Identification and characterization of crops through the analysis of spectral data with machine learning algorithms

Rigalli N.F., Montero Bulacio E., Romagnoli M., Terissi L., and Portapila M.

CIFASIS, Rosario, Santa Fe, Argentina
portapila@cifasis-conicet.gov.ar,
http://cifasis-conicet.gov.ar/

Abstract. This paper assesses the capability of an spectrometer used in field experiments of soybean, maize and wheat. The objective of this work is to select different wavelengths intervals of the spectral reflectance curve, within the range 632-1125 nm, as features for classification using machine learning methods. Two different classifications are presented, species selection and growth stage identification. For species classification accuracy of 92% is reached, while 99% is obtained for stage classification. In addition we propose a new index that outperforms analyzed established vegetation indices, which shows the potential advantage of using this type of devices.

Keywords: spectrometry, remote sensing, NIR, spectral feature selection

1 Introduction

The use of remote sensors in agronomy, for identification and characterization of crops, has been increasing in the recent years. For these purposes, different platforms have been employed for data collection at multiple scales. For instance, Virlet *et al.*[1] developed a terrestrial phenotyping platform based on thermal fluorescent chambers, chlorophyll fluorescence, hyperspectral cameras and 3D laser scanners, in order to produce a description of canopy development across the crops entire lifecycle. Sankaran *et al.* [2] predicted that the use of unmanned aerial vehicle (UAV) based technology will grow exponentially in the next few years, resulting in the development of robust aerial sensing-based crop phenotyping methods. Al-Yaari *et al.* [3] reported that remotely sensed satellite-based passive microwave soil moisture is strongly related to leaf area index. Multispectral cameras applied to any of these platforms are a good source of information, however, they measure the intensity of light at a few particular wavelengths.

The vegetation indices employed for crops characterization are generally computed from Red and near-infrared (NIR) spectrum zones [4]. The relationship between NIR and Red, Ratio Vegetation Index (RVI), was proposed in [5] to estimate the leaf area index (LAI) in forests. Tucker [6] defined the normalized

difference vegetation index (NDVI), which can be used to monitor the photosynthetically active biomass of plant canopies. Currently, NDVI is the most used index. To minimize soil brightness influences from the indices of spectral vegetation, Huete [7] applied the Soil-Adjusted Vegetation Index (SAVI) and added a correction factor according to the LAI and coverage. The modified SAVI (MSAVI) described in [8] replaces the correction factor of the SAVI by a function that increases the dynamic range and reduces the background effects of the soil. Spectrometers can measure spectral reflectance over an interval of the electromagnetic spectrum with high wavelength selectivity. This capability allows an in-depth analysis of the spectrum, for instance, Lorenzen *et al.*[9] used different intervals of the reflectance spectrum separately to detect leaves infected with mildew in barley. In addition, the traditional vegetation indices (i.e. NDVI, RVI, SAVI, MSAVI) can be computed from the spectral reflectance curve obtained with spectrometers [10, 11]. The high wavelength selectivity of the spectrometer also enables the design of new vegetation indices [6].

On the other hand, Machine Learning Algorithms (MLA) allow to predict and project the status of crops over vast extensions of land by using data gathered through remote sensing techniques. In Pantazi *et al.* [12] wheat yield is predicted using MLA through satellite computed NDVI and hyperspectral images. Yuan *et al.* [13] used Artificial Neural Networks (ANN), Random Forest (RF), and Support Vector Machine (SVM) regression to predict LAI in soybean (*Glycine max (L.) Merr.*) crops through hyperspectral remote sensing. Gao *et al.* [14] used RF to distinguish maize (*Zea mays*), *Convolvulus arvensis, Rumex sp.* and *Cirsium arvense* based on NIR hyperspectral images.

The objective of this paper is to select different wavelength intervals of the spectral reflectance curve, within the range 632-1125 nm, of soybean, maize and wheat (*Triticum aestivum L.*) crops for classification using machine learning methods. Data is obtained using an spectrometer from field experiments in the Pampas regions of Argentina. The crop classification technique proposed in this paper, based on the reflectance spectrum, is evaluated for two different applications, (i) species identification and (ii) growth stage detection of soybean crops. These numerical experiments are carried out considering SVM with linear kernel (SVMLIN), RF and ANN classification methods. Besides, in order to show the potential advantage of using a spectrometer for characterizing crops, a new vegetation index is proposed.

2 Materials and methods

2.1 Datasets

Soybean data was collected from plots in Conesa (33° 35' S , 60° 18'W) and Gilbert (32° 27'S, 58° 58' W). Maize data was obtained from plots in Santa Fe (31° 38' S, 60° 40' W). Wheat data was obtained from plots in Pergamino (33° 51' S, 60° 32' W). Experimental details are shown in table 1. All Field experiments were conducted following a randomized complete block design. All measurements were taken between 09:30 and 12:00.

Dataset 1: The spectral reflectance curve dataset for species classification contained all samples reported in table 1. Since we have different number of spectra curves per crop specie, the unbalanced dataset was balanced using bootstrap resampling [15].
Dataset 2: The dataset for stage classification consisted on 141 samples of soybean collected from Conesa and Gilbert (see table 1). Plants on 25.01.2017 were in vegetative stage, while on 13.01.2017 were in early reproductive stage (R1-2). Data gathered during March was from late reproductive stages (R5-6)[16].

Table 1. Experimental details

Crop	Site	C/MG	PDate	RD	PD	Soil	DCD	SCC
Maize	Santa Fe	-	06.01.2017	50	8	Sandy loam	16.02.2017	23
Soybean	Gilbert	5	16.11.2016	52	6,2	Vertisol	13.01.2017	18
							06.03.2017	24
Soybean	Conesa	5	11.11.2016	52	6,2	Vertic Argiudoll	25.01.2017	60
							01.03.2017	39
Wheat	Pergamino	SC	26.06.2017	20	200	Typic Argiudoll	15.09.2017	206
		LC	20.07.2017				15.09.2017	206

C/MG: Cycle/Maturity group, PDate: Planting dates, RD: Row distance [cm], PD: Plant density [plants/m^2], DCD: Data collection dates, SCC: Spectra curves collected, SC: short-cycle, LC: long-cycle

2.2 Measuring spectral reflectance curve

In this study, the spectral reflectance curve of the crop was obtained through a portable spectrometer (STS-NIR, Ocean Optics. Florida, USA). The spectrometer generates a vector of 1024 spectral intensities, which are measured from 632nm to 1125nm, with an optical resolution of 3nm and quantized to 14 bits.
In order to measure the spectral reflectance curve of crops, the spectra of a white reference material (WRM) is acquired after each crop measurement. Three measuring alternatives were initially proposed. For the first two alternatives the entrance slit was placed 3 cm above a leaf, one above the most recent fully developed leaf, and the other above the lower-most leaf. The third alternative had the entrance slit 50cm above the crop, perpendicular to the soil surface. After preliminary tests, we concluded that the third alternative provided better repeatability, adopting it as the method for all measurements in this study.
The integration time of the spectrometer was adjusted in order to avoid saturation of the WRM. The electrical noise was measured through the occlusion of the spectrometer's entrance slit. This electrical noise is then subtracted from the WRM and crop measurements. The crop reflectance was calculated as the ratio between energy reflected by the crop and energy incident on the crop (solar irradiation) [17] using equation 1, where $\boldsymbol{R}$ is the relative spectral reflectance for each wavelength, C represents the energy reflected by the crop, W represents the energy reflected by the WRM and B represents the electrical noise.

IV

$$\boldsymbol{R} = \frac{C - B}{W - B} \tag{1}$$

2.3 Spectral reflectance curve characterization

In figure 1 the wavelength ranges to be considered in this study are shown, from now on these ranges will be called "features". Figure 1 corresponds to a typical soybean spectral reflectance curve and is shown here only for the purpose to characterize the spectral curve, defining wavelength ranges and the corresponding features.

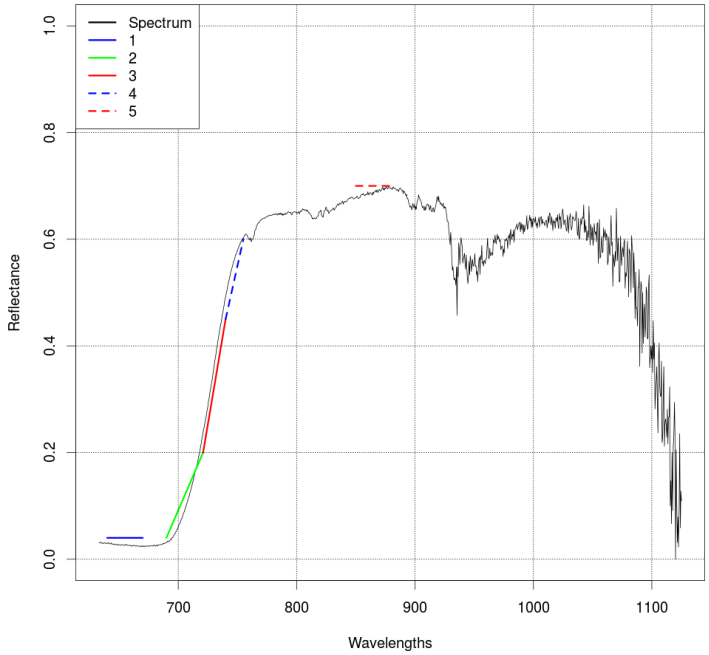


Fig. 1. Spectral reflectance curve of soybean (black solid line). Features: 1 solid blue line, 2 solid green line, 3 solid red line, 4 dashed blue line, 5 dashed red line.

Feature 1 (R) is obtained by computing the mean reflectance between 640 and 670 nm. The spectral band between 680 nm and 755 nm is evaluated through its first-order derivative. This interval is divided in three parts (680-721 nm, 721-740 nm and 740-755 nm) to evaluate the rate of change of the reflectance in these three wavelength ranges, resulting in Feature 2 (dA), Feature 3 (dB) and Feature 4 (dC), respectively. The derivatives are obtained by computing the slope for consecutive points, then the mean over each interval is calculated.

Feature 5 (NIR) is the mean for reflectance values between 850 and 880 nm, representing the near-infrared portion of the curve. Features 1 and 5 wavelengths were chosen to be equivalent to the wavelengths of the Operational Land Imager spectral bands from Landsat 8 [18]. Eight different combinations of features were considered for both experiments. In Table 2 these combinations are shown (S1 to S8).

To assess the performance of the features we propose in this work, we compare their classification results against classification derived from other datasets. For this purpose different established vegetation indices were selected, i.e. RVI [5], NDVI [6], SAVI [7] and MSAVI [8] (see equations 2, 3, 4 and 5 respectively), and classification performed. The coefficient L in equation 4 is the soil brightness correction factor which was set to 0.5 to accommodate the different conditions of the measured crops.

$$\text{RVI} = \frac{\text{NIR}}{\text{R}} \tag{2}$$

$$\text{NDVI} = \frac{\text{NIR-R}}{\text{NIR+R}} \tag{3}$$

$$\text{SAVI} = \frac{\text{NIR - R}}{\text{NIR+R+L}}(1+\text{L}) \tag{4}$$

$$\text{MSAVI} = \frac{2.\text{NIR} + 1 - \sqrt{(2.\text{NIR} + 1)^2 - 8(\text{NIR-R})}}{2} \tag{5}$$

2.4 Machine Learning methods and metrics

In this study the R software environment was used for implementation, training and validation of machine learning algorithms and statistical analysis. Overall classification accuracy was the metric chosen to evaluate model performance and 10-fold cross validation was used to provide a generalization measurement of this metric. The accuracy shown throughout this work is the average accuracy for the test data of all folds (fold collection). Feature importance is measured through the mean decrease Gini coefficient (MDG) [19]. Different feature sets for each machine learning algorithm were tested for statistical significance ($p<0.05$) with Student's t-test.

Artificial Neural Network Algorithm. This algorithm is based on layers of artificial neurons consisting commonly on an input layer, one or more hidden layers and an output layer [20]. ANN iterates over the training set updating weights until a stopping criterion is met. A single hidden layer neural network was proposed, the training process is done via the Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimization algorithm, a Quasi-Newton method that overcomes some of the limitations of plain gradient descent by seeking the second derivative (a

stationary point) of the objective function. The number of input neurons is set to be equal to the number of considered features. One-hot encoding is used for the output layer, each neuron representing a possible classification value. A grid search of 2, 3, 4 and 5 neurons in the hidden layer and weight decay values from 1.10^{-6} to 1.10^{-1} for species classification considering S8 was used to select the best set of parameters. The best performance was obtained with a hidden layer of 4 neurons and a weight decay factor of 1.10^{-3} which were chosen for all models.

Support Vector Machine Algorithm. This algorithm was developed based on statistical learning theory. It consists on obtaining an optimal hyperplane that separates classes and generalizes well [21]. The parameters to set include the kernel function and cost value. In this experiment a linear kernel was used with a cost value of 5 after a grid search from 1.10^{-1} to 1.10^{3}.

Random Forest This algorithm is based on training an ensemble of decision tree algorithms with bootstrapped samples and voting for the most probable class. The algorithm relies on using a random selection of features to split each node. Random Forest has the number of trees and number of features selected at each split. A large number of trees produce a limiting value of generalization error but do not produce overfit [22]. The minimum number of trees from which the accuracy stabilizes is 100, a number of 500 trees was chosen. During classification the subspace dimensionality for RF is rather small, in this work $\sqrt{p}$ was selected, where p is the total number of predictors.

3 Results

3.1 Species classification

In this section the dataset 1 is used to identify different species (soybean, maize and wheat) for all measurement dates, through three different MLA (as mentioned in section 2.4). The implemented procedure for defining a feature set consists in adding one by one different features from the field collected spectrum, as predictor variables for the MLAs. We started considering only two features (R + NIR, S1 in table 2) up to a collection of five features (S8 in table 2).
Accuracies and standard deviations for feature set 1 to 8 are presented in Table 2. Results are ordered from lower to higher accuracy. We can observe that the three tested MLAs follow the same accuracy sequence, from S1 to S8. Based on these results, S8 was selected as the feature set for both, species and stage classification.

Table 2. Species Classification Accuracy [%] and Standard Deviation (SD) for Support Vector Machine (SVMLIN), Random Forest (RF) and Artificial Neural Networks (ANN), using different feature sets.

Spectrum	Feature set	SVMLIN	SD	RF	SD	ANN	SD
S1	R + NIR	52.1	3.4	87.7	4.1	56.4	6.2
S2	R + NIR + dA	65.2	4.7	89.0	4.0	71.7	2.4
S3	R + NIR + dB	67.6	4.4	89.0	3.1	73.9	5.5
S4	R + NIR + dC	71.9	3.2	90.6	2.3	78.9	5.3
S5	R + NIR + dA + dB	71.7	3.9	90.7	2.8	82.2	3.0
S6	R + NIR + dB + dC	76.3	2.8	90.2	2.4	84.8	2.7
S7	R + NIR + dA + dC	86.3	2.1	92.3	2.3	91.2	2.8
S8	R + NIR + dA + dB + dC	87.2	2.6	92.8	2.3	92.7	4.1

In order to assess the S8 feature set performance for species classification a comparison with established vegetation indices is performed. Figure 2 shows the accuracy of each MLA (RF, ANN and SVMLIN) considering the S8 feature set and the established vegetation indices described in section 2.3 (NDVI, RVI, SAVI, and MSAVI). RF was the best-performing method for S8 and the 4 vegetation indices. S8 attained the highest accuracies for the three tested MLAs, reaching an accuracy of 92% for RF species classification, followed by the RVI index (89%). We checked that the difference between the RF-S8 and RF-RVI results are statistically significant (p=0.025). Besides, to describe the performance of the RF-S8 versus RF-RVI regarding actual and predicted classifications for each species, we present the confusion matrices for RF-S8 and RF-RVI (Table 3 and Table 4, respectively). Soybean and maize reached accuracies very close to 100% for both RF-S8 and RF-RVI. In the case of wheat RF-S8 reaches a 91% of accuracy, while RF-RVI attains a 77%, with a class error of 23%.

VIII

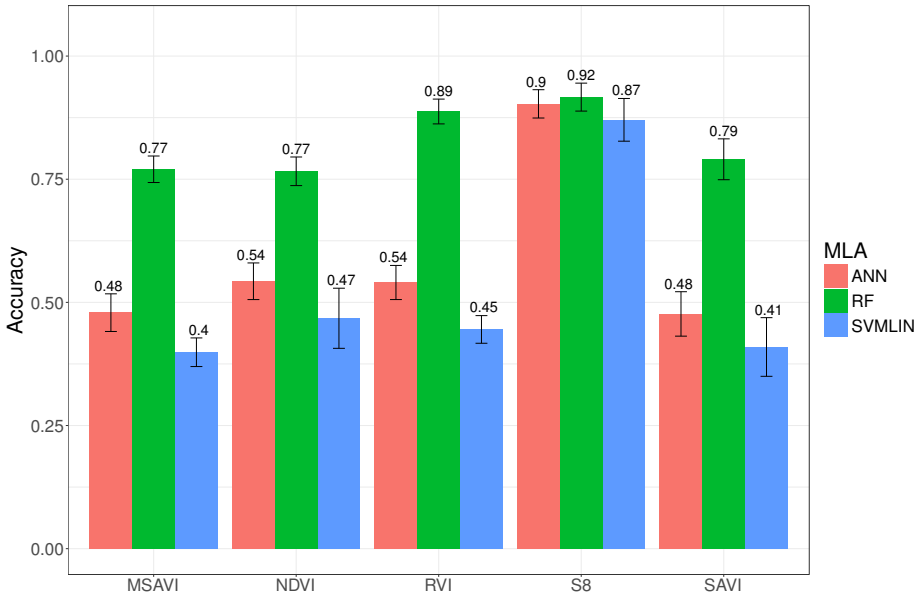


Fig. 2. Species classification accuracy for S8, and established vegetation indices (MSAVI, NDVI, RVI, SAVI), using ANN, RF and SVMLIN

Table 3. Confusion matrix for S8. Random Forest

True \ Predicted	Soybean	Maize	Wheat
Soybean	0.99	0.00	0.01
Maize	0.00	1.00	0.00
Wheat	0.09	0.00	0.91

Table 4. Confusion matrix for RVI. Random Forest

True \ Predicted	Soybean	Maize	Wheat
Soybean	0.98	0.00	0.02
Maize	0.00	0.98	0.02
Wheat	0.19	0.04	0.77

New index proposal. A potential advantage of using a field spectrometer is that it allows for the definition of new indices, based on the ability to analyse

the spectral reflectance curve with a wavelength resolution of about 1nm. In this section we propose a new index as the ratio between the two most important features in S8.
To measure how each feature contributes to S8 we performed a variable selection using Mean Decrease Gini coefficients (MDG). In Table 5 MDG for S8 is shown. Feature importance ranking is then used to define a new index, dCdA, which consists in the ratio between dC (MDG=179.80) and dA (MDG=235.43), as shown in equation 6.

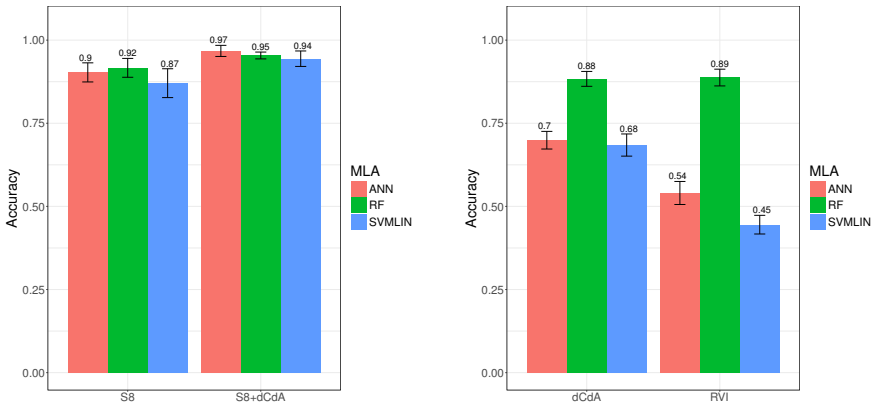
$$\mathrm{dCdA} = \frac{\mathrm{dC}}{\mathrm{dA}} \tag{6}$$

Table 5. Mean decrease Gini coefficient for S8. Random Forest method

Feature	MDG
dA	235.43
dC	179.80
Red	148.57
NIR	147.70
dB	111.83

In figure 3 results of species classification with dCdA are shown. In figure 3a the performance of S8 is compared to S8+dCdA. We observed that S8+dCdA for ANN increases accuracy up to 97%. Statistical analysis of S8 and S8+dCdA, for all the tested MLAs showed significant differences (p-SVMLIN=$2.5.10^{-4}$, p-ANN=$8.3.10^{-6}$, p-RF=0.0010).
To evaluate the contribution of dCdA, in Table 6 we show MDG coefficients for each feature making up S8+dCdA. Features ranking results indicate that dCdA is the most relevant component of S8+dCdA.
We also tested dCdA as a single index for species classification. For RF, dCdA performs as the second best index after RVI, but results between both indices are not statistically different ($p > 0.05$). Moreover, accuracies for dCdA are better than for any other vegetation index for ANN and SVMLIN. These results are shown in figure 3b.

X



(a) Accuracy in species prediction for S8 and S8+dCdA

(b) Accuracy in species prediction for dCdA as a single index and RVI.

Fig. 3. Performance of dCdA

Table 6. Mean decrease Gini coefficient for S8 + dCdA. Random Forest method

Feature	MDG
dCdA	313.21
dA	136.43
R	118.28
dC	91.83
NIR	87.82
dB	75.83

4 Stage classification

In this section the dataset 2 is used to classify different soybean growth stages. This dataset comprises samples of three different stages of soybean, vegetative, early reproductive and late reproductive stages.
The same three MLAs as in section 2.4 are tested here, and the spectrum feature set is S8. The accuracies attained by each MLA for S8 and the selected vegetation indices are presented in figure 4. Considering RF, S8 showed significant differences to the best performing index NDVI (p=0.0008). Regarding S8 results, ANN had the highest accuracy. RF was the best performing algorithm for all vegetative indices, with RVI as the best performing vegetation index.

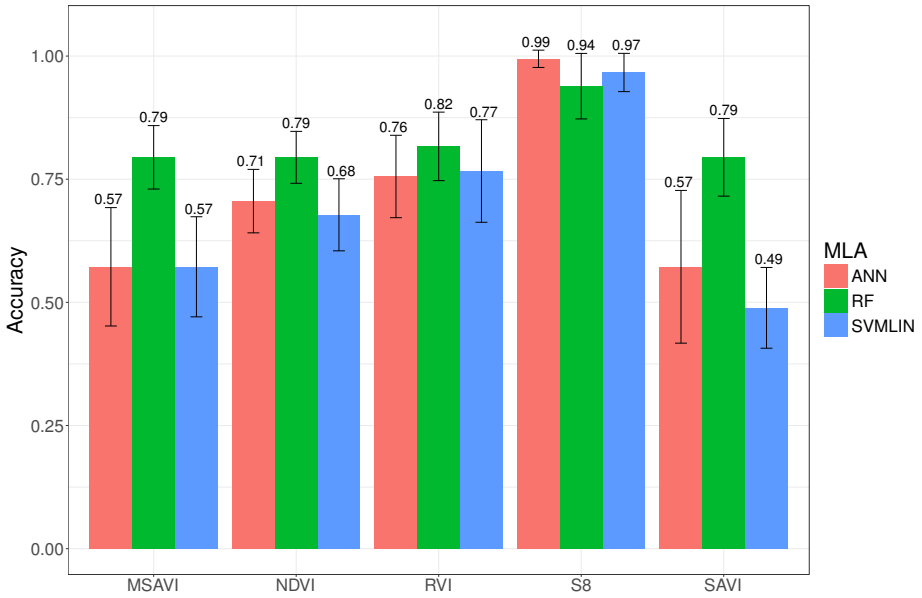
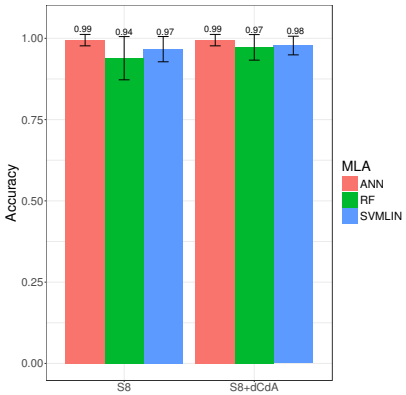
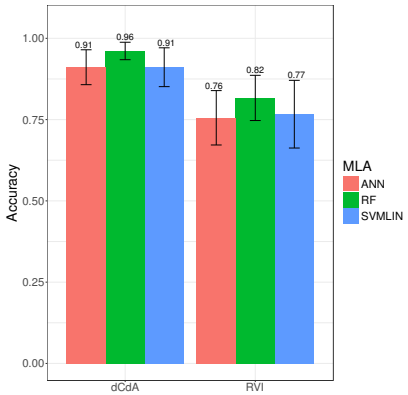


Fig. 4. Soybean growth stages classification accuracy for S8, and established vegetation indices (MSAVI, NDVI, RVI, SAVI), using ANN, RF and SVMLIN



(a) Accuracy in soybean growth stages classification for S8 and S8+dCdA



(b) Accuracy in soybean growth stages classification for dCdA and RVI

Fig. 5. Predicted accuracy of soybean growth stages considering the proposed index dCdA

The index proposed in the previous section is also assessed for stage classification. Since S8 already provided a very high accuracy, the addition of dCdA did not provide statistically different classification results (figure 5), $p > 0.05$ was obtained for the three MLAs. Performance of dCdA as a single index was compared to RVI, the best performing vegetation index (figure 5b). For RF, dCdA reaches a higher stage classification accuracy than RVI (96% accuracy for dCdA, 82% for RVI), with stylistically significant difference results ($p= 8.1.10^{-6}$). All three MLAs analysed followed the same behaviour.
In Table 7 we show the importance of dCdA as a feature, when we consider S8+dCdA. dCdA is ranked as the most important feature in the considered feature set.

Table 7. Mean Decrease Gini for S8 + dCdA. Random Forest method

Feature	MDG
dCdA	48.41
dC	26.19
R	19.33
dB	12.08
NIR	6.69
dA	6.66

5 Discussion and Conclusions

This work focuses on the assessment of a proposed feature set based on NIR spectral reflectance curves obtained from field experiments, for the classification of species (maize, soybean and wheat) and soybean growth stages (vegetative, early and late reproductive phases). Classification is done by applying three different MLAs, i.e. RF, ANN and SVMLIN.

Mean values (R, NIR), and first derivatives (dA, dB, dC) of the reflectance data are considered (we call them features) for certain ranges of wavelengths. These features are then grouped in sets, from S1 to S8 (Table 2), and classification accuracies are ranked. Results are presented for S8 since it leads to much better classification performance.

For species classification, RF is the best-performing MLA for S8 and for all the tested vegetation indices. The highest accuracy is reached by RF-S8 (92%), while the best performing vegetation index is the RF-RVI index (89%). Recently (June 2018), Gao *et al.*[14] used RF to classify among maize and three weed species using laboratory hyperspectral data. The authors achieved accuracies between 69% and 100% which are similar to the performance we obtained for RF-S8. Fletcher *et al.* [23] also used RF on spectral reflectance curves in the range of 350-2500 nm. From this range they proposed 16 features to identify between two weed species and soybean, obtaining overall accuracies between

93% and 100%. Fletcher *et al.* conducted greenhouse experiments and collected data from potted plants with a spectroradiometer. Our work achieved similar accuracies on less controlled conditions.

In this work, for stage classification the best result is obtained for ANN-S8 (99%), while the best performing vegetation index is RF-RVI with 82% of accuracy. It must be noted that differences between MLAs for S8 are not relevant, while differences between MLAs for any vegetation index may scale up to 30%. This fact shows that a collection of field spectral input data may be more representative of the classes to be grouped than information given by single vegetation indices. To the best of the author's knowledge no recent results have been published for soybean stage classification using spectral data and MLAs.

In addition we propose a new index, dCdA, as the ratio between the two best ranked features in S8. Accuracy for species classification obtained including this index as an additional feature to S8 were higher than those obtained without it. Moreover, dCdA as a single index, performs similarly to the best performing single vegetation index for species classification. For stage classification, dCdA outperforms every single vegetation index.

The use of spectrometry allows the selection of a feature set from the spectral reflectance curve for the classification of species and stage through machine learning algorithms. The achieved performance of the proposed feature set is similar or better when compared to established vegetation indices.

Based on both recent research (Burkart *et al* [24]) and the results presented in this work, we aim to extend this field spectroscopy technique, as a remote sensing analysis, from plot-level to farm-level by adding spectrometers to be carried onboard unmanned aerial vehicles (UAV). Compared to manual acquisition, UAVs will allow high throughput crop data collection, over larger areas.

References

[1] Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P., Hawkesford, M.J.: Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. Functional Plant Biology 44(1) (2017) 143–153

[2] Sankaran, S., Khot, L.R., Espinoza, C.Z., Jarolmasjed, S., Sathuvalli, V.R., Vandemark, G.J., Miklas, P.N., Carter, A.H., Pumphrey, M.O., Knowles, R.R., Pavek, M.J.: Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. European Journal of Agronomy 70 (oct 2015) 112–123

[3] Al-Yaari, A., Wigneron, J.P., Ducharne, A., Kerr, Y., de Rosnay, P., de Jeu, R., Govind, A., Al Bitar, A., Albergel, C., Muñoz-Sabater, J., Richaume, P., Mialon, A.: Global-scale evaluation of two satellite-based passive microwave soil moisture datasets (SMOS and AMSR-E) with respect to Land Data Assimilation System estimates. Remote Sensing of Environment 149 (jun 2014) 181–195

[4] Nijland, W., de Jong, R., de Jong, S.M., Wulder, M.A., Bater, C.W., Coops, N.C.: Monitoring plant condition and phenology using infrared sensitive consumer grade digital cameras. Agricultural and Forest Meteorology 184 (jan 2014) 98–106

[5] Jordan, C.F.: Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. Ecology 50(4) (jul 1969) 663–666

[6] Tucker, C.J.: Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment 8(2) (may 1979) 127–150
[7] Huete, A.R.: A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 25(3) (aug 1988) 295–309
[8] Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S.: A modified soil adjusted vegetation index. Remote Sensing of Environment 48(2) (may 1994) 119–126
[9] Lorenzen, B., Jensen, A.: Changes in leaf spectral properties induced in barley by cereal powdery mildew. Remote Sensing of Environment 27(2) (feb 1989) 201–209
[10] Antunes, M., Assad, E., Batista, G.: Varicacao das medidas espectrais tomadas com espectroradiometro ao longo do ciclo de crescimento da soja (Glycine max (L.) Merrill).. Anais do VII SBSR 3 (1993) 1–9
[11] Broge, N.H., Mortensen, J.V.: Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data. Remote Sensing of Environment 81(1) (jul 2002) 45–57
[12] Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R.L., Mouazen, A.M.: Wheat yield prediction using machine learning and advanced sensing techniques. Computers and Electronics in Agriculture 121 (feb 2016) 57–65
[13] Yuan, H., Yang, G., Li, C., Wang, Y., Liu, J., Yu, H., Feng, H., Xu, B., Zhao, X., Yang, X.: Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models. Remote Sensing 9(4) (mar 2017) 309
[14] Gao, J., Nuyttens, D., Lootens, P., He, Y., Pieters, J.G.: Recognising weeds in a maize crop using a random forest machine-learning algorithm and near-infrared snapshot mosaic hyperspectral imagery. Biosystems Engineering 170 (jun 2018) 39–50
[15] Efron, B., Tibshirani, R.: An introduction to the bootstrap. Chapman & Hall (1994)
[16] Fehr, W.R., Caviness, C.E., Burmood, D.T., Pennington, J.S.: Stage of Development Descriptions for Soybeans, Glycine Max (L.) Merrill1. Crop Science 11(6) (1971) 929
[17] Hansen, P.M., Schjoerring, J.K.: Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sensing of Environment 86(4) (2003) 542–553
[18] Barsi, J.A., Lee, K., Kvaran, G., Markham, B.L., Pedelty, J.A.: The spectral response of the Landsat-8 operational land imager. Remote Sensing 6(10) (oct 2014) 10232–10251
[19] Louppe, G., Wehenkel, L., Sutera, A., Geurts, P.: Understanding variable importances in forests of randomized trees. In: Advances in Neural Information Processing Systems 26. Curran Associates, Inc. (2013) 431–439
[20] Hassoun, M.H.: Fundamentals of Artificial Neural Networks. MIT Press (1995)
[21] Cortes, C., Vapnik, V.: Support Vector Networks. Machine Learning 20(3) (sep 1995) 273–297
[22] Breiman, L.: Random forests. Machine Learning 45(1) (2001) 5–32
[23] Fletcher, R.S., Reddy, K.N.: Random forest and leaf multispectral reflectance data to differentiate three soybean varieties from two pigweeds. Computers and Electronics in Agriculture 128 (oct 2016) 199–206
[24] Burkart, A., Cogliati, S., Schickling, A., Rascher, U.: A novel UAV-Based ultralight weight spectrometer for field spectroscopy. IEEE Sensors Journal 14(1) (jan 2014) 62–67