Scientific Research Publishing

Using Vegetation Indices as Input into Random Forest for Soybean and Weed Classification

Reginald S. Fletcher

Crop Production Systems Research Unit, Agricultural Research Service, United States Department of Agriculture, Stoneville, USA
Email: reginald.fletcher@ars.usda.gov

Abstract

Weed management is a major component of a soybean (*Glycine max* L.) production system; thus, managers need tools to help them distinguish soybean from weeds. Vegetation indices derived from light reflectance properties of plants have shown promise as tools to enhance differences among plants. The objective of this study was to evaluate normalized difference vegetation indices derived from multispectral leaf reflectance data as input into random forest machine learner to differentiate soybean and three broad leaf weeds: Palmer amaranth (*Amaranthus palmeri* L.), redroot pigweed (*A. retroflexus* L.), and velvetleaf (*Abutilon theophrasti* Medik). Leaf reflectance measurements were acquired from plants grown in two separate greenhouse experiments conducted in 2014. Twelve normalized difference vegetation indices were derived from the reflectance measurements, including advanced, green, green-red, green-blue, and normalized difference vegetation indices, shortwave infrared water stress indices, normalized difference pigment and red edge indices, and structure insensitive pigment index. Using the twelve vegetation indices as input variables, the conditional inference version of random forest (cforest) readily distinguished soybean and velvetleaf from the two pigweeds (Palmer amaranth and redroot pigweed) and from each other with classification accuracies ranging from 93.3% to 100%. The greatest errors were observed between the two pigweed classes, with classification accuracies ranging from 70% to 93.3%. Results suggest combining them into one class to increase classification accuracy. Vegetation indices results were equivalent to or slightly better than results obtained with sixteen multispectral bands used as input data into cforest. This research further supports using vegetation indices and machine learning algorithms such as cforest as decision support tools for weed identification.

Keywords

Normalized Difference Vegetation Index, Palmer Amaranth, Redroot Pigweed,

Velvetleaf, Remote Sensing

1. Introduction

Weed management is a major component of a crop production system. To implement weed management strategies, managers need tools to help them distinguish crop from weeds and at times, weeds from other weeds. The latter is important in determining which combination of chemical, mechanical, and/or biological control techniques to use for weed control. In 2012, agricultural producers in the United States (U.S.) spent $13.7 billion for agricultural chemicals [1]. Targeted chemical spraying reduces the amount of herbicide applied, thus decreasing cost and protecting the environment.

Palmer amaranth, redroot pigweed, and velvetleaf infestations occur in soybean fields throughout the eastern U.S. [2] [3] [4]. The weeds produce numerous seeds and establish large populations in soybean fields, thus reducing yields. Dense populations of Palmer amaranth and redroot pigweeds can damage agricultural equipment during harvesting. Populations of Palmer amaranth and redroot pigweed have become resistant to some commonly used herbicides, such as glyphosate. Velvetleaf is also difficult to control with glyphosate. Therefore, managers need effective tools to identify Palmer amaranth, redroot pigweed, and velvetleaf to implement the correct strategy to control them. Remote sensing systems measuring light reflectance properties of plants have shown good potential for weed crop discrimination, including soybean and weed discrimination [5] [6]. The success of remotely sensed data for crop and weed discrimination is dependent on the spectral sensitivity of the recording device and the algorithms used to process the data.

Vegetation indices derived from various mathematical combinations (*i.e.*, ratio, difference, normalized difference) of hyperspectral and multispectral data have shown promise as tools for agricultural application, including determining canopy water content and water stress of crops [7] [8] [9], assessing insect and disease infestations [10] [11] [12] [13] [14], differentiating crops from weeds [15]-[20], and assessing nutrient status of plants [21] [22]. The advantages of using vegetation indices over single wavebands include enhancing differences in the spectral properties of plants, while diminishing the influence of relief, nonphotosynthesizing elements of plants, atmosphere, soil background, shadow, and viewing and illumination geometry on spectral data [23] [24]. Gaps exist on using vegetation indices for crop weed discrimination and weed to weed discrimination, especially in soybean production systems.

Random forest has gain popularity as a tool in numerous disciplines including remote sensing. It is a nonparametric ensemble method that uses numerous classification trees to predict (regression) or determine (classification) the class of an unknown sample [25], hence the name random forest. The algorithm selects a random number of samples from the database provided by the analyst and then uses the samples to develop a decision tree. Sample collection is based on a bootstrap procedure with replacement.

The same process is repeated for each tree in the forest; a different model is constructed for each tree.

Samples randomly selected to derive the model for each tree are referred to as "in-bag" samples, and the samples not used to create the model are called "out-of-bag" samples. Typically, 2/3 and 1/3 of the samples are used as "in-bag" samples and "out-of-bag" samples, respectively. Based on the "in-bag" and "out-of-bag" concept, random forest does not need a separate testing set to evaluate the model.

To test classification or prediction accuracy, decision trees in which a sample is "out-of-bag" are predicted or classified with trees in which the sample is not "in-bag". The vote of each tree is tallied; then the unknown sample is assigned the class in which it received the most votes. For each tree, a subset of the variables is used to develop the split. Also, a variable importance ranking is derived with the algorithm. Analysts can evaluate the ranking and identify variables that are relevant to the model for prediction or classification problems, leading to reduction in the number of variables used for the analysis.

Random forest ranks very well among other classifiers [26]. Recently, it was ranked in the top ten of 100 classifiers tested for classification purposes [26]. It is capable of processing large datasets, can analyze numerous variables without deletion, is robust for analysis of datasets with missing variables, and has the ability to process unbalanced datasets. Models derived for classification can be used on other datasets. Light reflectance data recorded from sensors on-board satellite, airborne, and ground-based platforms have shown promise as input data for random forest to use for classification [27] [28] [29] [30] and regression [31] [32] problems.

Currently, no information is available on using vegetation indices derived from multispectral data as input into random forest for soybean weed discrimination. The objective of this study was to evaluate normalized difference vegetation indices as input into random forest to differentiate soybeans and three broad leaf weeds: Palmer amaranth, redroot pigweed, and velvetleaf. The study focused on leaf multispectral reflectance data of simulated World View 3 satellite sensor bands.

2. Materials and Methods

2.1. Experimental Setup

Two experiments were completed in 2014 in which 30 replicates of soybean variety 4928LL (LL-liberty link), Palmer amaranth, redroot pigweed, and velvetleaf were grown in a greenhouse located at USDA-ARS, Stoneville, MS. Planting dates were June 13, 2014, and August 28, 2014, for experiments one and two, respectively. Seeds of the different plants (*i.e.*, soybean variety 4928LL, Palmer amaranth, redroot pigweed, and velvetleaf) were sown into plugs. After emergence, the plant species were transferred to two liter pots. All plant species were exposed to a fourteen-hour photoperiod, and light was supplemented at the beginning and ending of the day with sodium vapor lamps. Day/night temperatures of the greenhouse was maintained at 28°C/24°C ± 3°C, respectively. The soybean variety used in the study had an indeterminate growth habit and

gray pubescence and was assigned to maturity group 4.9. The weed seeds were obtained from a seed bank maintained at the laboratory.

2.2. Leaf Reflectance Measurements

Leaf reflectance measurements were obtained with a plant contact probe attached to a spectroradiometer (Fieldspec 3, ASD Inc. Boulder, Colorado) sensitive to a spectral range of 350 to 2500 nm. The contact probe has its own light source allowing the user to collect data anytime during the day or night. Reflectance measurements were collected from the most recently matured leaf of each plant; each reading was an average of fifteen readings. The spectroadiometer was calibrated in fifteen minute intervals with a white spectralon panel. Leaf reflectance measurements were collected prior to the plants reaching 1 foot tall. The goal of weed management strategies is to detect and kill the weeds in the vegetative growth stages and prior to the seeds reaching full maturity levels.

2.3. Post Processing of Reflectance Measurements

Multispectral data simulating the spectral bands of the World View 3 satellite sensor were derived from the hyperspectral spectroradiometer data (Table 1). These spectral bands were chosen because they represented data in the visible, red edge, near infrared, and shortwave infrared regions of the light spectrum. Twelve vegetation indices were created with the multispectral spectral bands (Table 2). The selected indices have shown good potential for assessing leaf pigments [33]-[38], leaf internal structure [36], and leaf water content [39].

Table 1. Sixteen-spectral bands simulating World View 3 satellite sensor spectral bands. They were used as input variables into cforest for soybean and weed discrimination.

Band name	Acronym	Bandwidth
Coastal	C	400 - 450 nm
Blue	B	450 - 510 nm
Green	G	510 - 580 nm
Yellow	Y	585 - 625 nm
Red	R	630 - 690 nm
Red edge	RE	705 - 745 nm
Near infrared 1	NIR1	770 - 895 nm
Near infrared 2	NIR2	860 - 1040 nm
Shortwave infrared 1	SWIR1	1195 - 1225 nm
Shortwave infrared 2	SWIR2	1550 - 1590 nm
Shortwave infrared 3	SWIR3	1640 - 1680 nm
Shortwave infrared 4	SWIR4	1710 - 1750 nm
Shortwave infrared 5	SWIR5	2145 - 2185 nm
Shortwave infrared 6	SWIR6	2185 - 2225 nm
Shortwave infrared 7	SWIR7	2235 - 2285 nm
Shortwave infrared 8	SWIR8	2295 - 2365 nm

Table 2. Vegetation indices used as input into cforest for soybean and weed discrimination.

Vegetation index	Acronym	Formula for vegetation indices	Spectral bands[a] (nm)
Advanced Normalized Vegetation Index [18]	ANVI	$\text{ANVI1} = (\text{NIR1} - \text{C})/(\text{NIR1} + \text{C})$ $\text{ANVI2} = (\text{NIR1} - \text{B})/(\text{NIR1} + \text{B})$	C: 400 - 450, B: 450 - 510, NIR1: 770 - 895
Green Normalized Difference Vegetation Index [35]	GNDVI	$\text{GNDVI} = (\text{NIR1} - \text{G})/(\text{NIR1} + \text{G})$	G: 510 - 580, NIR1: 770 - 895
Green-Red Normalized Difference Vegetation Index [14]	GRNDVI	$\text{GRNDVI} = (\text{G} - \text{R})/(\text{G} + \text{R})$	G: 510 - 580, R: 630 - 690
Green-Blue Normalized Difference Vegetation Index (This paper)	GBNDVI	$\text{GBNDVI} = (\text{G} - \text{B})/(\text{G} + \text{B})$	G: 510 - 580, B: 450 - 510
Normalized Difference Vegetation Index [54]	NDVI	$\text{NDVI} = (\text{NIR1} - \text{R})/(\text{NIR1} + \text{R})$	R: 630 - 690, NIR1: 770 - 895
Shortwave Infrared Water Stress Index [8] [39]	SIWSI	$\text{SIWSI1} = (\text{NIR1} - \text{SWIR3})/(\text{NIR1} + \text{SWIR3})$ $\text{SIWSI2} = (\text{NIR1} - \text{SWIR6})/(\text{NIR1} + \text{SWIR6})$	NIR1: 770 - 895, SWIR3: 1640 - 1680, SWIR6: 2185 - 2225
Normalized Difference Pigment Index [33]	NPCI	$\text{NPCI} = (\text{R} - \text{C})/(\text{R} + \text{C})$	C: 400 - 450, R: 630 - 690
Normalized Difference Red Edge Index [53]	NDRE	$\text{NDRE} = (\text{NIR1} - \text{RE})/(\text{NIR1} + \text{RE})$	RE: 705 - 745, NIR1: 770 - 895
Structure Insensitive Pigment Index [34]	SIPI	$\text{SIPI1} = (\text{NIR1} - \text{C})/(\text{NIR1} + \text{R})$ $\text{SIPI2} = (\text{NIR1} - \text{B})/(\text{NIR1} + \text{R})$	C: 400 - 450, B: 450 - 510, R: 630 - 690, NIR1: 770 - 895

[a]C = coastal, B = blue, G = green, R = red, RE = red edge, NIR = near infrared, and SWIR = shortwave infrared.

Broadband and narrowband data have been used to develop vegetation indices. Therefore, the simulated band center wavelength closest to center wavelengths used in other studies were employed in developing each vegetation index. Periodically, two vegetation indices were developed for a designated index because the band centers were equidistant from the band centers used by other investigators. These include the advanced normalized difference vegetation index (ANVI), shortwave infrared water stress index (SIWSI), and structure insensitive pigment index (SIPI). The hsdar package (hyperspectral data analysis in R) of R [40] and base R packages [41] were used to develop the sixteen spectral bands and the twelve vegetation indices, respectively.

2.4. Classification

The conditional inference version of random forest (cforest) was used for the classifications. Researchers have suggested using the cforest version of random forest if the data are highly correlated [42]. Cforest is more stable in deriving variable importance values

in the presence of highly correlated variables, thus providing better accuracy in calculating variable importance [42].

Differences of cforest compared with the original version of random forest are as follows. Cforest employs conditional inference trees as base learners; random forest uses classification and regression trees as base learners [42] [43]. Cforest develops unbiased decision trees based on subsampling without replacement instead of using bootstrap samples. The algorithm uses the conditional permutation scheme described by Strobl *et al*. [42] [47] to determine variable importance ranking.

Two parameters were adusted to develop the cforest models, *mtry* and *ntree*. The former respresents the number of random variables to use in each tree; the latter characterizes the number of trees.

Model stability was tested using the following procedure [6] [44]: 1) develop model using the default *mtry* (5) and *ntree* (500) values, 2) tabulate and review variable of importance rankings, 3) adjust starting seed (*i.e.*, the random generator used as a starting point for sampling) and rerun the model using the default *mtry* and *ntree* values, 4) accept model if variable of importance rankings are consistent between the first and second runs; if not proceed to step five, 5) increase *ntree* by 500 and rerun model following steps one thru four. The procedure was repeated until the variable of importance rankings were consistent between the first and second runs. The party package [45] [46] [47] of R was used to complete the classifications and to obtain the variable importance readings.

For each date, two datasets were evaluated as input into the cforest algorithm: 1) twelve vegetation indices dataset and 2) sixteen band multispectral dataset. Overall, user's, and producer's accuracies, and the kappa coefficient, were tabulated to compare accuracies of the classifications [48].

3. Results

Identical accuracy results were obtained for the vegetation indices and the multispectral data classifications for the June 30, 2014 dataset (Table 3). Overall accuracy and the kappa coefficient were 90.8% and 0.878, respectively. User's and producer's accuracies ranged from 76.7% to 100%. The highest user's and producer's accuracies were obtained for the soybean class. The lowest user's and producer's accuracies were observed for the redroot pigweed and Palmer amaranth classes, respectively.

For the September 17, 2014 dataset, the vegetation indices achieved higher classification accuracies than the multispectral dataset with the differences ranging from 0.7% to 3.4% for user's, producer's, and overall accuracies (Table 3). Similar trends were observed in the classification accuracies of the vegetation indices and multispectral data. The best classification accuracies were achieved for the soybean class. Also, velvetleaf was tied for first for the producer's accuracy. The lowest user's and producer's accuracies were observed for the redroot pigweed and Palmer amaranth classes, respectively.

Variable importance rankings indicated that eight and nine of the twelve indices were important to the classification of the June and September vegetation indices datasets,

Table 3. Accuracy results for the cforest classifications for each dataset.

Accuracy measurements	Datasets			
	Vegetation indices (June 30, 2014)	Multispectral (June 30, 2014)	Vegetation indices (September 17, 2014)	Multispectral (September 17, 2014)
User's accuracy soybean	93.8%	93.8%	96.7%	93.3%
User's accuracy Palmer amaranth	92.0%	92.0%	80.7%	80.0%
User's accuracy redroot pigweed	84.4%	84.4%	78.8%	73.5%
User's accuracy velvetleaf	93.3%	93.3%	93.5%	90.3%
Producer's accuracy soybean	100%	100%	96.7%	93.3%
Producer's accuracy Palmer amaranth	76.7%	76.7%	70.0%	66.6%
Producer's accuracy redroot pigweed	93.3%	93.3%	86.7%	83.3%
Producer's accuracy velvetleaf	93.3%	93.3%	96.7%	93.3%
Overall accuracy	90.8%	90.8%	87.5%	84.2%
Kappa coefficient	0.878	0.878	0.833	0.789

respectively (Figure 1). SIWSI2 was ranked as the most important vegetation index used by the classification models. Its variable importance score was approximately 1.5 times greater than the second ranked vegetation index score.

The multispectral datasets variable importance rankings are summarized in Figure 1. Shortwave infrared bands had strong to moderate variable importance rankings for the June dataset. NIR2, G, RE, and NIR1 bands had moderate to low variable importance scores. All of the spectral bands were important to the September 17, 2014 multispectral dataset classification model because their variable importance rankings were distinguishable from the zero line. The highest and lowest rankings were assigned to G and C bands, respectively. Finally, more trees were needed to obtain stable variable importance rankings for the multispectral datasets compared with the vegetation indices datasets (Table 4). That aspect could be attributed to the multispectral bands datasets having more variables than the vegetation indices datasets, thus requiring more trees to be used for the stabilization process.
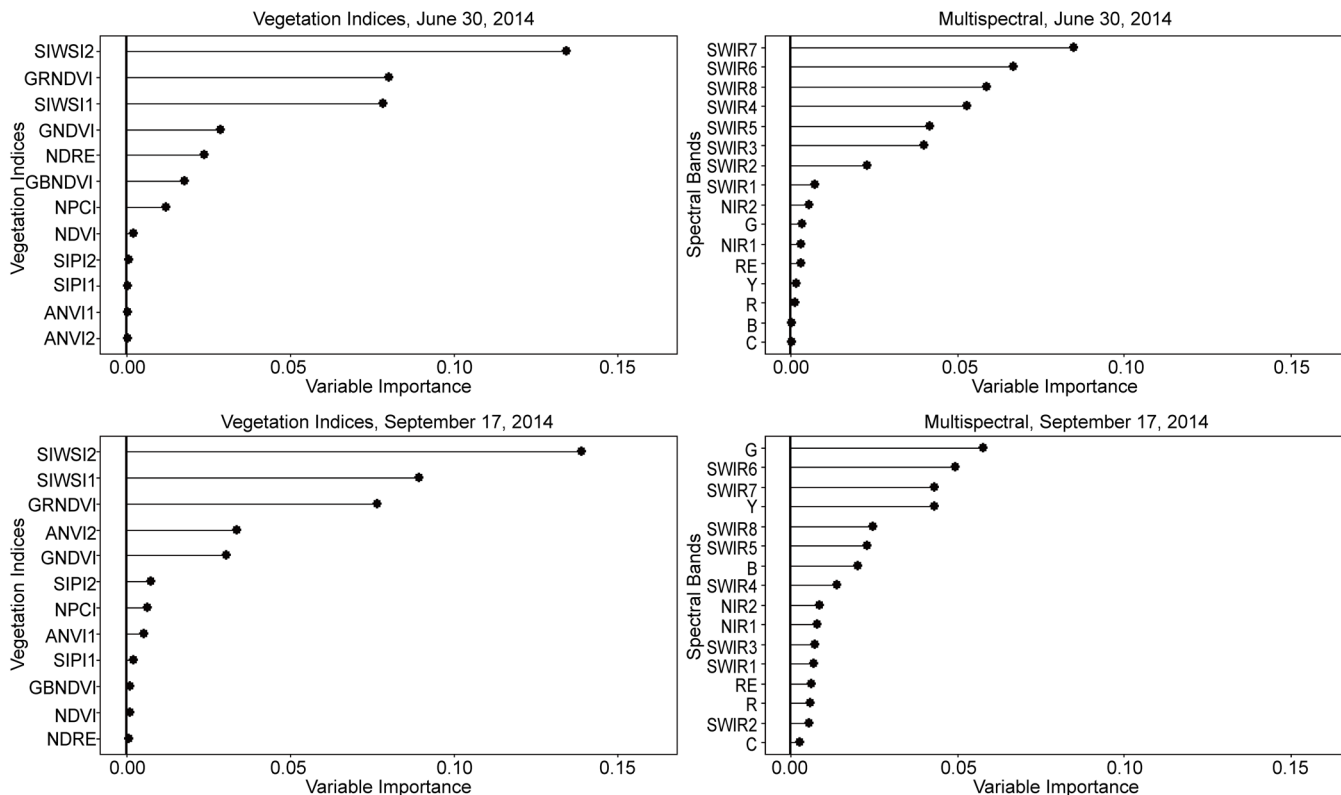
Figure 1. Variable importance rankings for vegetation indices and multispectral data used as input into cforest for classification.

Table 4. Parameters used for cforest classifications.

Dataset	Date	$mtry^a$	$ntree$
Vegetation indices	06/30/2014	5	1000
Multispectral	06/30/2014	5	1500
Vegetation indices	09/17/2014	5	500
Multispectral	09/17/2014	5	1000

[a]$mtry$ = number of predictors to use when splitting a node; $ntree$ = number of trees grown.

4. Discussions

Using vegetation indices as input variables for soybean and weed discrimination, cforest achieved classification accuracies that were equivalent to or slightly better than classification accuracies obtained with multispectral bands as input variables (Table 3). Kappa coefficients for the vegetation indices classifications indicated an almost perfect agreement (values within the range of 0.81 - 1.00, [49]) between reference and predicted data. Almost perfect to substantial agreement (*i.e.*, values within the range of 0.61 - 0.80, [49]) occurred between reference data and predicted data for the June 30 and September 17, 2014 multispectral datasets, respectively. Errors for soybean and velvetleaf classes were attributed to the former being misclassified as the latter and vice versa. The cforest algorithm using vegetation indices or multispectral bands as input

had problems in distinguishing between Palmer amaranth and redroot pigweed, suggesting combining them into one class.

Consistently, indices derived with SWIR and NIR bands, the G and NIR bands, and G and R bands were considered important in separating the plant species. SIWSI1 and SIWSI2 were ranked in the top three for vegetation indices variables when evaluating both datasets (Figure 1). Those indices are sensitive to moisture content in plant leaves [50] [51], suggesting leaf succulence was important in crop weed separation and weed to weed separation. The SWIR band serves as a measure of water content, leaf internal structure, and dry matter; whereas, the NIR band serves as measurement of only leaf internal structure and dry matter [50]. When combined into an index, SIWSI provides a measurement of leaf water content expressed as equivalent water thickness.

Also, GRNDVI was ranked within the top three indices for each date supporting the role that plant pigment played in separating the plant species. Other pigment indices having relevant variable importance values on both dates were GNDVI and NPCI. NDVI, the most wildly used vegetation index throughout the literature [52] [53] [54], was ranked eighth and eleventh according to the variable importance tabulated for the June 30, 2014 and September 17, 2014 datasets respectively, indicating its weak relevance for discriminating between the plant species.

The vegetation indices were derived using the multispectral bands (Table 1). Thus, the variable importance rankings of the former were easily understood by evaluating the variable importance rankings of the latter. SWIR bands 2 - 8 were a necessity for differentiating between the plant species. Therefore, vegetation indices derived with one those bands would be ranked higher on the variable importance scale than a vegetation index excluding SWIR bands (*i.e.*, SWIR bands 2 - 8). The findings of this study are similar to those of [5], which indicated shortwave-infrared data were important for soybean weed discrimination. Their study focused on the following weeds discrimination from soybean: hemp sesbania [*Sesbania exaltata* (Raf.) Rydb. ex. A.W. Hill], palm leaf morning glory (*Ipomoea wrightii* Gray), pitted morning glory (*Ipomoea lacunose* L.), prickly sida (*Sida spinose* L.), sicklepod [*Senna obtusifolia* (L.) H.S. Irwin and Barnaby], and small flower morning glory [*Jacquemontia tamnifolia* (L.) Griseb.].

Overall, this study demonstrated that vegetation indices derived from leaf reflectance data can be used as input into cforest for differentiating soybean, velvetleaf, and two pigweeds. It is important to note that the data were collected using pure leaf spectra and not canopy spectra. For canopy spectra, some differences will occur due to in-canopy shadowing, leaf orientation, and differences in leaf area. One of the strengths of vegetation indices is that they are influenced less by shadowing and sensor viewing angle than multispectral bands, indicating the benefit of using them at the canopy level.

5. Conclusion

Vegetation indices derived from spectral data simulating World View 3 bands showed promise as input variables into cforest for discriminating soybean and three weed species. Indices consisting of shortwave infrared bands were important for plant species

differentiation. Also, the analyst should consider combining redroot pigweed and Palmer amaranth into one class, pigweed. If different herbicides were going to be used to treat the pigweeds and the velvetleaf plants, the results indicate that the cforest algorithm could use vegetation indices derived from leaf light reflectance data to accurately identify pigweeds, velvetleaf, and soybean plants. This research further supports using vegetation indices and machine learning algorithms such as cforest as decision support tools for weed and crop differentiation.

Acknowledgements

References

[1]   Bunge, J. (2014) For Weed Control, Farmers Widen Their Arsenal of Herbicides. *The Wall Street Journal*.
http://www.wsj.com/articles/SB10001424052702303847804579481641717350038

[2]   Spencer, N.R. (1984) Velvetleaf, Abutilon Theophrasti (Malvaceae), History and Economic Impact in the United States. *Economic Botany*, **38**, 407-416.
http://dx.doi.org/10.1007/BF02859079

[3]   Ulloa, S.M., Datta, A. and Knezevic, S.Z. (2010) Growth Stage-Influenced Differential Response of Foxtail and Pigweed Species to Broadcast Flaming. *Weed Technology*, **24**, 319-325. http://dx.doi.org/10.1614/WT-D-10-00005.1

[4]   Ward, S.M., Webster, T.M. and Steckel, L.E. (2013) Palmer Amaranth (*Amaranthus palmeri*): A Review. *Weed Technology*, **27**, 12-27. http://dx.doi.org/10.1614/WT-D-12-00113.1

[5]   Gray, C.J., Shaw, D.R. and Bruce, L.M. (2009) Utility of Hyperspectral Reflectance for Differentiating Soybean (*Glycine max*) and Six Weed Species. *Weed Technology*, **23**, 108-119. http://dx.doi.org/10.1614/WT-07-117.1

[6]   Fletcher, R.S. (2015) Testing Leaf Multispectral Reflectance Data as Input into Random Forest to Differentiate Velvetleaf from Soybean. *American Journal of Plant Sciences*, **6**, 3193-3204. http://dx.doi.org/10.4236/ajps.2015.619311

[7]   Jackson, T.J., Chen, D.Y., Cosh, M., Li, F.Q., Anderson, M., Walthall, C., Doriaswamy, P. and Hunt, E.R. (2004) Vegetation Water Content Mapping Using Landsat Data Derived Normalized Difference Water Index for Corn and Soybeans. *Remote Sensing of Environment*, **92**, 475-482. http://dx.doi.org/10.1016/j.rse.2003.10.021

[8]   Chen, D.Y., Huang, J.F. and Jackson, T.J. (2005) Vegetation Water Content Estimation for Corn and Soybeans Using Spectral Indices Derived from MODIS Near- and Short-Wave Infrared Bands. *Remote Sensing of Environment*, **98**, 225-236.
http://dx.doi.org/10.1016/j.rse.2005.07.008

[9]   Cheng, T., Riano, D., Koltunov, A., Whiting, M.L., Ustin, S.L. and Rodriguez, J. (2013) Detection of Diurnal Variation in Orchard Canopy Water Content Using MODIS/ASTER

Airborne Simulator (MASTER) Data. *Remote Sensing of Environment*, **132**, 1-12.
http://dx.doi.org/10.1016/j.rse.2012.12.024

[10] Genc, H., Genc, L., Turhan, H., Smith, S.E. and Nation, J.L. (2008) Vegetation Indices as Indicators of Damage by the Sunn Pest (Hemiptera: Scutelleridae) to Field Grown Wheat. *African Journal of Biotechnology*, **7**, 173-180.

[11] Kumar, J., Vashisth, A., Sehgal, V.K. and Gupta, V.K. (2010) Identification of Aphid Infestation in Mustard by Hyperspectral Remote Sensing. *Journal of Agricultural Physics*, **10**, 53-60.

[12] Calderon, R., Navas-Cortes, J.A., Lucena, C. and Zarco-Tejada, P.J. (2013) High-Resolution Airborne Hyperspectral and Thermal Imagery for Early Detection of Verticillium Wilt of Olive Using Fluorescence, Temperature, and Narrow-Band Spectral Indices. *Remote Sensing of Environment*, **139**, 231-245. http://dx.doi.org/10.1016/j.rse.2013.07.031

[13] Ashourloo, D., Mobasheri, M.R. and Huete, A. (2014) Developing Two Spectral Disease Indices for Detection of Wheat Leaf Rust (*Puccinia triticina*). *Remote Sensing*, **6**, 4723-4740. http://dx.doi.org/10.3390/rs6064723

[14] Ranjitha, G., Srinivasan M.R. and Rajesh, A. (2014) Detection and Estimation of Damage Caused by *Thrips tabaci* (Lind) of Cotton Using Hyperspectral Radiometer. *Agrotechnology*, **3**, 1-5.

[15] Lamb, D.V. and Weedon, M. (1998) Evaluating Accuracy of Mapping Weeds in Fallow Fields Using Airborne Imaging: *Panicum effesum* in Oil-Seed Rape Stubble. *Weed Research*, **38**, 443-451. http://dx.doi.org/10.1046/j.1365-3180.1998.00112.x

[16] Lamb, D.V., Weedon, M. and Rew, L.J. (1999) Evaluating the Accuracy of Mapping Weeds in Seedling Crops using Airborne Digital Imagery: *Avena* spp. in Seedling Triticale. *Weed Research*, **39**, 481-492. http://dx.doi.org/10.1046/j.1365-3180.1999.00167.x

[17] Smith, A.M. and Blackshaw, R.E. (2003) Weed-Crop Discrimination Using Remote Sensing: A Detached Leaf Experiment. *Weed Technology*, **17**, 811-820. http://dx.doi.org/10.1614/WT02-179

[18] Peña-Barragãn, J.M., López-Granados, F., Jurado-Expósito, M. and Garciá-Torres, L. (2006) Spectral Discrimination of *Ridolfia segetum* and Sunflower as Affected by Phenological Stage. *Weed Research*, **46**, 10-21. http://dx.doi.org/10.1111/j.1365-3180.2006.00488.x

[19] Gómez-Casero, M.T., Castillejo-González, I.L., García-Ferrer, A., Peña-Barragán, J.M., Jurado-Expósito, M., García-Torres, L. and López-Granados, F.L. (2010) Spectral Discrimination of Wild Oat and Canary Grass in Wheat Fields for Less Herbicide Application. *Agronomy for Sustainable Development*, **30**, 689-699. http://dx.doi.org/10.1051/agro/2009052

[20] De Castro, A.I., Jurado-Expósito, M., Gómez-Casero, M.T. and López-Granados, F. (2012) Applying Neural Networks to Hyperspectral and Multispectral Field Data for Discrimination of Cruciferous Weeds in Winter Crops. *The Scientific World Journal*, **2012**, Article ID: 630390. http://dx.doi.org/10.1100/2012/630390

[21] Li, F., Gnyp, M.L., Jia, L., Miao, Y., Yu, Z., Koppe, W., Bareth, G., Chen, X. and Zhang, F. (2008) Estimating N Status of Winter Wheat Using a Handheld Spectrometer in the North China Plain. *Field Crop Research*, **106**, 77-85. http://dx.doi.org/10.1016/j.fcr.2007.11.001

[22] Bausch, W.C. and Khosla, R. (2010) Quick Bird Satellite versus Ground-Based Multi-Spectral Data for Estimating Nitrogen Status of Irrigated Maize. *Precision Agriculture*, **11**, 274-290. http://dx.doi.org/10.1007/s11119-009-9133-1

[23] Huete, A. and Justice, C. (1999) MODIS Vegetation Index (MOD 13) Algorithm Theoretical Basis Document. https://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf

[24] Hatfield, J.L. and Prueger, J.H. (2010) Value of Using Different Vegetative Indices to Quantify Agricultural Crop Characteristics at Different Growth Stages under Varying Management Practices. *Remote Sensing*, **2**, 562-578. http://dx.doi.org/10.3390/rs2020562

[25] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. http://dx.doi.org/10.1023/A:1010933404324

[26] Fernández-Delgado, M., Cernadas, E., Barro, S. and Amorim, D. (2014) Do We Need Hundreds of Classifiers to Solve Real World Classification Problems? *Journal of Machine Learning Research*, **15**, 3133-3181.

[27] Pal, M. (2005) Random Forest Classifier for *Remote Sensing* Classification. *International Journal of Remote Sensing*, **26**, 217-222. http://dx.doi.org/10.1080/01431160412331269698

[28] Gislason, P.O., Benediktsson, J.A. and Sveinsson, J.R. (2006) Random Forests for Land Cover Classification. *Pattern Recognition Letters*, **27**, 294-300. http://dx.doi.org/10.1016/j.patrec.2005.08.011

[29] Lawrence, R.L., Wood, S.D. and Sheley, R.L. (2006) Mapping Invasive Plants Using Hyperspectral Imagery and Breiman Cutler Classifications (Random Forest). *Remote Sensing of Environment*, **100**, 356-362. http://dx.doi.org/10.1016/j.rse.2005.10.014

[30] Adam, E.M., Mutanga, O., Rugege, D. and Ismail, R. (2012) Discriminating the Papyrus Vegetation (*Cyperus papyrus* L.) and Its Co-Existent Species Using Random Forest and Hyperspectral Data Resampled to HYMAP. *International Journal of Remote Sensing*, **33**, 552-569. http://dx.doi.org/10.1080/01431161.2010.543182

[31] Ismail, R. and Mutanga, O. (2010) A Comparison of Regression Tree Ensembles: Predicting *Sirex noctilio* Induced Water Stress in *Pinus patula* Forests of KwaZulu-Natal, South Africa. *International Journal of Applied Earth Observation and Geoinformation*, **12**, S45-S51. http://dx.doi.org/10.1016/j.jag.2009.09.004

[32] Abdel-Rahman, E.M., van den Berg, M., Way, M.J. and Ahmed, F.B. (2009) Hand-Held Spectrometry for Estimating Thrips (*Fulmekiola serrata*) Incidence in Sugarcane. *IEEE International Geoscience and Remote Sensing Symposium* (*IGARSS*), **4**, 268 -271. http://dx.doi.org/10.1109/igarss.2009.5417322

[33] Penuelas, J., Gamon, J., Freeden, A., Merino, J. and Field, C. (1994) Reflectance Indices Associated with Physiological Changes in Nitrogen and Water Limited Sunflower Leaves. *Remote Sensing of Environment*, **48**, 135-146. http://dx.doi.org/10.1016/0034-4257(94)90136-8

[34] Penuelas, J., Baret, F. and Filella, I. (1995) Semi-Empirical Indices to Assess Carotenoids/Chlorophyll a Ratio from Leaf Spectral Reflectance. *Photosynthetica*, **31**, 221-230.

[35] Gitelson, A.A., Kaufman, Y.J. and Merzlyak, M.N. (1996) Use of a Green Channel in Remote Sensing of Global Vegetation from EOS-MODIS. *Remote Sensing of Environment*, **58**, 289-298. http://dx.doi.org/10.1016/S0034-4257(96)00072-7

[36] Araus, J.L., Casadesus, J. and Bort, J. (2001) Recent Tools for the Screening of Physiological Traits Determining Yield. In: Reynolds, M.P., Ortiz-Monasterio, J.I. and McNabb, A., Eds., *Application of Physiology in Wheat Breeding*, CIMMYT, International Maize and Wheat Improvement Center Apdo, 59-77.

[37] Motohka, T., Nasahara, K.N., Oguma, H. and Tsuchida, S. (2010) Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. *Remote Sensing*, **2**, 2369-2387. http://dx.doi.org/10.3390/rs2102369

[38] Eitel, J.U.H., Vierling, L.A., Litvak, M.E., Long, D.S., Schulthess, U., Ager, A.A., Krofcheck, D.J. and Stoscheck, L. (2011) Broadband, Red-Edge Information From Satellites Improves Early Stress Detection in a New Mexico Conifer Woodland. *Remote Sensing of Environment*, **115**, 3640-3646. http://dx.doi.org/10.1016/j.rse.2011.09.002

[39] Gao, B.C. (1996) NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. *Remote Sensing of Environment*, **58**, 257-266. http://dx.doi.org/10.1016/S0034-4257(96)00067-3

[40] Lehnert, L.W., Meyer, H. and Bendix, J. (2016) Hsdar: Manage, Analyse and Simulate Hyperspectral Data in R. R Package Version 0.4.1.

[41] R Core Team (2015) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[42] Strobl, C., Hothorn, S. and Zeileis, A. (2009) Party on! A New, Conditional Variable Importance Measure for Random Forests Available in the Party Package. Technical Report Number 050, Department of Statistics, University of Munich, Munich.

[43] Hothorn, T., Hornik, K. and Zeileis, A. (2006) Unbiased Recursive Portioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, **15**, 651-674. http://dx.doi.org/10.1198/106186006X133933

[44] Strobl, C., Malley, J. and Tutz, G. (2009) An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests. *Physiological Methods*, **14**, 323-348. http://dx.doi.org/10.1037/a0016973

[45] Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A. and Van Der Laan, M. (2006b) Survival Ensembles. *Biostatistics*, **7**, 355-373. http://dx.doi.org/10.1093/biostatistics/kxj011

[46] Strobl, C., Boulesteix, A.L., Zeileis, A. and Hothorn, T. (2007) Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *BMC Bioinformatics*, **8**, 25. http://www.biomedcentral.com/1471-2105/8/25
http://dx.doi.org/10.1186/1471-2105-8-25

[47] Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T. and Zeileis, A. (2008) Conditional Variable Importance for Random Forests. *BMC Bioinformatics*, **9**, 307.
http://www.biomedcentral.com/1471-2105/9/307
http://dx.doi.org/10.1186/1471-2105-9-307

[48] Congalton, R. and Green, K. (2009) Assessing the Accuracy of Remotely Sensed Data Principles and Practices. 2nd Edition, Lewis, Boca Raton.

[49] Landis, J.R. and Koch, G.G. (1977) An Application of Hierarchical Kappa-Type Statistics in the Assessment of Majority Agreement Among Multiple Observers. *Biometrics*, **33**, 363-374. http://dx.doi.org/10.2307/2529786

[50] Ceccato, P., Flassee, S., Tarantola, S., Jacquemoud, S. and Gregoire, J.M. (2001) Detecting Vegetation Leaf Water Content using Reflectance in the Optical Domain. *Remote Sensing of Environment*, **77**, 22-33. http://dx.doi.org/10.1016/S0034-4257(01)00191-2

[51] Fensholt, R. and Sandholt, I. (2003) Derivation of a Shortwave Infrared Water Stress Index from MODIS Near- and Shortwave Infrared Data in a Semiarid Environment. *Remote Sensing of Environment*, **87**, 111-121. http://dx.doi.org/10.1016/j.rse.2003.07.002

[52] Wójtowicz, M., Wójtowicz, A. and Piekarczyk, J. (2016) Application of Remote Sensing Methods in Agriculture. *Communications in Biometry and Crop Science*, **11**, 31-50.

[53] Barnes, E.M., Clarke, T.R., Richards, S.E., Colaizzi, P.D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R.J., Li, H. and Moran, M.S. (2000) Coincident Detection of Crop Water Stress, Nitrogen Status, and Canopy Density Using Ground-Based Multispectral Data. 5*th International Conference on Precision Agriculture*, Bloomington, 16-19 July 2000, 1-15.

[54] Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D.W. (1974) Monitoring Vegetation Systems in the Great Plains with ERTS. 3*rd Earth Resource Technology Satellite (ERTS)*, **1**, 48-62.

Scientific Research Publishing

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work

Submit your manuscript at: http://papersubmission.scirp.org/
Or contact ajps@scirp.org