Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Identifying relevant hyperspectral bands using Boruta: a temporal analysis of water hyacinth biocontrol

Na’eem Hoosen Agjee
Riyad Ismail
Onisimo Mutanga

SPIE.

Na’eem Hoosen Agjee, Riyad Ismail, Onisimo Mutanga, “Identifying relevant hyperspectral bands using Boruta: a temporal analysis of water hyacinth biocontrol,” *J. Appl. Remote Sens.* **10**(4), 042002 (2016), doi: 10.1117/1.JRS.10.042002.

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

Identifying relevant hyperspectral bands using Boruta: a temporal analysis of water hyacinth biocontrol

Na'eem Hoosen Agjee,* Riyad Ismail, and Onisimo Mutanga
University of KwaZulu-Natal, School of Environmental Sciences, King George V Avenue,
Durban 4041, South Africa

Abstract. Water hyacinth plants (*Eichhornia crassipes*) are threatening freshwater ecosystems throughout Africa. The *Neochetina* spp. weevils are seen as an effective solution that can combat the proliferation of the invasive alien plant. We aimed to determine if multitemporal hyperspectral data could be utilized to detect the efficacy of the biocontrol agent. The random forest (RF) algorithm was used to classify variable infestation levels for 6 weeks using: (1) all the hyperspectral bands, (2) bands selected by the recursive feature elimination (RFE) algorithm, and (3) bands selected by the Boruta algorithm. Results showed that the RF model using all the bands successfully produced low-classification errors (12.50% to 32.29%) for all 6 weeks. However, the RF model using Boruta selected bands produced lower classification errors (8.33% to 15.62%) than the RF model using all the bands or bands selected by the RFE algorithm (11.25% to 21.25%) for all 6 weeks, highlighting the utility of Boruta as an all relevant band selection algorithm. All relevant bands selected by Boruta included: 352, 754, 770, 771, 775, 781, 782, 783, 786, and 789 nm. It was concluded that RF coupled with Boruta band-selection algorithm can be utilized to undertake multitemporal monitoring of variable infestation levels on water hyacinth plants. © *2016 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.10.042002]

Keywords: water hyacinth; biocontrol; hyperspectral; random forest; Boruta; recursive feature elimination.

Paper 16122SS received Feb. 14, 2016; accepted for publication Apr. 18, 2016; published online Jun. 1, 2016.

1 Introduction

Water hyacinth [*Eichhornia crassipes* (Mart.) Solms-Laubach (Pontederiaceae)] is an invasive alien plant (IAP) species that is threatening the integrity of freshwater ecosystems throughout South Africa.[1] Native to Brazil, water hyacinth is a free-floating vascular plant that forms extensive monocultures on the surface of freshwater bodies.[2] This noxious weed is able to proliferate profusely and propagate rapidly, thereby successfully expanding the range of its distribution into non-native areas.[3,4] In South Africa, the success of the invader lies in the absence of natural enemies and the prevalence of nutrient-enriched waters that are suitable for their development.[5] Once established in non-native areas, the water hyacinth poses many negative ecological impacts, including the reduction of native biodiversity, impairment of water quality, and reduction in catchment water yields.[6,7] Therefore, it is necessary to control water hyacinth growth and proliferation in order to maintain the vital ecosystem services that humans depend on.

The use of biocontrol agents is seen as a sustainable solution in managing and controlling the spread of water hyacinth plants.[8] In South Africa, *Neochetina eichhorniae* and *Neochetina bruchi* weevils are the primary control agents used to combat the spread of water hyacinth plants.[9,10] Herbivory by *Neochetina* spp. weevils inflicts both morphological and physiological damage to water hyacinth plants. Adult weevils feed by forming rectangular scars on the surface of the leaf, thus exposing the leaf to desiccation.[11,12] Over time, extensive feeding damage causes the petioles of the plant to become thinner and "spindly" along with reduction in plant growth

*Address all correspondence to: Na'eem Hoosen Agjee, E-mail: agjeen@gmail.com

1931-3195/2016/$25.00 © 2016 SPIE

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

and reproduction.[13,14] Eventually, after severe damage, plants become completely water-logged, gradually sink to the bottom of the water body, and decompose. However, the efficacy of *Neochetina* spp. weevils is reported to be variable.[15] The weevil's reproductive ability, host quality, cold temperatures, dynamically changing water hyacinth growth, and interferences from herbicide application, all influence the weevils' behavior as an effective biocontrol agent.[15] Consequently, it is imperative to detect and monitor weevil infestation levels over time in order to gain valuable insight into the efficacy of the weevil.

However, regular monitoring of variable infestation levels over large spatial scales poses a challenge. Field-based surveys are unsustainable as a long-term data-acquisition solution, as they are labor intensive, expensive, and time consuming to conduct.[16,17] Numerous studies have advocated the utility of multitemporal remote sensing paradigm to monitoring vegetation.[18,19] Consequently, laboratory-based hyperspectral studies (i.e., spectroscopic) are useful as a first step toward investigating the potential of remote sensing to detect variable infestation levels and its utility for multitemporal monitoring. Similar studies have investigated the potential of detecting insect defoliation on IAP species; however, the application of detecting *Neochetina* spp. weevil damage on water hyacinth plants over time is unique. In addition, most studies have primarily focused on detecting biocontrol damage on terrestrial IAP species.[20,21,22,23] This study extends the current body of research by attempting to detect biocontrol damage on an aquatic plant. Detecting biocontrol damage on aquatic plants is complex in nature. The reflectance spectra of aquatic canopies are combined with the reflectance spectra of background surface water and atmospheric vapor. Hyperspectral data have high spectral resolution, offering the potential to discriminate variable infestation levels within aquatic/semiaquatic environments.[20] Importantly, hyperspectral data can be collected synoptically and nondestructively, ensuring that plant characteristics or variable infestation levels can be captured easily and repeatedly over the progression of the stressor. This enables the identification of key spectral regions and wavebands that could potentially discriminate variable infestation levels over the duration of the stressor.[24] Therefore, the ability to detect variable infestation levels using different wavebands at different stages of the weevil infestation is critical in gaining insight into the efficacy of the weevil.

Analyzing hyperspectral data to extract meaningful information poses various challenges. The high dimensionality ($n < p$) of the dataset: (1) decreases the classifiers ability to generalize accurately and (2) hampers the identification of important bands due to the strong correlation between adjacent bands.[25] One approach to mitigating these challenges is to employ band selection as a precursor to the classification process, with the aim of improving classification accuracies. Through band selection, an optimal subset of bands from the original dataset is selected with the redundant bands being discarded.[24,26] Importantly, there is no loss of information during the band selection process, thus warranting its implementation.[27] Therefore, band selection assists in producing a simplistic and interpretable model. Furthermore, band selection also assists in avoiding over-fitting of the data and improves generalization of unseen observations.[28] An investigation into the utility of existing band selection approaches is therefore necessary to identify wavebands that could potentially discriminate variable weevil infestation levels.

Random forest (RF) is an ensemble classifier that constructs numerous classification trees with the final classification of an observation based on the majority vote of all the trees.[29,30] RF is a robust algorithm that is capable of consistently achieving high classification accuracies.[30–32] In addition, RF also provides two internal measures of variable importance during the classification process, known as the mean decrease in accuracy and the Gini importance.[29] The mean decrease in accuracy variable importance measure is the most reliable and widely used importance measure in most hyperspectral studies.[30,32] However, since RF does not primarily undertake feature selection, most studies have implemented either a forward, backward, or best-first search feature selection approach to reduce data dimensionality and improve classification accuracies.[25,33,34] These approaches select an optimal subset of bands or bands with a high RF variable importance measures, thus producing low-misclassification rates. For example, Ismail and Mutanga[35] explored discriminating the early stages of *Sirex noctilio* infestations using RF coupled with a backward variable selection algorithm. Results showed that the RF classifier using five bands selected by the backward variable selection technique produced the lowest misclassification error rate of 6.14%. Similarly, Dye et al.[36] examined the utility of RF to predict

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

Pinus patula age and compared the forward and backward variable selection methods. Results showed that the forward variable selection method obtained the highest model accuracy ($R^2 = 0.60$) while utilizing only a small subset of nonredundant bands ($n = 9$). This study extends the current body of research by comparing two different band selection approaches and their effect on RF classification accuracies. In this paper, we compare the recursive feature elimination (RFE) algorithm which adopts a minimal optimal approach with the Boruta algorithm which adopts an all relevant feature selection approach.

Boruta is an all relevant embedded feature selection algorithm, that uses the RF algorithm to identify all strongly and weakly relevant bands in an effort to reduce data dimensionality and improve classification accuracies.[37] Boruta is unique, as it uses a statistical measure (Z-score) to deem bands important or unimportant, thus providing a statistical criterion for band selection. Therefore, bands selected by Boruta will not necessarily have a high RF mean decrease in accuracy. This differentiates Boruta from other feature-selection algorithms such as the RFE algorithm,[38] that selects a ranked subset of bands based on the mean decrease in accuracy as determined by RF. A limited number of studies have implemented Boruta for feature selection with the aim of improving hyperspectral classification accuracies. Poona and Ismail[24] used Boruta selected spectroscopic bands for the asymptomatic detection of *Fusarium circinatum* stress. Results showed a decrease in error rates, with the out-of-bag (OOB) error decreasing from 16.67% to 14.00% and the cross-validation (CV) error decreasing from 18.67% to 16.00%, when using RF with the Boruta algorithm. These results were achieved using just 107 wavebands of an original 1769 wavebands, which is equivalent to a 93.95% reduction in dimensionality.

In light of the above, this study aims to determine if multitemporal hyperspectral data can be utilized to detect the efficacy of biocontrol agents. More specifically, the objectives of this study were: (1) to compare bands selected by the RFE and Boruta algorithm, (2) to understand how bands selected by both feature selection algorithms varied over the progression of the weevil infestation, (3) to determine if band selection improves the classification accuracy of variable infestation levels, and (4) to determine if these improved models provide insight into the efficacy of the weevil as a water hyacinth biocontrol measure.

2 Materials and Methods

2.1 *Experimental Procedure*

Healthy water hyacinth (*Eichhornia crassipes*) plants ($n = 180$) were collected from the Amanzimtoti River, KwaZulu-Natal Province, South Africa (30 deg 03′ 29.44″ S; 30 deg 52′ 38.53″ E) and were transported to a laboratory at the University KwaZulu-Natal. Water hyacinth plants that were collected were of the same size (i.e., phenostage five)[13] as well as free of any biocontrol agents or biocontrol damage. At the laboratory, individual circular plastic containers (55 cm in diameter each) were filled with 20 L of water. Thereafter, nitrogen (potassium nitrate: 7.5 mg N L^{-1}) and phosphorus (dihydrogen orthophosphate: 1.37 mg P L^{-1}) were added to each container to simulate conditions found within highly eutrophic environments.[39] Commercial iron chelate (13% Fe) was also added to each container at a concentration of 11.2 mg Fe L^{-1}.[1,40] The nutrient medium was replaced on a weekly basis to maintain a constant nutrient concentration for plant growth and development. Fifteen water hyacinth plants were placed in each container creating a dense mat within each container. Water hyacinth plants were then acclimated to the surrounding environment for 1 week prior to weevil exposure.[10]

After the acclimation period, each water hyacinth plant was cleaned of all debris, dead leaves, dead petioles, and daughter plants were removed to maintain the original stocking density. Four *Neochetina* spp. infestation levels: zero adult male weevils per plant, two adult male weevils per plant, four adult male weevils per plant, and six adult male weevils per plant were considered in order to study the efficacy of variable infestation levels on plant spectral characteristics.[6] Four damage classes (no damage, low damage, medium damage, and high damage) representing the four *Neochetina* spp. infestation levels or variable infestation levels were then used for all subsequent classification procedures. The experiment was setup in a complete random design

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

with one control and the three treatments, each replicated three times. After the weevils were introduced into each container, each container was covered with a mesh (3 m × 1.5 m; 2 mm × 2 mm mesh cell size) to prevent weevils from leaving the containers. The number of male weevils in each container was maintained by replacing dead weevils on a weekly basis.[41] Water hyacinth plants were exposed to 1 week of weevil herbivory prior to plants being sampled for canopy reflectance spectra.

2.2 *Canopy Reflectance Measurements*

A FieldSpec® 3 spectroradiometer[42] was used to collect reflectance data of weevil herbivory. The ASD is a portable spectrometer that uses a fiber optic cable for reflectance measurements and a personal computer for data logging. The spectrometer has a spectral range of 350 to 2500 nm with a sampling interval of 1.4 nm in the 350 to 1000 nm range and 2 nm in the 1000 to 2500 nm range. Reflectance measurements were taken at an ambient air temperature of 21°C. All reflectance measurements were taken within a black box to account for any background reflectance. The fiber optic cable with a 10-deg field of view was pointed 0.5 m above each container with one 50-W halogen lamp across the container providing the only illumination. The spectrometer was calibrated by measuring a "white reference" reading using a spectralon panel before sample reflectance measurements were taken.

To measure canopy reflectance, the mesh from each container was removed and the container placed on the target platform. Each container was rotated 45-deg eight times with reflectance measurements being captured at the center of each container. Four reflectance measurements were captured at each rotation of the container (per container $n = 32$, per treatment $n = 96$). Each spectrum was inspected for quality purposes and was discarded if the spectrum was problematic. After reflectance measurements were captured, the mesh on each container was replaced. The first set of reflectance measurements were taken after 1 week of infestation, thereafter, on a weekly basis over a period of 5 weeks. Prior to analysis, reflectance spectra captured at each rotation were averaged and atmospheric water absorption bands (1350 to 1450 nm, 1773 to 2020 nm, 2400 to 2500 nm) were removed.

2.3 *Statistical Analysis*

Initially, the RF classifier was implemented to classify variable infestation levels using the entire hyperspectral dataset. Thereafter, the Boruta band selection algorithm was implemented to identify all relevant bands that discriminate variable infestation levels. Additionally, for comparative purposes, the RFE algorithm was also implemented to select a subset of bands based on the ranked order of bands as determined by the RF mean decrease in accuracy. Lastly, the RF classifier was implemented using the Boruta selected bands and bands selected by the RFE algorithm to determine if band selection improves classification accuracies. The RF classification algorithm, Boruta band selection algorithm, and RFE algorithm are discussed in greater detail in the sections below.

2.3.1 *Random forest algorithm*

RF is an ensemble classifier that fits many classification trees to a dataset and then combines the predictions from all the trees to classify the input data. Initially, bagging is performed to create new training datasets (bootstrap samples) by randomly resampling the original training dataset with replacement. For each bootstrap sample generated, an OOB sample is created, which is one-third of the original dataset that does not occur in the bootstrap sample. Subsequently, an individual classification tree is grown for each bootstrap sample to a maximum size without pruning. To improve the prediction accuracy and reduce the correlation between individual classification trees, RF introduces randomness to the classification process. Randomness is introduced by selecting a random sample of training sets for growing each classification tree and by randomly selecting a subset of bands at each node to determine the best split. The final classification of a given observation is determined by applying the majority rule over the votes of the individual trees.[29]

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

RF inherently provides two internal measures of variable importance for each predictor variable during the classification process. In this study, the mean decrease in accuracy was used. Variable importance is calculated by randomly permuting each band in the OOB sample and passing the new OOB sample down each tree. The difference in OOB error between using the original OOB sample and permuted OOB sample is calculated and averaged over all the trees. The loss in performance can be used as a ranking index to determine the importance of each band.

The mean decrease in accuracy of feature X^j is defined as

$$\mathrm{VI}(X^j) = \frac{1}{\mathrm{ntree}} \sum_t (\mathrm{errOOB}_{t^j} - \mathrm{errOOB}_t), \tag{1}$$

where the summation is over all trees t of the RF and ntree is the number of trees, and OOB_t is the features not included in the bootstrap sample used to construct t, errOOB_t is the misclassification error rate of a single tree t on the sample OOB_t, OOB_{t^j} is the permuted sample from randomly permuting the values X^j in $\mathrm{OOB_t}$ and errOOB_{t^j} is the error of the predictor t on the perturbed sample.

The RF algorithm is defined by two hyperparameters: the number of trees to be grown (ntree) and the number of predictors (mtry) to be sampled at each node. In this study, hyperparameter optimization was conducted; however, the default ntree and mtry hyperparameters as determined by the RF algorithm were used for classification as they presented the best results. Studies have implemented RF using the default ntree and mtry values showed that the sensitivity of the user defined parameters is minimal and the default values are often a good choice.[31,35] The default mtry value is calculated as the square root of the number predictors in the dataset.[43] Since the "default" mtry value is calculated, it can be used for very high-dimensional datasets. The RF software library[43] developed in the *R* statistics package version 2.5.1[44] was used for all analysis.

2.3.2 *Recursive feature elimination algorithm*

The RFE algorithm identifies an optimal subset of bands by ranking variables based on the RF variable importance measure (i.e., mean decrease in accuracy). The algorithm iteratively constructs numerous RF models. At each iteration, variables with the smallest variable importance (as determined by the mean decrease in accuracy) are discarded. The subset of variables that yields the smallest OOB error rate is then selected. The varSelRF software library[45] developed in the *R* statistics package version 2.5.1[44] was used for band selection.

2.3.3 *The Boruta algorithm*

Boruta is an all relevant feature selection algorithm, i.e., embedded with the RF classification algorithm and uses calculated Z-scores as a measure of band importance. Initially, the algorithm replicates each band thus creating "shadow" bands by randomly permuting the observations of replicated bands. Thereafter, an RF classification is performed, the importance of all the bands computed (calculated as the Z-scores) and the maximum Z-score among the shadow bands determined. The frequency that the importance of a band is higher than the maximum Z-score among the shadow bands is counted. A band is deemed "important" when the frequency is significantly higher than the expected value whereas if a band is significantly lower, it is deemed "unimportant" and removed. In this study, as recommended by Poona and Ismail,[24] the ntree hyperparameter of the Boruta algorithm was optimized by varying the ntree value from 500 to 2500 by 500. The ntree value that produced the lowest error was then utilized for subsequent analysis. The Boruta software library[46] developed in the *R* statistics package version 2.5.1[44] was used for band selection.

2.3.4 *Accuracy assessment*

The accuracy of the classifier was evaluated by computing the OOB estimation of error for each week. The OOB error rate is defined as

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

$$\mathrm{errOOB} = \left(\frac{1}{\mathrm{ntree}}\right)\sum_{i=1}^{\mathrm{ntree}}[y_i - g_{\mathrm{OOB}}(X_i)]^2, \tag{2}$$

where y_i the i'th element of the training dataset (X), g_{OOB} is the aggregated prediction based on the random trees, and (X_i) is the bootstrap sample.

Each classification tree, i.e., grown on a bootstrap sample is used to classify the observations in its corresponding OOB sample.[43] The OOB estimation of error was then calculated as the misclassified proportion of the OOB sample.[47–50] To further test the validity of the classifiers, a 10-fold CV was performed. The sample was split into 10 stratified subsamples. Each subsample was successively used as a validation dataset, while the remaining subsamples were used as training datasets. A model is then fitted to the training dataset and the error computed using the test dataset.[51]

3 Results

3.1 *Visual Observations of Neochetina spp. Weevil Infestations*

Figure 1 graphically shows the extent of biocontrol damage on water hyacinth plants infested with varying densities of *Neochetina* spp. weevils over a period of 6 weeks. Plants infested with two weevils were observed as vigorous and healthy even after 4 weeks of infestation (Fig. 1). However, water hyacinth plants infested with four and six weevils exhibited moderate and severe damage after 3 weeks of infestation (Fig. 1). Mortality was visible in plants infested with four and six weevils per plant after an infestation period of 4 weeks. In general, water hyacinth plants infested with six weevils exhibited a reduction in the production of new leaves, which was followed by a decrease in plant size. Over time, water hyacinth plants lost their vigor and buoyancy; they became limp and brown in color. The bases of the petioles were severely eaten, resulting in leaves wilting through desiccation. Over time, leaves became brown losing their photosynthetic ability and eventually fall of the petiole.

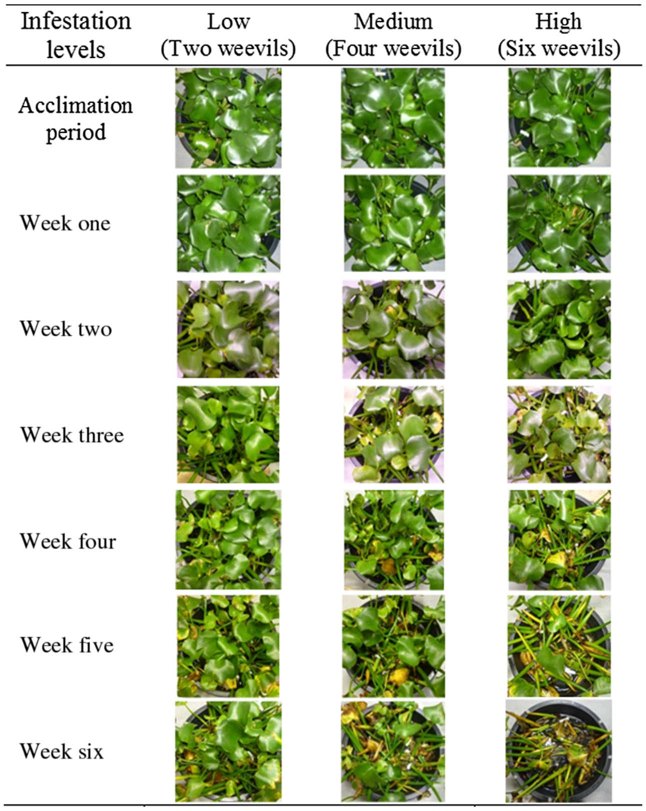

Fig. 1 Damage on water hyacinth (*E. crassipes*) plants infested with varying levels of *Neochetina* spp. weevils over 6 weeks.

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

3.2 *Spectral Reflectance of Neochetina spp. Weevil Infestations*

The spectral reflectance of each treatment for 6 weeks is shown in Fig. 2. Generally, the control and low treatments (i.e., two weevils) exhibit a higher spectral reflectance than the medium (i.e., four weevils) and high (i.e., six weevils) treatments. However, for weeks 4 and 5 the spectral reflectance of the low and medium treatments was higher than the control and high infestation levels.

3.3 *Random Forest Classification Using all Hyperspectral Bands*

RF error rates using all the hyperspectral bands ($n = 1808$) is presented in Table 1. Results show that RF error rates decreased from week 1 (OOB error= 32.29%; CV error= 34.38%) to week 5 (OOB error= 12.50%; CV error= 15.62%). The highest overall classification accuracy using all the bands was achieved for week 5 (OOB error= 12.50%; CV error= 15.62%). These results indicate that the optimum time to discriminate variable infestation levels would be after 5 weeks of infestation.

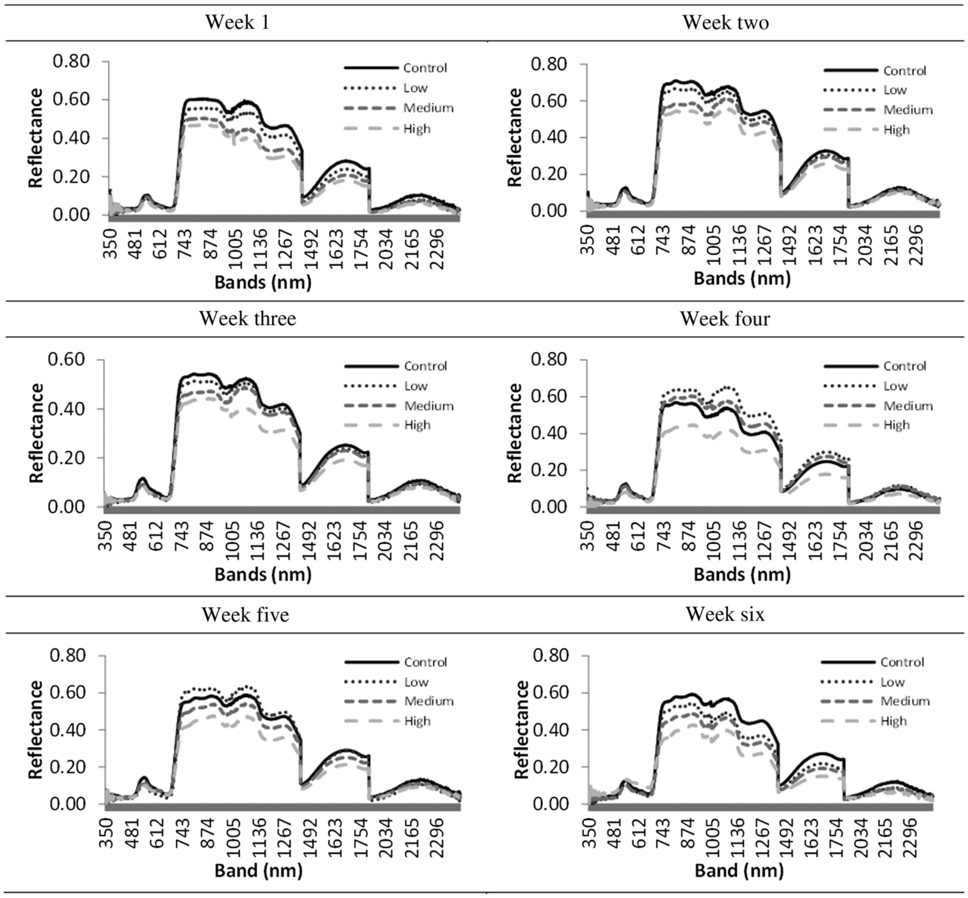

Fig. 2 Spectral reflectance for each treatment for each week under laboratory conditions. Spectral features between 1350 to 1450 nm, 1773 to 2020 nm, and 2400 to 2500 nm were removed due to excessive noise.

Table 1 RF classification results using all hyperspectral bands ($n = 1808$) for all 6 weeks.

Week	1	2	3	4	5	6
OOB error (%)	32.29	29.17	27.08	19.79	**12.50**	19.79
CV error (%)	34.38	32.29	26.04	22.92	**15.62**	25.00

Note: Bold values represent the lowest OOB error and CV error which was achieved for week 5.

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

3.4 *Band Selection*

3.4.1 *Comparison between random forest variable importance and bands selected by the recursive feature elimination algorithm*

Figure 3 shows the mean decrease in accuracy as determined by RF and bands selected by the RFE algorithm for all 6 weeks. The RFE algorithm selected bands that have a high mean decrease in accuracy (Fig. 3).

3.4.2 *Comparison between random forest variable importance and Boruta selected bands*

Figure 4 compares the mean decrease in accuracy as determined by RF and bands selected by Boruta. In contrast to the bands selected by the RFE algorithm, not all bands selected by Boruta had a high mean decrease in accuracy. Boruta selected all relevant bands which included bands that had a high and low-mean decrease in accuracy from all spectral regions.

3.4.3 *A temporal analysis of bands selected by the recursive feature elimination algorithm and the Boruta algorithm*

The RFE and Boruta algorithms were able to substantially reduce the dimensionality of the hyperspectral dataset by <92% for each week. The number of bands selected by the RFE and Boruta algorithms per spectral region for 6 weeks is shown in Fig. 5. The total number

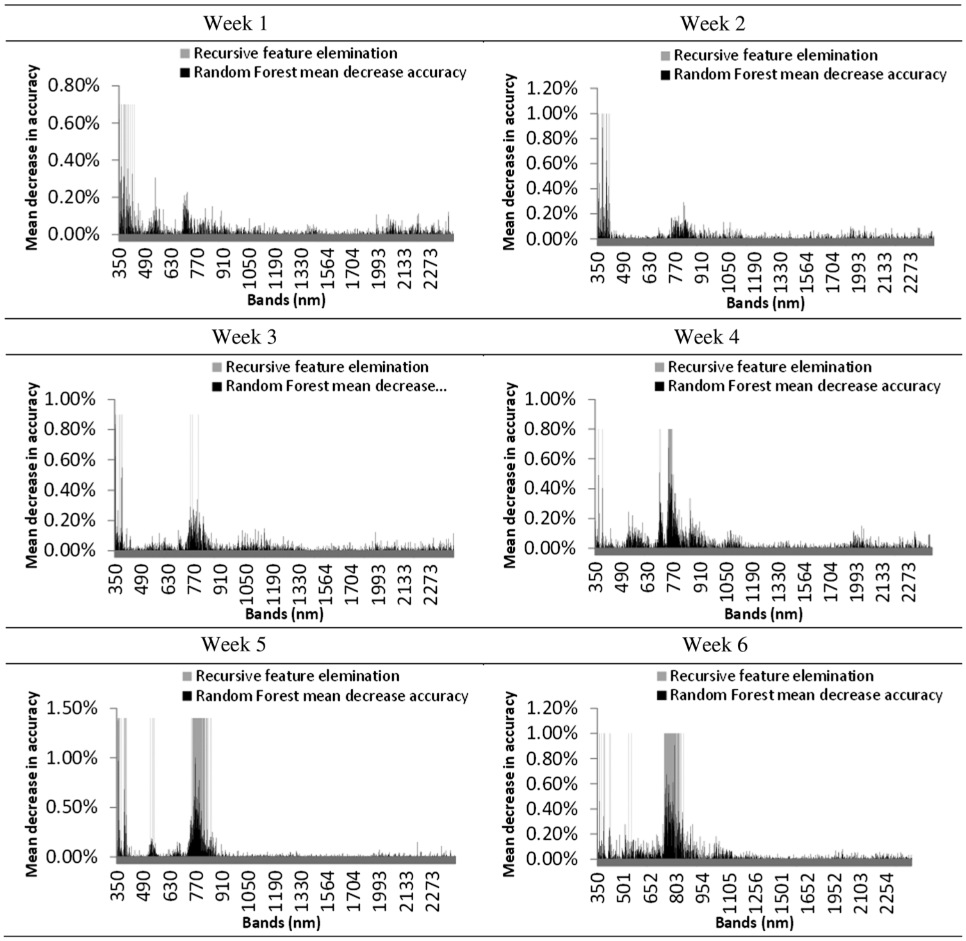

Fig. 3 Band importance as determined by RF variable importance and bands selected by the RFE algorithm for all 6 weeks.

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

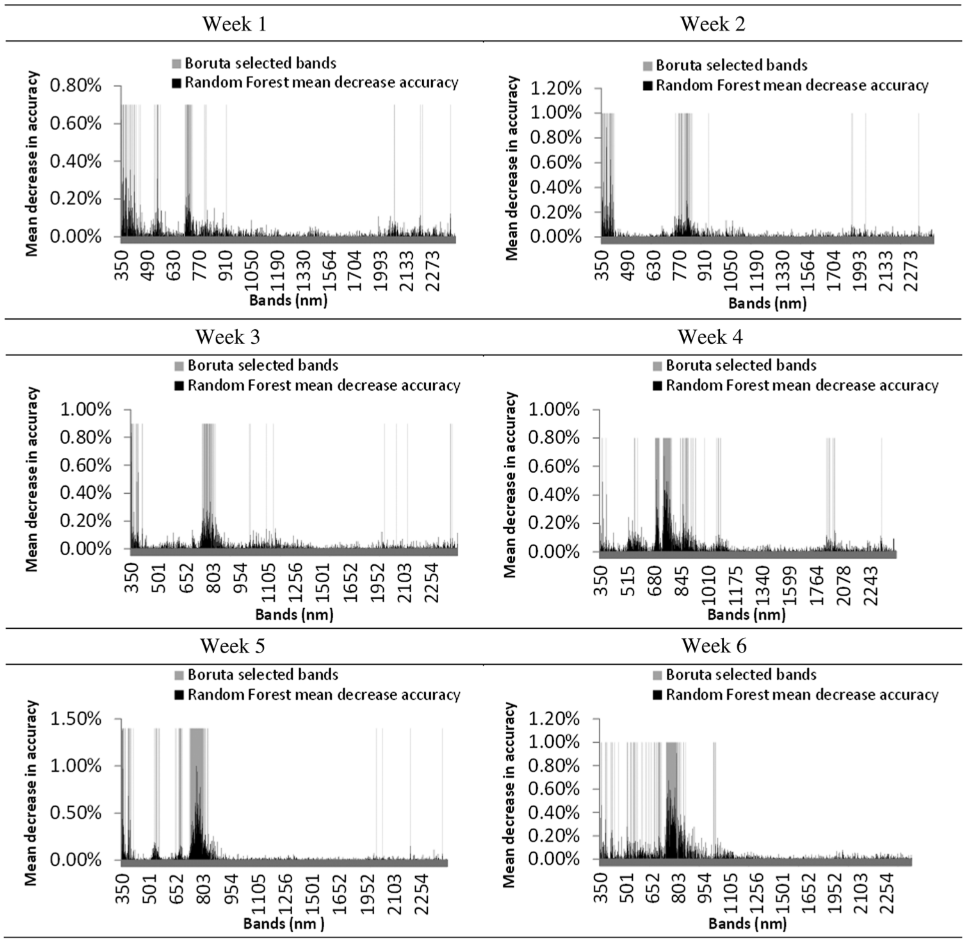

Fig. 4 Band importance as determined by RF variable importance and Boruta selected bands for all 6 weeks.

of bands selected by the RFE and Boruta algorithms increases over the progression of the stressor. It is evident from Fig. 5 that the RFE algorithm selected bands primarily from the visible region for weeks 1 and 2; however, from week 3 to 6, bands were selected from the visible and near-infrared regions. In contrast, Boruta selected bands from the visible region and the near-infrared region from week one until week six. From week 3, the number of bands selected by both feature selection algorithms from the near-infrared region was more numerous than the number of bands selected from the visible region. Boruta consistently selected bands 352, 754, 770, 771, 775, 781, 782, 783, 786, and 789 nm for 5 out of the 6 weeks. The RFE algorithm selected bands located at 755 and 766 nm for 4 out of 6 weeks. Based on these results, it is evident that the discrimination of variable infestation levels is highly dependent on the near-infrared region especially during weeks 4, 5, and 6.

3.4.4 *Random forest classification using bands selected by the recursive feature elimination algorithm and Boruta band selection algorithm*

Table 2 shows the RF error rates using bands selected by the RFE algorithm and Boruta for 6 weeks. The RF algorithm using bands selected by both band selection algorithms yielded better classification accuracies for all 6 weeks when compared with using all the bands (Tables 1 and 2). RF error rates improved by up to 16.04% and 16.67% using RFE and Boruta selected bands, respectively. Overall, the RF algorithm using Boruta selected bands consistently yielded the best classification results for all 6 weeks (Table 2). It is evident that the RF algorithm using Boruta selected bands achieved the lowest RF error rates from week 1. The lowest RF error rate using Boruta selected bands was achieved for week 5 (OOB error= 8.33%; CV error= 11.46%).

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

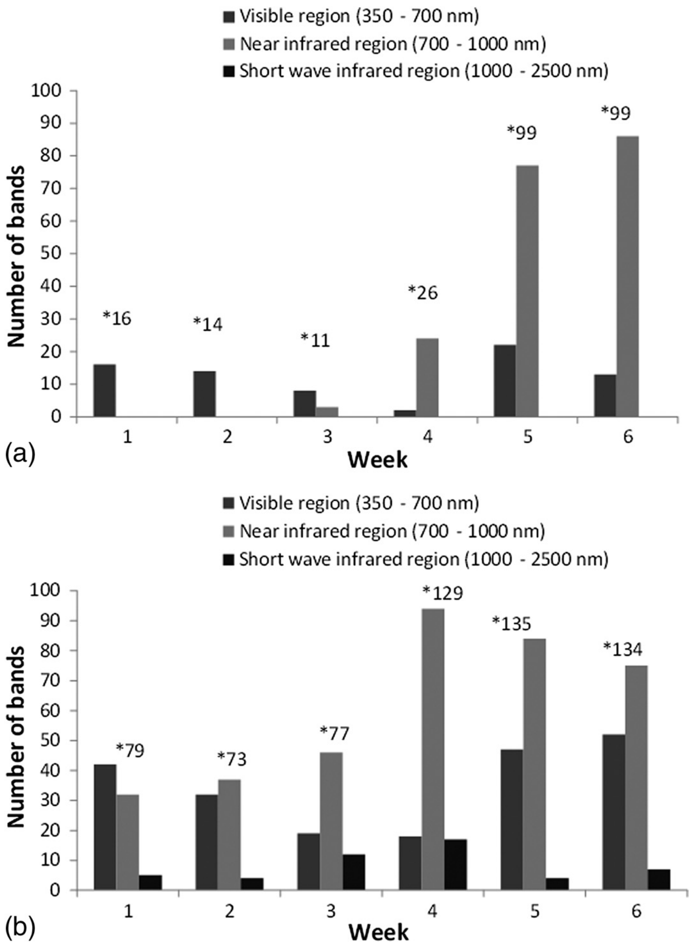

Fig. 5 The number of bands selected by (a) RFE algorithm and (b) Boruta per spectral region (visible region, near-infrared region, and shortwave infrared region) per week.

Table 2 RF classification results using bands selected by the RFE algorithm and Boruta selected bands for all 6 weeks.

	Week	1	2	3	4	5	6
Recursive feature elimination	OOB error (%)	16.25	14.71	13.24	13.24	**11.25**	21.25
	CV error (%)	19.56	16.23	15.70	15.70	**14.42**	25.38
	No. of bands selected	16	14	11	26	99	99
Boruta	OOB error (%)	15.62	13.54	12.50	10.42	**8.33**	15.62
	CV error (%)	20.83	15.62	14.58	13.54	**11.46**	16.67
	No. of bands selected	79	73	77	129	135	134

Note: Bold values represent the lowest OOB error and CV error which was achieved for week 5.

RF error rates improved between 0.63% (week 1) and 2.92% (week 5) using Boruta selected bands as compared to using RFE bands.

4 Discussion

Over the past decade, numerous biocontrol programs have been initiated to suppress and eventually eradicate water hyacinth infestations from freshwater bodies throughout Africa.[52–55]

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

However, the challenge faced by most biocontrol programs is differentiating and monitoring variable infestation levels to establish the efficacy of biocontrol agents.[53] Long term post release evaluations are of critical importance to gather strong evidence that the biocontrol agent is in fact responsible for the observed changes in the weed population and not influenced by other extraneous factors.[56] Additionally, biocontrol programs are dynamic in nature, constantly needing to be adapted to the prevailing conditions of the infestation. The ability to classify and monitor variable infestation levels remotely is important to inform environmental managers that wish to adopt a proactive approach in eradicating water hyacinth plants. It is envisaged that further releases of water hyacinth biocontrol agents could be tailored to the requirements of the receiving environment. Recently, remote sensing technologies have emerged as an efficient means of data collection to accurately evaluate[20,21] and monitor vegetation condition.[18,19] In this study, we hope to provide a remote sensing framework that integrates hyperspectral remotely sensed data with band selection techniques and machine learning algorithms to classify and monitor variable infestation levels on water hyacinth plants.

4.1 *Classification Using all Hyperspectral Bands*

RF error rates using all hyperspectral bands were low (Table 1) confirming its utility as a robust classification algorithm to classify variable infestation levels on water hyacinth plants. The low classification errors obtained in this study compare favorably with that of Adelabu et al. who implemented RF on the entire hyperspectral dataset to classify three defoliation levels in Mopane woodlands.[57] RF is able to achieve low-error rates by using a subset of randomly selected variables to grow each classification tree and a random subset of variables to split the subspace at each node, therefore, avoiding issues associated with $n < p$.[29] The randomly selected bootstrap samples and randomly selected subset of variables ensures that not only bands of low predictive performance are selected. RF combines many weak trees in an ensemble producing a strong classifier that achieves low error rates. Whilst numerous studies have advocated the utility of RF, most studies classifying insect defoliation have opted to implement RF at a static point in time.[31,32,57,58,59] Consequently, this study is unique by implementing RF over time to determine if RF can consistently achieve low error rates when dealing with high dimensional datasets. In this study, RF error rates decreased by 19.79% between week 1 and week 5. Similarly, Ismail and Poona[24] observed a decrease in RF error rates by 11.34% between week 1 and week 3 when classifying asymptomatic *Fusarium circinatum* stress on *Pinus radiate* seedlings. The consistent improvement in classification performance achieved in this study clearly demonstrates the algorithms' utility for analyzing multitemporal datasets. The consistent improvement in classification performance could be attributed to the spectral reflectance of each infestation level becoming more distinctive over the progression of the stressor. Researchers have noted that weevil herbivory negatively affects the plant's ability to complete photosynthesis causing chlorophyll to deteriorate and absorb light energy less efficiently, therefore increasing reflectance in the visible region.[22] In addition, extensive feeding damage destroys the cellular leaf structure of defoliated plants consequently altering their spectral reflectance in the near-infrared region.[22] The high-spectral resolution of hyperspectral data to detect differences in leaf reflectance between variable infestation levels coupled with the robust nature of the RF algorithm are key factors in achieving low-error rates. Overall, the results from this study positively illustrate the performance of RF as a classification algorithm to classify variable infestation levels on water hyacinth plants.

4.2 *Important Hyperspectral Bands*

The RFE and Boruta band selection algorithms consistently selected a subset of important bands each week successfully reducing the redundancy and dimensionality of the hyperspectral dataset (Fig. 5). However, in this study, the RFE and Boruta algorithms were primarily implemented to determine if the bands selected using a minimal optimal approach differs from an all relevant feature selection approach.[60] The key distinguishing feature between the two approaches is that RFE algorithm selects an optimal subset of bands based on RF mean decrease in accuracy, whereas Boruta selects both strongly and weakly relevant bands.[37,60] On comparing the

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

bands selected by the RFE algorithm and Boruta, it was found that Boruta selected more bands than the RFE algorithm demonstrating Boruta as an all relevant band selection algorithm. The results obtained in this study are consistent with the results achieved by Kursa and Rudnicki[60] and Kursa et al.[61] who also implemented Boruta to reduce data dimensionality. In this study, Boruta selected all relevant bands across the spectrum. This is important and advantageous as it provides a better holistic understanding of the damage experienced by the plant from the initial stages of the infestation. In contrast, the RFE algorithm primarily selected bands from the visible region at the initial stages of the infestation thereafter from the near-infrared region. This may skew the health assessment of water hyacinth plants under stress, as recessive physiological effects may go undetected providing an incorrect assessment of water hyacinth health status and severity of infestation. Despite the numerous advantages of adopting an all relevant band selection approach, most hyperspectral studies have opted to implement band selection techniques that select an optimal subset of bands based on RF mean decrease in accuracy.[33,62,63] Dye et al. [36] highlighted one of the limiting factors of the backward variable selection is that the method assumes that the prediction accuracy does not decrease as the number of bands increases. Therefore, good predictors could be removed early on by the wrapper. Boruta avoids such issues by selecting all relevant bands ensuring that all explanatory variables are retained. Overall, the results from this study confirm the utility of Boruta as a superior band selection technique to select all relevant bands and help alleviate high-dimensional complexity.

4.3 *Random Forest Classification Accuracies Using Bands Selected by the Recursive Feature Elimination and Boruta*

RF error rates improved when using RFE and Boruta selected bands as compared to using all hyperspectral bands for all weeks (Table 2) clearly supporting the utility of band selection techniques to improve classification accuracies. In this study, results showed that RF error rates decreased from the first week of infestation when using Boruta selected bands or RFE bands as compared to using the entire hyperspectral dataset (Tables 1 and 2). The results achieved in this study are consistent with the results of Adam et al.,[64] Adjorlolo et al.[34], and Poona and Ismail[24] who observed increased classification performance on a reduced hyperspectral dataset. Poona and Ismail[24] achieved lower RF error rates when using Boruta selected bands as compared with using all hyperspectral bands. Importantly, in this study, RF error rates using RFE and Boruta bands were compared to determine if classification accuracies are influenced by two different band selection approaches. RF error rates using Boruta selected bands were lower as compared with using RFE bands (Table 2) highlighting the benefit of adopting an all relevant band selection approach prior to classification. Selecting all relevant bands resulted in a decrease in error rates of between 0.63% (week 1) and 2.92% (week 5) over classification accuracies achieved using RFE bands. Even though, an improvement of 0.63% is relatively low, RF using Boruta bands still achieved a low-error rate of 15.62%. Classifying variable infestation levels accurately at the initial stages of the infestation is critical in determining if weevil populations are beginning to establish and inflict biocontrol damage on water hyacinth plants. This improvement could be attributed to fewer bands resulting in less noise, enabling the model to limit the use of redundant bands thus improving overall classification model and accuracy.[65] Overall, the results from this study illustrate the potential of RF coupled with Boruta band-selection algorithm in classifying biocontrol damage on water hyacinth plants.

References

1. J. A. Coetzee, M. J. Byrne, and M. P. Hill, "Impact of nutrients and herbivory by Eccritotarsus catarinensis on the biological control of water hyacinth, Eichhornia crassipes," *Aquat. Bot.* **86**(2), 179–186 (2007).
2. W. T. Penfound and T. T. Earle, "The biology of the water hyacinth," *Ecol. Monogr.* **18**(4), 447–472 (1948).
3. R. Verma, S. P. Singh, and R. A. J. Ganesha, "Assessment of changes in water hyacinth coverage of water bodies in northern part of Bangalore city using temporal remote sensing data," *Curr. Sci.* **84**(6), 795–804 (2003).

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

4. C. A. Kull and H. Rangan, "Acacia exchanges: wattles, thorn trees, and the study of plant movements," *Geoforum.* **39**(3), 1258–1272 (2008).
5. M. C. Law, "Willingness to pay for the control of water hyacinth in an urban environment of South Africa," MCom Thesis, Rhodes University, Grahamstown, South Africa (2007).
6. R. A. Goyer and J. D. Stark, "The impact of Neochetina eichhorniae on water hyacinth in Southern Louisiana," *J. Aquat. Plant Manage.* **22**, 57–61 (1984).
7. D. Jianqing et al., "Water hyacinth in China: its distribution, problems and control status," in *Biological and Integrated Control of Water Hyacinth Eichhornia Crassipes*, M. H. Julien et al., Eds., pp. 29–32, Australian Centre for International Agricultural Research, Canberra (2001).
8. M. M. Jiménez and M. A. Gómez Balandra, "Integrated control of Eichhornia crassipes by using insects and plant pathogens in Mexico," *Crop Prot.* **26**(8), 1234–1238 (2007).
9. M. H. Julien, "Biological control of water hyacinth with arthropods: a review to 2000," in *Biological and Integrated Control of Water Hyacinth Eichhornia Crassipes*, M. H. Julien et al., Eds., pp. 8–20, Australian Centre for International Agricultural Research, Canberra (2001).
10. A. Bownes, M. P. Hill, and M. J. Byrne, "Evaluating the impact of herbivory by a grasshopper, *Cornops aquaticum* (Orthoptera: Acrididae), on the competitive performance and biomass accumulation of water hyacinth, *Eichhornia crassipes* (Pontederiaceae)," *Biol. Control.* **53**(3), 297–303 (2010).
11. R. Van Driesche, *Biological Control of Invasive Plants in the Eastern United States*, USDA Forest Services, West Viginia (2002).
12. P. J. Moran, "Feeding by water hyacinth weevils (Neochetina spp.) (Coleoptera: Curculionidae) in relation to site, plant biomass, and biochemical factors," *Environ. Entomol.* **33**, 346–355 (2004).
13. T. D. Center et al., "Biological control of water hyacinth under conditions of maintenance management: can herbicides and insects be integrated?" *Environ. Manage.* **23**(2), 241–256 (1999).
14. M. H. Julien, M. W. Griffiths, and A. D. Wright, *Biological Control of Water Hyacinth. The Weevils Neochetina Bruchi and N. eichhorniae: Biologies, Host Ranges, and Rearing, Releasing and Monitoring Techniques for Biological Control of Eichhornia Crassipes*, Australian Centre for International Agricultural Research, Canberra (1999).
15. J. A. Coetzee and M. P. Hill, "Biological control of water hyacinth—the South African experience," *Bull. OEPP/EPPO Bull.* **38**(3), 458–463 (2008).
16. M. Zhang, X. Liu, and M. O'Neill, "Spectral discrimination of Phytophthora infestans infection on tomatoes based on principal component and cluster analyses," *Int. J. Remote Sens.* **23**(6), 1095–1107 (2002).
17. E. L. Hestir et al., "Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem," *Remote Sens. Environ.* **112**(11), 4034–4047 (2008).
18. Z. Qin and M. Zhang, "Detection of rice sheath blight for in-season disease management using multispectral remote sensing," *Int. J. Appl. Earth Obs. Geoinf.* **7**(2), 115–128 (2005).
19. J. Franke and G. Menz, "Multi-temporal wheat disease detection by multi-spectral remote sensing," *Precis. Agric.* **8**(3), 161–172 (2007).
20. J. H. Everitt et al., "Using remote sensing to assess biological control of saltcedar," *South Western Entomol.* **32**(2), 93–103 (2007).
21. S. Ge et al., "Multiple-level defoliation assessment with hyperspectral data: integration of continuum-removed absorptions and red edges," *Int. J. Remote Sens.* **32**(21), 6407–6422 (2011).
22. R. S. Fletcher, "Applying broadband spectra to assess biological control of saltcedar in west Texas," *Geocarto Int.* **29**(4), 383–399 (2013).
23. J. H. Everitt et al., "Using hyperspectral data to assess biological control damage of giant salvinia," *Geocarto Int.* **28**(6), 502–516 (2013).
24. N. Poona and R. Ismail, "Using Boruta-selected spectroscopic wavebands for the asymptomatic detection of Fusarium circinatum stress," *J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7**(9), 3764–3772 (2014).

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

25. J. C. Chan and D. Paelinckx, "Evaluation of random forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery," *Remote Sens. Environ.* **112**(6), 2999–3011 (2008).
26. A. Hapfelmeier and K. Ulm, "A new variable selection approach using Random forests," *Comput. Stat. Data Anal.* **60**, 50–69 (2013).
27. A. G. K. Janecek et al., "On the relationship between feature selection and classification accuracy," *J. Mach. Learn. Res.* **4**, 90–105 (2008).
28. M. N. Anwar, "Complexity measurements for dealing with class imbalance in classification modeling," PhD Thesis, Institute of Fundamental Science Massey University, New Zealand (2012).
29. L. Breiman, "Random forests," *Mach. Learn.* **45**(1), 5–32 (2001).
30. D. R. Cutler et al., "Random forests for classification in ecology," *Ecology* **88**(11), 2783–2792 (2007).
31. R. L. Lawrence, S. D. Wood, and R. L. Sheley, "Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest)," *Remote Sens. Environ.* **100**, 356–362 (2006).
32. V. F. Rodriguez-Galiano et al., "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogramm. Remote Sens.* **67**, 93–104 (2012).
33. E. M. Adam et al., "Discriminating the papyrus vegetation (*Cyperus papyrus* L.) and its co-existent species using random forest and hyperspectral data resampled to HYMAP," *Int. J. Remote Sens.* **33**(2), 552–569 (2012).
34. C. Adjorlolo et al., "Spectral resampling based on user-defined inter-band correlation filter: C_3 and C_4 grass species classification," *Int. J. Appl. Earth Obs. Geoinf.* **21**, 535–544 (2013).
35. R. Ismail and O. Mutanga, "Discriminating the early stages of Sirex noctilio infestation using random forest and shortwave infrared (SWIR) wavelengths," *Int. J. Remote Sens.* **32**, 4249–4266 (2011).
36. M. Dye, O. Mutanga, and R. Ismail, "Examining the utility of random forest and AISA Eagle hyperspectral image data to predict Pinus patula age in KwaZulu-Natal, South Africa," *Geocarto Int.* **26**(4), 275–289 (2011).
37. C. D. Duro, S. E. Franklin, and M. G. Dubé, "Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests," *Int. J. Remote Sens.* **33**(14), 4502–4526 (2012).
38. R. Diaz-Uriarte and S. Alvarez de Andres, "Gene selection and classification of microarray data using random forest," *BMC Bioinfor.* **7**(3), 1–13 (2006).
39. A. Bownes, M. P. Hill, and M. J. Byrne, "Assessing density-damage relationships between water hyacinth and its grasshopper herbivore," *Entomol. Exp. Appl.* **137**(3), 246–254 (2010).
40. P. G. Soti and J. C. Volin, "Does water hyacinth (*Eichhornia crassipes*) compensate for simulated defoliation? Implications for effective biocontrol," *Biol. Control.* **54**(1), 35–40 (2010).
41. M. O. Bashir, Z. E. El Abjar, and N. S. Irving, "Observations on the effect of the weevils Neochetina eichhorniae Warner and Neochetina bruchi Hustache on the growth of water hyacinth," *Hydrobiologia* **110**(1), 95–98 (1984).
42. ASD, *Handheld Spectroradiometer: User's Guide Version 4.05*, Analytical Spectral Devices, Boulder (2005).
43. A. Liaw and M. Wiener, *Random Forest: Breiman and Cutler's Random Forests for Classification and Regression*, R package version 4.6-7 (2013).
44. R Development Core Team, *R: A Language and Environment for Statistical Computing* (2012).
45. R. Diaz-Uriarte, *Variable Selection Using Random Forests*, R package version 0.7-3 (2014).
46. M. B. Kursa and R. W. Rudnicki, *Boruta: A Wrapper Algorithm for All-relevant Feature Selection*, R package version 2.1.0 (2013).
47. M. Pal, "Random forest classifier for remote sensing classification," *Int. J. Remote Sens.* **26**(1), 217–222 (2005).

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

48. M. B. Garzon et al., "Predicting habitat suitability with machine learning models: the potential area of Pinus sylvestris L. in the Iberian Peninsula," *Ecol. Modell.* **197**(3–4), 383–393 (2006).
49. A. M. Prasad, L. R. Iverson, and A. Liaw, "Newer classification and regression tree techniques: bagging and random forests for ecological prediction," *Ecosystems* **9**, 181–199 (2006).
50. J. Peters et al., "Random forest as a tool for ecohydrological distribution modelling," *Ecol. Modell.* **207**(2–4), 304–318 (2007).
51. R. Geneur, J. Poggi, and C. Tuleau-Malot, "Variable selection using random forests," *Pattern Recognit. Lett.* **31**(14), 2225–2236 (2010).
52. O. Ajuonu et al., "Impact of the weevils *Neochetina eichhorniae* and *N. bruchi* (Coleoptera: Curculionidae) on water hyacinth, *Eichhornia crassipes* (Pontederiaceae), in Benin, West Africa," *Afr. Entomol.* **11**(2), 153–161 (2003).
53. G. Mbati and P. Neuenschwander, "Biological control of three floating water weeds, *Eichhornia crassipes*, *Pistia stratiotes*, and *Salvinia molesta* in the Republic of Congo," *BioControl* **50**(4), 635–645 (2005).
54. G. M. Sikoyo and L. Goldman, "Assessing the assessments: case study of an emergency action plan for the control of water hyacinth in Lake Victoria," *Int. J. Water Res. Dev.* **23**(3), 443–455 (2007).
55. J.A. Coetzee et al., "A review of the biological control programmes on *Eichhornia crassipes* (C. Mart.) Solms (Pontederiaceae), *Salvinia molesta* D. S. Mitch. (Salviniaceae), *Pistia stratiotes* L. (Araceae), *Myriophyllum aquaticum* (Vell.) Verdc. (Haloragaceae) and *Azolla filiculoides* Lam. (Azollaceae) in South Africa," *Afr. Entomol.* **19**, 451–468 (2011).
56. L. Morin et al., "Review of approaches to evaluate the effectiveness of weed biological control agents," *Biol. Control.* **51**(1), 1–15 (2009).
57. S. Adelabu et al., "Spectral discrimination of insect defoliation levels in Mopane woodland using hyperspectral data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7**(1), 177–186 (2014).
58. P. O. Gislason, J. A. Benediktsson, and J. R. Sveinsson, "Random forest for land cover classification," *Pattern Recognit. Lett.* **27**(4), 294–300 (2006).
59. S. Adelabu, O. Mutanga, and E. Adam, "Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels," *ISPRS J. Photogramm. Remote Sens.* **95**, 34–41 (2014).
60. M. B. Kursa and W. R. Rudnicki, "Feature selection with the Boruta package," *J. Stat. Sci.* **36**(11), 1–12 (2010).
61. M. B. Kursa, A. Jankowski, and W. R. Rudnicki, "Boruta-A system for feature selection," *Fundam. Inf.* **101**, 271–285 (2010).
62. E. Adam et al., "Field spectrometry of papyrus vegetation (Cyperus papyrus L.) in swamp wetlands of St. Lucia, South Africa," in *Proc. IEEE Int. Geoscience and Remote Sensing Symp.*, pp. 260–263 (2009).
63. N. Chehata, L. Guo, and C. Mallet, "Airborne Lidar feature selection for urban classification using random forests," in *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, **38**(*Part 3/W8*), 207–212 (2009).
64. E. Adam, O. Mutanga, and R. Ismail, "Determining the susceptibility of *Eucalyptus nitens* forests to *Coryphodema tristis* (cossid moth) occurrence in Mpumalanga, South Africa," *Int. J. Geog. Inf. Sci.* **27**(10), 1924–1938 (2013).
65. F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.* **42**(8), 1778–1790 (2004).

Na'eem Hoosen Agjee is a doctoral student at the University of KwaZulu-Natal. He received his BSc, honors, and master's degrees in environmental science from the University of KwaZulu-Natal in 2009, 2010, and 2012, respectively. His current research interests include invasive biology, hyperspectral remote sensing, machine learning algorithms, and image processing.

Riyad Ismail received his MSc degree in GIS (cum laude) and his PhD in remote sensing from the University of KwaZulu-Natal, South Africa. He has over 15 years of experience in

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

implementing spatial technologies (GIS, GPS, and remote sensing) at commercial, academic, and research institutions. He was recently appointed as a senior research associate at the University of KwaZulu-Natal and is currently employed as a principal research officer at Sappi forests.

Onisimo Mutanga received his PhD in hyperspectral remote sensing of tropical grass quality and quantity from Wageningen University-ITC, Wageningen, The Netherlands, in 2004. He is a professor with the School of Agriculture, Earth, and Environmental Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa. He has graduated seven PhD and 14 master's students. He has authored more than 67 articles and has several conference proceedings and book chapters. His research interest includes ecological remote sensing.

Downloaded From: http://remotesensing.spiedigitallibrary.org/ on 10/16/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx