PALMER AMARANTH (*AMARANTHUS PALMERI* S. WATSON) IDENTIFICATION USING HYPERSPECTRAL IMAGING TECHNOLOGY

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Cristiano Manuel Barros da Costa

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Program:
Agricultural and Biosystems Engineering

November 2020

Fargo, North Dakota

North Dakota State University
Graduate School

Title

PALMER AMARANTH (*AMARANTHUS PALMERI* S. WATSON) IDENTIFICATION USING HYPERSPECTRAL IMAGING TECHNOLOGY

By

Cristiano Manuel Barros Da Costa

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Xin Sun
Chair

Yu Zhang

Thomas Peters

John Nowatzki

Approved:

11/17/2020
Date

Kenneth J. Hellevang
Department Chair

ABSTRACT

Palmer amaranth is a troublesome weed in modern day agriculture. Timely identification, along with adoption of site-specific weed management practices, will enable farmers to reduce Palmer amaranth control costs and improve efficacy. The feasibility of collecting hyperspectral imagery to identify Palmer amaranth and soybean was evaluated in the greenhouse and field. Hyperspectral images were collected across 224 spectral bands on palmer amaranth and soybean twice weekly from the one to three-leaf growth stage in three different runs (28 replications per run) temporally separated in a greenhouse. Partial least squares-discriminant analysis and soft independent modelling of class analogy models were developed to identify Palmer amaranth and soybean plants and had cumulative variations of 60% and 85%, and predictive abilities of 60% and 82%, respectively. This study concluded that hyperspectral imaging could be a potential tool to decipher Palmer amaranth from soybean plants.

ACKNOWLEDGEMENTS

I would like to acknowledge the United States Department of Agriculture (USDA) for providing funding for this research project. I would like to thank my adviser, Dr. Xin Sun, and research assistant professor, Dr. Yu Zhang, for all the support provided during these two years. To John Stenger: thank you for your huge contribution to the data collection and related work. To all my colleagues: I would also like to thank the help provided during data collection.

I would like to thank Dr. Tom Peters for helping me establish contact with the extension specialists in Nebraska where we collected the field data. I also want to thank Dr. Amit J. Jhala from the University of Nebraska- Lincoln for letting us collect data in his research plots.

I would like to thank John F. Nowatzki for his advice during these past two years.

To all my committee members: thank you for being part of this project.

DEDICATION

I would like to dedicate this work to my family.

To my Portuguese family who taught me the best values that I could ever have and worked so hard so that I could have a chance at a better life than they had an opportunity to have.

To my American family, the Schindele's, who always motivated me to move forward and get my Master of Science degree at North Dakota State University. Without them, I would not be where I am today.

To my girlfriend and life partner, one of the best people I have ever met, who helped me through this process and makes me grow every day.

This work is for all of them who I consider family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

GR.....Glyphosate-resistant.

GS .....Glyphosate-susceptible.

NDSU.....North Dakota State University.

NDVI.....Normalized vegetation index.

OTSU .....Otsu's binary automatic threshold method.

P .....PCA generated loadings.

PC.....Principal component.

PCA.....Principal component analysis.

PLS.....Partial least squares.

PLS-DA.....Partial least squares-discriminant analysis.

$Q^2Y_{(cum)}$ .....Predictive ability.

$R^2Y_{(cum)}$ .....Cumulative variation.

RAM .....Random-access memory.

SIMCA.....Soft independent modelling of class analogy.

T .....PCA generated scores.

X.....Matrix containing the spectra for calibration.

Y.....Response matrix containing the information about the species.

LIST OF APPENDIX FIGURES

1. INTRODUCTION

Palmer amaranth (*Amaranthus palmeri* S. Watson) is an invasive weed species and an aggressive competitor with crops throughout their growing seasons. Palmer amaranth, a native to the Sanoran desert region, was documented in the midwest region of the United States since 2001 (Ward et al., 2013) and was first confirmed in North Dakota in 2018 (Peters and Jenks, 2018). Palmer amaranth grows 5 to 8 cm per day to a height over 2 m tall with many branches, which is intensely competitive to slow-growing crops, and occupies large field resources by dominating the water and fertilizer available to crops (Davis et al., 2015; Horak and Loughin, 2000). Palmer amaranth populations have evolved resistance to several modes of action of herbicides (Ikley and Jenks, 2019). Therefore, Palmer amaranth has become a significant threat to crop growth and productivity.

Palmer amaranth herbicide resistance makes the control of this species challenging, since chemical control is the most common and efficient method of weed control in the United States. Mechanical weed control is a good alternative for control of herbicide-resistant species; nevertheless, it is not a very efficient alternative since it is generally slower. The efficiency of mechanical weeding can be improved if the problematic areas in the field can be mapped. Precision agriculture techniques associated with imagery processing are capable of identifying different weed species, which consequently increases weed control efficiency.

Hyperspectral imaging technology has been used in multiple areas in agriculture because of its plentiful spectral information compared with other imaging technologies. Hyperspectral imaging systems can collect broad wavelength reflectance profiles from an object. Many studies reported that hyperspectral imaging is a powerful technology for detecting weeds since different plant species show different profiles of spectrum reflectance (Farooq et al., 2019; Herrmann et

al., 2013; Huang et al., 2016; Matzrafi et al., 2017; Mink et al., 2020; Okamoto et al., 2007; Pantazi et al., 2016; Suzuki et al., 2008; Zhang et al., 2019).

Palmer amaranth is morphologically similar to other pigweed species. In the early growth stages, the only differentiating factor is the petiole length, which is usually longer than the leaf blade. This characteristic is not present in other pigweeds. Therefore, shape recognition methods can not be used to accurately classify this species. Hyperspectral imaging has the advantage of providing more information than any other imaging technology.

This research intends to test if hyperspectral imaging technology is a viable solution to identify Palmer amaranth species. The specific objectives included: (1) developing a high-resolution hyperspectral imaging system to collect hyperspectral imagery; and (2) classifying Palmer amaranth and soybean species using a hyperspectral imaging system in greenhouse and field conditions.

2. LITERATURE REVIEW

2.1. Palmer amaranth

Palmer amaranth is an invasive weed species native to the desert regions of the southwest United States and northern Mexico (Sauer, 1957). Palmer amaranth is a fast-growing weed and is highly competitive with row crops, including corn and soybean. A 1989 survey of agricultural professionals by the Southern Weed Science Society identified Palmer amaranth as a troublesome weed in crop production in the southern United States (Webster and Coble, 1997). By 2009, Palmer amaranth was ranked as the most troublesome cotton weed in nine states and the second most problematic weed in soybean production (Webster and Nichols, 2012). By 2014, Palmer amaranth had become one of the most concerning and economically important weed species in corn, cotton, and soybean in the United States (Norsworthy et al., 2014). Palmer amaranth interferes with crop yield, especially in row crops. A density of 10 Palmer amaranth plants per 9.1 m^2 area reduced cotton yield by 54% (Morgan et al., 2001). Palmer amaranth densities of 0.33 to 10 plants m^{-1} row length can reduce soybean yield by 68% (Klingaman and Oliver, 1994). Corn yield is also affected by the presence of Palmer amaranth. Palmer amaranth at densities of 1 to 6 plants m^{-1} row length resulted in a corn grain yield loss of 18% to 38% (Liphadzi and Dille, 2006).

Palmer amaranth was first identified in North Dakota in 2018 (Peters and Jenks, 2018). North Dakota had 5.6 million acres of soybean planted in 2019 (USDA National Agricultural Statistics Service, 2019); therefore, the spreading of Palmer amaranth may severely impact the profitability of North Dakota farms.

2.1.1. Biology

Palmer amaranth is a summer annual pigweed species in the *amaranthaceae* family. Palmer amaranth and its relative, waterhemp [Amaranthus tuberculatus (Moq.) J.D. Sauer] are dioecious species, meaning separate male and female plants. As a dioecious species, Palmer amaranth is an obligate outcrosser. Its pollination occurs by wind, leading to increased genetic diversity (Franssen et al., 2001). Palmer amaranth is a prolific seed producer, reliably able to produce at least 100,000 seeds per plant when it competes with a crop and able to produce nearly 600,000 seeds per plant in non-competitive scenarios (Sellers et al., 2003). Male plants produce many pollen grains capable of traveling 800 m and remaining viable up to 120 h (personal communication with Thomas Peters, NDSU). Therefore, this species has a great capability to adapt to different environmental conditions and herbicides.

Palmer amaranth's emergence period extends throughout the crop season and can occur after crop harvest (Webster and Nichols, 2012; Legleiter and Johnson, 2013). For example, researchers in Indiana observed Palmer amaranth emergence from early May until mid-September (Legleiter and Johnson, 2013). This extended emergence period forces producers to manage Palmer amaranth throughout the year, unlike other summer annual weeds that are typically managed only through early summer.

Palmer amaranth's season-long emergence characteristic combined with its rapid growth (up to 10 cm per day) and prolific seed production make it a problematic weed in agronomic cropping systems. Its photosynthetic capacity and utilization of water and light are greater than other C4 plants (Downton, 1975). Palmer amaranth has wider leaves than other pigweed species, including waterhemp (Mallard, 2009), which creates a sunlight interception advantage over its counterparts. Likewise, Palmer amaranth outcompetes most other plant species for light because

its leaves track the sunlight, thereby remaining perpendicular to the direct solar rays (Ehleringer, 1983). Finally, Palmer amaranth is drought tolerant since it can increase its leaf solute concentration, allowing the stoma to remain open longer during a period of drought (Mulroy and Rundel, 1977). Palmer amaranth has plant dry weight, leaf area, height, growth rate, and water-use efficiency greater than other pigweeds, including waterhemp, redroot pigweed, and tumble pigweed (Horak and Loughin, 2000). This weed species will compete with the intended crop by growing taller than the crop canopy. Palmer amaranth can produce and accumulate large quantities of biomass and grow up to 2 m in height.

2.1.2. Herbicide-resistance in Palmer amaranth

Palmer amaranth's reproductive habits also have encouraged herbicide resistance. Palmer amaranth populations have evolved resistance to herbicide groups with multiple mechanisms of action, including acetolactate synthase inhibitors (Group 2), mitotic inhibitors (Group 3), synthetic auxin inhibitors (Group 4), photosystem II inhibitors (Group 6), 5-enolpyruvylshikimate-3-phosphate synthase inhibitors (Group 9), protoporphyrinogen oxidase inhibitors (Group 14), and 4-hydroxyphenylpyruvate dioxygenase inhibitors (Group 27) (Ikley and Jenks, 2019). Some Palmer amaranth biotypes have also been found to be resistant to multiple modes-of-action herbicides (Heap, 2015).

2.1.3. Identification

Palmer amaranth identification can be difficult, especially in early growth stages. Misidentification perhaps was one of the reasons for the success of Palmer amaranth spreading across the United States. The first, and often critical, step to managing Palmer amaranth (or any weed) is to scout and identify the species that exists in each agronomic field.

It is easy to misidentify Palmer amaranth because it looks similar to three other common amaranth species: redroot pigweed, Powell amaranth, and waterhemp. The resemblance is especially strong during the seedling stages of growth. Only redroot and Powell amaranth have hairs (pubescence) on their stems and leaf surfaces. The fine hairs will be most noticeable on the stems towards the newest growth. Palmer amaranth and common waterhemp do not have hair on any surface. Looking for pubescence is a quick and easy way to differentiate redroot and Powell amaranth from waterhemp and Palmer amaranth. The leaf shapes of amaranth can vary quite a bit within a single species and is subjective; however, there are general shapes that distinguish the species. Waterhemp leaves are generally long, linear, and lanceolate, whereas Palmer amaranth leaves are wider and ovate to diamond-shaped. Redroot and Powell amaranth leaves are similar to Palmer amaranth leaves and have a round to ovate shape; however, they have hairs while Palmer amaranth and common waterhemp leaves do not.

The petiole is the stem-like structure that connects the leaf blade to the main stem. In Palmer amaranth, the petioles (especially on older leaves) will be as long (or longer) than the leaf blade itself. The petioles of waterhemp, on the other hand, will be shorter than their long, lance-shaped leaves. An easy way to determine the petiole length is to simply pull a leaf and petiole off a plant and bend the petiole back over the leaf blade to compare the petiole and leaf blade lengths. This is the most consistent and reliable characteristic that differentiates Palmer amaranth from common waterhemp, and it is most evident on the oldest leaves of the plant (lowest on the stem).

The flowering structure/inflorescence is also a differentiating factor. The flowering structure in Palmer amaranth plants is open, non-branched, can reach up to 90 cm long, and frequently is on the top of the plant, whereas the flowering structure in waterhemp is multi-

branched and located near the top of the plant and at the tips of branches. The inflorescence of a mature female Palmer amaranth plant is sharp and prickly, while the inflorescence of a mature female waterhemp plant is soft and smooth to the touch (Ikley and Jenks, 2019).

2.2. Hyperspectral technology in agriculture

2.2.1. Hyperspectral imagery

Camera and imaging developments have improved the quality of imagery used in precision agriculture. Today there are cameras with better spatial resolution that can capture many bands compare to the standard RGB (red, green, and blue) cameras. Multispectral imaging, having capabilities beyond the RGB imaging, provides more data to the user. In agriculture, for example, multispectral imaging made the study of plant health possible (Lowe et al., 2017). However, hyperspectral imaging brought new capabilities, as its high number of continuous narrow bands provide a larger quantity of data related to the material characteristics (Qureshi et al., 2019).

Hyperspectral systems include a spectrograph that records the reflectance in a wide range of the spectrum, from ultraviolet to near-infrared into a digital sensor (Bock et al., 2010). Data acquired by multispectral and hyperspectral cameras have three-dimensional structures consisting of two spatial dimensions and one spectral dimension. These structures are usually called ‘hypercubes’ or ‘datacubes’ (Figure 1) (Qureshi et al., 2019). Processing techniques applied in analyzing hyperspectral data are applied in every layer in the spectral dimension, which requires a high level of computational power.

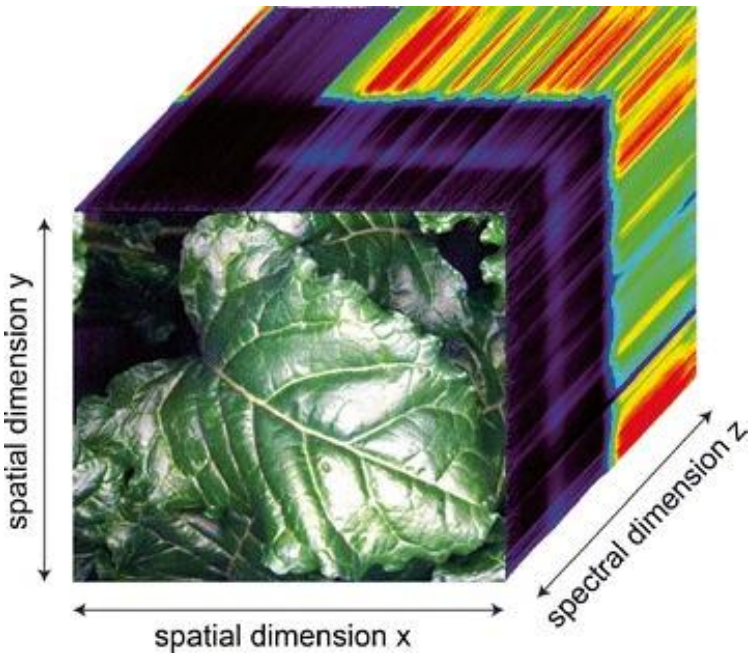


Figure 1. Hyperspectral data are organized in a “hypercube” with three dimensions, two spatial dimensions, and one spectral dimension.

2.2.2. Palmer amaranth identification using remote sensing techniques

Palmer amaranth morphology is similar to other pigweeds, especially in early growth stages. Therefore, identifying this species using morphology characteristics can result in species misidentification (Legleiter and Johnson, 2013).

Some cameras can visualize light reflectance wavelengths that cannot be processed by the human eye, which provides more information that can be used to facilitate Palmer amaranth identification. Weed species have been identified using multispectral and hyperspectral reflectance, associated with image processing techniques (Wang et al., 2019). However, limited studies were found on Palmer amaranth species identification through imaging technologies. A study compared different accessions of Palmer amaranth and pitted morning-glory (*Ipomoea lacunosa* L.) (Gray et al., 2007). The authors used two analysis techniques, linear discriminant analysis and best spectral-band combination analysis, using data collected with a hand-held hyperspectral radiometer. The results of this work suggest that there are slight spectral

reflectance differences between accessions; however, the genetic diversity within species made the classification difficult. On the other hand, the differentiation between Palmer amaranth and pitted morning glory had 100% accuracy in their differentiation (Gray et al., 2007). These results demonstrate the potential of the hyperspectral data for weed species differentiation.

A study by Andersen et al. (1985) identified differences in the spectral reflectance using an airborne platform to collect data. The authors demonstrated that their methods were able to distinguish climbing milkweed [*Sarcostemma cyancboides* (Decne.) Schltr.] in orange (*Citrus sinensis* L.) trees; ragweed parthenium (*Parthenium hysterophorus* L.) in carrot (*Daucus carota* L., var. sativa 'Long Imperator'); johnsongrass (*Sorghum halepense* L.) in cotton (*Gossypium hirsutum* L.); London rocket (*Sisymbrium irio* L.) in cabbage; and Palmer amaranth (*Amaranthus palmeri* S. Watson) in cotton. A more recent study aimed to distinguish glyphosate-resistant (GR) Palmer amaranth from glyphosate susceptible (GS) Palmer amaranth plants using hyperspectral imagery (Reddy et al., 2014). It was found that GS plants have higher light reflectance in the visible region and lower light reflectance in the infrared region of the spectrum relative to the GR plants. The most significant reflectance wavelength ranges to separate these two groups of plants were from 400 to 500 nm, 650 to 690 nm, 730 to 740 nm, and 800 to 900 nm. The authors concluded that hyperspectral imagery could distinguish GR and GS Palmer amaranth plants before any glyphosate application (Reddy et al., 2014). Multispectral reflectance data have also been used to identify soybean, sunflower (*Helianthus pumilus* L.), redroot pigweed (*Amaranthus retroflexus* L.), and velvetleaf (*Abutilon theophrasti* Medik) species based on their leaf shape (Fletcher and Reddy, 2016). Shortwave-infrared wavelengths were identified as the most important variables used for discrimination of soybeans and pigweeds, such as Palmer amaranth (Fletcher and Reddy, 2016).

3. MATERIALS AND METHODS

3.1. Greenhouse experiment for Palmer amaranth hyperspectral image data collection

Palmer amaranth and soybean were grown in the spring, 2019, in the Waldron Greenhouse at North Dakota State University (NDSU). Palmer amaranth plants were grown from seed provided by Kansas State University. The soybean variety used was AG08X8 RRXtend (Asgrow, Bayer Crop Science, St. Louis, MO). The experiment was a randomized complete block design with two replications and consisted of three different runs. Each run was 28 replications. The first run was seeded on April 10, 2019, the second run was seeded on April 17, 2019, and the third run was seeded on April 24, 2019. The experiment design accounted for temporal and spatial variance in the plant's growth.

Seeds were planted in labeled cone cells SC10 U (Ray Leach, United States) with a cell diameter of 3.8 cm and a depth of 19 cm. The cones contained a commercial potting mix (SunshineMix #1; Sun Gro Horticulture, Bellevue, WA). Plants were irrigated manually in each cell, daily. Each cell contained a central drainage hole at the bottom, as well as four lateral holes. Cone cells were arranged (Figure 2) in 70 × 30.5 ×17.5 cm racks that could hold 98 cells.

PA	Soy	PA	PA	Soy	PA	Soy
Soy	PA	Soy	Soy	PA	Soy	PA
Soy	PA	Soy	Soy	Soy	PA	PA
PA	Soy	PA	PA	PA	Soy	Soy
PA	Soy	PA	Soy	PA	PA	Soy
Soy	PA	Soy	PA	Soy	Soy	PA
Soy	Soy	Soy	PA	PA	Soy	PA
PA	PA	PA	Soy	Soy	PA	Soy

Figure 2. Experiment design for the third run.

The greenhouse temperature was kept at around 21 C. Natural light was supplemented by 400-watt high-pressure sodium bulbs set for 12 hour day lengths. Soybean emerged approximately 14 days after seeding and Palmer amaranth approximately three days longer (Figure 3). Soybean emerged uniformly, but Palmer amaranth emergence was highly variable.

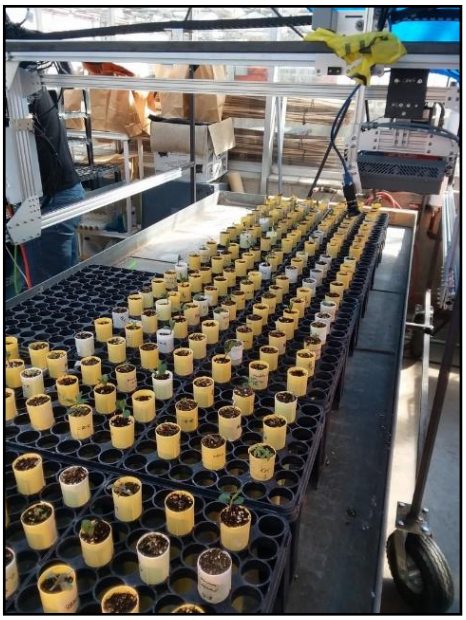


Figure 3. Palmer amaranth plants at the germination stage, May 2019.

Data collection began when Palmer amaranth and soybean were at the seedling stage, generally around 2.5 cm tall. The hyperspectral data acquisition occurred twice weekly in each run until the plants reached a height of about 13 cm. Data collection for the first run started on April 26, 2019, the second run on May 03, 2019, and the third run on May 10, 2019. The data collection in every run lasted about three weeks.

3.2. Field experiment for Palmer amaranth hyperspectral image data collection

Field data experiments were conducted near Carleton and Clay Center, Nebraska, in July 2019. The field at Carleton, was a commercial soybean field, with an indigenous Palmer amaranth density at 80 plants per 10 m^2 (Figure 4). The second field was at the U.S Meat Animal Research Center, Clay Center, NE. The field was divided into soybean plots, with indigenous

Palmer amaranth density at 80 plants per 10 m^2 (Figure 5). Canopy hyperspectral reflectance was collected in both locations within a three-day period (July 10 to July 12, 2019).

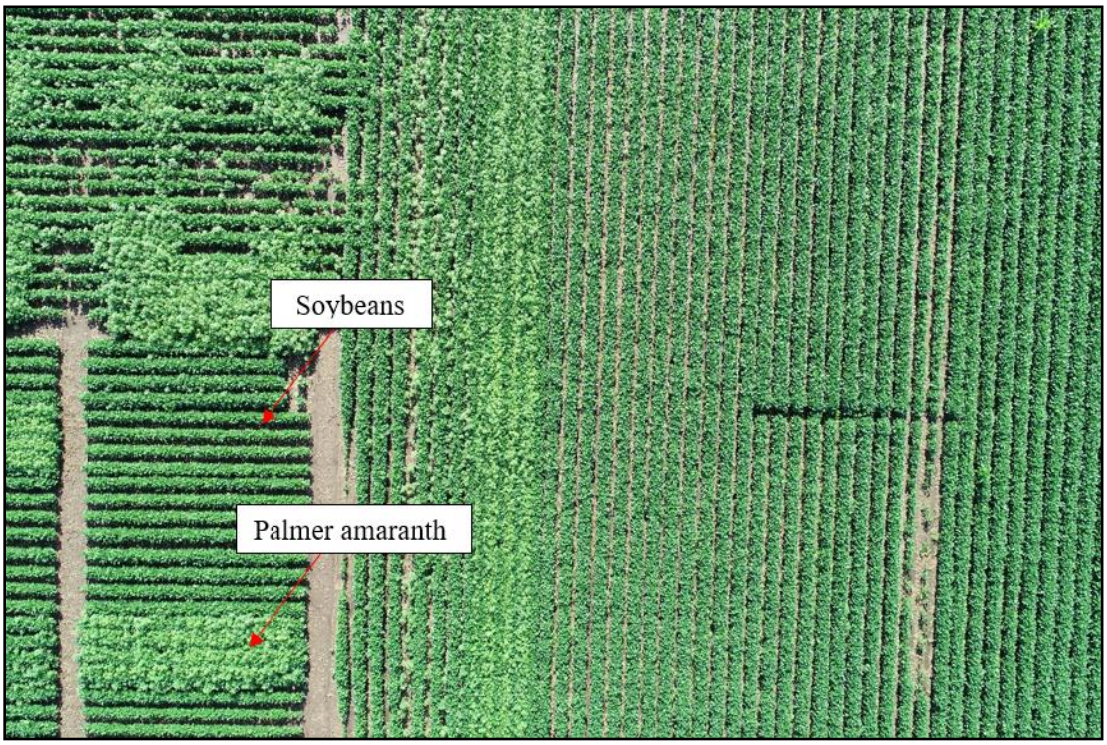


Figure 4. Palmer amaranth indigenous population in a soybean field, Carleton, Nebraska, July 2019.

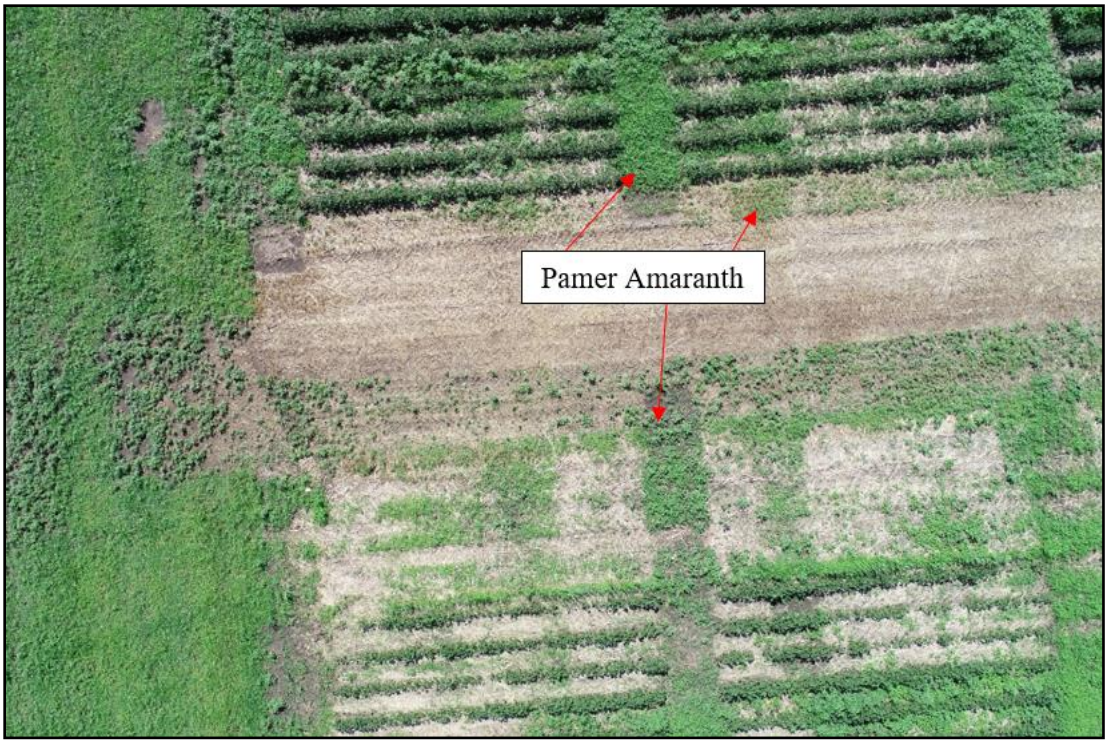


Figure 5. Palmer amaranth population in soybean plots, Clay Center, Nebraska, July 2019.

3.3. Hyperspectral imaging system

An attachable hyperspectral imaging system was developed to conduct data collection under different conditions. The hyperspectral imaging system was composed of the imaging and scanning units. The imaging unit included a push-broom hyperspectral camera model FX-10 (SPECIM, Finland) in the visible and near-infrared with a spectrum ranging from 400 nm to 1000 nm with 224 bands and a frame rate of 30 frames per second (Figure 6). The spatial sensor resolution was 1024 pixels and the field of view was 38 degrees. This device was equipped with an integrated shutter, which allowed the dark reference to be collected without the need to block the light manually. The white reference collection was carried out using a Teflon board with a light reflection of 99%. Six halogen lamps were integrated in the imaging unit.

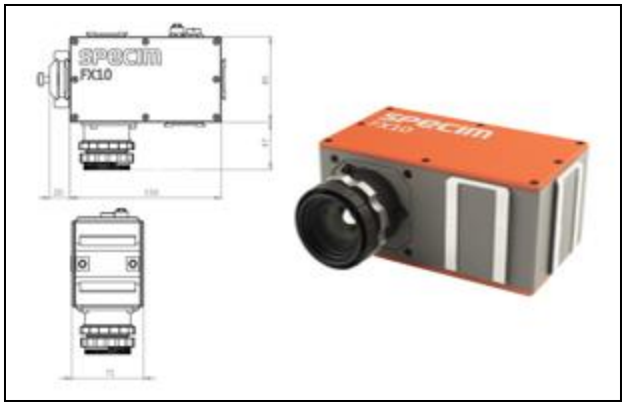


Figure 6. Hyperspectral camera model FX-10 (SPECIM, Finland).

A threaded rod with a width of 12 mm and a length of 1500 mm, powered by a motor stepper-motor (MDI1FRD17A4-EQ, Schneider Electric Intelligent Motion Systems, USA), was responsible for the scanning unit movement. The longest scanning distance was designed at 1.5 m, encompassing the different crop row options in the field, e.g. 0.38-cm row or 0.76-cm row. LUMO Recorder (SPECIM, Finland) software was used for data collection.

In the greenhouse experiments, the hyperspectral imaging system was attached to a 3-dimensional scanning platform, which allowed the movement of the hyperspectral imaging system in 3-dimensional directions, according to the position and size of the target plants and scanning area. The bottom platform frame was made of steel to reduce vibration when operating. Wheels were incorporated into the platform to make it easier to move over the bench and field plots (Figure 7).

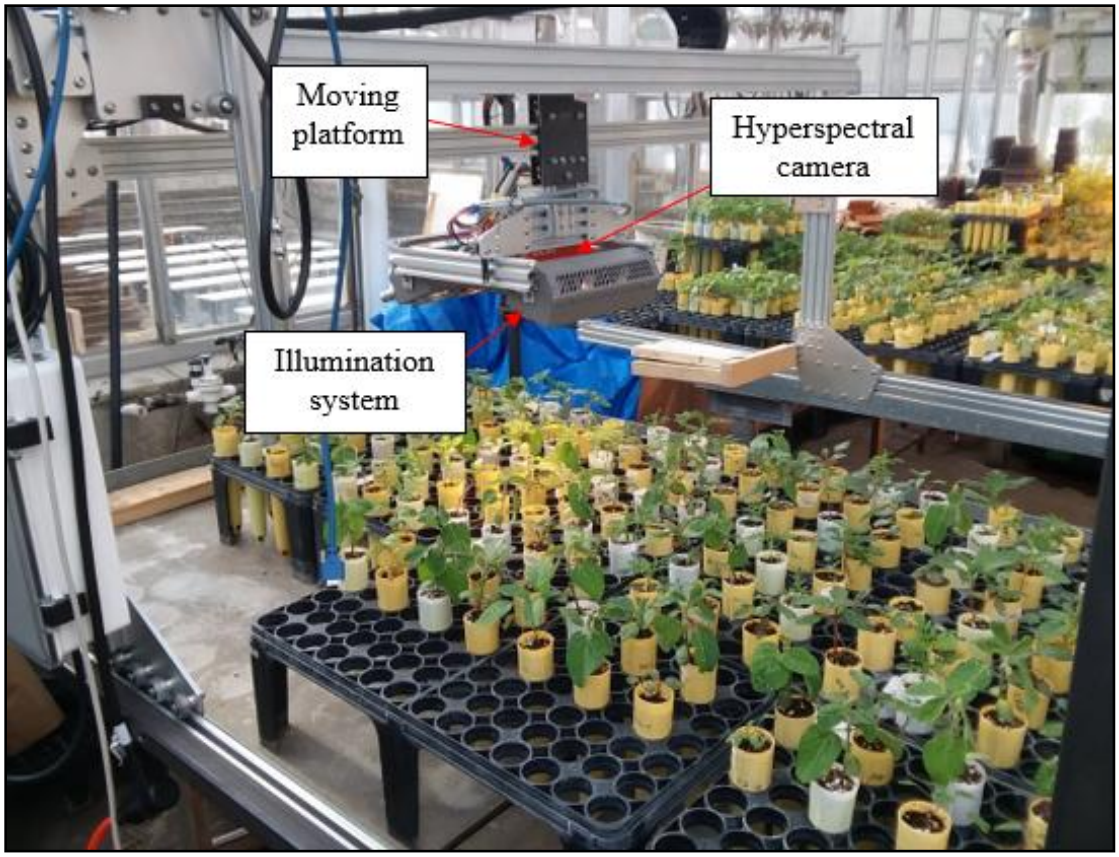


Figure 7. Scanning platform collecting data, May 2019.

A second platform was developed to collect hyperspectral reflectance in Nebraska, as the greenhouse platform was large and had fragile components that would have to be transported in a trailer during long travel. Therefore, a second smaller and more portable platform was developed. The second platform consisted of an aluminum axis built into a steel stand. This platform did not have any wheels. The operator had to carry it to move to a different location. The platform was positioned perpendicular to the crop rows during the data collection (Figure 8). Soybean and Palmer amaranth plants were captured in the same scan to reduce the variability among different pictures caused by ambient conditions.

Figure 8. Field data collection platform in action, Nebraska, July 2019.

3.4. Hyperspectral image data pre-processing

The output from the image acquisition software was a raw format file. This file was processed to eliminate irrelevant information caused by external factors, such as the influence of nuisance signals due to illumination effects and detector sensitivity.

These misleading characteristics could be presented in function of the wavelength or as variations in the spatial domain. Spectral calibration was necessary to correct these differences in the spectral and spatial domains (Mishra et al., 2017). The image calibration and correction from reflectance to pseudo-absorbance were conducted in the software Evince (Prediktera, Sweden), according to equation (1):

$$I_{\lambda,n} = -log_{10}\left[\left(\frac{S_{\lambda,n}-B_{\lambda,n}}{W_{\lambda,n}-B_{\lambda,n}}\right)\right] \quad (1)$$

where n is the pixel index variable of the reorganized hypercube; $I_{\lambda,n}$ is the absorbance intensity of pixel n at wavelength λ; $S_{\lambda,n}$ is the sample pixel n at wavelength λ; $B_{\lambda,n}$ is the dark reference pixel n at wavelength λ; and $W_{\lambda,n}$ is the white reference pixel n at wavelength λ.

The calibrated images from both plant species were organized in a single mosaic, which facilitated the processing methods used. Pixels associated with the background were removed from the mosaic image to reduce noise in the data. The background pixels in the mosaic were eliminated by calculating the Normalized Vegetation Index (NDVI) to enhance the features of plant vegetation attributes, which makes the hyperspectral image background segmentation more accurate. A binary automatic threshold using the OTSU's method was conducted. This is a method used to separate two classes of pixels and is commonly used in agronomic applications (Otsu, 1979). This method calculates the optimum threshold based on minimizing intra-class variance. It is easy to compute, stable, and does not require prior information (Torres-Sánchez et al., 2015). The binary layer was applied to the dataset as a mask, and each pixel of the mask was multiplied by each spectral layer of the data cube (Figure 9).

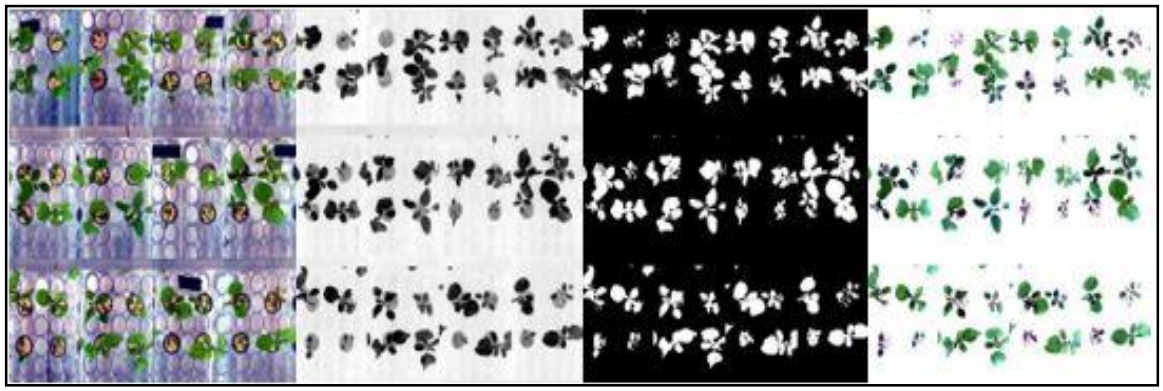


Figure 9. Pre-processing sequence: a) Original mosaic b) Mosaic after normalized difference vegetation index calculation c) Mosaic after threshold method d) Final pre-processed mosaic.

This technique assigned the value of zero to the pixels that did not represent any vegetation. Values of zero were ignored by the software when the final dataset was exported. All

segmentation was done in the software ENVI (L3 HARRIS Geospatial, USA). Data were mean-centered, as the mean value of a variable was subtracted from each individual value (Burger, 2006). Standard normal variate and multiplicative scatter correction techniques were applied to correct baseline shifts and slopes in the data (Mishra et al., 2017).

3.5. Palmer amaranth and soybean classification methods

It was necessary to associate spectral characteristics with each vegetation type to distinguish plant species. These spectral characteristics are usually related to chemical and physical properties. Due to the hyperspectral data's complexity, the information is typically mixed in spectral and spatial dimensions, which means that the chemical or physical characteristics are not associated with a single wavelength, but instead with multiple wavelengths. This complexity of hyperspectral data requires multivariate data analysis and reduction techniques to reduce the data's dimensionality.

Principal component analysis (PCA) is a powerful tool to reduce the dimensions of large datasets and is widely used in hyperspectral data processing.

The dataset was unfolded, so the three-dimensional matrix became a two-dimensional matrix with columns corresponding to wavebands and rows corresponding to pixels. The PCA generated matrices of scores (T) and loadings (P) from each pixel's spectrum. A score plot was generated, in which similar pixels were plotted close together. The PCA score plots were used to visualize possible clusters and eliminate outliers.

Partial least squares-discriminant analysis (PLS-DA) modelling was used to classify the different plant species. This type of modelling has been used in hyperspectral imaging classification with favorable results (Amigo et al., 2015; Herrmann et al., 2013). A PLS-DA is a supervised class-modeling technique that uses a PLS algorithm to predict whether a sample

belongs to a determined class. To create a calibration method, a matrix (X), containing the spectra for calibration, and a corresponding matrix (Y), containing the information about the species, were created. Due to the high amount of samples (pixels) in the dataset, about 50% of the pixels were used to train the model, unlike the common 70 to 75% in published studies (Eddy et al., 2014). Training a model with 70% of the current dataset was not possible due to limitations of the Random-Access Memory (RAM) in the hardware. The software required the dataset to be arranged in a single mosaic, which required high RAM capacity. The plants in the calibration mosaic were selected manually and identified as Palmer amaranth or soybean (Figure 10). This dataset was used to train both machine learning models. The PLS-DA models were calculated with partial evenly spread cross-validation. The model was applied to the testing portion of the dataset, and a classification image was generated.

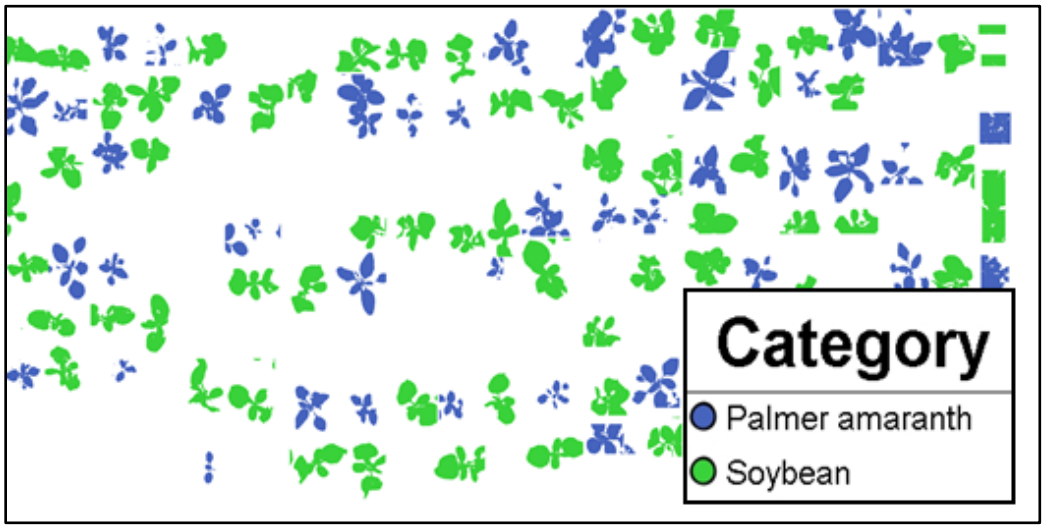


Figure 10. Portion of the calibrated dataset, which consisted of manually classified plants.

A soft independent modelling of class analogy (SIMCA) method was also tested to classify both plant species (Shirzadifar et al., 2018). Under this method, the classification was based upon each classified observation's distance to a PCA model created from a certain class. A

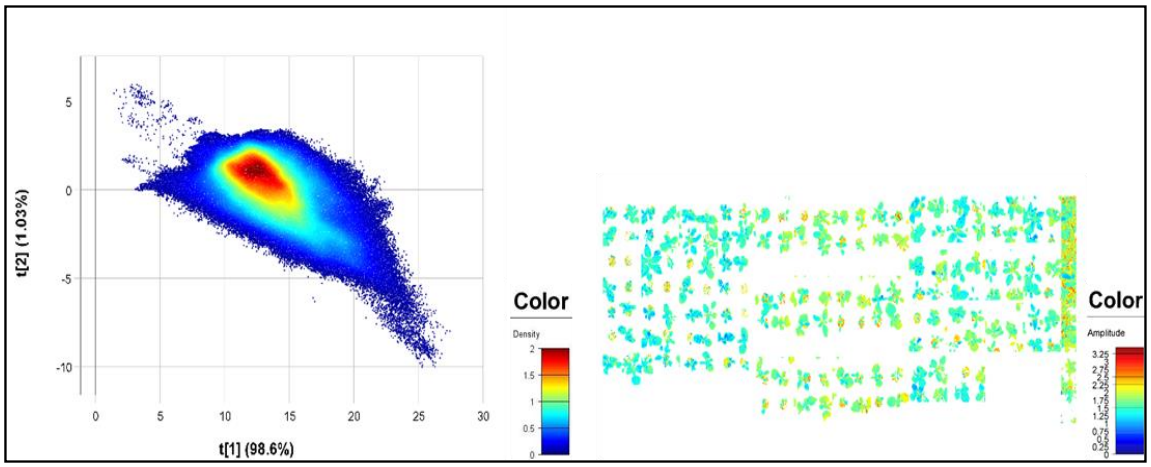
critical distance was calculated, where 95% of the model's observations are considered to belong to the class (Shirzadifar et al., 2018).

The PLS-DA and SIMCA classification models were developed to find the best method to discriminate Palmer amaranth from soybean species.

4. RESULTS AND DISCUSSION

4.1. Principal component analysis for species classification

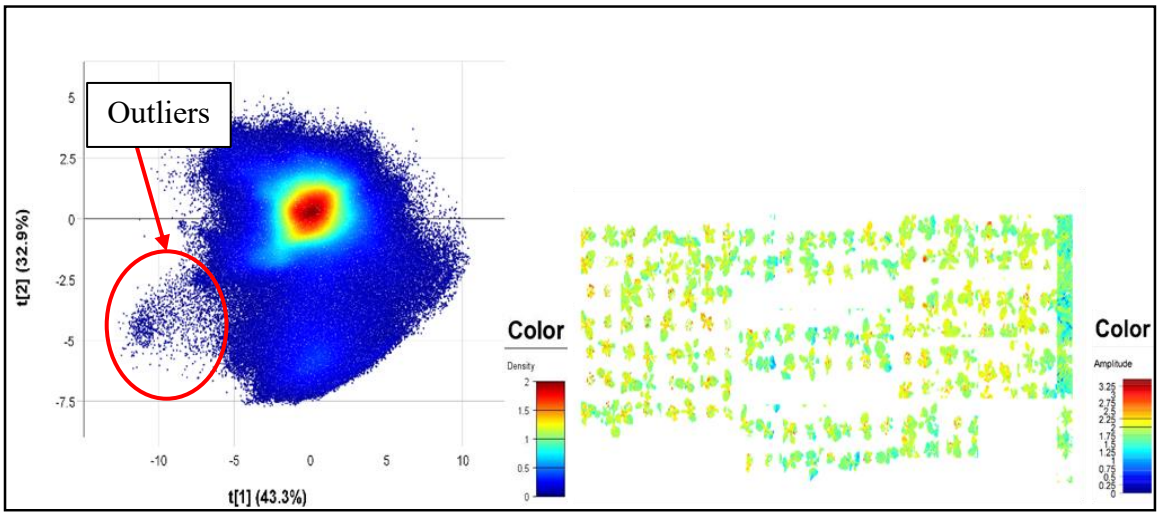
Plants were separated from the background particles (soil particles and pixels containing irrelevant information) using the segmentation procedure explained in previous sections. A PCA was applied to the raw data (without any previous pre-processing) to explore the data and identify potential clusters (Figure 11). The scored plots were shaded according to the density of points (density of pixels per component in the score plot). The score images were shaded according to the first principal component (PC1) score values, where the amplitude represented the variance within the first component in the score image. Some outliers were visible in the score plot, which suggested that the segmentation method used did not remove all the unnecessary data. The results did not show any clusters in the scores plot, and the score images did not demonstrate any difference among species (Sendin et al., 2019). Some variability in the score images was noticeable. However, this was due to differences in the data collection times rather than species, which suggested that further pre-processing techniques were needed to eliminate the external variability from the data.



a) b)

Figure 11. Principal component analysis of the raw data: a) Score plot of the first and second components b) Score images of the first component.

The pre-processing techniques of mean centering, multiplicative scatter correction, and standard normal variate were applied, and a PCA analysis was conducted a second time to explore the new dataset. As shown in the score images in Figure 12 (a), the techniques applied removed some variability previously seen in Figure 11 (a). However, there were still no clusters in Figure 12 (a), which indicated that a supervised classification method would be required to classify plant species. There was no presence of evident individual clusters; however, outliers could be identified in the score plots since they were further from the main cluster, as shown in Figure 12 (a). These outliers were selected and removed from the dataset.

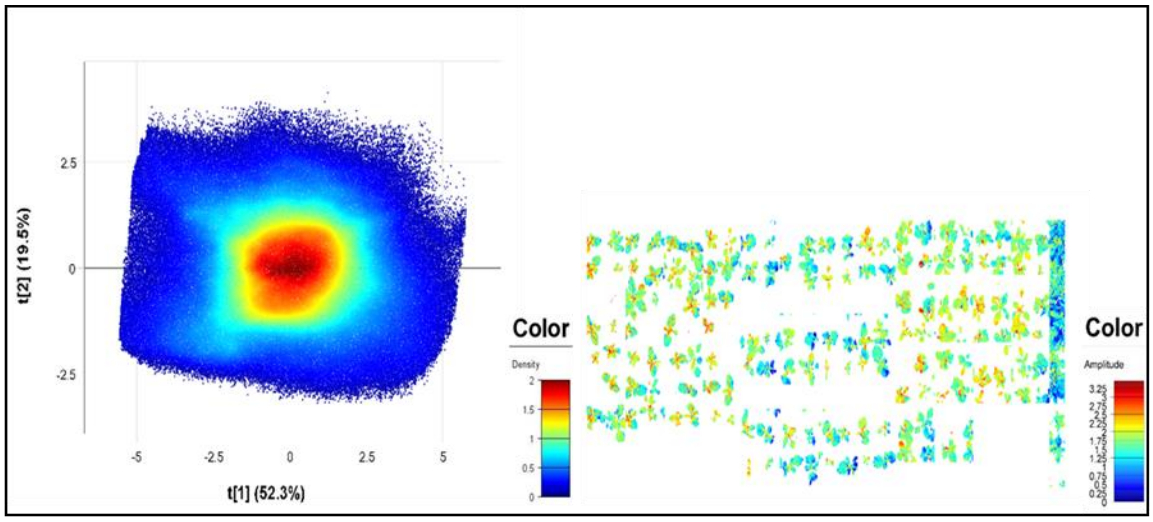


a) b)

Figure 12. Principal component analysis of the pre-processed data: a) Score plot of the first and second components b) Score images of the first component.

The outliers were removed through the selection of points that were farther away from the main density of points in the score plot (Figure 12 (a)). This selection removed small particles of the background that had not been removed by the previous segmentation techniques. In the score plot, shown in Figure 13 (a), the cluster was not as condensed as the cluster in Figure 12 (a). However, the cluster separation was not due to species classification differences; rather, this separation occurred due to the positions of the leaves on the plant, as the light from the

illumination system was reflected differently according to each leaf position on the plant. This can be seen in the score image in Figure 13 (b), in which different colors represent different parts of each plant.



a) b)

Figure 13. Principal component analysis of the pre-processed data with the outliers removed: a) Score plot of the first and second components b) Score images of the first component.

A PCA method was used in two different studies to separate kernels from the background (Manley et al., 2009; Sendin et al., 2019). This method was effective in separating the background from the particles of interest; however, in this method, the complete dataset had to be imported into the software Evince (Prediktera, Sweden), which made processing much slower compared with the method used in this paper. In this research, only the relevant data was imported.

The PCA method was used to remove noise and fine particles in the background that were not removed by the threshold method. Principal component analysis was also used as an exploratory tool; however, it was not possible to visualize any cluster that could represent plant species in any of the components. Therefore, the use of supervised classification techniques using training datasets was necessary for further analysis. The number of components for the

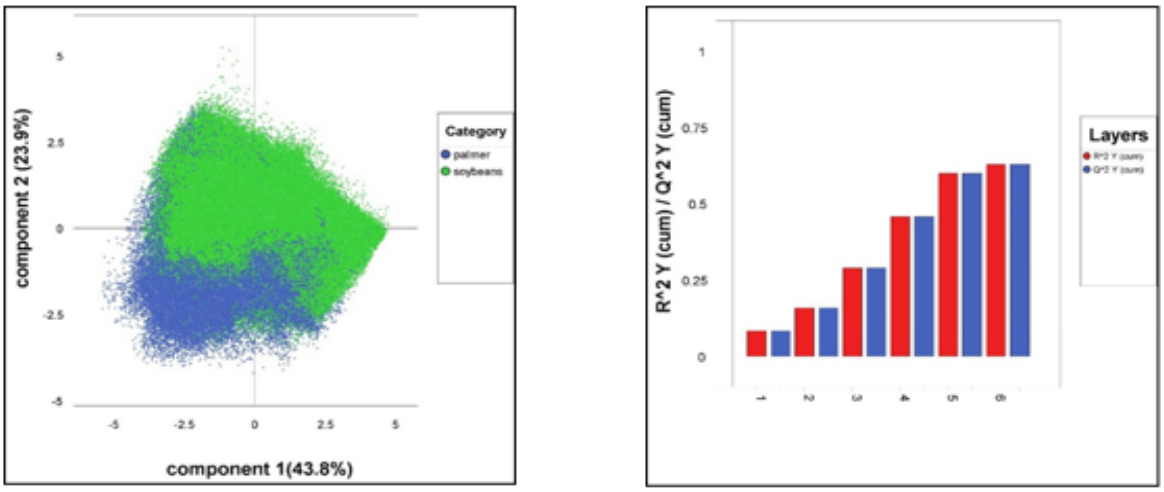
classification models was decided according to the slope of the cumulative variance. When the cumulative variance stopped increasing significantly, the components were not considered important. Six components were used in both models.

4.2. Partial least squares discriminant analysis (PLS-DA) for species classification

This study aimed to identify two different plant species; therefore, it can be considered a discrimination problem. A PLS-DA has been used in similar discrimination problems with high rates of success (Herrmann et al., 2013; McGoverin et al., 2011; Sendin et al., 2019).

A PLS-DA is based on conventional PLS regression; however, different categories (categorical variables) are organized in a response matrix (Y) that indicates the class category (Rodionova et al., 2016); in this case, the plant species. A pixel-wise approach was chosen; each pixel in the image was classified as Palmer amaranth or soybean. In different studies, object-wise classification achieved better results because the objects classified were usually low and flat, with a small variability related to the vertical dimension (Sendin et al., 2018). In this study, object-wise classification did not work as expected due to the variability associated with the vertical dimension, which made the plant segmentation more difficult and less accurate.

The first two PLS factors were able to differentiate Palmer amaranth and soybean species. Blue and green pixels are located in different areas of the score plot. There was also some overlap in the score plot, which was predictable since redundancy of spectral information is expected in a pixel-wise classification (Figure 14 (a)). The PLS-DA model was able to classify both species with a predictive ability ($Q^2Y_{(cum)}$; cross-validated explained variance, calculated from the ratio of the residual cross-validated variance and the original variance of the dataset) of 60%, using six factors. The cumulative variation ($R^2Y_{(cum)}$; total variance explained by the model) for the same number of PLS factors was also 60% (Figure 14 (b)).



a) b)

Figure 14. Palmer amaranth vs soybean classification: a) Score plot (first and second factors) b) Cumulative variation and predictive ability for the first six factors.

Loading line plots were used to identify the most significant wavelengths for each model (Figure 15). Peaks in the plot represent a higher level of significance for the model.

The wavelength region from 530 to 680 nm was relevant for the first component of the model, showing positive peaks that were positively correlated to the component and negative peaks, which were inversely correlated to the component (Figure 15). The region above 800 nm did not contribute to the model discrimination, as indicated by the values near zero. There were also some positive peaks in the second PLS factor; however, as the factor position increased, the noise increased, and the peaks became smaller. In general, the wavelength region between 500 and 700 nm contributed to species classification, as they contained the major peaks in the first two PLS factors.

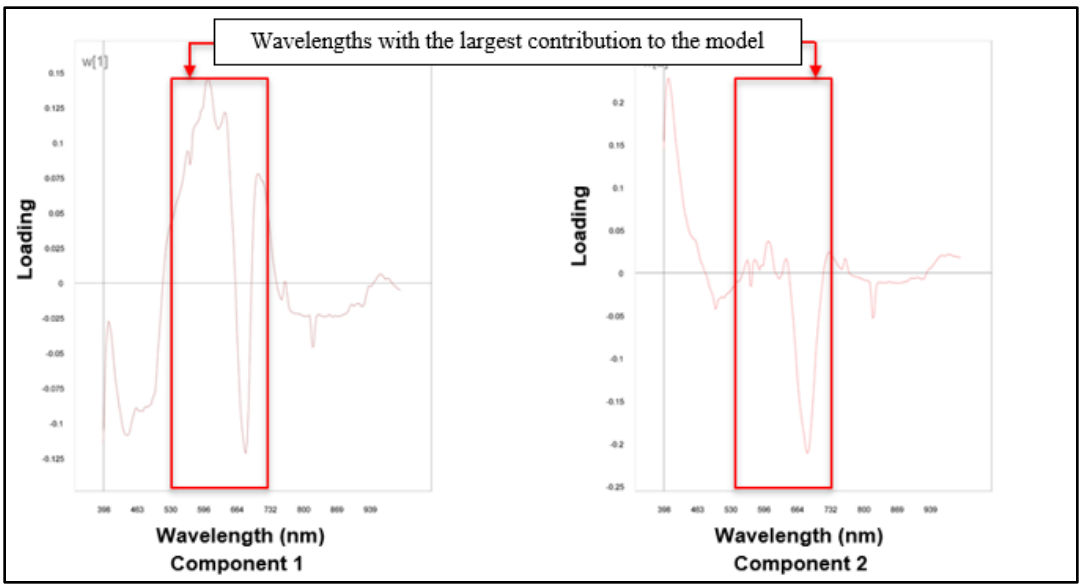


Figure 15. Loading line plots for the first and second components of the Partial least squares-discriminant analysis model.

The model was tested on an independent dataset. Plants were identified as PA or SOY (Figure 16). Some pixels within each plant were not identified correctly because of variability within the same plant due to leaf angle and necrotic lesions on some leaves. Any particles that were not classified as either of the species were considered “No class”.

Some pixels in PA plants were classified as SOY and vice-versa in the pixel-wise classification. Several pixels in the classification images were shaded parts of the plants or corresponded to leaf edges. Since both species had shaded leaves, and those pixels have similar reflectance properties, there were several pixels with similar spectral information in both classes, which reduced the classification capacity of the machine learning models. Consequently, it was technically impossible to obtain a classification in which 100% of the pixels were classified correctly. However, in most cases, there was a big difference in the total amount of pixel classified as PA or SOY in the same plant. The pixel-wise predictive ability of the model was 60%, while Figure 16 showed that most of the plants were classified correctly, considering the average number of pixels classified as PA and SOY in each plant.

Based on pixel-wise classification alone, the $Q^2Y_{(cum)}$ of 60% would not be high enough for some applications where species identification is necessary. Therefore, a different model was also tested in order to find the one that would best fit our application.

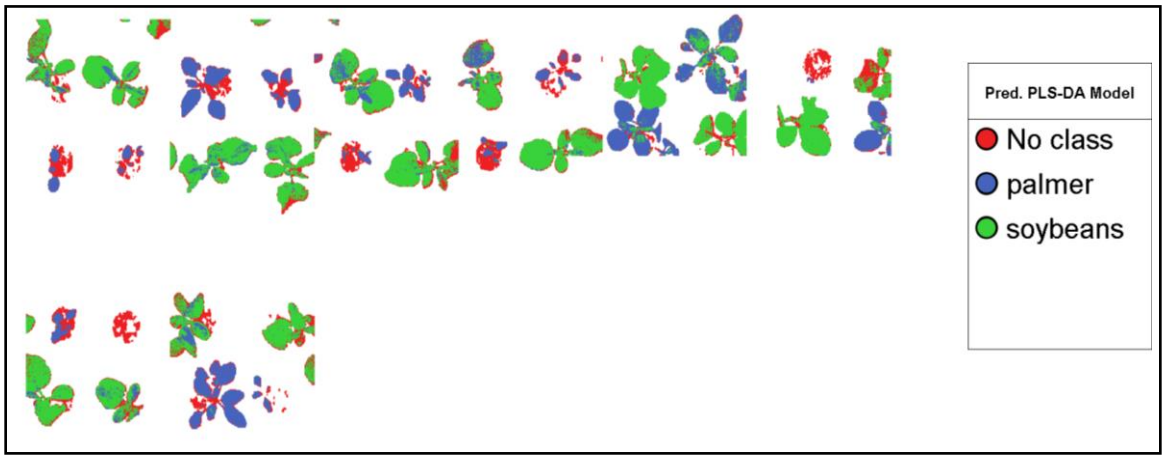


Figure 16. Classification image of the Partial least squares – discriminant analysis model, applied to the testing dataset.

4.3. Soft independent modelling of class analogy (SIMCA) for species classification

A SIMCA is a supervised classification technique in which a PCA is developed for each class individually, within its calculated boundary. For a given class, a hyper-plane is created according to the principal components and the mean orthogonal distance of the training data samples from the hyper-plane is used to determine the critical distance for the classification. New observations are projected into each PC model, and the residual distances are calculated. The observation is assigned to the model if its residual distance is below the class' statistical limit. A SIMCA has commonly been used for chemometrics analysis (Rodionova et al., 2016; Shirzadifar et al., 2018). Models from SIMCA were also used to classify three weed species using canopy reflectance with an accuracy of 100% (Shirzadifar et al., 2018).

A SIMCA model was applied using the same training dataset as an alternative to distinguish PA from SOY species. The Coomman's plot showed the discrimination between

species achieved by the SIMCA model. There was some classification overlap on the plot; however, species were generally organized in clusters (Figure 17). A big portion of the pixels was below the critical value (represented by the red lines), which meant that most of the pixels were attributed to a specific class (Palmer amaranth or soybean).

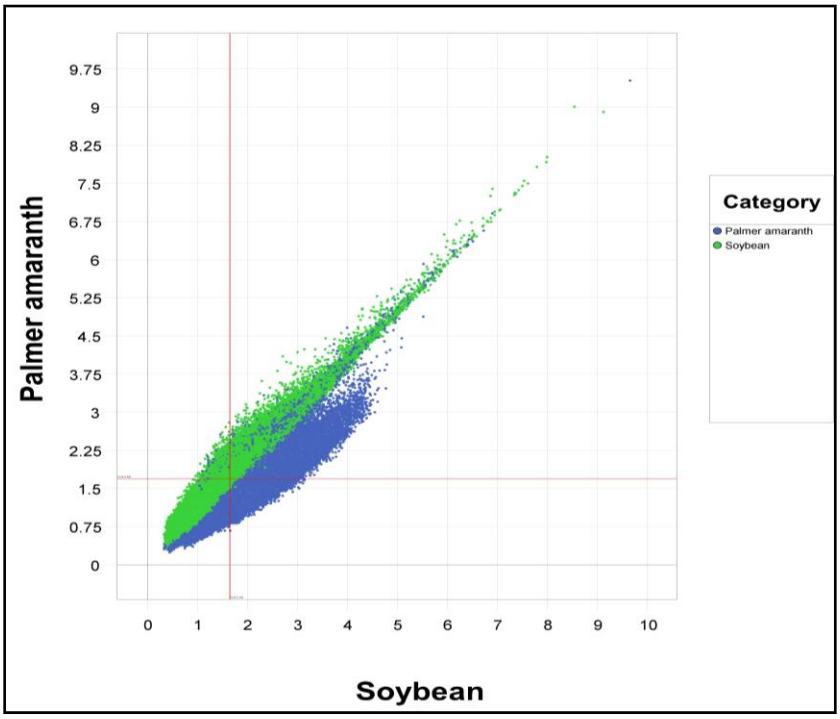


Figure 17. Coomman's plot of the soft independent modelling of class analogy classification model applied to the testing dataset. Most of the pixels were attributed successfully to one of the classes.

The $Q^2Y_{(cum)}$ of the SIMCA model using six components was 82%. The $R^2Y_{(cum)}$ for the same number of components was 85% (Figure 18). The SIMCA model performed better compared to the PLS-DA model. These results are in agreement with previous research using the SIMCA model for plant species identification (Shirzadifar et al., 2018).

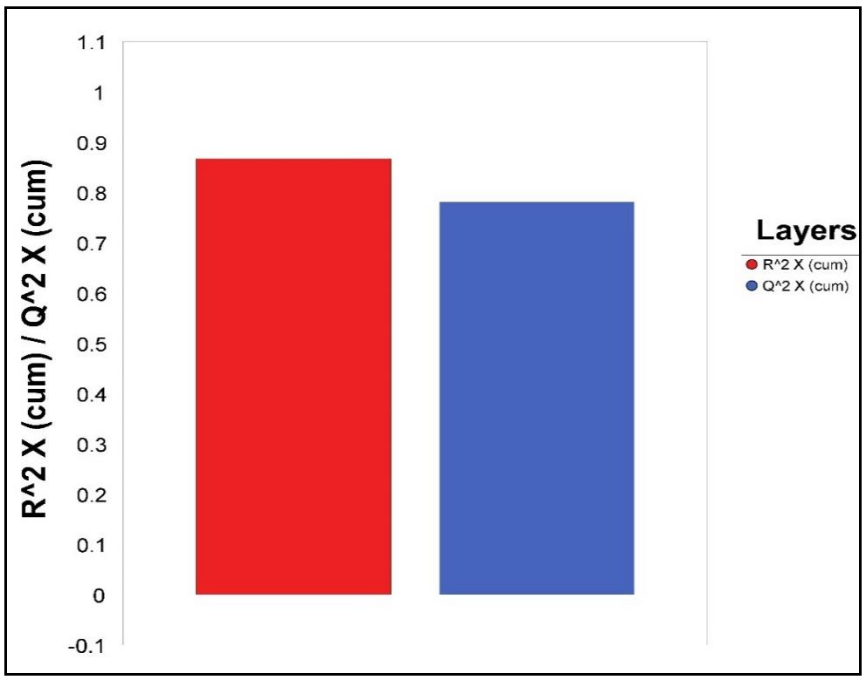


Figure 18. Cumulative variation ($R^2Y_{(cum)}$) and predictive ability ($Q^2Y_{(cum)}$), SIMCA model.

The SIMCA model was applied to an independent dataset. There were some pixels that were considered "No class" because they belonged to the background and were not included in the training dataset. The model was able to classify the dataset with an accuracy of 0.90 on a pixel basis (Table 1). These results were in line with the results obtained from the model cross-validation in Figure 16 and confirmed SIMCA models' ability for PA and SOY classification.

Table 1. Error matrix of the SIMCA model applied to the independent dataset.

		Classification (pixels)			
		Soybean	Palmer amaranth	No class	Total
Ground Truth	Soybean	189581	19509	22394	231484
	Palmer amaranth	11761	75428	9862	97051
	No class	x	y	32256	328535

Several studies have been conducted in weed identification using hyperspectral data, yet, none have classified Palmer amaranth and soybean species. Okamoto et al. (2007) developed a portable hyperspectral imaging system, which captured spectral images in the range of 400 to

1000 nm with 10 nm resolution. Their research focused on identifying wild buckwheat, field horsetail, green foxtail, and common chickweed in sugar beet crop and achieved a classification success rate of 75% to 80% using wavelength transformation and stepwise variable selection method. The processing methods used by Okamoto et al. (2007) were different from the methods used in this study; furthermore, the species classified were different.

Suzuki et al. (2008) distinguished goosefoot pigweed, small crabgrass, field horsetail, and pearlwort from soybean by employing a hyperspectral imaging system. Reddy et al. (2014) characterized Palmer amaranth glyphosate-resistant and susceptible populations with 95% accuracy using a laboratory hyperspectral imaging system. Herbicide resistance is highly related to the chemical properties of the plants; therefore, it can be classified using the plant's reflectance characteristics. In opposition, differences between plant species can be more difficult to identify.

Gao et al. (2018) separated Canada thistle, rumex, and field bindweed from maize crop using a snapshot mosaic hyperspectral imaging sensor, which provided 25 bands from 600 to 871 nm. Zhang et al. (2019) developed a line-scanning visual and near-infrared hyperspectral imaging system that acquired images of samples in a laboratory environment. The authors used this system to differentiate barnyard grass and weedy rice from the rice crop and indicated that six wavelengths (415, 561, 687, 705, 735, and 1007 nm) showed the potential to discriminate the two weeds and rice. Both studies conducted by Gao et al. (2018) and Zhang et al. (2019) used hyperspectral data in weed identification; however, none of them had the same objective of classifying Palmer amaranth and soybean using hyperspectral imaging.

The use of hyperspectral imaging in plant species identification has been the focus of several researchers. However, no study has been conducted on the identification of Palmer

amaranth and soybean in the same manner as described in this paper. In this study, a customized scanning platform was also developed to collect hyperspectral imaging data of Palmer amaranth and soybean in greenhouse and field conditions. In addition to this, PLS-DA and SIMCA models have not been extensively used in species classification using hyperspectral imaging. The need for soybean and Palmer amaranth identification and the necessity to explore machine learning methods applied to plant species identification made this study necessary.

5. CONCLUSIONS

A platform capable of collecting hyperspectral images was developed to facilitate the data collection in greenhouse and field conditions. Palmer amaranth and soybean hyperspectral imaging data were collected. Machine learning models were created to discriminate both species. The background and noise were removed from the dataset using threshold techniques and PCA methods. Models developed using PLS-DA and SIMCA machine learning methods were used to classify plants as either Palmer amaranth or soybean. The PLS-DA model achieved a cumulative variation ($R^2Y_{(cum)}$) of 60% and predictive ability ($Q^2Y_{(cum)}$) of 60%, while the SIMCA model achieved $R^2Y_{(cum)}$ of 85% and $Q^2Y_{(cum)}$ of 82%. Hyperspectral imaging showed the potential for species identification in greenhouse and field conditions. These findings have the potential to be instrumental in the development of new weed control techniques. These species identification models, associated with a vision system and a mechanical or chemical weeding tool, allow site-specific weed management techniques, which would increase weed control efficiency.

REFERENCES

USDA National Agricultural Statistics Service. (2019). Retrieved August 10, 2020, from https://data.nal.usda.gov/dataset/nass-quick-stats.

Amigo, J. M., Babamoradi, H., Elcoroaristizabal, S. (2015). Hyperspectral image analysis. A tutorial. *Analytica Chimica Acta*, *896*, 34–51. https://doi.org/10.1016/j.aca.2015.09.030

Andersen, R. N., Menges, R. M., Conn, J. S. (1985). Variability in Velvetleaf (Abutilon theophrasti) and Reproduction Beyond Its Current Range in North America . *Weed Science*, *33*(4), 507–512. https://doi.org/10.1017/s0043174500082746

Beckie, H. J. (2006). Herbicide-Resistant Weeds: Management Tactics and Practices. *Weed Technology*, *20*(3), 793–814.

Bock, C. H., Poole, G. H., Parker, P. E., Gottwald, T. R. (2010). Plant Disease Severity Estimated Visually, by Digital Photography and Image Analysis, and by Hyperspectral Imaging. *Critical Reviews in Plant Sciences*, *29*(2), 59–107. https://doi.org/10.1080/07352681003617285

Burger, J. E. (2006). Hyperspectral NIR image Analysis. PhD diss. Sweden: Swedish University of Agricultural Sciences.

Davis, A. S., Schutte, B. J., Hager, A. G., Young, B. G. (2015). Palmer Amaranth (Amaranthus palmeri) Damage Niche in Illinois Soybean Is Seed Limited . *Weed Science*, *63*(3), 658–668. https://doi.org/10.1614/ws-d-14-00177.1

Downton, W. (1975). The occurrence of C4 photosynthesis among plants. *Photosynthetica*.

Eddy, P. R., Smith, A. M., Hill, B. D., Peddle, D. R., Coburn, C. A., Blackshaw, R. E. (2014). Weed and crop discrimination using hyperspectral image data and reduced bandsets. *Canadian Journal of Remote Sensing*, *39*(6), 481–490. https://doi.org/10.5589/m14-001

Ehleringer, J. (1983). Ecophysiology of Amaranthus palmeri, a sonoran desert summer annual. *Oecologia*, *57*(1), 107–112. https://doi.org/10.1007/BF00379568

Farooq, A., Jia, X., Hu, J., Zhou, J. (2019). Multi-Resolution Weed Classification via Convolutional Neural Network and Superpixel Based Local Binary Pattern Using Remote Sensing Images. *Remote Sensing*.

Fletcher, R. S., Reddy, K. N. (2016). Random forest and leaf multispectral reflectance data to differentiate three soybean varieties from two pigweeds. *Computers and Electronics in Agriculture*, *128*, 199–206. https://doi.org/10.1016/j.compag.2016.09.004

Franssen, A. S., Skinner, D. Z., Al-Khatib, K., Horak, M. J., Kulakow, P. A. (2001). Interspecific hybridization and gene flow of ALS resistance in Amaranthus species. *Weed Science*, *49*(5), 598–606. https://doi.org/10.1614/0043-1745(2001)049[0598:ihagfo]2.0.co;2

Gao, J., Nuyttens, D., Lootens, P., He, Y., Pieters, J. G. (2018). Recognising weeds in a maize crop using a random forest machine-learning algorithm and near-infrared snapshot mosaic hyperspectral imagery. *Biosystems Engineering*, *170*, 39–50. https://doi.org/https://doi.org/10.1016/j.biosystemseng.2018.03.006

Gray, C. J., Shaw, D. R., Bond, J. A., Iv, D. O. S., Oliver, L. R. (2007). Assessing the Reflective Characteristics of Palmer Amaranth (Amaranthus palmeri) and Pitted Morningglory (Ipomoea lacunosa) Accessions. *Weed Science*, 293–298. https://doi.org/10.1614/WS-06-041.1

Heap I. (2015). International survey of herbicide-resistant weeds. Retrieved February 6, 2019, from http://www.weedscience.org/in.asp

Herrmann, I., Shapira, U., Kinast, S., Karnieli, A., Bonfil, D. J. (2013). Ground-level hyperspectral imagery for detecting weeds in wheat fields. *Precision Agriculture*, *14*(6), 637–659. https://doi.org/10.1007/s11119-013-9321-x

Horak, M. J., Loughin, T. M. (2000). Growth Analysis of Four Amaranthus Species. *Weed Science*, *48*(3), 347–355. Retrieved from http://www.jstor.org/stable/4046301

Huang, Y., A.Lee, M., Reddy, N. (2016). Ground-based hyperspectral remote sensing for weed management in crop production. *International Journal of Agricultural and Biological Engineering*. https://doi.org/10.3965/j.ijabe.20160902.2137

Ikley, J., Jenks, B. (2019). Identification, Biology and Control of Palmer Amaranth and Waterhemp in North Dakota. Retrieved August 10, 2020, from https://www.ag.ndsu.edu/publications/crops/identification-biology-and-control-of-palmer-amaranth-and-waterhemp-in-north-dakota

Klingaman, T. E., Oliver, L. R. (1994). Palmer Amaranth (*Amaranthus palmeri*) Interference in Soybeans (*Glycine max*). *Weed Science*, *42*(4), 523–527. https://doi.org/10.1017/S0043174500076888

Legleiter, T., Johnson, B. (2013). Palmer Amaranth Biology, Identification, and Management. Retrieved January 2, 2020, from https://www.extension.purdue.edu/extmedia/ws/ws-51-w.pdf

Liphadzi, K. B., Dille, J. A. (2006). Annual weed competitiveness as affected by preemergence herbicide in corn. *Weed Science*, *54*(1), 156–165. https://doi.org/DOI: 10.1614/WS-05-156R1.1

Lowe, A., Harrison, N., French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods*, *13*(1), 80. https://doi.org/10.1186/s13007-017-0233-z

Mallard, D. (2009). Management of glyphosate resistent palmer amaranth in LibertyLink ® soybeans.MS Thesis. United States: University of Tennessee.

Manley, M., Williams, P., Nilsson, D., Geladi, P. (2009). Near infrared hyperspectral imaging for the evaluation of endosperm texture in whole yellow maize (Zea maize L.) Kernels. *Journal of Agricultural and Food Chemistry*, *57*(19), 8761–8769. https://doi.org/10.1021/jf9018323

Matzrafi, M., Herrmann, I., Nansen, C., Kliper, T., Zait, Y., Ignat, T., … Eizenberg, H. (2017). Hyperspectral Technologies for Assessing Seed Germination and Trifloxysulfuron-methyl Response in Amaranthus palmeri (Palmer Amaranth). *Front Plant Sci*, *8*, 474. https://doi.org/10.3389/fpls.2017.00474

McGoverin, C. M., Engelbrecht, P., Geladi, P., Manley, M. (2011). Characterisation of non-viable whole barley, wheat and sorghum grains using near-infrared hyperspectral data and chemometrics. *Analytical and Bioanalytical Chemistry*, *401*(7), 2283–2289. https://doi.org/10.1007/s00216-011-5291-x

Mink, R., Linn, A. I., Santel, H.-J., Gerhards, R. (2020). Sensor-based evaluation of maize (Zea mays) and weed response to post-emergence herbicide applications of Isoxaflutole and Cyprosulfamide applied as crop seed treatment or herbicide mixing partner. *Pest Management Science*, *76*(5), 1856–1865. https://doi.org/10.1002/ps.5715

Mishra, P., Asaari, M. S. M., Herrero-Langreo, A., Lohumi, S., Diezma, B., Scheunders, P. (2017). Close range hyperspectral imaging of plants: A review. *Biosystems Engineering*, *164*, 49–67. https://doi.org/https://doi.org/10.1016/j.biosystemseng.2017.09.009

Morgan, G. D., Baumann, P. A., Chandler, J. M. (2001). Competitive Impact of Palmer Amaranth (Amaranthus palmeri) on Cotton (Gossypium hirsutum) Development and Yield. *Weed Technology*, *15*(3), 408–412. https://doi.org/10.1614/0890-037x(2001)015[0408:ciopaa]2.0.co;2

Mulroy, T. W., Rundel, P. W. (1977). Annual Plants: Adaptations to Desert Environments. *BioScience*, *27*(2), 109–114. https://doi.org/10.2307/1297607

Norsworthy, J. K., Griffith, G., Griffin, T., Bagavathiannan, M., Gbur, E. E. (2014). In-Field Movement of Glyphosate-Resistant Palmer Amaranth (*Amaranthus palmeri*) and Its Impact on Cotton Lint Yield: Evidence Supporting a Zero-Threshold Strategy. *Weed Science*, *62*(2), 237–249. https://doi.org/10.1614/WS-D-13-00145.1

Okamoto, H., Murata, T., Kataoka, T., Hata, S. I. (2007). Plant classification for weed detection using hyperspectral imaging with wavelet analysis. *Weed Biology and Management*, *7*(1), 31–37. https://doi.org/10.1111/j.1445-6664.2006.00234.x

Otsu, N. (1979). Threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern*, *SMC-9*(1), 62–66. https://doi.org/10.1109/tsmc.1979.4310076

Pantazi, X.-E., Moshou, D., Bravo, C. (2016). Active learning system for weed species recognition based on hyperspectral sensing. *Biosystems Engineering*, *146*, 193–202. https://doi.org/10.1016/J.BIOSYSTEMSENG.2016.01.014

Peters, T., Jenks, B. (2018). Palmer amaranth confirmed in N.D. Retrieved October 10, 2020, from https://www.ag.ndsu.edu/news/newsreleases/2018/aug-27-2018/palmer-amaranth-confirmed-in-n-d

Qureshi, R., Uzair, M., Khurshid, K., Yan, H. (2019). Hyperspectral document image processing: Applications, challenges and future prospects. *Pattern Recognition*, *90*, 12–22. https://doi.org/https://doi.org/10.1016/j.patcog.2019.01.026

Reddy, K. N., Huang, Y., Lee, M. A., Nandula, V. K., Fletcher, R. S., Thomson, S. J., Zhao, F. (2014). Glyphosate-resistant and glyphosate-susceptible Palmer amaranth (Amaranthus palmeri S. Wats.): hyperspectral reflectance properties of plants and potential for classification. *Pest Management Science*, *70*(12), 1910–1917. https://doi.org/doi:10.1002/ps.3755

Rodionova, O. Y., Titova, A. V., Pomerantsev, A. L. (2016, April 1). Discriminant analysis is an inappropriate method of authentication. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. https://doi.org/10.1016/j.trac.2016.01.010

Sauer, J. (1957). Recent Migration and Evolution of the Dioecious Amaranths. *Evolution*, *11*(1), 11–31. https://doi.org/10.1111/j.1558-5646.1957.tb02872.x

Sellers, B. A., Smeda, R. J., Johnson, W. G., Kendig, J. A., Ellersieck, M. R. (2003). Comparative growth of six Amaranthus species in Missouri. *Weed Science*, *51*(3), 329–333. https://doi.org/10.1614/0043-1745(2003)051[0329:cgosas]2.0.co;2

Sendin, K., Manley, M., Baeten, V., Fernández Pierna, J. A., Williams, P. J. (2019). Near Infrared Hyperspectral Imaging for White Maize Classification According to Grading Regulations. *Food Analytical Methods*, *12*(7), 1612–1624. https://doi.org/10.1007/s12161-019-01464-0

Sendin, K., Manley, M., Williams, P. J. (2018). Classification of white maize defects with multispectral imaging. *Food Chemistry, 243*, 311–318. https://doi.org/10.1016/j.foodchem.2017.09.133

Shirzadifar, A., Bajwa, S., Mireei, S. A., Howatt, K., Nowatzki, J. (2018). Weed species discrimination based on SIMCA analysis of plant canopy spectral data. *Biosystems Engineering*. https://doi.org/10.1016/j.biosystemseng.2018.04.019

Suzuki, Y., Okamoto, H., Kataoka, T. (2008). Image Segmentation between Crop and Weed using Hyperspectral Imaging for Weed Detection in Soybean Field. *Environment Control in Biology*, *46*(3), 163–173. https://doi.org/10.2525/ecb.46.163

Torres-Sánchez, J., López-Granados, F., Peña, J. M. (2015). An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. *Computers and Electronics in Agriculture*, *114*, 43–52. https://doi.org/10.1016/j.compag.2015.03.019

Wang, A., Zhang, W., Wei, X. (2019, March 1). A review on weed detection using ground-based machine vision and image processing techniques. *Computers and Electronics in Agriculture*. Elsevier B.V. https://doi.org/10.1016/j.compag.2019.02.005

Ward, S. M., Webster, T. M., Steckel, L. E. (2013). Palmer Amaranth (*Amaranthus palmeri*): A Review . *Weed Technology*, *27*(1), 12–27. https://doi.org/10.1614/wt-d-12-00113.1

Webster, T. M., Coble, H. D. (1997). Changes in the Weed Species Composition of the Southern United States: 1974 to 1995. *Weed Technology*, *11*(2), 308–317. https://doi.org/10.1017/S0890037X00043001

Webster, T. M., Nichols, R. L. (2012). Changes in the Prevalence of Weed Species in the Major Agronomic Crops of the Southern United States: 1994/1995 to 2008/2009. *Weed Science, 60*(02), 145–157. https://doi.org/10.1614/ws-d-11-00092.1

Zhang, Y., Gao, J., Cen, H., Lu, Y., Yu, X., He, Y., Pieters, J. G. (2019). Automated spectral feature extraction from hyperspectral images to differentiate weedy rice and barnyard grass from a rice crop. *Computers and Electronics in Agriculture, 159,* 42–49. https://doi.org/10.1016/j.compag.2019.02.018

APPENDIX. PLANT MOSAICS

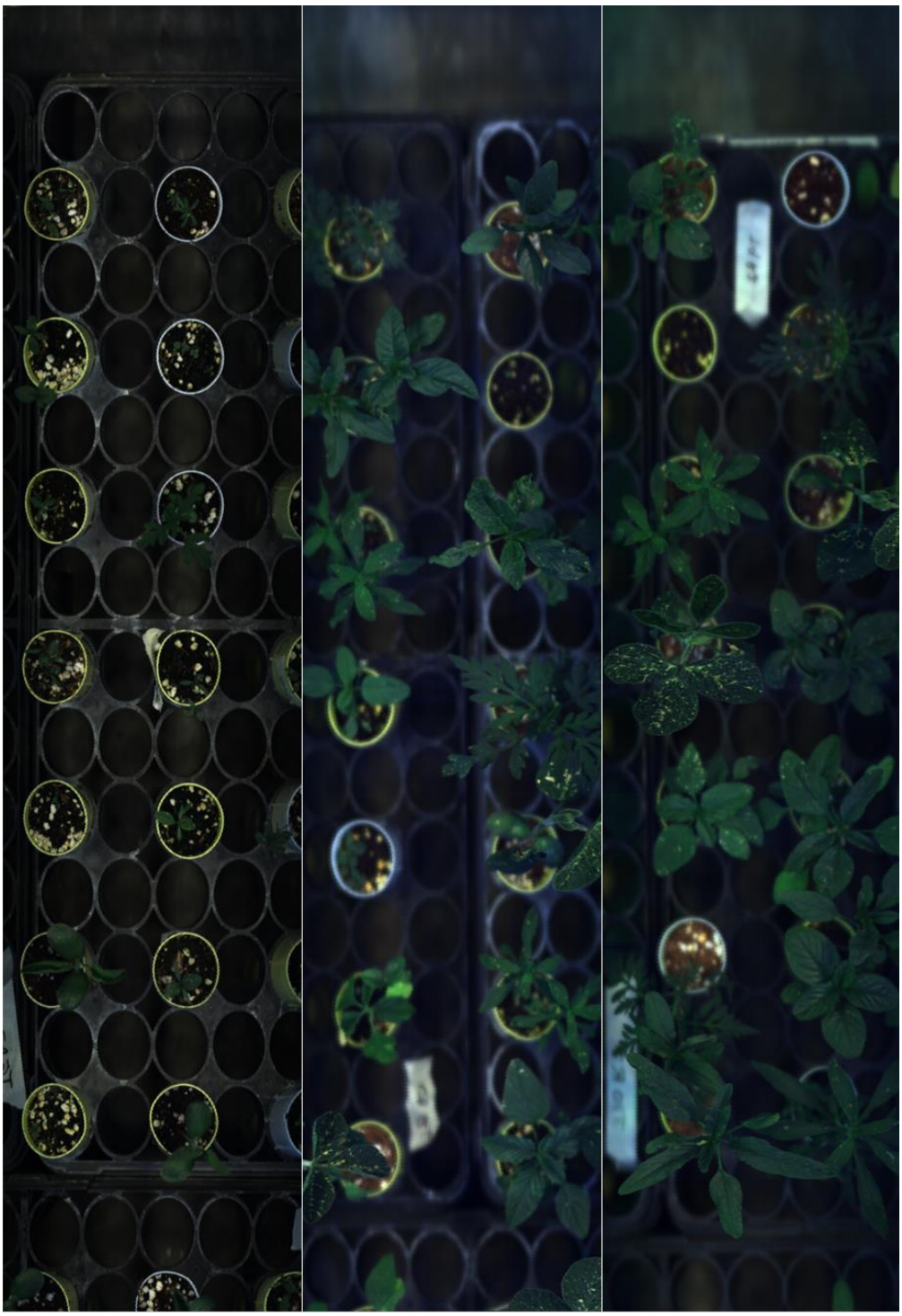


Figure A1. Mosaic of different plant species, at different stages, grown in the greenhouse, 2019.

Figure A2. Mosaic of plant images after the segmentation technique, which removed the background and unnecessary data.

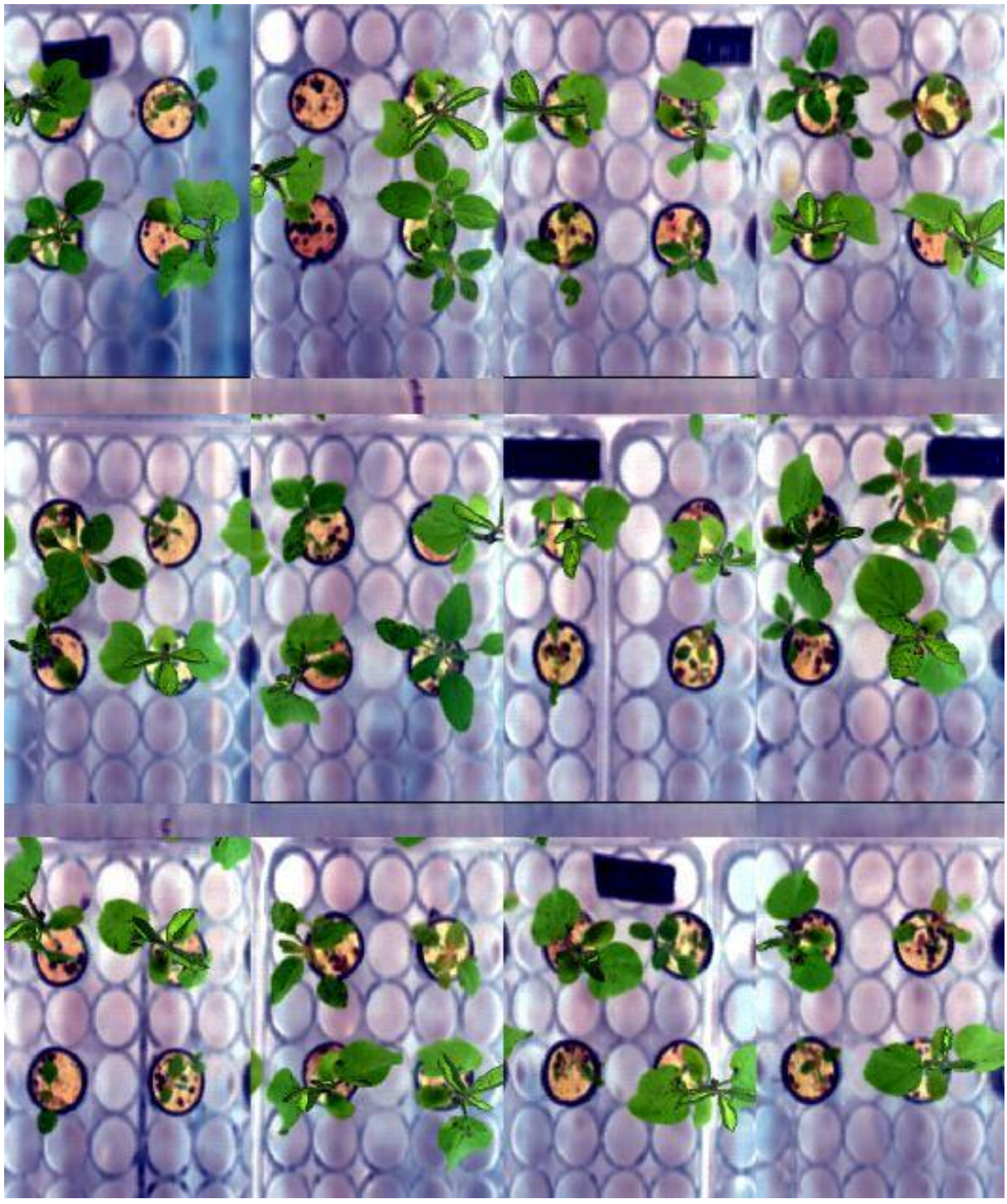


Figure A3. Mosaic of soybeans and Palmer amaranth plants.

Figure A4. Sample of field data collected in Carleton, NE. The wind had an important contribution to the data quality. Leaves were moving during data collection, generating low quality images.