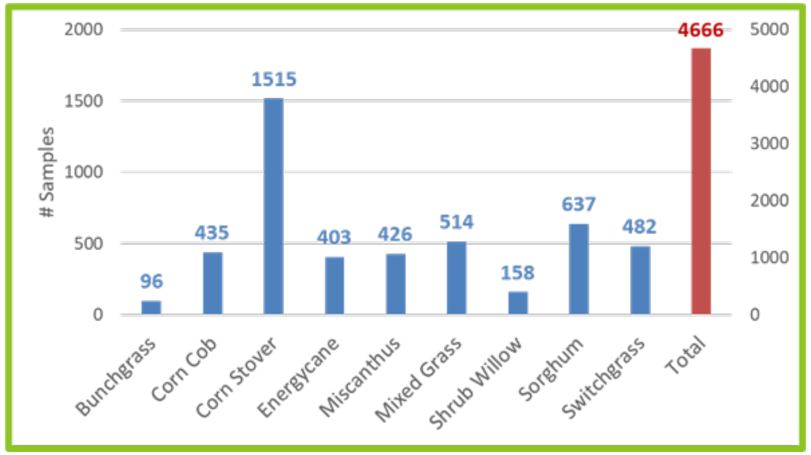
Manager: Dr. Jesse Douglas Carrie

Department Manager: Dr. Julie Bowen

Idaho National Laboratory

Regional Feedstock Partnership Overview

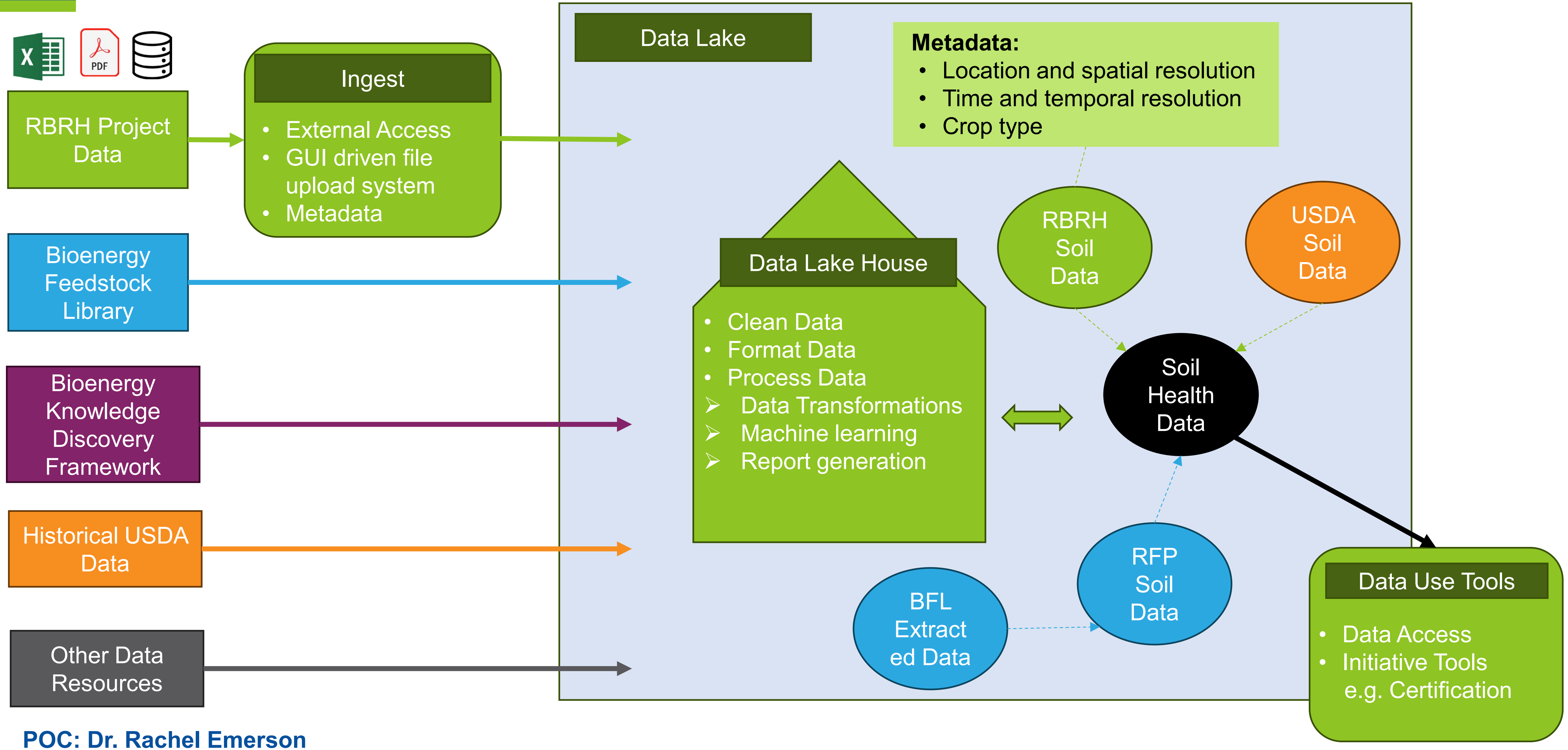




https://bioenergylibrary.inl.gov/Home/Home.aspx

- **U.S. DOE & Sun Grant Institute initiated in 2007** to address information gaps associated with the sustainable and reliable production of a billion-tons of biomass annually
- Key accomplishments:
  1. **Demonstrated production potential** of diverse herbaceous & woody feedstocks across the U.S. **for 5-7 growing seasons**
  2. **Validated *Billion Ton* estimates** and resulted in national yield potential maps
  3. Demonstrated **improved yields of new varieties/cultivars** of biomass sorghum, energycane, hybrid poplars, and shrub willows
  4. **130+ scientific publications** and numerous presentations; **workforce development** in post-docs/students
- INL contributions in landscape management, harvest, collection, storage, preprocessing, quality data, sample and data management of ag. residues and energy crops
- Additional resources across the supply chain are necessary to mobilize energy crops to meet BT23 predictions and SAF Grand Challenge targets

Regional Biomass Resource Hubs Data System



POC: Dr. Rachel Emerson

Role of Machine Learning in Agriculture

Resource Optimization

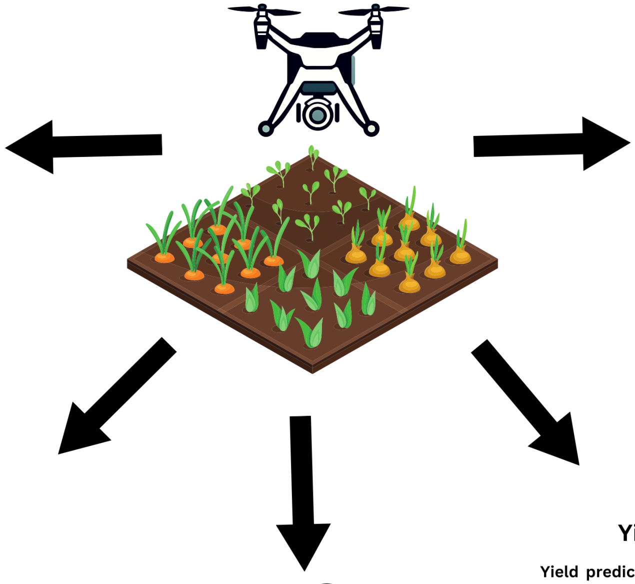
- Water Use Efficiency
- Nutrient Management

Growth Monitoring and Yield Forecasting

- Real-Time Monitoring
- Early Stress Detection
- Predicting Corn Yield
- Field-Wide Insights

Post-Harvest Quality Control

- Quality Assessment
- Optimized Processing



Crop Health Monitoring

Drone-based HSI captures multi-wavelength spectral data to detect early signs of crop stress, nutrient deficiencies, and diseases. It identifies subtle changes like water stress or chlorophyll degradation, enabling timely interventions that improve crop management and yields. Machine learning models (PLSR, RF, CNN) enhance the precision of crop health assessments.

Disease Detection

HSI detects early signs of crop diseases by capturing spectral changes before visible symptoms appear. Drones monitor large areas for issues like fungal infections or pest infestations, enabling timely interventions to prevent disease spread. Machine learning models (SVM, CNN) classify healthy and diseased plants with high accuracy

Soil Analysis

Drone-based HSI provides non-invasive soil analysis, mapping key properties like moisture, organic matter, nutrients, and contamination. This real-time data helps optimize irrigation and fertilization, ensuring efficient resource use. Machine learning models (SVM, DT) enhance large-scale soil health assessments for improved fertility and sustainability

Yield Prediction

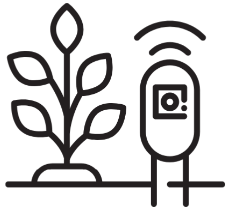
Yield prediction is a key application of drone-based HSI, capturing spectral data throughout crop growth. This data helps monitor plant development and predict yields with high precision. Machine learning models (RF, ANNs) analyze the data, providing accurate, timely yield forecasts essential for resource planning and optimizing supply chains

Water Management

In water-scarce regions, drone-based HSI detects early plant water stress, providing real-time data on water uptake. This enables precision irrigation, optimizing water use and reducing waste. Machine learning models (RF, DT) are used to analyze the data for efficient irrigation management.

Machine Learning for Water Use Efficiency in Corn Crops

Optimized Irrigation & Water Use Efficiency

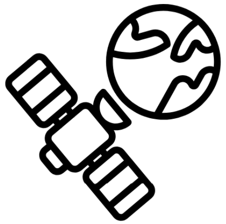


W-Tens Tool: Soil Moisture Monitoring

- **Model:** Tensiometer-based ML Threshold Model
- **Function:** Tracks soil moisture levels at various depths
- **Outcome:** Reduces over-irrigation; optimizes schedules based on real-time data

W-Mod Tool: Soil Water Balance Simulation

- **Model:** Based on soil dynamics & Richards Equation, Thermal growth models (e.g., Growing Degree Days)
- **Function:** Tracks soil moisture, root depth, soil, and weather
- **Outcome:** Comprehensive soil-water balance for irrigation timing & volume



IRRISAT® Tool: Remote Sensing and ET Prediction

- **Model:** ML-enhanced FAO-56 & Penman-Monteith for ET calculation, Sentinel-2 satellite data
- **Function:** Predicts crop water needs from ET & soil moisture deficit
- **Outcome:** Precise irrigation recommendations to save water, protect crop health

Bonfante, A., et al. "LCIS DSS—An irrigation supporting system for water use efficiency improvement in precision agriculture: A maize case study." Agricultural Systems 176 (2019): 102646.

Machine Learning for Nutrient Management in Corn Crops

Adaptive Nutrient Recommendations

- **Tool:** Nutrient Expert®
- **Approach:** 4R Stewardship (Right type, rate, time, placement)
- **Outcome:** Site-specific, sustainable fertilizer use

Environmental Sustainability

- **Enhanced NUE:** Less nutrient leaching, reduced runoff
- **Outcome:** Improved soil health, lower environmental impact

Yield and Efficiency Gains

- **ML Model:** Random Forest (RF)
- **Focus:** Identifies critical factors for N, P, K uptake
- **Outcome:** Average yield increase of 3.5 t/ha

Cost Savings and Profitability

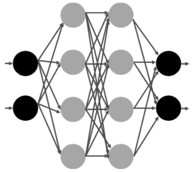
- **Precision Application:** Reduced fertilizer waste
- **Outcome:** Lower input costs, higher farmer profitability

Timsina, Jagadish, et al. "Improved nutrient management in cereals using Nutrient Expert and machine learning tools: Productivity, profitability and nutrient use efficiency." *Agricultural Systems* 192 (2021): 103181.

Machine Learning for monitoring growth parameters in Corn Crops

Using UAVs and Deep Learning

- **Tools:** UAVs with multispectral & RGB sensors
- **Data:** High-resolution spectral and spatial imagery
- **Outcome:** Detailed field data for growth analysis



Deep Learning Models

- **Model:** YOLOv5
- **Functions:** Counts plants, assesses density & growth
- **Outcome:** Automated, accurate plant monitoring

Impacts on Field Management

- **Insights:** Optimal seeding, planting depth, fertilization
- **Result:** Enhanced corn emergence and yield potential

Vegetation Indices for Growth Monitoring

- **Indices:** NDVI & NDRE
- **Function:** Non-invasive health and vigor tracking
- **Outcome:** Real-time monitoring of growth status

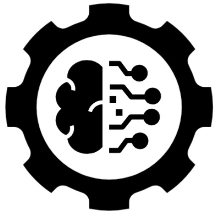
Biophysical Parameter Mapping

- **Measurements:** Plant height & density from UAVs
- **Technique:** Otsu Thresholding for plant/background separation
- **Outcome:** Accurate parameter extraction (NDVI, NDRE, height)

Shao, Guomin, et al. "Prediction of maize crop coefficient from UAV multisensor remote sensing using machine learning methods." Agricultural Water Management 276 (2023): 108064.

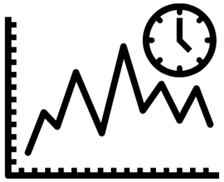
Xiao, Juan, et al. "Enhancing assessment of corn growth performance using unmanned aerial vehicles (UAVs) and deep learning." Measurement 214 (2023): 112764.

Machine Learning for yield prediction in Corn Crops



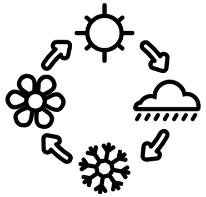
Hybrid Modeling Approach

- **Combination:** APSIM crop simulation + ML techniques
- **Features:** Soil moisture, drought stress, crop phenology
- **Outcome:** Enhanced yield prediction accuracy



ML Models and Ensembles

- **Models:** Random Forest, LightGBM, XGBoost, and ensembles
- **Focus:** Optimized ensembles for reduced bias
- **Outcome:** Improved prediction accuracy



Improvement with In-Season Data

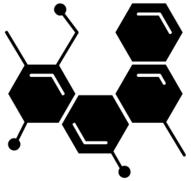
- **Data:** Early-season weather data (up to June)
- **Accuracy:** RRMSE as low as 9.2%
- **Outcome:** Early yield forecasts for management adjustments

Impact on Precision Agriculture

- **Support:** Timely resource-use decisions
- **Scalability:** Enhances profitability at multiple scales
- **Outcome:** Optimized resource management

Shahhosseini, Mohsen, Guiping Hu, and Sotirios V. Archontoulis. "Forecasting corn yield with machine learning ensembles." Frontiers in Plant Science 11 (2020): 1120.
Shahhosseini, Mohsen, et al. "Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt." Scientific reports 11.1 (2021): 1606.

Post-Harvest Biomass characterization challenges



Complex Biomass Composition

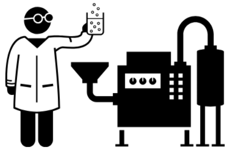
- **Components:** Cellulose, hemicellulose, lignin, etc.
- **Challenge:** Diverse compounds make analysis difficult
- **Outcome:** Complicates precise characterization

Sampling Challenges

- **Issue:** Biomass heterogeneity in large-scale production
- **Difficulty:** Obtaining representative samples
- **Outcome:** Impacts reliability of characterization

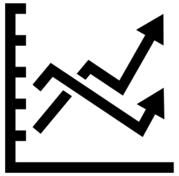
Analytical Limitations

- **Method:** Traditional techniques like wet chemistry
- **Challenge:** Incomplete capture of composition
- **Outcome:** Limited understanding of biomass structure



Sample Pretreatment

- **Process:** Drying, grinding, extraction steps
- **Issue:** Introduces variability in samples
- **Outcome:** Affects characterization consistency



Structural Variability

- **Factors:** Varies by species, growth, location
- **Challenge:** Limits universal characterization approaches
- **Outcome:** Increases complexity in analysis

Cousins, Dylan S., et al. "Near-infrared spectroscopy can predict anatomical abundance in corn stover." *Frontiers in Energy Research* 10 (2022): 836690.s

Treatment analysis of Corn Stover for Biofuel Production

Sample Preparation

Alkaline Pretreatment (NaOH):

- Sodium hydroxide treatment to break down lignin and hemicellulose.
- Removes structural barriers for improved enzymatic hydrolysis.

Enzymatic Hydrolysis:

- Cellulase enzymes are added to pretreated samples to break down cellulose into glucose.
- Both pretreated and untreated samples undergo hydrolysis for comparison.

Post-Treatment Analysis:

- Glucose yields are measured to evaluate the effectiveness of treatment.
- Fractions are compared to determine which anatomical parts respond best to treatment.

Biofuel Production Optimization:

- Results inform adjustments to treatment strategies for maximizing sugar yield from different anatomical fractions.

Preprocessing of the Corn Bales

Bale Grinder
(5 to 15 cm pieces)

→ Disc screen
(Separate the larger and smaller fractions)

→ Oscillating screen
(Refining the smaller fractions)

→ Spudnik air separation
(Separating the heavier and lighter fractions)

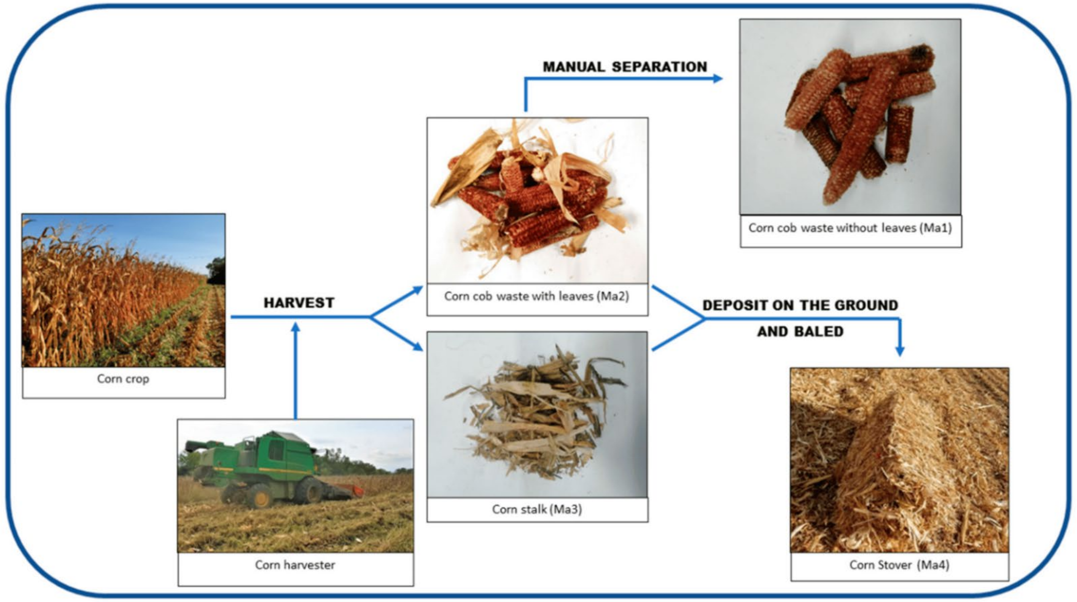
↓ Oven Drying
(105 degrees C for 24 hours)

← Cooling in Desiccator

← Weigh final mass

← Calculate the moisture content

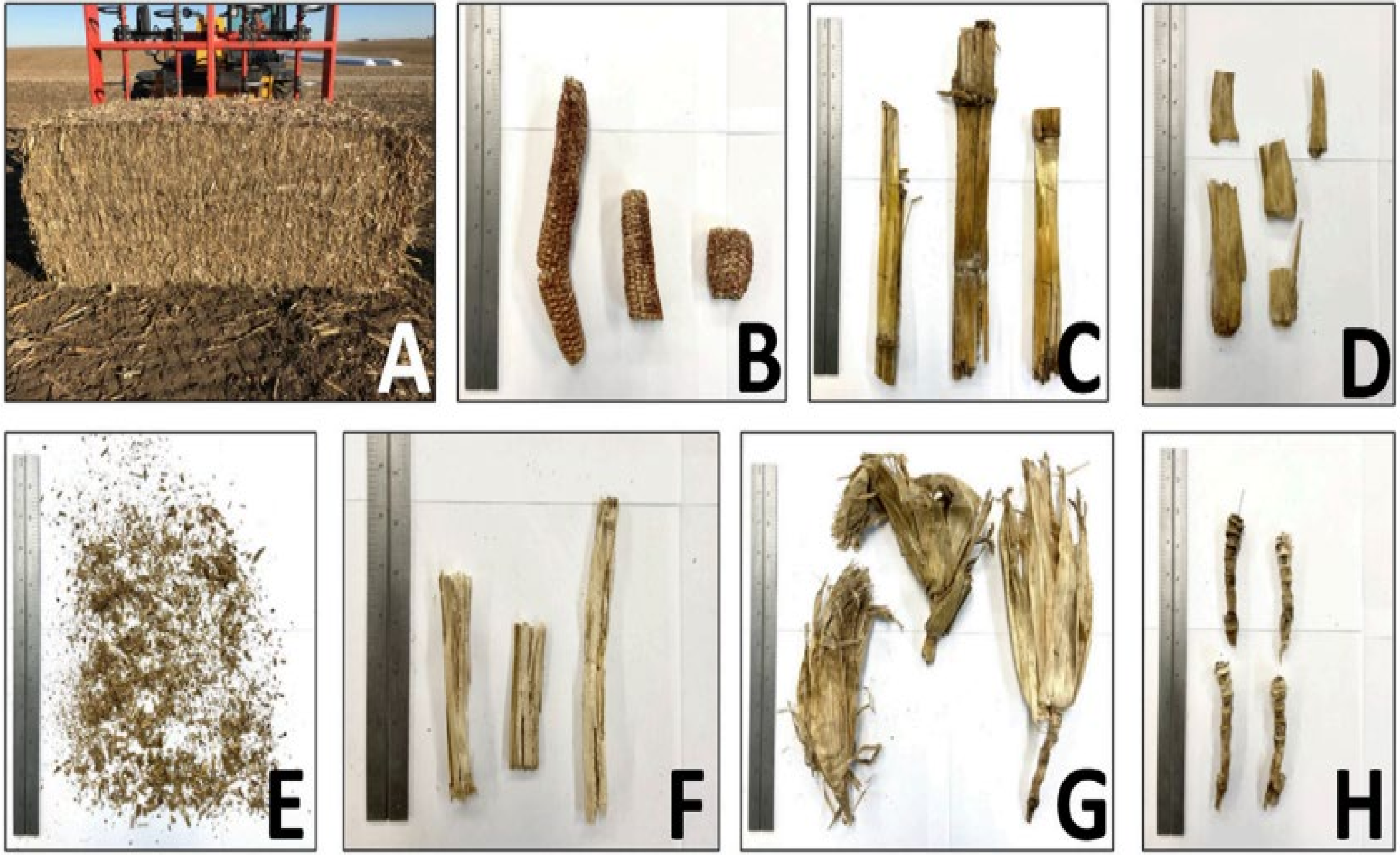
↓ Select the anatomical parts with least moisture content
(~10%)



Anatomical parts of Corn Stover (Sample Preparation)

A: corn stover rectangular bale
B: corn cob
C: corn stalk
D: corn sheath
E: corn leaf fractions pulverized during harvesting operations
F: corn pith fractions isolated from stalk fractions
G: corn husk with attached shank
H: isolated corn shank attached to stalk and husk plant fractions

- Samples were milled **(#20 standard mesh**) and equilibrated at **varying relative humidity levels** for moisture content prediction.



Chemical Composition Analysis

Wet chemistry techniques are used to determine exact proportions of chemical components in the biomass

Table 1: Measurement Techniques for Chemical Composition of Corn Stover

Component	Method	Description
Glucan	NREL/TP-510-42618, NREL/TP-510-48087	Hydrolysis with acids, quantified via HPLC
Xylan	NREL/TP-510-42618, NREL/TP-510-48087	Hydrolysis with acids, quantified via HPLC
Klason Lignin	NREL/TP-510-42618, NREL/TP-510-48087	Hydrolysis with acids, quantified via HPLC
Acetate	Chemical assays	Measured breakdown products from hemicellulose
Ash	Combustion (burning biomass)	Weighing inorganic residue after combustion
Moisture Content	Oven drying	Drying in oven and weighing mass difference
Water Extractives	Solvent extraction (water)	Washing biomass with water and measuring soluble compounds
Ethanol Extractives	Solvent extraction (ethanol)	Washing biomass with ethanol and measuring soluble compounds

Table 2: Measurement Techniques for Anatomical Composition of Corn Stover

Anatomical Fraction	Method	Process	Measurement
Cob	Manual separation	Separated by hand, visually inspected	Weighed and recorded
Husk	Manual separation	Separated by hand, visually inspected	Weighed and recorded
Leaf	Manual separation	Separated by hand, visually inspected	Weighed and recorded
Sheath	Manual separation	Separated by hand, visually inspected	Weighed and recorded
Stalk (Rind)	Manual separation	Separated by hand, visually inspected	Weighed and recorded
Stalk (Pith)	Manual separation	Separated by hand, visually inspected	Weighed and recorded

Pre-processing of NIR Absorbance Spectra

Purpose:

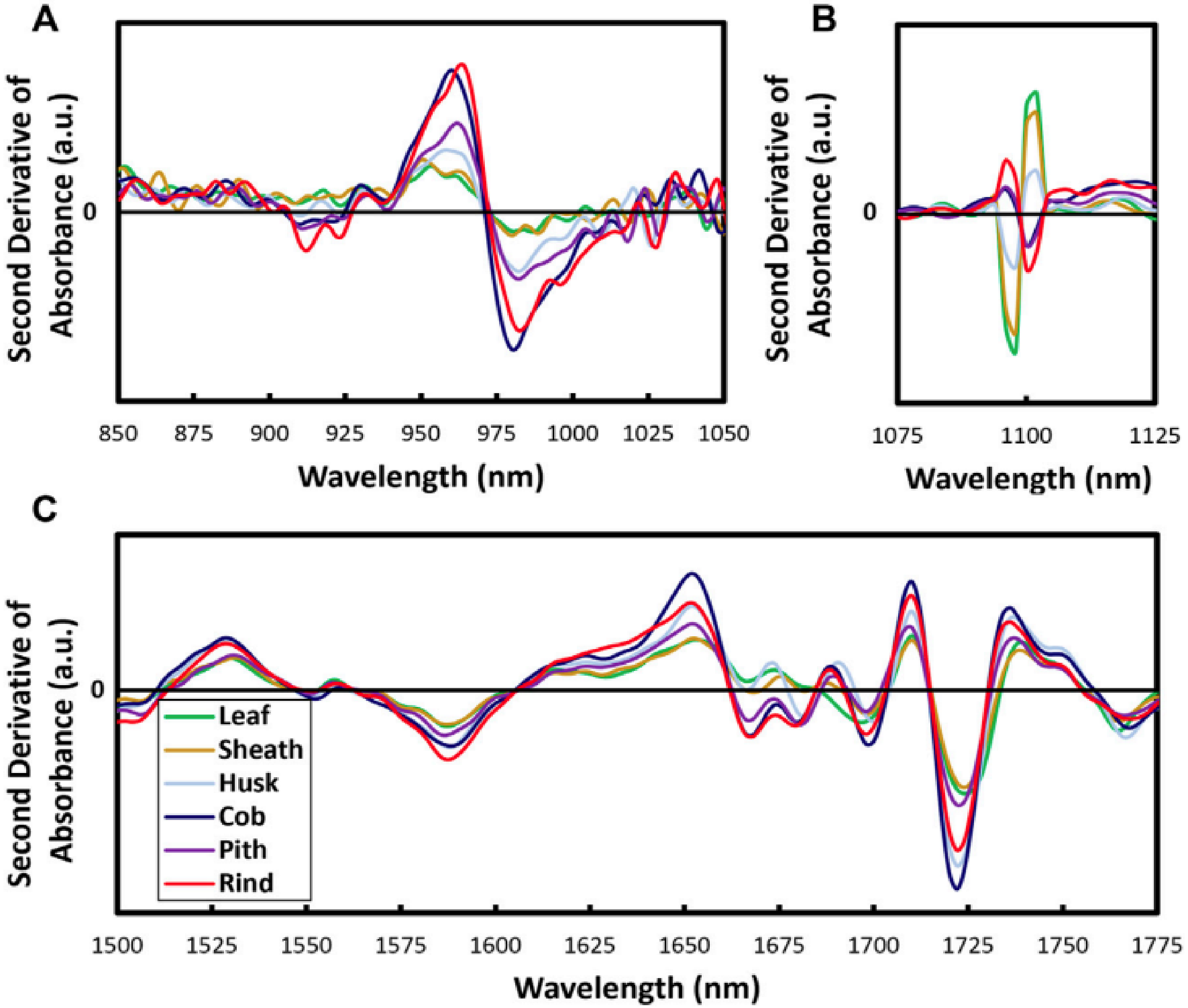
➢ Enhance the signal-to-noise ratio in the NIR spectra and remove irrelevant variations caused by sample handling or equipment.

Sample Pre-processing:

• **Standard Normal Variate (SNV):**
➢ Corrects for scattering effects and improves spectral quality by normalizing each spectrum.

• **Multiplicative Scatter Correction (MSC):**
➢ Adjusts for differences in path length and sample scattering.

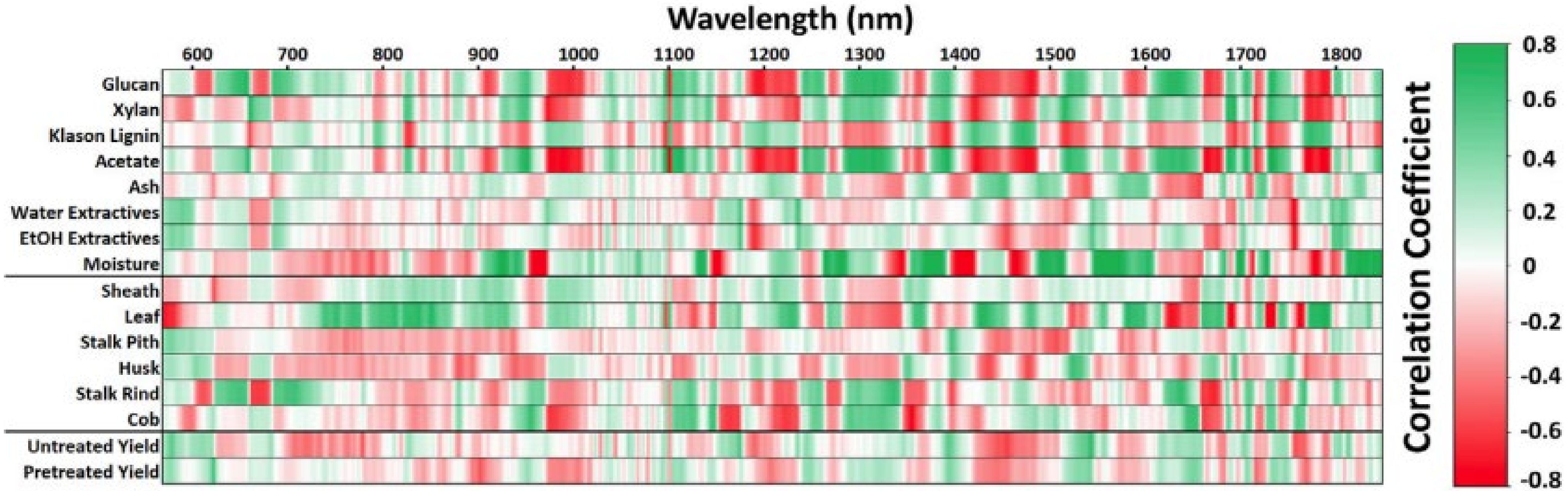
• **Second Derivative Transformation:**
➢ Enhances spectral features, especially for overlapping peaks, by focusing on changes in absorbance with respect to wavelength.



Exploratory Data Analysis

- For each sample,
  - ➢ The chemical composition and anatomical composition is recorded
  - ➢ Correlation analysis is carried between the processed spectra and the composition of each sample

- **Main purpose: Helps in identifying the important wavelengths for predicting each chemical component in corn stover**



Predictive models for Chemical and Anatomical Composition of Corn Stover

Purpose of Predictive Models:

Objective:

To predict the chemical composition (e.g., glucan, xylan, lignin) and anatomical composition (e.g., cob, husk, stalk) of corn stover using Near-Infrared Spectroscopy (NIRS) data.

Goal:

Enable real-time, non-destructive analysis of biomass for efficient biofuel production.

Predictive models for chemical and anatomical composition

Data Input:

- NIR spectral data (570–1850 nm).

Ground Label:

- Chemical composition (e.g., glucan, xylan) and anatomical fractions (e.g., cob, husk) determined through wet chemistry and manual sorting.

Training the models:

- 70% of data used for training the models.
- Preprocessing includes **Standard Normal Variate (SNV)**, **Multiplicative Scatter Correction (MSC)**, and **Second Derivative** transformation to improve model accuracy.

Testing the models:

- 30% of data used for testing, with **cross-validation** to avoid overfitting.

Table 3: Comparison of Predictive Models for Chemical and Anatomical Composition

Model	Strengths	Weaknesses	R^2 for Chemical Composition	R^2 for Anatomical Composition
Gaussian Process Regression	High accuracy, captures uncertainty	Computationally intensive	Up to 88%	**Up to 95%**
Partial Least Squares	Simple, handles multicollinearity	Lower accuracy than GPR	**84-88%**	92%
Neural Networks	Models complex relationships	Requires large datasets, lower performance	~80%	~85%

Predictive models for detecting anatomical parts in complex biomass mixtures

- **Problem Significance**: Biomass parts (cob, husk, stalk, leaf) have similar colors and textures, complicating automated separation. Accurate identification is crucial for efficient industrial processing like biofuel production.

- **Initial Approach and Challenges**:

➢ **Pixel-Wise Segmentation**: Used U-Net and Meta's Detectron2 (FCN-50 and FCN-101) for classification.

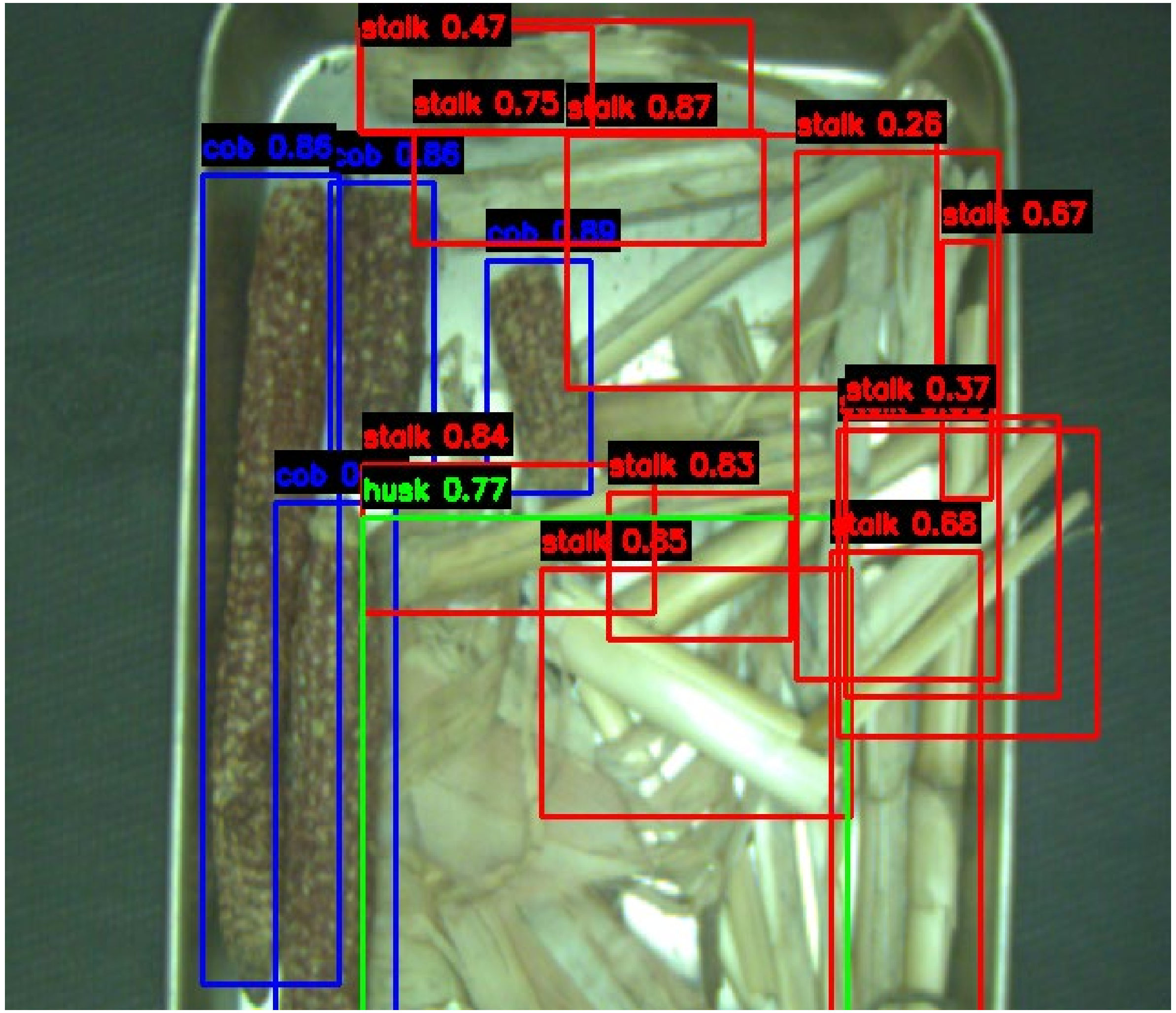
➢ **Issue**: High color and texture similarity led to poor model performance, with difficulty in distinguishing parts.

- **Solution with YOLOv8**:

➢ Reframed as an **object detection** problem, treating each part as a distinct object.

➢ **Result**: YOLOv8 provided improved accuracy by focusing on object boundaries rather than subtle color differences.

➢ **Future Direction**: Consider **volumetric weight analysis** of mixed samples for enhanced part separation in complex biomass.



Conclusion:

• **Optimized Resource Use**: LCIS DSS and Nutrient Expert® reduced water and fertilizer waste, supporting sustainable corn production.

• **Early Yield Forecasting**: Hybrid models (e.g., APSIM + ML) provided accurate early-season yield predictions, aiding timely decisions.

• **Biofuel Optimization with NIR Spectroscopy:** NIR spectroscopic data enabled real-time analysis of corn stover composition, improving biofuel yield by targeting high-cellulose components.

• **Enhanced Accuracy**: YOLOv8 improved biomass part detection (cob, husk, stalk, leaf) for precise characterization, addressing color and texture challenges.

• **Future Directions**: Volumetric analysis and real-time in-field detection offer promising next steps for advancing ML in agriculture.

Thank you!

Any questions?