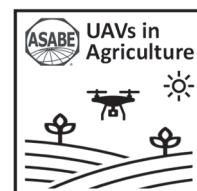# PALMER AMARANTH (*AMARANTHUS PALMERI* S. WATSON) AND SOYBEAN (*GLYCINE MAX* L.) CLASSIFICATION IN GREENHOUSE USING HYPERSPECTRAL IMAGING AND CHEMOMETRICS METHODS

Cristiano Costa[1], Yu Zhang[1], Kirk Howatt[2], Billy Ram[1],
John Stenger[1], John Nowatzki[1], Sreekala Bajwa[3], Xin Sun[1,*]

[1] Department of Agricultural and Biosystems Engineering, North Dakota State University, Fargo, North Dakota, USA.
[2] Department of Plant Sciences, North Dakota State University, Fargo, North Dakota, USA.
[3] College of Agriculture, Montana State University, Bozeman, Montana, USA.
[*] Correspondence: xin.sun@ndsu.edu

HIGHLIGHTS

- Hyperspectral image processing was used to classify Palmer amaranth and soybean species.
- Chemometrics methods (PCA, PLS-DA, and SIMCA) were used to extract features and establish classification models.

ABSTRACT. *Herbicide-resistant weed species are one of the largest threats to modern agriculture, as ineffective weed control results in significant yield losses or increased costs through alternatives such as mechanical methods. Palmer amaranth (*Amaranthus palmeri S. Watson) has been one of the most troublesome weeds. Its identification through the adoption of site-specific weed management systems will help farmers select more appropriate control options to reduce costs and improve efficacy, resulting in increased farm revenue. In this study, a pixel-wised method was evaluated for the classification of Palmer amaranth and soybean (*Glycine max L.). A pushbroom hyperspectral imagery acquisition system was used to collect imagery from 224 spectral bands ranging from 400 to 1000 nm. Greenhouse experiments were conducted in three different runs. Greenhouse-grown plants were evaluated to generate predictive models from paired samples generated with 56 replications in each run. Data collection occurred weekly when Palmer amaranth plants were between approximately 2.5 and 12.7 cm tall. Partial least squares discriminant analysis (PLS-DA) and soft independent modeling of class analogy (SIMCA) were tested to classify Palmer amaranth and soybean. Half of the dataset (Palmer amaranth = 42, soybean = 42) was used to train the models, and the other half (Palmer amaranth = 42, soybean = 42) was used to test model performance. Preliminary results showed that the PLS-DA model with PLS factors as input had a cumulative variation ($R^2Y_{(cum)}$) of 60% and predictive ability ($Q^2Y_{(cum)}$) of 60%. The SIMCA model showed a cumulative variation of 85% and a predictive ability of 82%. Overall, this study illustrated the capability of hyperspectral imagery to classify Palmer amaranth and soybean, which will increase the efficiency of weed control in modern agriculture.*

Keywords. Chemometrics methods, Hyperspectral imaging, Palmer amaranth classification.

almer amaranth (*Amaranthus palmeri* S. Watson) is an invasive weed species and an aggressive competitor with crops throughout the growing season. In the early 20th century, Palmer amaranth was native to the Sonoran Desert region and was documented in the U.S. in 2001 (Ward et al., 2013). In 2018, Palmer amaranth was confirmed in North Dakota (NDSU Extension, 2018). Palmer amaranth biotypes have developed resistance to different modes of chemical action, which makes their control using common herbicides difficult. Individual plants of the species may grow 5 to 8 cm per day, exceeding 2 m in height with many branches. The species is intensely competitive with slow-growing crops for water and fertilizer (Davis et al., 2015; Horak and Loughin, 2000). Palmer amaranth has a strong ability to germinate and emerge throughout the growing season. Its reproductive biology and wind-pollinated habit (Sosnoskie et al., 2012) favor a genetically variable population formation because its pollen can travel long distances. Culpepper (2012) demonstrated that pollen from glyphosate-resistant male plants could be transferred by wind up to 300 m to glyphosate-susceptible females, which resulted in the survival of progeny when sprayed with glyphosate. Most Palmer amaranth found in crop fields shows resistance or tolerance to glyphosate, resulting in

increased chemical usage and increased difficulties in weed control. All these features result in reduced crop yields and biomass content.

Soybean (*Glycine max* L.) is a prominent crop in North Dakota. In 2019, an estimated 5.9 million acres were planted to soybean, which accounted for 7.5% of the soybean acreage in the U.S. (NASS, 2019). Palmer amaranth has been reported to present a large interference to soybean growth and yield. A large reduction in both biomass and seed yield of two soybean cultivars in Arkansas was observed in a study of Palmer amaranth in a soybean field (Monks et al., 1988). A density of 10 plants m^{-1} resulted in a soybean yield reduction of 68% (Klingaman and Oliver, 1994). Studies have demonstrated that Palmer amaranth caused economic yield loss in many other crop species, including cotton (Berger et al., 2015; Norsworthy et al., 2014), corn (Massinga et al., 2001; Massinga and Currie, 2002; Wiggins et al., 2015), and sorghum (Besançon et al., 2017; Hay et al., 2019).

Palmer amaranth has become a major threat to crop growth and productivity. Instead of uniform application of herbicide, site-specific weed management is an effective solution to apply herbicide based on field scouting results. Because resistance management should always include more than one herbicide mode of action, site-specific scouting enables optimum selection of complementary herbicides to maximize weed control and application timing, which can reduce the number of applications to a field, reduce the total amount of herbicide applied in a season, and increase efficacy to limit selection pressure for the development of herbicide resistance as well as provide better control of invading resistant weeds. To achieve this goal, a new, fast, and accurate field scouting method needs to be developed to increase scouting efficiency and provide accurate information to site-specific weed management systems. In contrast to conventional scouting methods, which are based on walking certain parts of the field and making decisions based on the observed portions of the field, which might miss new weed infestations that cannot be easily detected, remote sensing may provide a suitable alternative to address these issues.

Hyperspectral imaging technology has been used in multiple areas in agriculture for its high specificity in comparison with conventional RGB imagery. Different objects on the ground reflect different spectra of sunlight due to their specific components. Hyperspectral imaging systems have the ability to collect wide reflectance profiles from an object and differentiate the object from its surroundings. In the visible region (400-700 nm), leaf pigmentation controls the plant spectra (Huete, 2004). The leaf and canopy structures of different species show strong spectral reflectance in the near-infrared (NIR) and red-edge regions (Gausman, 1985). This is an important indicator for identifying plant species (Shapira et al., 2013). Many studies have reported that hyperspectral imaging is a powerful technology for detecting weeds because different species of plants showed different profiles of spectral reflectance (Farooq et al., 2019; 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). Okamoto et al. (2007) developed a portable hyperspectral imaging system that captured spectral images in the range of 400 to 1000 nm with 10 nm resolution and successfully identified buckwheat (*Fallopia convolvulus*), field horsetail (*Equisetum arvense*), green foxtail (*Setaria viridis* (L.) P. Beauv.), and common chickweed (*Stellaria media* (L.) Vill.) in a sugar beet (*Beta vulgaris*) field. Suzuki et al. (2008) distinguished goosefoot pigweed, small crabgrass, field horsetail (*Equisetum arvense*), and pearlwort from soybean in a field by employing the hyperspectral imaging system developed by Okamoto et al. (2007).

Research on Palmer amaranth identification using hyperspectral imaging systems was investigated in the past decade. Huang et al. (2016) identified and characterized glyphosate-resistant Palmer amaranth and Italian ryegrass (*Lolium perenne* L. sp. *multiflorum* (Lam.) Husnot) using a laboratory hyperspectral imaging system and achieved a high accuracy at 90%. The hyperspectral imaging system in our study has been improved and is available for data collection in the field, compared with the system used by Huang et al. (2016). Gao et al. (2018) recognized Canada thistle (*Cirsium arvense*), *Rumex*, and field bindweed (*Convolvulus arvensis*) in a maize crop using a snapshot mosaic hyperspectral imaging sensor that provided 25 bands from 600 to 871 nm. Zhang et al. (2019) developed a line-scanning visible and NIR hyperspectral imaging system that acquired 675 × 512 (spatial × spectral) images of samples in the laboratory.

Hyperspectral imaging combined with chemometrics methods is a promising tool to acquire more spectral information from plants and presents strong advantages for identification of plant species (Astor et al., 2014; Farooq et al., 2019; Wendel and Underwood, 2016). In this study, hyperspectral imaging technology was applied to classify Palmer amaranth and soybean based on chemometrics methods.

MATERIALS AND METHODS

GREENHOUSE EXPERIMENT

A greenhouse experiment was conducted in the spring of 2019 in Waldron Greenhouse at North Dakota State University (NDSU) to collect data on Palmer amaranth in soybean crops. The weeds and crops were seeded in SC7 Conetainers (21 cm deep with 3.81 cm diameter, Ray Leach) in a potting mix (Sunshine Mix #1, SunGrow Horticulture, Agawam, Mass.). The plants were irrigated once daily. Sodium vapor lamps were used 24 h per day to provide supplementary illumination to accelerate the plant growth. The experiment was organized in three different runs with 56 replications (Palmer amaranth = 28, soybean = 28) per run. Each replication contained paired samples in a block size of two. A randomized experimental design was used to minimize the variations due to the greenhouse microclimate. The three runs were seeded on April 10, April 17, and April 24, respectively. Data collection started when the plants were at the seedling stage, generally 2.5 cm tall. Hyperspectral data were collected twice per week in each run until the plants reached a height of 12.7 cm.

HYPERSPECTRAL IMAGING SYSTEM

A hyperspectral imaging system was developed in this study to conduct data collection in different conditions. The

hyperspectral imaging system was composed of an imaging unit and a scanning unit. The imaging unit included a pushbroom hyperspectral camera (model FX-10, Specim, Oulu, Finland) working in the visible and NIR (VNIR) region. The spectrum ranged from 400 to 1000 nm with 224 spectral bands. The full width at half maximum (FWHM) was 5.5 nm. Line capture occurred at 30 frames per second. The spatial sensor resolution was 1024 pixels with a field-of-view of 38°. The spectral resolution of this camera was 5.5 nm. This device was equipped with an integrated shutter, which allowed collection of the dark reference without the need to block the light manually. White reference collection was performed using a Teflon board with reflection of 99% of the light. Six halogen lamps (MR16-12V-50W) were positioned at the sides of the camera in the imaging unit.

A linear rail with a screw drive having a diameter of 1.2 cm and length of 150 cm powered by a hybrid stepper motor (MDI1FRD17A4-EQ, Intelligent Motion Systems, Marlborough, Conn.) was used for movement of the scanning unit. The longest scanning distance was designed as 150 cm, which encompassed different crop row configurations, e.g., 15 and 30 cm crop row spacings. LUMO Recorder software (Spectral Vision, Middleton, Wisc.) was used for data collection. In the greenhouse experiments, the hyperspectral imaging system was mounted on a 3D scanning platform, which allowed movement of the hyperspectral imaging system in three directions based on the position and size of the target plants and scanning area (fig. 1). The bottom frame of the platform was made of steel to reduce vibration when the platform was in operation. Wheels were incorporated into the design to make it easier to move the platform over the bench in the greenhouse.

PREPROCESSING OF HYPERSPECTRAL DATA

The method used to process the hyperspectral images is shown in figure 2. The output from the image acquisition software was a raw format file. These data needed to be processed to eliminate irrelevant information due to external factors such as the influence of nuisance signals due to illumination effects and detector sensitivity. These misleading characteristics can be presented as a function of the wavelength but also as variations in the spatial domain. To correct these differences in the spectral and spatial domains, a spectral calibration and correction from reflectance to pseudo-absorbance was conducted in document software (Evince Version 2020, Prediktera, Umeå, Sweden) according to equation 1 (Mishra et al., 2017):

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

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 processing. Pixels associated with the background were removed from the mosaic image to reduce noise. The background pixels in the mosaic were eliminated by calculating the normalized difference vegetation index (NDVI) to enhance the features of plant vegetation attributes, which made the hyperspectral image background separation for Palmer amaranth more accurate in this study. Binary automatic threshold was conducted using Otsu's method (Otsu, 1979), a commonly used
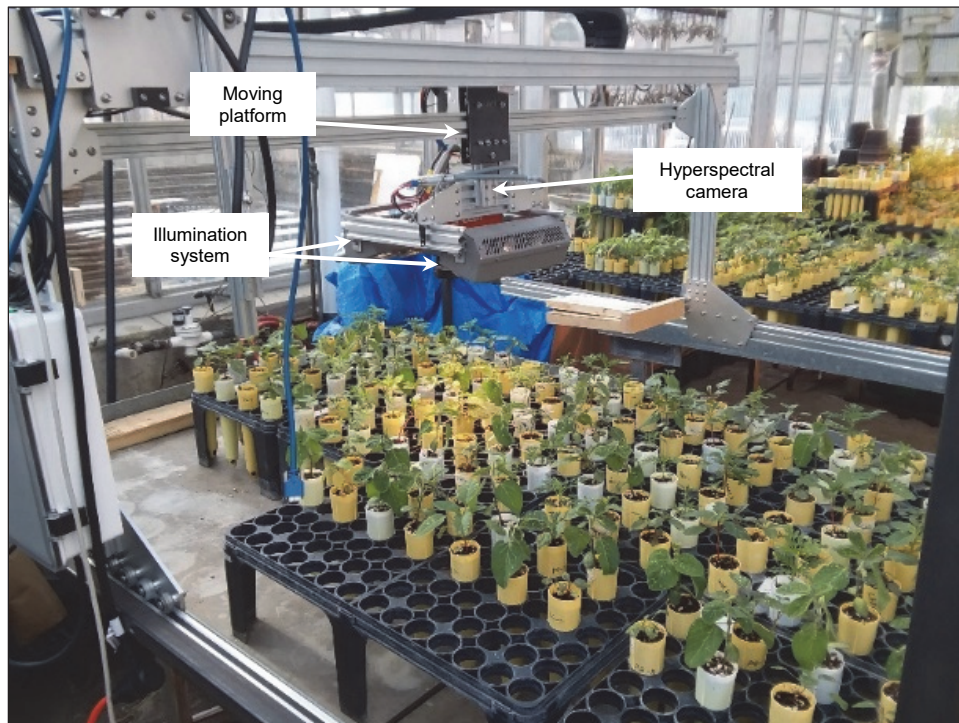

Figure 1. Hyperspectral image data collection for Palmer amaranth and soybean in the greenhouse.
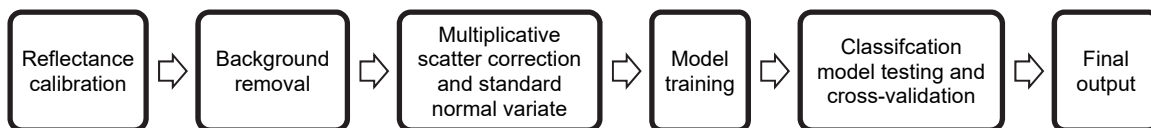
Figure 2. Hyperspectral image data processing flowchart for Palmer amaranth and soybean classification.
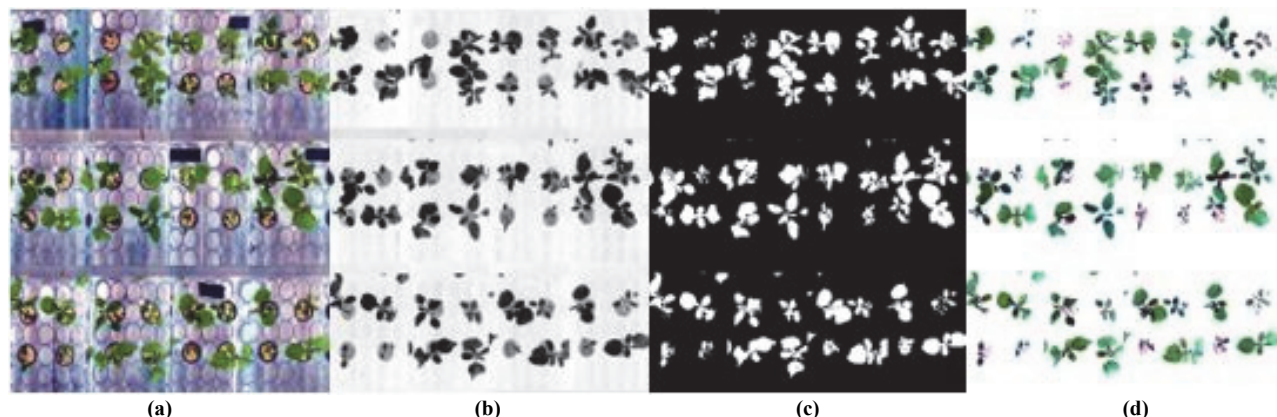

(a)   (b)   (c)   (d)

Figure 3. Preprocessing sequence for Palmer amaranth and soybean hyperspectral images: (a) original mosaic, (b) mosaic after NDVI, (c) mosaic after Otsu, and (d) final preprocessed mosaic.

method in agronomic applications to separate two classes of pixels. 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 in the dataset as a mask, in which each pixel of the mask was multiplied by each spectral layer of the data cube, as shown in figure 3.

This technique assigned values of zero to pixels that were not part of the vegetation. The zero values were ignored by the software when the final dataset was exported. Segmentation was carried out using ENVI software (ver. 5.6, Harris Geospatial Solutions, Boulder, Colo.). The data were mean-centered by subtracting the mean value of a variable from each individual value (Burger, 2006). Standard normal variate (SNV) transformation, which corrects the dataset to baseline shifts and slopes, was also applied to the data (Mishra et al., 2017).

CLASSIFICATION

To distinguish vegetation species, it was necessary to associate spectral characteristics with each species. These spectral characteristics are usually related to chemical and physical properties. Due to the complexity of the hyperspectral data, 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 or pixel but instead with multiple pixels and wavelengths. This complexity of hyperspectral data requires the use of multivariate data analysis and reduction techniques due to the redundancy of the data.

Principal component analysis (PCA) was used to reduce the dimensions of the hyperspectral data cube to a two-dimensional matrix with columns corresponding to wavebands and rows corresponding to pixel values. The PCA method was also used to remove noise and fine particles in the background that were not removed by the threshold method. The

number of components for the classification models was determined according to the slope of the cumulative variance. When the cumulative variance stopped increasing significantly, further components were not considered important.

Partial least squares discriminant analysis (PLS-DA) is a supervised class-modeling technique that uses a PLS algorithm to predict whether a sample belongs to a determined class. For the calibration method, two matrices were created: a matrix (X) containing the spectra for calibration and a corresponding matrix (Y) containing the pixels in the model that were ideally identified as belonging to the discriminated class. In this study, each pixel of an image was defined as belonging to a determined class. To reduce the computational demand, each image was separated into a training dataset and a testing dataset; 50% of the pixels (512 pixels) were used to train the model, and the other 50% of the pixels were used for testing the model (Eddy et al., 2014). To further reduce the computational demand, the PLS-DA models were calculated with partial evenly spread cross-validation in which only a fraction of all possible "leave-one-out" validation tests were evaluated. PLS-DA is based on conventional PLS regression; however, different categories (categorical variables) were organized in the corresponding matrix (Y) that indicated the class category, in this case, the plant species (Rodionova et al., 2016). A pixel-wise approach was chosen so that each pixel in the image was classified as Palmer amaranth or soybean.

Soft independent modeling of class analogy (SIMCA), a statistical method for supervised classification, was also tested to classify both plant species. SIMCA is a powerful classification algorithm because it employs PCA as a dimension reduction method (Borregaard et al., 2000). In this method, classification is based on the distance of each classified observation to a PCA model created from a certain class. The PLS-DA and SIMCA classification models were compared to find the best method to discriminate Palmer amaranth from soybean.

To evaluate the performance of the PLS-DA and SIMCA models, the predictive ability ($Q^2Y_{(cum)}$) and cumulative variation ($R^2Y_{(cum)}$) were used as the main parameters to compare the two models. The larger the value of $Q^2Y_{(cum)}$, the better the predictive ability of the model, and the larger the value of $R^2Y_{(cum)}$, the stronger the explanatory ability.

RESULTS

PRINCIPAL COMPONENT ANALYSIS

Plants were separated from the background using the segmentation procedure explained in previous sections. To explore the data and identify potential clusters, PCA was applied to the raw data without preprocessing (fig. 4). The score plots were shaded according to the density of points, while the score images were shaded according to the scores of the first principal component (PC1). Some outliers were visible in the score plots, which suggested that the segmentation did not remove all the unnecessary data. The results did not show any clusters in the score plots, and the score images did not demonstrate any difference among species (Sendin et al., 2019). Some variability in the score plots was noticeable due to the time differences in the data collections rather than the species, which suggested that further preprocessing techniques were needed to eliminate that external variability from the data.

Mean centering and SNV were applied as preprocessing techniques, and PCA was conducted again to remove the outliers in figure 4a. There were still no clusters after preprocessing (fig. 5a), which indicated that a supervised classification method was needed to classify the plant species. Even without the presence of evident individual clusters, outliers could be identified in the score plots because they were farther from the main cluster. These outliers were selected and removed from the dataset with the image processing software (Evince Version 2020, Prediktera, Umeå, Sweden).

The outliers were removed from the score plot through the selection of points that were farther away from the centroid area (fig. 6a). The cluster was not as condensed when using the preprocessed data in comparison with the raw data. However, that cluster separation was not due to species classification but rather to the positions of the leaves on the plants, as the light from the illumination system was reflected differently according to each leaf position. Evidence of this can be observed in the score image in figure 6b, where different parts of each plant are represented by different colors.

PARTIAL LEAST SQUARES DISCRIMINANT ANALYSIS

As shown in figure 7a, the first two PLS components were able to differentiate Palmer amaranth and soybean species; however, there was some overlap in the score plot, which is
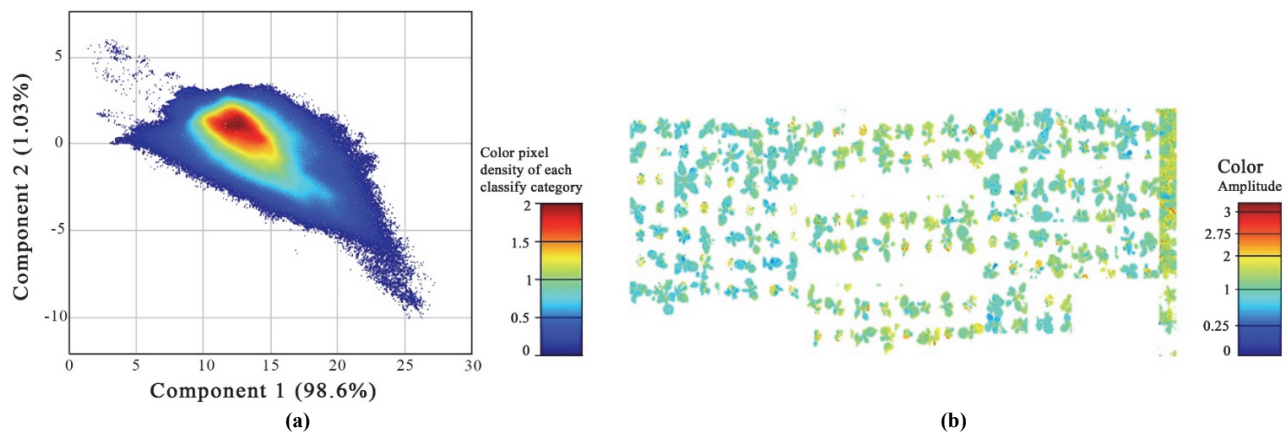

Figure 4. Principal component analysis of raw hyperspectral data for Palmer amaranth and soybean: (a) PCA score plot of first and second components, and (b) Palmer amaranth and soybean image using first component in PCA.
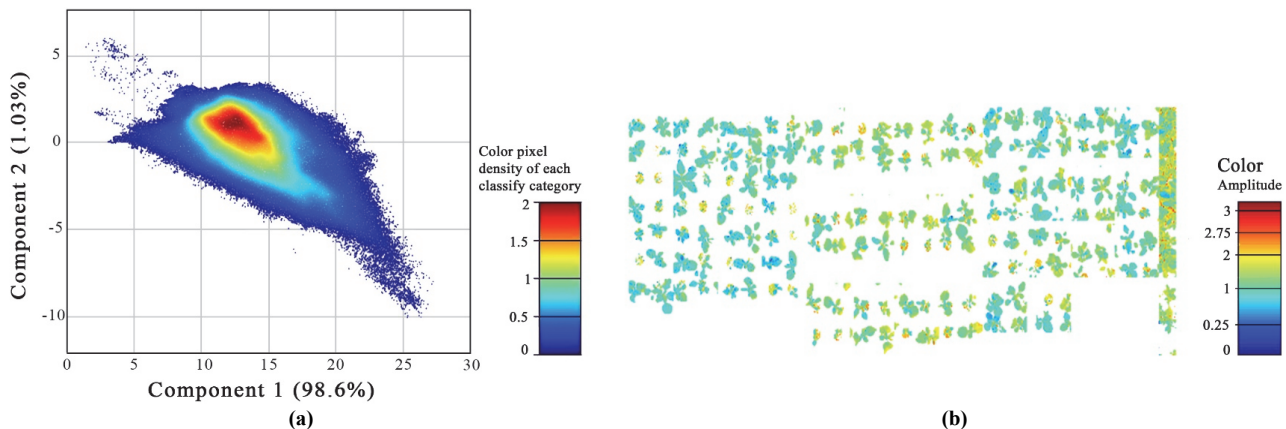

Figure 5. Principal component analysis of preprocessed data for Palmer amaranth and soybean: (a) PCA score plot of first and second components, and (b) Palmer amaranth and soybean image using first component in PCA.
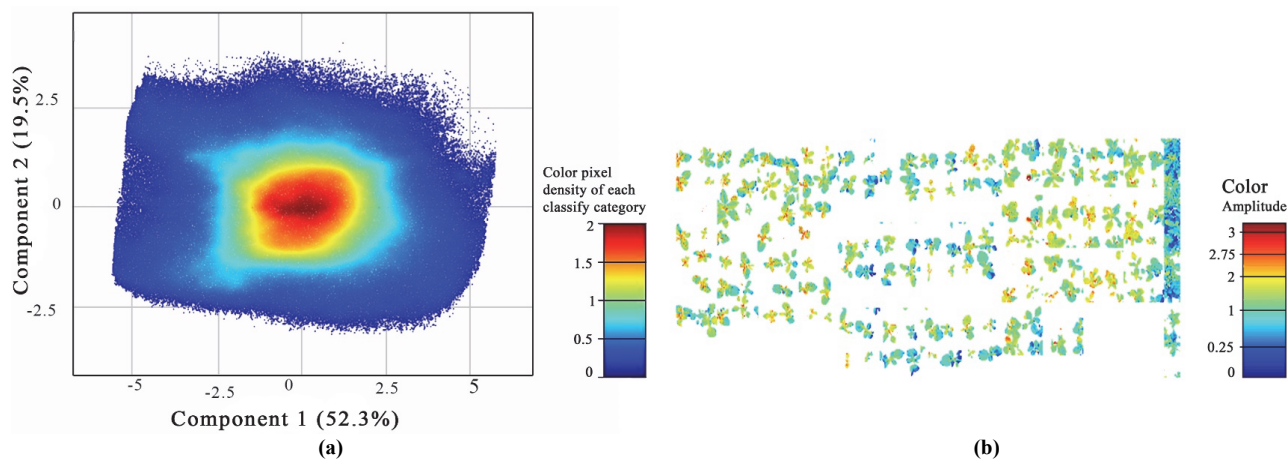
Figure 6. Principal component analysis of preprocessed data for Palmer amaranth and soybean with outliers removed: (a) PCA score plot of first and second components, and (b) Palmer amaranth and soybean image using first component in PCA.
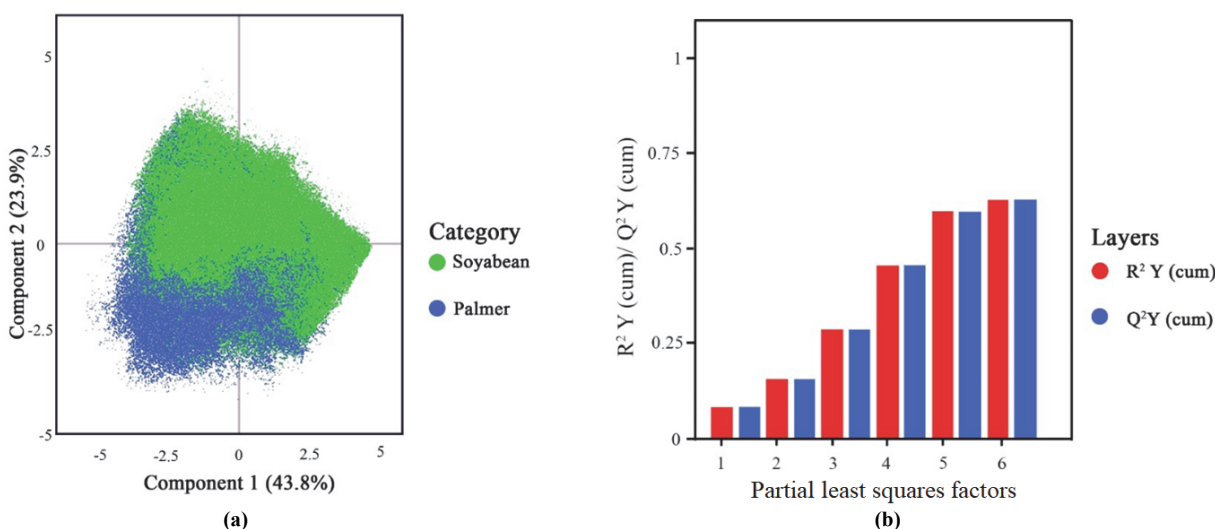

Figure 7. Comparison of PLS components for Palmer amaranth and soybean: (a) score plot (components 1 vs. components 2), and (b) coefficient of determination(R^2Y) and cross-validated coefficient of determination (Q^2Y) by factor number.

predictable considering that redundancy of information is common in pixel-wise classification. Figure 7b shows that the PLS-DA model was able to classify both species with a predictive ability ($Q^2Y_{(cum)}$) of 60% (i.e., the cross-validated explained variance, which was calculated from the ratio of the residual cross-validated variance and the original variance of the dataset) when using six factors. The cumulative variation ($R^2Y_{(cum)}$) was also 60% (i.e., the total variance explained by the model) for the same number of PLS factors.

The wavelength region from 530 to 680 nm was relevant for the first component of the model, showing positive peaks, which were positively correlated to the component, and negative peaks, which were inversely correlated to the component (fig. 8). The region above 800 nm did not contribute as much to the model discrimination, as indicated by values near zero. For the second PLS component, one positive peak occurred at 398 nm. There were also some positive peaks in the region from 530 to 680 nm; however, as the component position increased, the noise also increased, and the peaks became smaller. In general, the visible region between 398 and 700 nm contributed to the classification of Palmer amaranth and soybean, as that region contained the

major peaks for the first two PLS components and represented almost 70% of the variability in the data.

The model was tested with the testing dataset. As shown in figure 9, the plants were identified as Palmer amaranth or soybean. However, some pixels within each plant were not identified correctly because of variability within the plant due to leaf angle and necrotic lesions on some leaves. Any particles that were not classified as either of the two plant species were considered "no class".

Some pixels in Palmer amaranth plants were classified as soybean, 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. Both plant species had this problem, which originated from pixels with similar spectral information in different classes in the classification results. However, there was a large difference in the total number of pixels classified as Palmer amaranth or soybean in the same plant in most cases. The model's predictive ability was 60%, and figure 9 shows that most of the plants were classified correctly, considering the average number of pixels classified as Palmer amaranth and soybean in each plant.
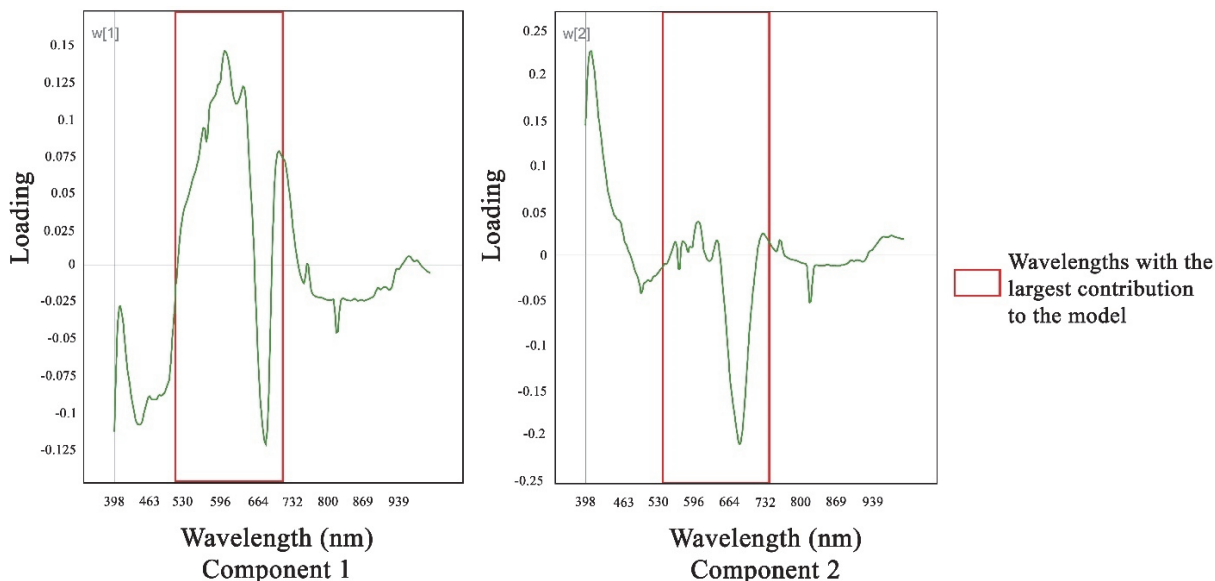
Figure 8. Component correlation loading plot from PLS-DA model for Palmer amaranth and soybean classification.

Figure 9. PLS-DA classification image using testing dataset for Palmer amaranth and soybean.

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

SOFT INDEPENDENT MODELING OF CLASS ANALOGY

The predictive ability ($Q^2X_{(cum)}$) of the SIMCA model using six components was 82% (fig. 10). The cumulative variation ($R^2X_{(cum)}$) for the same number of components was 85%. These results indicated a better performance for the model SIMCA compared with the PLS-DA model.

The SIMCA model was applied to the testing dataset, and the results are shown in table 1. Some pixels were not classified because they belonged to the background and were not included in the training dataset. The model was able to classify the dataset with accuracies of 80.7% on a pixel basis and 86% based on a per plant classification. The per plant classification results were obtained when using the model to classify plants based on the groups of pixels representing each plant. If most of the pixels were classified as Palmer

amaranth, then the plant was classified as Palmer amaranth, while the pixel classification was based on single pixels. The results were in line with the results obtained from the cross-validation in figure 10.
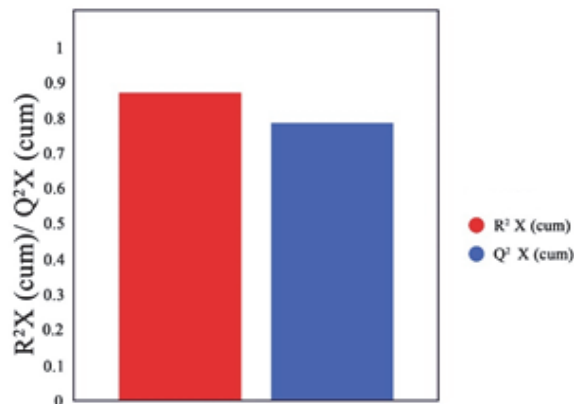

Figure 10. Cumulative variation ($R^2X_{(cum)}$) and predictive ability ($Q^2X_{(cum)}$) of the SIMCA model for classification of Palmer amaranth and soybean data using components extracted by PCA.

Table 1. Error rate of SIMCA model applied to testing dataset for classification of Palmer amaranth (PA) and soybean.

		Classification			Total
		Soybean	PA	No class	(pixel count)
Ground truth	Soybean	81.90%	8.43%	9.67%	231,484
	PA	12.12%	77.72%	10.16%	97,051
					328,535

DISCUSSION

Several studies have used hyperspectral data for weed identification, but none has classified Palmer amaranth and soybean species. Okamoto et al. (2007) developed a portable hyperspectral imaging system that captured spectral images in the range of 400 to 1000 nm with 10 nm resolution. Their study focused on identifying wild buckwheat, field horsetail, green foxtail, and common chickweed in sugar beet and achieved classification success rates of 75% to 80% using wavelength transformation and stepwise variable selection methods. The processing methods used by Okamoto et al. (2007) were different from the methods used in this study; furthermore, the classified species were different.

Gao et al. (2018) separated Canada thistle, *Rumex*, and field bindweed from maize using a snapshot mosaic hyperspectral imaging sensor, which provided 25 bands from 600 to 871 nm. Zhang et al. (2019) developed a line-scanning VNIR hyperspectral imaging system that acquired images of samples in a laboratory environment. They used their system to differentiate barnyard grass and weedy rice from rice and indicated that six wavelengths (415, 561, 687, 705, 735, and 1007 nm) showed potential to discriminate the two weeds and rice. The studies conducted by Gao et al. (2018) and Zhang et al. (2019) used hyperspectral data for weed identification, which provided a proof-of-concept for hyperspectral imaging technology in weed identification research. However, Gao et al. (2018) used only 25 bands in the visible and near-infrared region for data collection. In our research, spectral information was acquired from 224 bands in the 400-1000 nm range, which provided more spectral features for weed identification.

The SIMCA method is commonly used in data mining, especially in data classification (Tsugawa et al., 2011; Smidt et al., 2008; Gemperline et al., 1989; Xu et al., 2013). Shawky et al. (2020) constructed a SIMCA model to authenticate, detect, and quantify plant adulterants used in fraudulent saffron stigmas, including safflower, pomegranate fruit peel, calendula flower, paprika, curcuma, hibiscus, saffron stamens, and exhaustively extracted saffron stigmas. They also use PLS-DA to compare the results. Their study found that SIMCA class modeling was successful for authentication of saffron, while PLS-DA effectively discriminated authentic pure samples from intentionally adulterated samples with adulterants as low as 10 mg g^{-1}, including saffron stamens and exhaustively extracted saffron stigmas. Todorova and Atanassova (2016) used NIR diffuse reflectance spectroscopy with the SIMCA method to distinguish three major soil types. Their results showed that soil samples could be distinguished according to their classes, and the SIMCA models correctly classified 100% of the samples in both the calibration and validation sets.

The research presented in this article also showed the superiority of the SIMCA method compared to other classification methods (e.g., PLS-DA) for distinguishing weeds from crop plants. Although the SIMCA method has great potential for classification of weed species, more experimental validation should be conducted in field conditions to test the robustness of the SIMCA method. In addition, improvement of model accuracy should be investigated by inputting more weed and crop hyperspectral data into the SIMCA model in both greenhouse and field conditions.

CONCLUSION

In this study, hyperspectral imaging technology paired with chemometrics methods was adopted to identify Palmer amaranth from soybean plants. A movable platform capable of collecting hyperspectral images was developed to facilitate data collection in greenhouse conditions. The background and noise were removed using Otsu's method, and PCA, PLS-DA, and SIMCA were used to classify Palmer amaranth plants and soybean plants. The PLS-DA model achieved a cumulative variation (R^{2}Y$_{(cum)}$) of 60% and predictive ability (Q^{2}Y$_{(cum)}$) of 60%, while the SIMCA model achieved R^{2}X$_{(cum)}$ of 85% and Q^{2}X$_{(cum)}$ of 82%. In the PLS-DA model, the visible region from 398 to 700 nm contributed to the classification of Palmer amaranth and soybean plants. Although the chemometrics methods and hyperspectral imaging technology used in this study achieved promising results for the classification of Palmer amaranth and soybean, improvements still need to be considered, such as noise reduction and the robustness of the model. In future studies, field experiments need to be conducted to improve the classification accuracy and robustness of the classification model.

REFERENCES

Astor, T., Dalmayne, J., Prentice, H. C., Eklundh, L., Purschke, O., Schmidtlein, S., & Hall, K. (2014). Classification of grassland successional stages using airborne hyperspectral imagery. *Remote Sensing, 6*(8), 7732-7761. https://doi.org/10.3390/rs6087732

Berger, S. T., Ferrell, J. A., Rowland, D. L., & Webster, T. M. (2015). Palmer amaranth (*Amaranthus palmeri*) competition for water in cotton. *Weed Sci., 63*(4), 928-935. https://doi.org/10.1614/WS-D-15-00062.1

Besançon, T. E., Heiniger, R. W., Weisz, R., & Everman, W. J. (2017). Grain sorghum and Palmer amaranth (*Amaranthus*

palmeri) response to herbicide programs and agronomic practices. *Weed Tech., 31*(6), 781-792. https://doi.org/10.1017/wet.2017.53

Borregaard, T., Nielsen, H., Nørgaard, L., & Have, H. (2000). Crop-weed discrimination by line imaging spectroscopy. *J. Agric. Eng. Res., 75*(4), 389-400. https://doi.org/10.1006/jaer.1999.0519

Burger, J. E. (2006). Hyperspectral NIR image analysis data exploration, correction, and regression. PhD diss. Uppsala, Sweden: Swedish University of Agricultural Sciences. Retrieved from https://pub.epsilon.slu.se/1200/1/Jim_Burger_thesis.pdf

Culpeper, L. (2012). Changes in cotton weed management practices following the development of glyphosate-resistant Palmer amaranth. *Proc. Beltwide Cotton Conf.* Cordova, TN: National Cotton Council.

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 Sci., 63*(3), 658-668. https://doi.org/10.1614/WS-D-14-00177.1

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 J. Remote Sensing, 39*(6), 481-490. https://doi.org/10.5589/m14-001

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, 11*(14), article 1692. https://doi.org/10.3390/rs11141692

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

Gausman, H. W. (1985). Plant leaf optical properties in visible and near-infrared light. In *Graduate studies / Texas Tech University.* Lubbock, TX: Texas Tech Press.

Gemperline, P. J., Webber, L. D., & Cox, F. O. (1989). Raw materials testing using soft independent modeling of class analogy analysis of near-infrared reflectance spectra. *Anal. Chem., 61*(2), 138-144. https://doi.org/10.1021/ac00177a012

Hay, M. M., Albers, J. J., Dille, J. A., & Peterson, D. E. (2019). Control of atrazine-resistant Palmer amaranth (*Amaranthus palmeri*) in double-crop grain sorghum. *Weed Tech., 33*(1), 115-122. https://doi.org/10.1017/wet.2018.102

Horak, M. J., & Loughin, T. M. (2000). Growth analysis of four *Amaranthus* species. *Weed Sci., 48*(3), 347-355. https://doi.org/10.1614/0043-1745(2000)048[0347:GAOFAS]2.0.CO;2

Huang, Y., Lee, M., & Reddy, N. (2016). Ground-based hyperspectral remote sensing for weed management in crop production. *Intl. J. Agric. Biol. Eng., 9*(2), 98-109. https://doi.org/10.3965/j.ijabe.20160902.2137

Huete, A. R. (2004). Chapter 11: Remote sensing for environmental monitoring. In J. F. Artiola, I. L. Pepper, & M. L. Brusseau (Eds.), *Environmental monitoring and characterization* (pp. 183-206). Amsterdam, Netherlands: Elsevier Academic. https://doi.org/10.1016/B978-012064477-3/50013-8

Klingaman, T. E., & Oliver, L. R. (1994). Palmer amaranth (*Amaranthus palmeri*) interference in soybeans (*Glycine max*). *Weed Sci., 42*(4), 523-527. https://doi.org/10.1017/S0043174500076888

Massinga, R. A., & Currie, R. S. (2002). Impact of Palmer amaranth (*Amaranthus palmeri*) on corn (*Zea mays*) grain yield and yield and quality of forage. *Weed Tech., 16*(3), 532-536.

https://doi.org/10.1614/0890-037X(2002)016[0532:IOPAAP]2.0.CO;2

Massinga, R. A., Currie, R. S., Horak, M. J., & Boyer, J. (2001). Interference of Palmer amaranth in corn. *Weed Sci., 49*(2), 202-208. https://doi.org/10.1614/0043-1745(2001)049[0202:IOPAIC]2.0.CO;2

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*, article 474. https://doi.org/10.3389/fpls.2017.00474

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 Mgmt. Sci., 76*(5), 1856-1865. https://doi.org/10.1002/ps.5715

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

Monks, D. W., & Oliver, L. R. (1988). Interactions between soybean (*Glycine max*) cultivars and selected weeds. *Weed Sci., 36*(6), 770-774. https://doi.org/10.1017/S0043174500075809

NASS. (2019). Acreage. Washington, DC: USDA National Agricultural Statistics Service. https://usda.library.cornell.edu/concern/publications/j098zb09z?locale=en

NDSU Extension. (2018). Palmer amaranth confirmed in MacIntosh County, ND. *Farm and Ranch Guide* (31 August 2018). Retrieved from https://www.agupdate.com/farmandranchguide/news/crop/palmer-amaranth-confirmed-in-mcintosh-county-nd/article_3e135b5e-aace-11e8-809a-976b47d30ee7.html

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 Sci., 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 Biol. Mgmt., 7*(1), 31-37. https://doi.org/10.1111/j.1445-6664.2006.00234.x

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Trans. Systems Man Cybernetics, 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. *Biosyst. Eng., 146*, 193-202. https://doi.org/10.1016/j.biosystemseng.2016.01.014

Rodionova, O. Y., Titova, A. V., & Pomerantsev, A. L. (2016). Discriminant analysis is an inappropriate method of authentication. *TrAC Trends Anal. Chem., 78*, 17-22. https://doi.org/10.1016/j.trac.2016.01.010

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

Shapira, U., Herrmann, I., Karnieli, A., & Bonfil, D. J. (2013). Field spectroscopy for weed detection in wheat and chickpea fields. *Intl. J. Remote Sens., 34*(17), 6094-6108. https://doi.org/10.1080/01431161.2013.793860

Shawky, E., El-Khair, R. M. A., & Selim, D. A. (2020). NIR spectroscopy-multivariate analysis for rapid authentication, detection, and quantification of common plant adulterants in saffron (*Crocus sativus* L.) stigmas. *LWT - Food Sci. Tech., 122*, article 109032. https://doi.org/10.1016/j.lwt.2020.109032

Smidt, E., Meissl, K., Schwanninger, M., & Lechner, P. (2008). Classification of waste materials using Fourier transform infrared spectroscopy and soft independent modeling of class analogy. *Waste Mgmt., 28*(10), 1699-1710. https://doi.org/10.1016/j.wasman.2007.08.003

Sosnoskie, L., Webster, T., & Culpepper, A. (2012). Palmer amaranth pollen viability. Athens, GA: University of Georgia Cooperative Extension. Retrieved from http://ugacottonnews.org/vault/rer/2007/p43.pdf

Suzuki, Y., Okamoto, H., & Kataoka, T. (2008). Image segmentation between crop and weed using hyperspectral imaging for weed detection in soybean field. *Environ. Control. Biol., 46*(3), 163-173. https://doi.org/10.2525/ecb.46.163

Todorova, M. H., & Atanassova, S. L. (2016). Near-infrared spectra and soft independent modeling of class analogy for discrimination of Chernozems, Luvisols, and Vertisols. *J. Near-Infrared Spectros., 24*(3), 271-280. https://doi.org/10.1255/jnirs.1223

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. *Comput. Electron. Agric., 114*, 43-52. https://doi.org/10.1016/j.compag.2015.03.019

Tsugawa, H., Tsujimoto, Y., Arita, M., Bamba, T., & Fukusaki, E. (2011). GC/MS based metabolomics: Development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). *BMC Bioinfo., 12*(1), article 131. https://doi.org/10.1186/1471-2105-12-131

Ward, S. M., Webster, T. M., & Steckel, L. E. (2013). Palmer amaranth (*Amaranthus palmeri*): A review. *Weed Tech., 27*(1), 12-27. https://doi.org/10.1614/WT-D-12-00113.1

Wendel, A., & Underwood, J. (2016). Self-supervised weed detection in vegetable crops using ground-based hyperspectral imaging. *Proc. IEEE Intl. Conf. on Robotics and Automation (ICRA)* (pp. 5128-5135). https://doi.org/10.1109/ICRA.2016.7487717

Wiggins, M. S., McClure, M. A., Hayes, R. M., & Steckel, L. E. (2015). Integrating cover crops and POST herbicides for glyphosate-resistant Palmer amaranth (*Amaranthus palmeri*) control in corn. *Weed Tech., 29*(3), 412-418. https://doi.org/10.1614/WT-D-14-00145.1

Xu, C., Jia, X., Xu, R., Wang, Y., Zhou, Q., & Sun, S. (2013). Rapid discrimination of Herba Cistanches by multi-step infrared macro-fingerprinting combined with soft independent modeling of class analogy (SIMCA). *Spectrochim. Acta A, 114*, 421-431. https://doi.org/10.1016/j.saa.2013.05.024

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. *Comput. Electron. Agric., 159*, 42-49. https://doi.org/10.1016/j.compag.2019.02.018