Multiclass Classification on Soybean and Weed Species Using a Customized Greenhouse Robotic and Hyperspectral Combination System

Mohammed Raju Ahmed, Ph.D.

Postdoctoral Research Associate
Precision Agriculture laboratory
Department of Agricultural and Biosystems Engineering
North Dakota State University, Fargo, ND

Presentation Outline

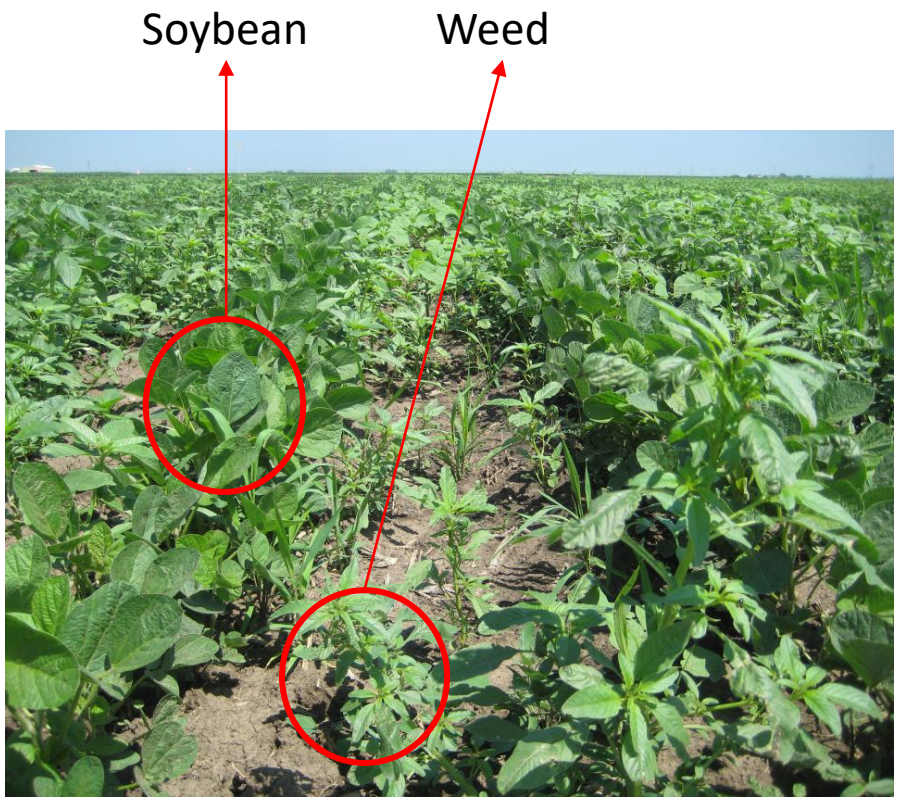
- **Introduction**
  - Background
  - Objectives
- **Martials and Methods**
  - Soybean and weed plants preparation
  - Hyperspectral image acquisition
  - Image analysis
- **Result and Discussion**
  - Model performance evaluation
  - Waveband information
- **Conclusion**
  - Summary of the research
  - Prospects direction

Introduction

Introduction

Research Background

- Soybean production was 339 million metric tons in 2020 globally (Anon 2020).
- Weeds multiple negative effects on soybean:
  - rivalry for water, light, and soil nutrients,
  - difficulty in harvesting operations, and
  - increase the diseases and pest risk.
- 37% soybean production loss occurs due to weeds alone while diseases and pests are accounted for 22% losses (Oerke and Dehne 2004).
- The most common weeds: horseweed, kochia, ragweed, redroot pigweed, waterhemp
- To adopt precision agriculture for controlling weeds, identification of weed and crop is crucial.



Anon. 2020. “International: World Soybean Production.” Retrieved November 29, 2021
Oerke, E. C., & Dehne, H. W. (2004). Safeguarding Production—Losses in Major Crops and the Role of Crop Protection. Crop Prot.,23(4):275–85.

Introduction

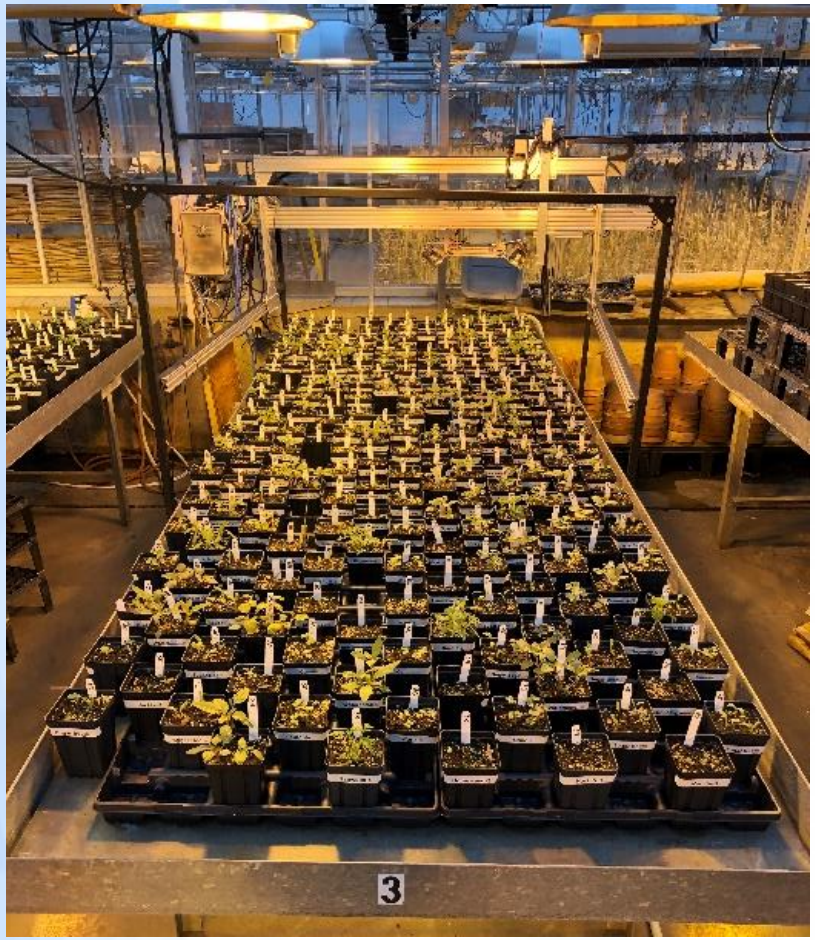
Objectives

- To investigate the potential of hyperspectral imaging for weed and crop identification in greenhouse.
- Customize robotic hyperspectral data collection scanning platform performance evaluation.
- Classification Model development for soybean and 5 types of weed and prime waveband investigation.

Materials and Methods

Materials and Methods

Sample preparation

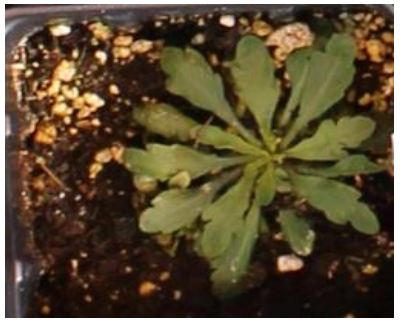



Soybean


Waterhemp


Redroot pigweed


Kochia


Horseweed

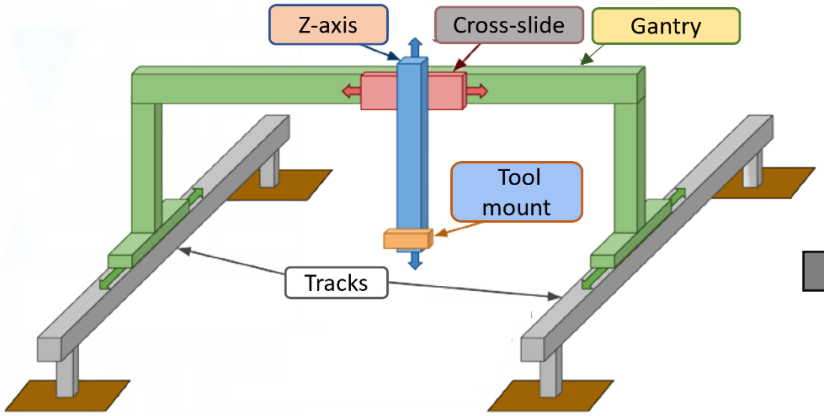

Ragweed

- Experiment environment: Greenhouse (NDSU)
- Used pots for soybean: 150
- Used pots for weeds: 100 for each weed (total 500)
- Temperature and humidity: ambient
- Image acquisition: after 21 days (plant height 10-12 cm)

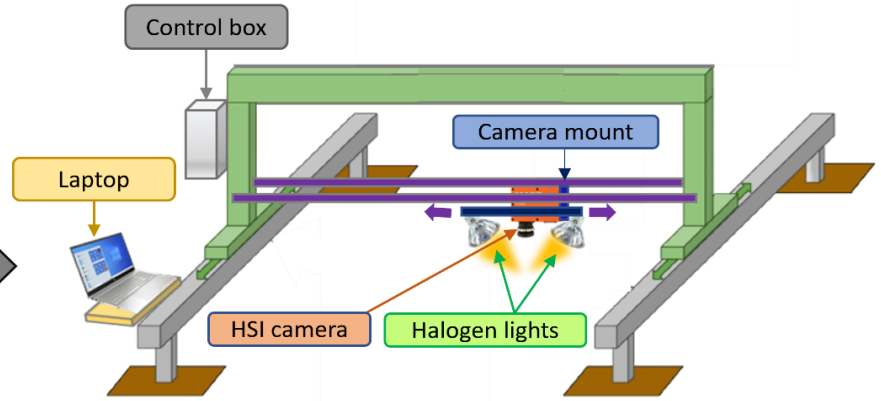
Materials and Methods

HSI image acquisition

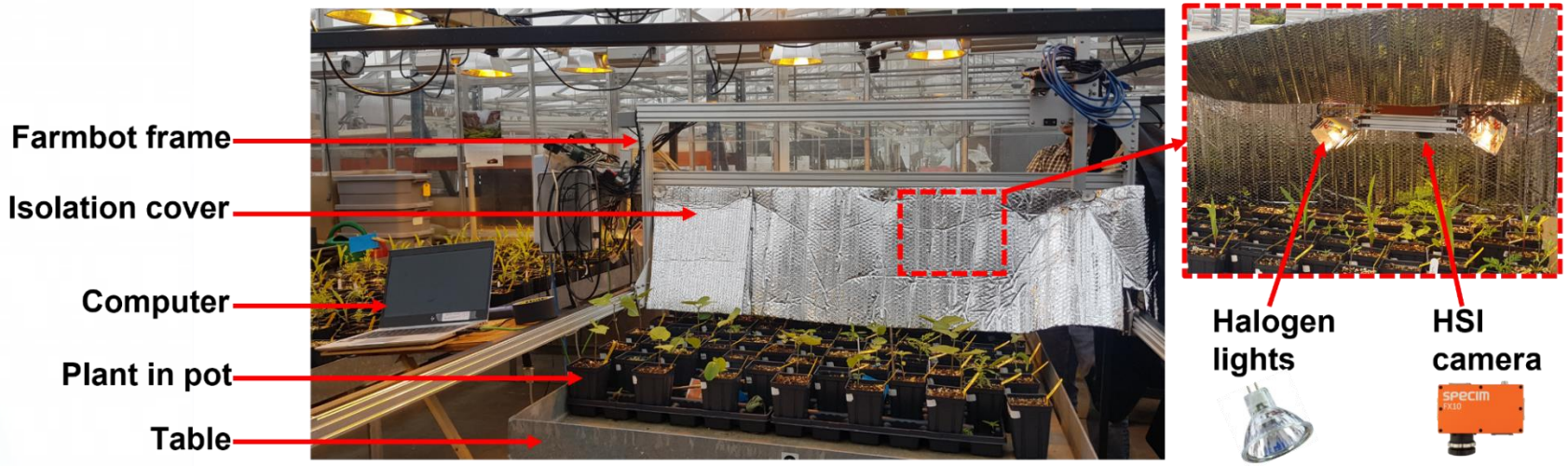
Lab based scanner



Original version (Farmbot)



Customized version

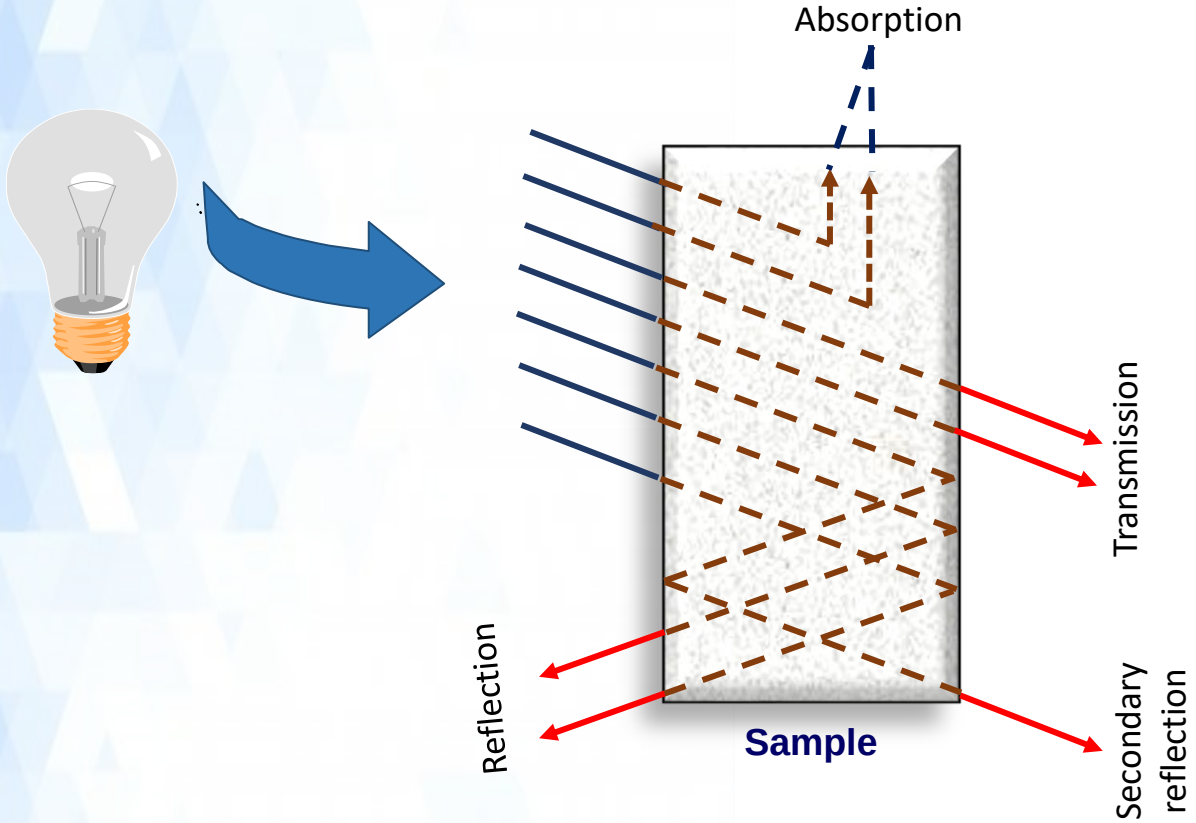


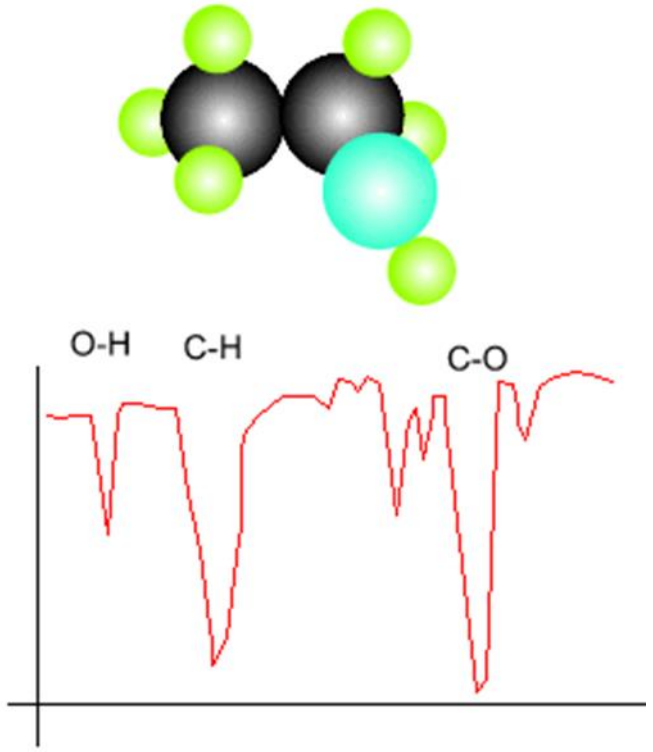
Camera: Specim FX10
Wavelength range: 400-1000 nm
Illumination: halogen bulbs
Control: Lumo Scanner software

Farmbot (model: Genesis v1.6, *FarmBot Inc, San Luis Obispo, CA, USA*)

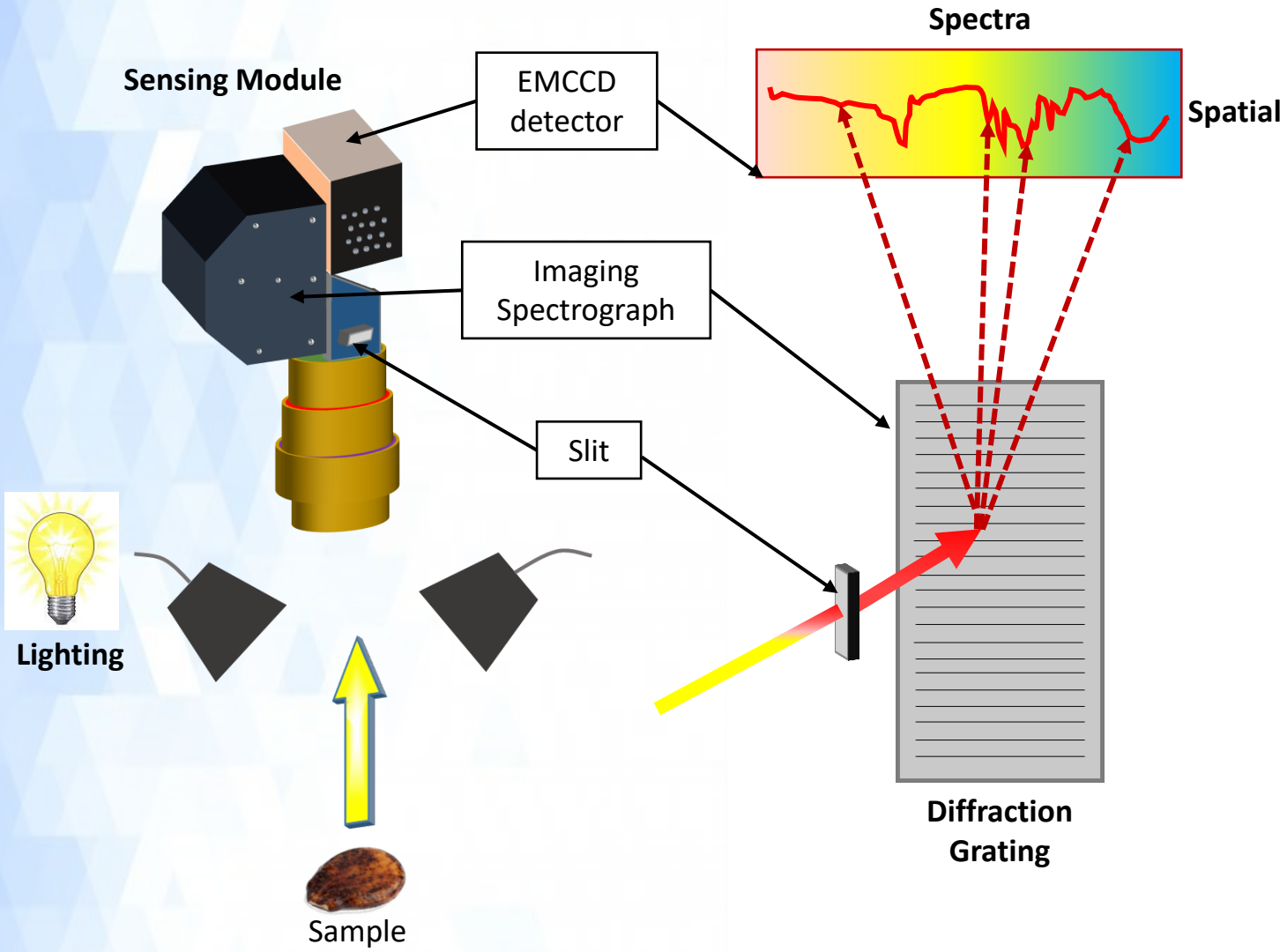
Materials and Methods

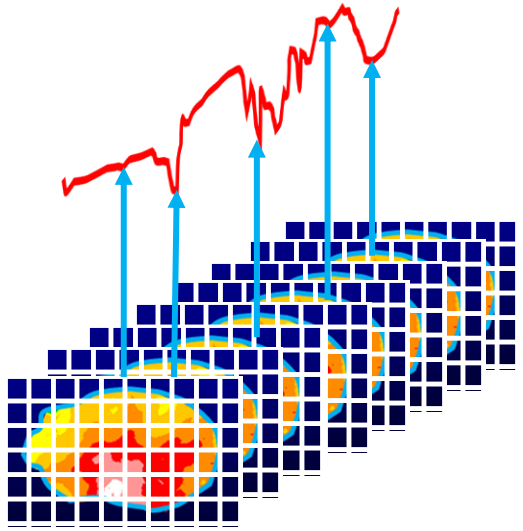
- Spectroscopy is the study of light interaction with matter.





Materials and Methods





3D hypercube

Materials and Methods

Spectral data extraction

- Soybean plant images: 252
- Weed images: horseweed-149, kochia-156, ragweed-151, redroot pigweed-118, waterhemp-157 (total 731)

- Image correction was by white and dark images following the equation:

$$I = \frac{I_0 - D}{W - D}$$

I - is the calibrated image,
I_0 - is the raw hyperspectral image,
W - represents a mean value of the white reference, and
D - represents a mean value of the dark reference.

- Median filter (window size: 3 × 3) was applied on each band image to reduce the noise and produce clean image.
- Image background was removed, and region of interest (ROI) was manually selected.

- Applied preprocessing methods are:
  - mean normalization, maximum normalization, range normalization,
  - multiplicative scatter correction (MSC),
  - standard normal variate (SNV),
  - Savitzky–Golay first derivatives, Savitzky–Golay second derivatives, and
  - data smoothing

Results and discussion

Results and discussion

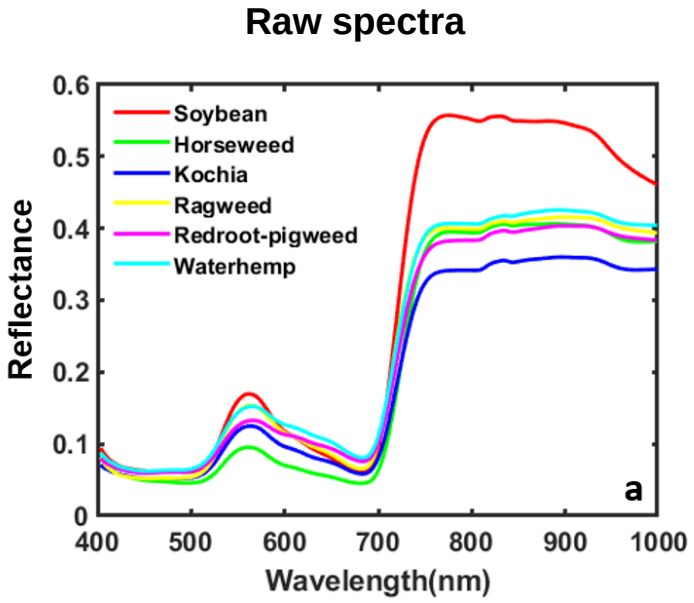
PLS-DA model (6 class)

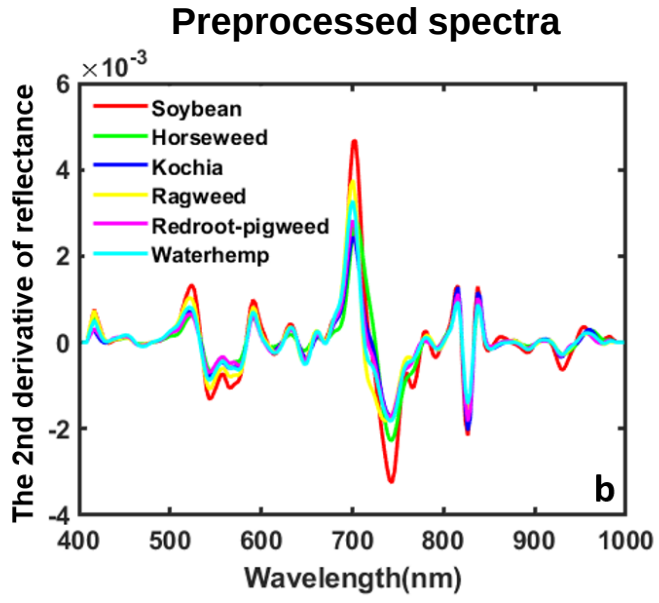
Preprocessing methods	Calibration accuracy (%)	Prediction accuracy (%)	LVs [a]
Mean normalization	88.2	82.6	15
Max normalization	81.5	77.4	15
Range normalization	82.4	77.8	15
MSC [b]	81.8	75.1	15
SNV [c]	80.3	76.2	15
Savitzky–Golay 1st [d]	86.4	81.6	9
Savitzky–Golay 2nd [e]	**91.2**	**86.2**	**10**
Raw [f]	74.7	68.4	12

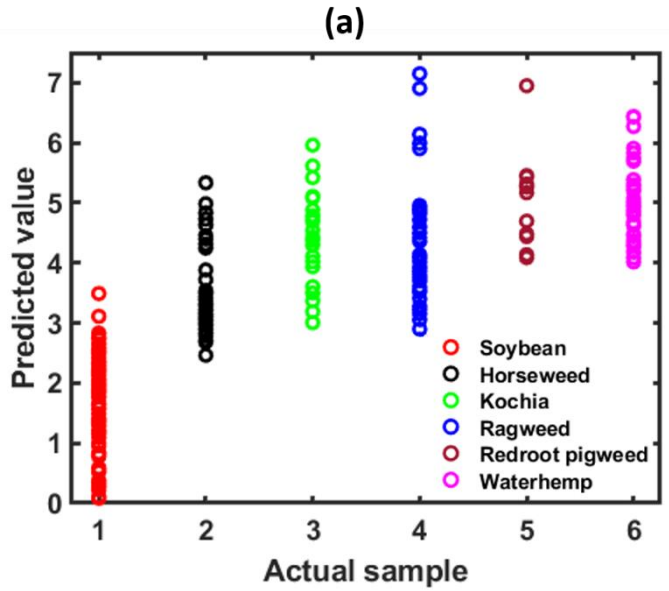
[a] LVs: Latent variables;[b] MSC: Multiplicative Scatter Correction; [c] SNV: Standard Normal Variate;[d] Savitzky-Golay 1st derivation;[e] Savitzky-Golay 2nd derivation and[f] Raw: Raw data model.

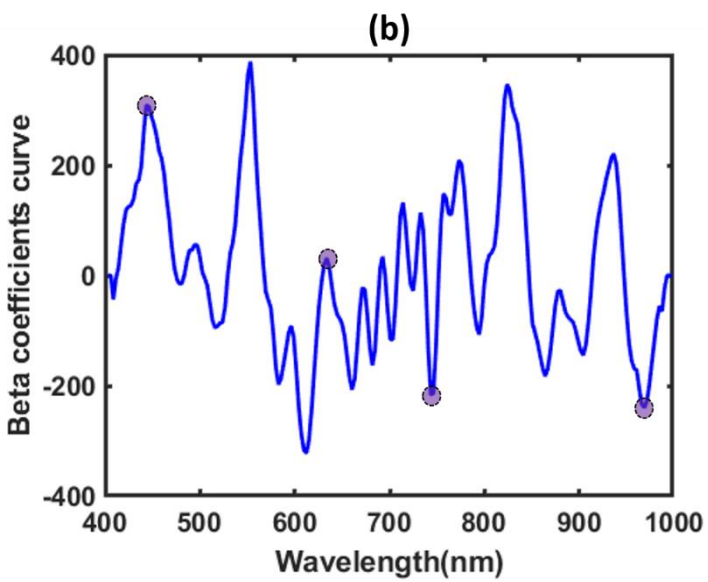
Important wavebands:

- 443 nm -α-caroteniod (C-H),
- 633 nm –chlorophyll (C-H)
- 743 nm –chlorophyll (C-H),
- 968 nm -moisture (O-H)







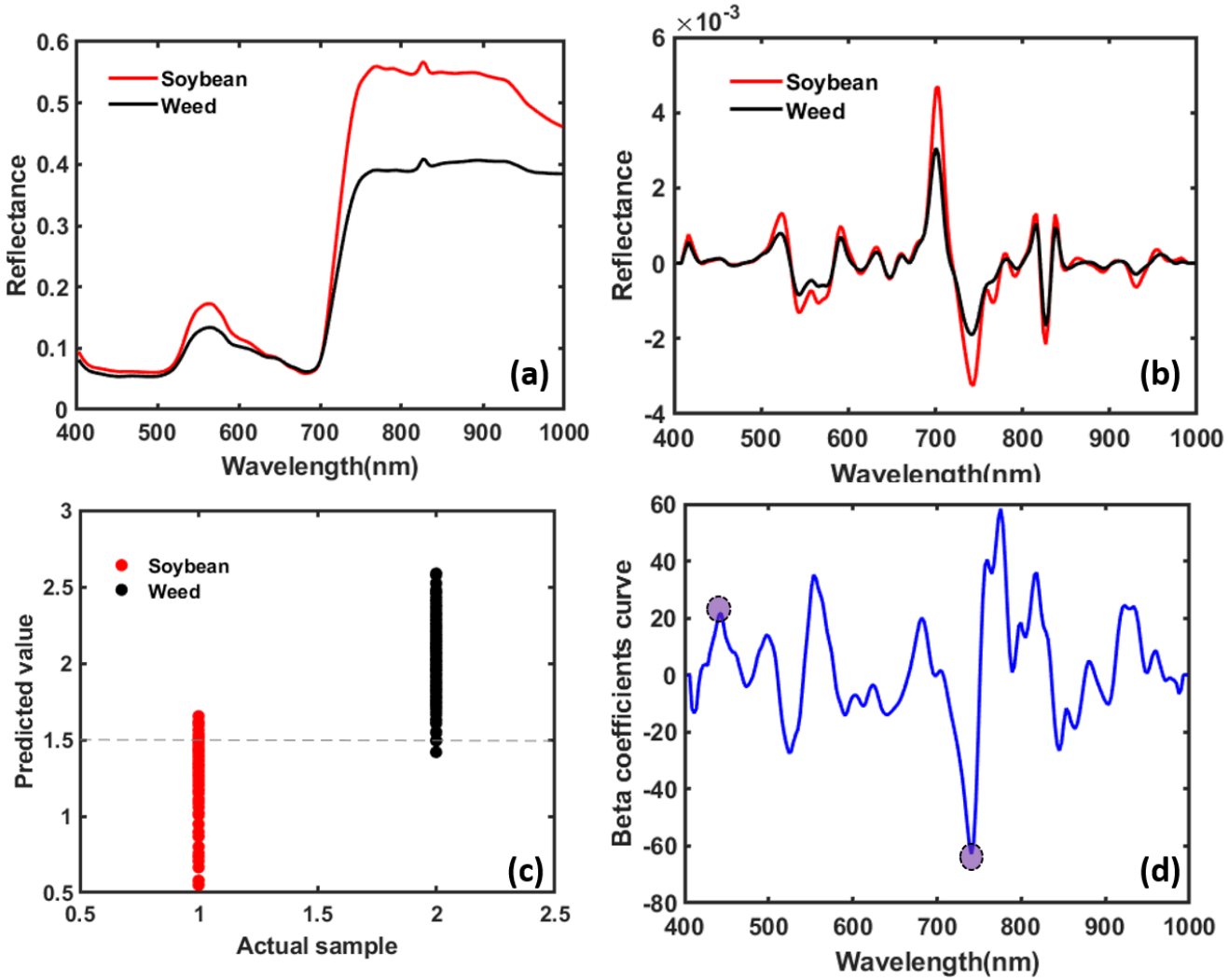


Results and discussion

PLS-DA model (2 class)

Preprocessing methods	Calibration accuracy (%)	Prediction accuracy (%)	LVs [a]
Mean normalization	90.3	84.4	14
Max normalization	83.7	81.4	15
Range normalization	81.5	77.7	15
MSC [b]	88.6	84.1	13
SNV [c]	81.2	75.8	15
Savitzky–Golay 1st [d]	88.7	83.3	10
Savitzky–Golay 2nd [e]	**93.4**	**89.4**	**10**
Raw [f]	77.6	71.3	15

[a] LVs: Latent variables; [b] MSC: Multiplicative Scatter Correction; [c] SNV: Standard Normal Variate; [d] Savitzky-Golay 1st derivation; [e] Savitzky-Golay 2nd derivation and [f] Raw: Raw data model.



Important wavebands:

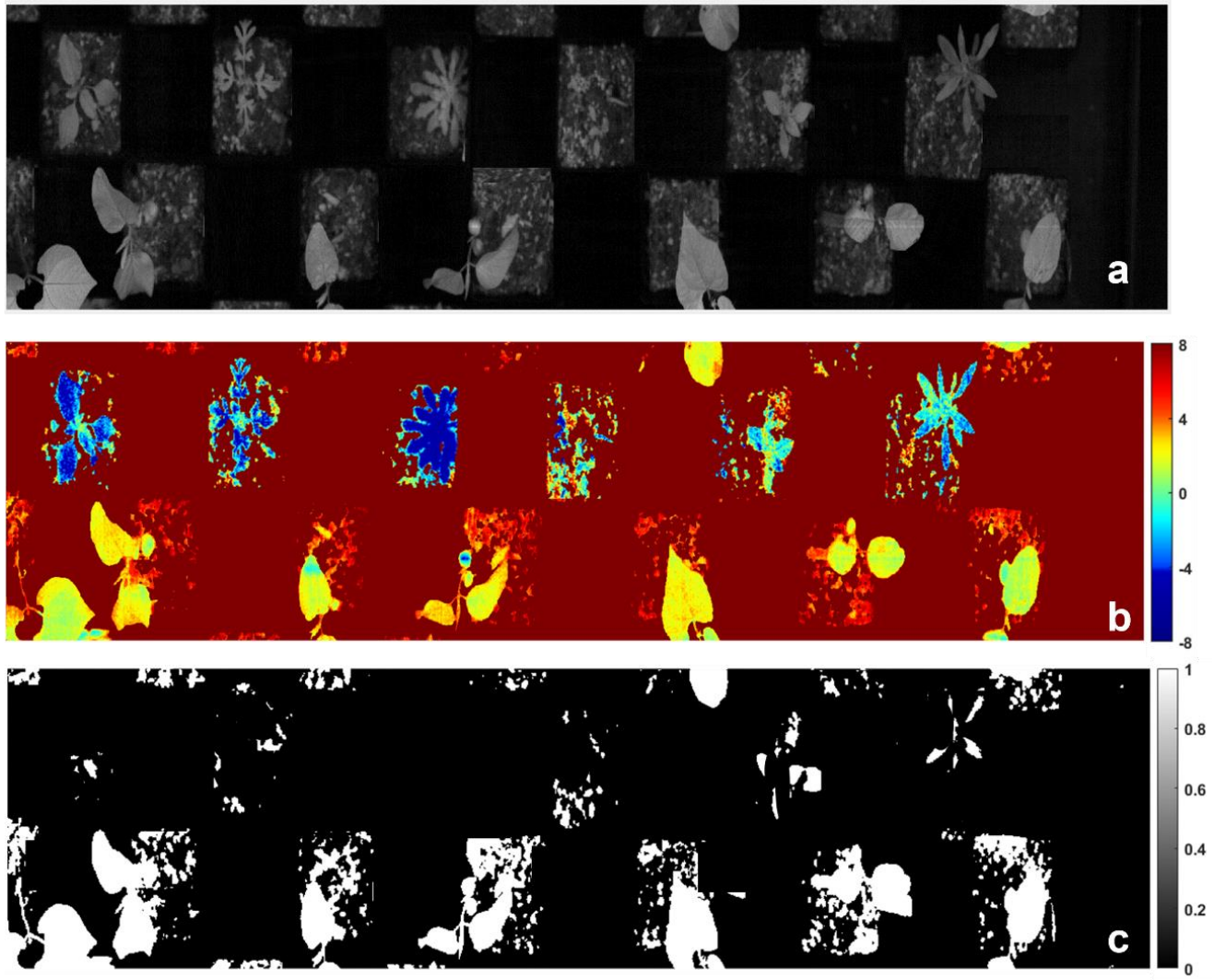
- 443 nm -α-caroteniod (C-H),
- 743 nm –chlorophyll (C-H),

Results and discussion

Chemical image using Savitzky-Golay 2nd derivative

Chemical image=

raw image × Beta coefficient curve



Conclusion

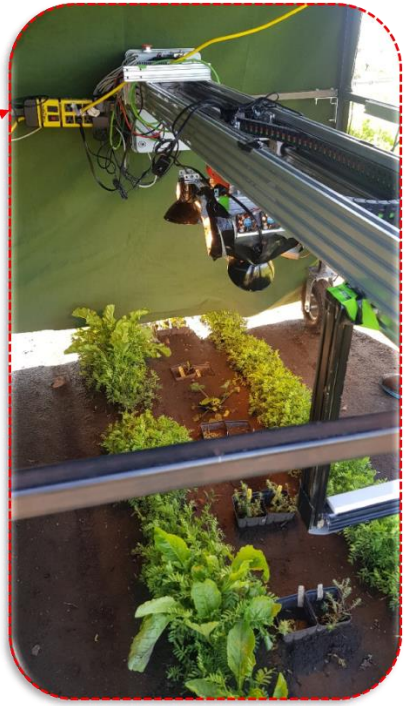
Conclusion

Summary of the study

- HSI was used to identify the soybean plants among five types of weeds where a semi-automatic robotic platform was used for image acquisition.
- Developing a multiclass PLS-DA model combine with Savitzky–Golay second derivatives the highest accuracy was found 86.2% and a higher accuracy (89.4%) was obtained for binary class
- The best wavebands were observed from the beta coefficient from 443 nm to 968 nm.

Prospects direction

Experiment performs in the real open field



Thank you