Mapping Bugweed (*Solanum mauritianum*) Infestations in *Pinus patula* Plantations Using Hyperspectral Imagery and Support Vector Machines

Jonathan Tom Atkinson, Riyad Ismail, and Mark Robertson

***Abstract*—The invasive plant known as bugweed (*Solanum mauritianum*) is a notorious invader of forestry plantations in the eastern parts of South Africa. Not only is bugweed considered to be one of five most widespread invasive alien plant (IAP) species in the summer rainfall regions of South Africa but it is also one of the worst invasive alien plants in Africa. It forms dense infestations that not only impacts upon commercial forestry activities but also causes significant ecological and environment damage within natural areas. Effective weed management efforts therefore require robust approaches to accurately detect; map and monitor weed distribution in order to mitigate the impact on forestry operations. The main objective of this research was to determine the utility of support vector machines (SVMs) with a 272-waveband AISA Eagle image to detect and map the presence of co-occurring bugweed within mature *Pinus patula* compartments in KwaZulu Natal. The SVM when utilized with a recursive feature elimination (SVM-RFE) approach required only 17 optimal wavebands from the original image to produce a classification accuracy of 93% and True Skills Statistic of 0.83. Results from this study indicate that (1) there is definite potential for using SVMs for the accurate detection and mapping of bugweed in commercial plantations and (2) it is not necessary to use the entire 272-waveband dataset because the SVM-RFE approach identified an optimal subset of wavebands for weed detection thus enabling improved data processing and analysis.**

***Index Terms*—AISA eagle, recursive feature elimination, support vector machines, weed detection.**

1. Introduction

In recent years there has been a rapid decline in the number of newly afforested areas in South Africa with the forestry industry largely limited to a fixed production area [1]. This is largely due to a decrease in the number of suitable sites for afforestation coupled with greater reluctance by the Department of Agriculture, Forestry and Fisheries (DAFF) in granting new water permit licences for afforestation [2].

It has therefore become increasingly important, not only from an environmental but also economically sustainable perspective, for forestry managers to ensure that forest productivity within existing planted areas is maximised. Thus any agents which present a significant threat to commercial forest sustainability and could lead to a decline in timber productivity need to be identified and mitigated [3]. One such threat is the occurrence of non-native plants, or weeds, within the plantations which can negatively impact upon the growth and productivity of the commercial species [4]. One species of weed that warrants concern, particularly in the province of KwaZulu-Natal (KZN) is *Solanum mauritianum* (bugweed). According to the National Environmental Management: Biodiversity Act of 2003, (NEMBA) bugweed is a declared category 1b invader weed [5]. As such, one of the directives of NEMBA regarding the regulation of category 1b invader weeds is that, due to their high invasive potential, these plants require compulsory control as part of an invasive species control programme. NEMBA thus imposes a legal obligation upon all landowners to actively locate and regulate the predominance and limit the spreading of category 1b species such as bugweed occurring on their land. The hardiness and resilience of this species has already established it as a major declared weed [5] of natural ecosystems, forestry plantations, riverine habitats and conservation areas and the weed can become quite ubiquitous if not controlled [6]. Bugweed is also extremely resilient and opportunistic and competes fiercely for resources, often suppressing the growth or even displacing the surrounding vegetation [7]. This is particularly evident in commercial forest plantations where it has been reported to stunt the growth of certain *Pinus* species [8]. From a fire management perspective, the presence of bugweed provides undesirable under canopy fuel loads within commercial forestry stands during extreme uncontrolled fire events. This can consequently lead to increased costs of fire protection and suppression and greater overall severity of wild fire damage as a result of additional fuel load material [9]. Forestry managers therefore require accurate as well as timely spatial information on bugweed occurrence in order to ascertain the severity of invasion and contain small infestations before they get too large and expensive to eradicate. Consequently, the ability to develop accurate and spatially explicit techniques for early weed detection, mapping and monitoring at all stages of occurrence is regarded as of high priority for commercial forestry management [10]. In light of these challenges the use of digitally analysed, remotely-sensed data for the recognition and quantification of IAPs could not only result in a more time-efficient approach to classification, and thus weed detection, but the technology could also potentially reduce weed management costs [11]. Optical remote sensing technologies, more specifically hyperspectral sensors [12],

This work was supported in part by Sappi Southern Africa Pty Ltd-Forest Division.

J. T. Atkinson and M. Robertson are with the Center for Environmental Management, University of Pretoria, Pretoria, Gauteng 0002, South Africa.

R. Ismail is with the Department of Geography, University of KwaZulu Natal, Berea 4041, South Africa.

[13], have the capacity to rapidly and synoptically exploit the unique spectral, phenological and structural characteristics of plants such as bugweed [14]. Due to their excellent spectral resolution, hyperspectral sensors are extremely well suited to map the abundance of weed species over large spatial extents [15]. Consequently, there has been an increase in the number of studies that have applied hyperspectral image analysis to detect weed species including Brazilian pepper [13], leafy spurge [16], spotted knapweed [17] and tamarisk [18]. More recently, [19] have shown that in diverse environments where areas are particularly susceptible to weed invasion, hyperspectral imagery is essential for weed detection and mapping. The study found that sensors that are less sensitive to varying spectral and environmental conditions will not be able to properly detect infestations resulting in classifications that may confound species variability and detectability.

Unlike broadband multispectral remote sensing platforms, the continuous nature of the spectra extracted from the hyperspectral imagery allows for discrimination of more subtle differences between individual species [20]. This ultimately enables classification to occur at both an in-depth biochemical and structural level which would otherwise not be possible with the coarse bandwidths acquired by multispectral sensors [13]. [21] points out that the inability of multispectral sensors to adequately detect canopy level occurrence in complex, layered forested environments renders them especially ineffective in providing a biophysically-based approach for mapping occurrence. Indeed, [21] endorses this assertion by showing that the generally high classification errors associated with the use of multispectral sensors for damage discrimination impose operational limitations on their use by forestry companies. Interestingly, studies of direct weed detection using hyperspectral sensors have often been limited to scenarios that do not necessarily require the spectral precision or fully utilise the capabilities offered by hyperspectral sensors [20]. For example, [22] noted that weed locality and physiology within particular homogeneous landscapes would enable fairly accurate detection, discrimination and classification, even at a multispectral level. However, even though multispectral data may be more cost effective and more accessible than hyperspectral data, both classification and spatial resolution may be inadequate for regional or site-specific weed management activities [23]. However, detecting bugweed within a commercial forestry environment presents a unique set of challenges. Bugweed can be identified as shrubs or small trees ranging in height from 2–12 m having rounded canopies of distinctive well-developed pale to dark green foliage [8]. The high degree of tree uniformity as well as dense canopy closure in commercial forests often results in the inability to distinguish small to medium bugweed plants, which occur amongst other tree species. Consequently, hyperspectral sensors provide the necessary detail to identify the unique spectral signatures of this species relative to a backdrop of *P. patula* trees [18]. More specifically, this study focussed on separating bugweed for *P. patula*. However the study did not look at further separating the spectral characteristics of bugweed from other understory species. In terms of understory vegetation within the plantation, bugweed was considered to be the most dominant weed species with detection therefore directly dependent on how well the sensor was able to capture the reflected radiation from the canopy of the dominant weed.

One of the fundamental processes of deriving land cover information from remotely sensed imagery is image classification [24], the outcome of which is a thematic map of the original image depicting the pixels assigned to a particular class [25]. Recently, a new generation learning algorithm, namely support vector machines (SVM) [26], has increasingly been used for the classification of hyperspectral imagery [27]. SVMs used for classification are based on finding the optimal separation surface, [erred to as a hyperplane, between classes by identifying the most representative training samples, or support vectors, on either side of the hyper plane [28]. Studies have shown that SVMs are capable of producing higher classification accuracies than more widely used pattern recognition models such as maximum likelihood and neural network classifiers [29] and [30]. Another attractive feature of SVMs is that they have proven to be particularly expedient when used in studies dealing with homogenous classes with a limited number of training samples available [29]. Additionally, SVMs seem to be robust to the effects of the Hughes phenomenon [31] or *curse of dimensionality* that is with a limited number of training samples the classification rate decrease as the data dimensionality increases [32]. A difficult task even for techniques dedicated to processing hyperspectral data such as Spectral Angle Mapping or spectral unmixing [28].

However, although SVMs are effective at classifying data, they do not directly provide the user with an indication of feature importance [29]. Feature selection (FS) is a dimension reduction approach and while SVMs are known to be robust to dimensionality the application of FS with SVMs improves classification accuracy [30] and expedites subsequent data processing [15]. FS does not alter the original representation of the variables but merely enables a model, such as SVMs, to selectively focus on relevant variables whilst ignoring the contribution of irrelevant or redundant (noisy) variables [31]. Advantages of FS include reduced data storage requirements, improved model prediction performance, reducing the costs of future measurements and improving data or model understanding [32]. FS is therefore considered to be an integral component in bridging the gap between research and operational remote sensing applications [37]. For a review of FS see [38] and [31].

The aim of this research was to evaluate support vector machines (SVM) could be used to detect the presence of co-occurring bugweed trees within mature *P. patula* plantations using hyperspectral imagery. With regards to this study, the SVM algorithm and FS methods were used to produce the smallest subset of AISA Eagle wavelengths that would allow for the accurate classification of bugweed reflectance spectra. The overall objective of the study was to demonstrate, for the first time, the practicality and utility of using SVMs to identify the presence of bugweed in commercial *P. patula* plantations.

2. Materials and methods

A. Study Site Description

The study area is located in the Sappi Hodgsons plantation (Centroid: $30^\circ\ 29'56\mathrm{E}$ and $29^\circ\ 13'42''\mathrm{S}$), and is situated ap-

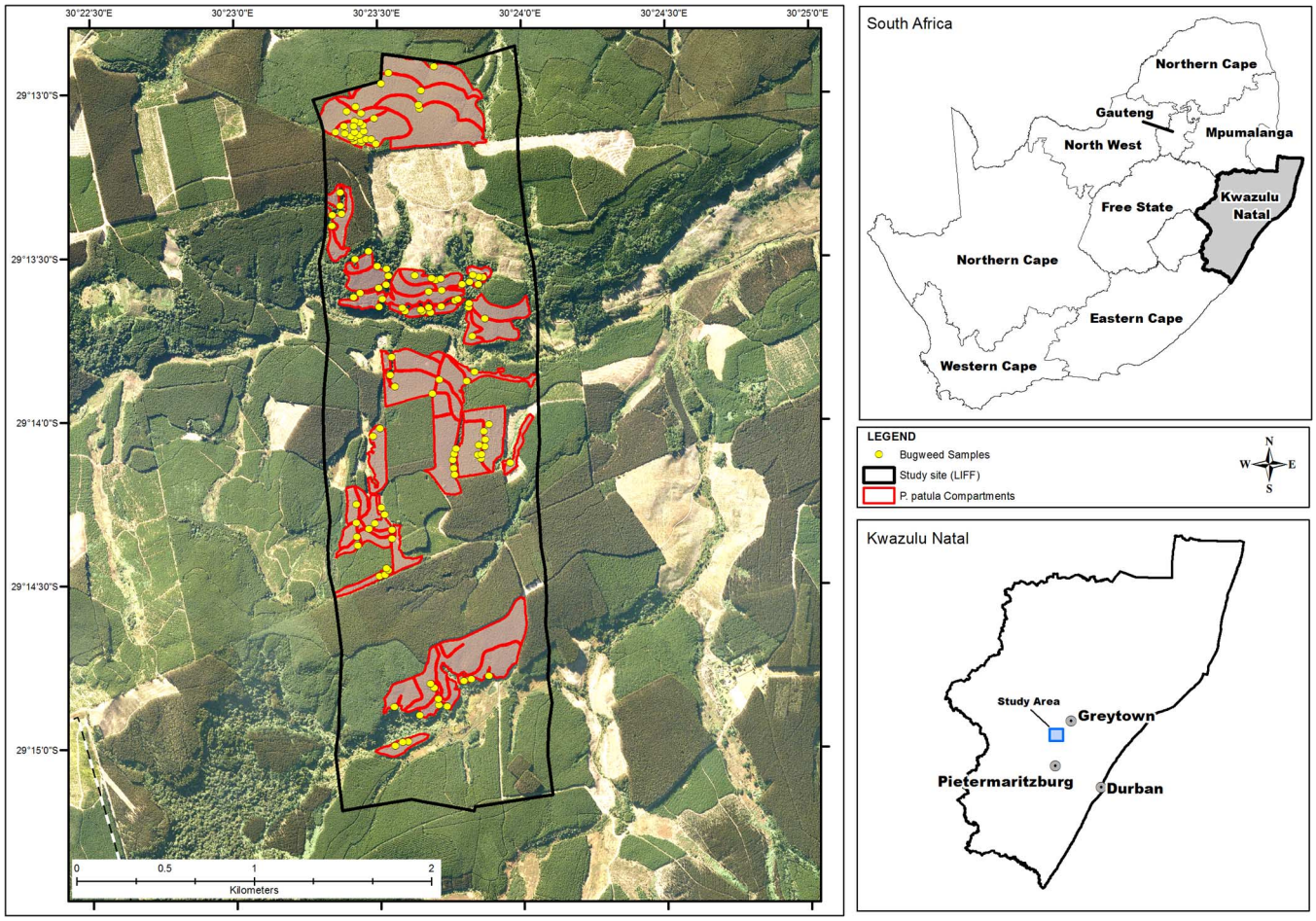

Figure 1 : Location of study area with the shaded areas representing the *Pinus patula* compartments and the dots indicating where bugweed samples were collected.

proximately 50 km north of Pietermaritzburg in the KwaZulu-Natal midlands (Fig. 1). The topography of the region ranges from gently sloping to moderately undulating with slopes between 1000–1400 m above sea level [39]. The climates is cool with mean annual temperatures in the region of 15.9° C north of Pietermaritzburg in the KwaZu and mean annual precipitation in the region of 1015 mm. Areas that are not occupied by commercial timber species are characterized by vegetation types such as the Ngongoni veld of the Natal mist-belt and Southern tall grassveld [39]. The majority of compartments within the site consist of *P. patula* trees which range in age from 1 to 22 years and form part of a pulpwood management regime. The prevalence of non-native plants has become a serious problem in the plantation where weed species such as bugweed have become increasingly prolific. Infestations are particularly evident along riparian zones, previously disturbed areas such as grasslands and indigenous forest areas which border commercial compartment stands [40].

B. Image Acquisition

The hyperspectral imagery was acquired using the Airborne Imaging Spectrometer for Applications (AISA). The AISA Eagle sensor is a pushbroom sensor consisting of a hyperspectral sensor head, data logger, GPS unit and irradiance sensor. The sensor operates in the visible (400 nm–700 nm) as well as the near-IR portion (701 nm–2000 nm) of the spectrum [44]. The AISA Eagle sensor samples wavelengths 400–900 nm using 272 bands at a spectral resolution (bandwidth) of 2–4 nm and spatial resolution of 2.4 m [42]. The imagery was acquired on 11 March 2009 under cloudless conditions at 11:45 a.m. A fixed wing, light aircraft was used to collect the imagery at a mean GPS flight altitude of 2728.42 m. The 272 band image dataset, with an initial spectral range of 393.23 nm–994.09 nm, was spectrally resampled to 4.9 nm in line with the spectral binning options identified from the technical specifications for the AISA Eagle sensor [42]. Spectral binning [43], [44] was employed as a resampling method to eliminate redundant and damaged variables in the dataset and spectrally resampling the imagery allowed (a) for rapid image analysis by reducing data redundancy and (b) removed bands which contained a high degree of noise. After binning, the resulting dataset had been reduced to 110 bands with a spectral range of 400.00 nm–901.40 nm eliminating bands greater than 905 nm due to excessive noise.

The AISA imagery was geometrically registered (RMSE $<$ 1.0 pixels with 3rd order polynomial approximation) using 20 ground control points (GCPs) [19]. The GCPs were selected from high resolution colour (RGB) aerial photographs of the same region collected in April 2009 [30]. The aerial photographs, having an estimated ground accuracy of 10 cm

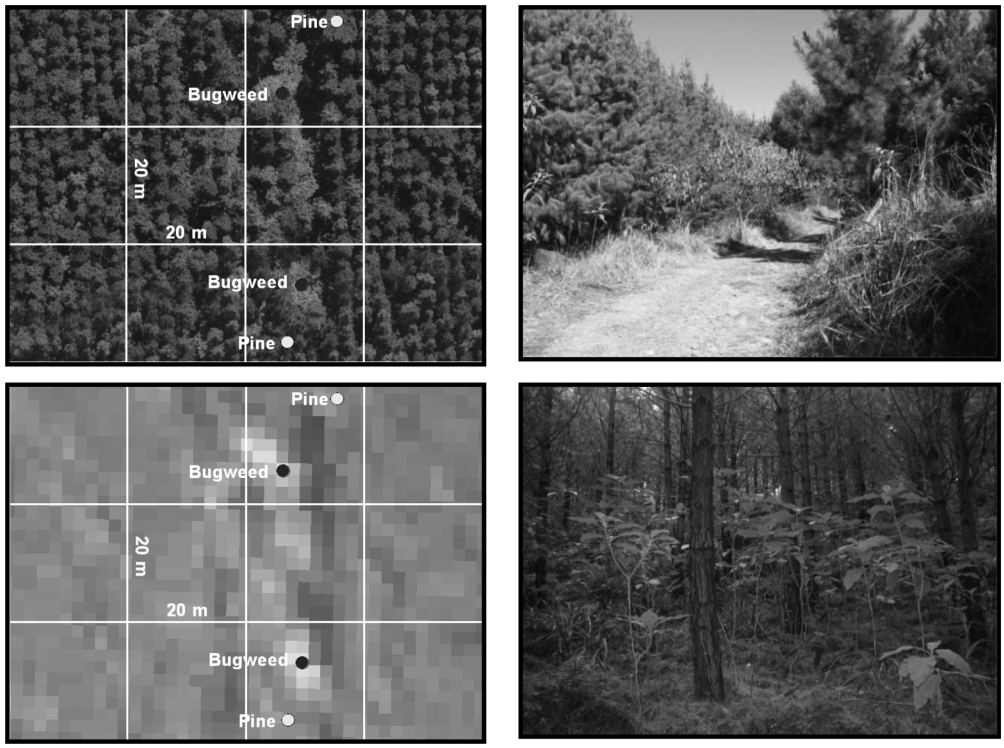

Figure 2 : 10 cm resolution aerial imagery (a) and 2.4-m AISA Eagle imagery (b). The photographs show examples of the bugweed that were sampled between (c) and within compartments (d).

($\mathrm{RMSE} < 1.0$ pixels with 1st order polynomial approximation) were geometrically registered using topographical features such as roads, streams and cadastral boundaries [45] provided by Sappi Forests and were referenced to the Universal Transverse Mercator projection (WGS 84 datum, UTM Zone 36S) using the Environment for Visualization software [46]. The AISA imagery was converted to reflectance using the empirical line method [47], [48] within ENVI 4.7 [46]. The empirical line method matches the image radiance data to field reflectance spectra of two materials of contrasting colour (usually black and white). The image radiance spectra were then regressed with field reflectance spectra to determine a linear transformation from radiance to reflectance. The gain and offset curves for the image spectra are then used to derive the average ground reflectance for the entire image [49]. This then allows for the accurate retrieval and interpretation of the hemispherical-conical reflectance factor [50].

C. Bugweed Reference Data

A purposive sampling approach [51] was used for the identification of bugweed within the pine compartments using the high resolution colour (RGB) imagery. Bugweed occurrence was based on photographic interpretation of the high resolution airborne imagery [18], [52] and, [25] and subsequent field verification was carried out using a GPS. More specifically, a 20 m $\times$ 20 m digital grid was super-imposed onto the high resolution aerial imagery [53] and bugweed occurring within mature *P. patula* stands, having a canopy greater than 5.7 m_2 (1 pixel) and ranging in age from 7 to 20 years, were then visually selected [52]. Fig. 2 illustrates how successful bugweed samples were identified within the *P. patula* compartments and located at least 20 m from other bugweed samples [54]. Juvenile *P. patula* trees, or trees younger than seven years, were excluded from the sampling procedure because the bugweed occurrence within these young compartments was both limited and infrequent i.e. there were not enough bugweed samples to statistically represent the juvenile age group. Another reason for the low occurrence of bugweed within the juvenile pine compartments is that the weeds are easily detected by the foresters and are subsequently removed.

Two sample point features per grid cell were recorded with the first point representing the bugweed and the second point representing the pine trees. Subsequently, a total of 240 tree samples were collected (120 bugweed samples and 120 *P. patula* samples). The spectral reflectance signatures for the 240 sample points were then extracted in a GIS using ArcMap [55] with the resampled imagery then input into the R Project for Statistical Computing [56] and used in the classification process. Field visits were also conducted between June 2009 and July 2009 to confirm that bugweed from the imagery was present within the study site. In order to develop the model, the data were partitioned into an equal number of bugweed and *P. patula* samples in the test and training datasets to provide training $(\mathrm{n} = 60)$ and testing $(\mathrm{n} = 60)$ datasets.

D. Support Vector Machines

The support vector machine (SVM) algorithm uses a supervised machine learning technique that is based on statistical learning theory [57] and Vapnik's Structural Risk Minimization principle [26]. SVMs used for image classification are based on finding the optimal separation surface, [erred to as a hyperplane, between classes (i.e. bugweed and *P. patula*) by identifying the

most representative training samples, or support vectors, on either side of the hyperplane [28].

The optimal hyperplane is therefore the one that separates the classes with the maximum distance between the separating margin and the data points (support vectors) on the plane with the least generalization error and is known as the optimum separating hyperplane [24]. In a two-class linearly separable classification problem SVM employs an optimization technique to select the optimal separating boundary (hyperplane) from the infinite number of linear decision hyperplanes. In cases where the training data are not linearly separable a "kernel trick" is used to project the data into a hyper dimension, or feature space, where the kernel can then simulate the optimal separation of the classes [25].

There are a large number of standard and customised kernels available and two most commonly cited kernels in remote sensing literature are the radial basis function (RBF) and polynomial kernels, [25] However, for this study the linear kernel function was employed to classify the occurrence of bugweed from *P. patula* trees. The decision was based on results obtained from a preliminary run of the SVM model using the polynomial, RBF and linear kernel functions [57] and after comparing the computational time of each model as well as the classification accuracies it was observed that the linear model produced the highest overall accuracy in the fastest computational time. For a review of suitable kernel selection and parameter optimization of SVMs, see [59] and [60].

The process of training the SVM model for this study was adapted from the procedure outlined by [61]. For training the SVM classifier using the linear kernel only the regularization parameter C (a penalty parameter) requires optimization. Changing the kernel parameter is equivalent to selecting a particular feature space whereas tuning the C *variable*corresponds to weighting the slack or penalty variables for the SVM [62]. Hence the C parameter is selected by the user to balance out the competing criteria of margin maximization and error minimization [25]. The higher the value of C, the higher the penalty associated with misclassified samples [58]. In order to determine which C values will produce the best classification result an optimum parameter search must be performed on the training dataset [61]. Common approaches of determining the optimal value(s) of C is to implement a search utilising k*-fold* cross validation [63]. The method exhaustively searches for the optimal C parameters over a defined parameter range and reports the k*-fold* cross validation classification error for each parameter [62]. The parameter combination that produces the best cross validation accuracy is then selected as the most optimum for the classification problem. Consequently, each instance of the entire training subset is predicted at least once so that cross validation accuracy is the percentage of data correctly classified [61]. Naturally the C parameter with the lowest cross validated error is then selected.

E. Feature Selection and Variable Ranking

The AISA Eagle hyperspectral sensor is capable of simultaneously acquiring data from more than a hundred narrow spectral bands (data channels) ranging from the visible to infrared portions of the electromagnetic spectrum [64]. More spectral bands include more information. However, as [65] point out, dealing with such a large number of narrow band channels presents problems in the acquisition phase (noise), storage and transmission phases (data size) and processing phase (complexity). Consequently, this limits robust statistical estimations and often results in overfitting of the training data leading to poor generalization capabilities of the classifier [31]. Approaches that are able to circumvent these challenges by processing a subset of relevant bands which best characterize a particular feature while limiting the effects of dimensionality, are essential to remote sensing. One such approach is FS. For this study an adaptation of the SVM Recursive Feature Extraction (RFE) algorithm proposed by [66] was utilized to select the most important subset of bands that provided the best classification accuracy.

The SVM-RFE utilizes all the hyperspectral bands and then successfully eliminates bands from the dataset based on their influence on the SVM algorithm. As the SVM is trained using the linear kernel, each iteration of the model eliminates bands with the smallest ranking criterion. The ranking criterion corresponds to the vector weights of the decision hyperplane assigned by the SVM algorithm (see [32] for a detailed discussion on SVM-RFE). However, the SVM-RFE approach outlined in this study uses forward feature selection (FFS) instead of backward feature elimination (BFE) proposed by [32] as a search strategy. FFS begins with an empty subset of variables and progressively adds relevant variables into larger and larger subsets. BFE starts with the set of all variables and progressively eliminates the least relevant variables [31]. Since BFE starts by evaluating all bands in the dataset, it is computational more demanding than FFS. Consequently, using FFS to building classifiers, when there are a large number of features (for example hyperspectral bands) in the dataset, is much faster [38]. Additionally, the SVM-RFE algorithm was modified so that at each stage of the FFS process, the C parameter of the SVM linear kernel is optimized as well. The optimal subset of band's are then selected based on the prediction error as calculated by a 10-fold cross validation (CV). Once the bands with the highest accuracy are identified using the modified SVM-RFE procedure they are utilized in the library e1017 Library within the R statistical software to implement the SVM algorithm.

F. Accuracy Assessment

The most widely accepted, and perhaps most effective, way to represent classification accuracy is by means of a presence/absence model. The performance of the model is usually summarized in an error matrix that cross tabulates the observed and predicted presence/absence patterns [67] (Table I). With reference to this study the performance of the optimal subset of bands was evaluated both numerically (overall accuracy) as well as statistically (true skills statistic, specificity and sensitivity). The overall accuracy is interpreted as the total number of correctly classified pixels divided by the total number of sample pixels analyzed within the error matrix. In a similar way, the accuracies of individual categories can also be represented. The sensitivity (*sens*) represents the probability that a sample pixel will be correctly classified to a particular category and includes the error of

Table I : Error matric showing derived measures of classification accuracy

		validation dataset		
		presence	absence	row total
model	presence	True Positive (TP)	False Positive (FP)	*TP + FP*
	absence	False Negative (FN)	True Negative (TN)	*FN + TN*
	column total	*TP + FN*	*FP + TN*	Total

omission which occurs when a pixel is not included into a category it does belong to i.e. false negative [68]. The sensitivity can be defined as:

$$\boldsymbol{Sens} = \frac{\boldsymbol{TP}}{(\boldsymbol{TP} + \boldsymbol{FN})} \tag{1}$$

The specificity (spec) is a measure of how reliable the classified map actually is and represents the probability that a sample pixel classified on the image represents the same category on the ground and includes the error of commission which occurs when a pixel is classified to a category that it does not belong to (false positive) [69]. The specificity may be represented as:

$$\boldsymbol{Spec} = \frac{\boldsymbol{FP}}{(\boldsymbol{FP} + \boldsymbol{TN})} \tag{2}$$

Additionally, the True Skill Statistic (TSS) [68] was used as a measure to evaluate the model's agreement with the reference data and can be defined as:

$$\boldsymbol{TSS} = (\boldsymbol{sens} + \boldsymbol{spec}) - 1 \tag{3}$$

The TSS is very similar to Cohen's kappa statistic but has the advantage of correcting for dependency on prevalence whilst still maintaining all the advantages of kappa. Consequently, the TSS is able to account for errors of commission and omission in one statistic, and just like kappa, the TSS values also range from -1 to $+1$. The TSS provides a good indication of the extent to which the percentage of correctly classified pixels in the error matrix is as a result of true agreement or chance agreement, with TSS values approaching one indicating true agreement and values approaching zero, chance agreements [68].

3. Results

A. Band Selection and Ranking

The modified SVM-RFE method was able to identify the optimal subset of variables with the lowest cross-validated error for bugweed detection as shown is Fig. 3. Results showed that best accuracy (CV $\text{accuracy} = 97\%$) was achieved by using a subset of 17 bands from the original 110 bands resulting in a 85% decrease in the number of bands required for analyses. As shown in Fig. 3, model accuracy subsequently decreased and remained constant from 17 bands onwards. By evaluating the CV error for each band combination in Fig. 3, it is evident that using all the bands does not improve the model's predictive accuracy. Rather there is an optimal subset of bands that produce the best

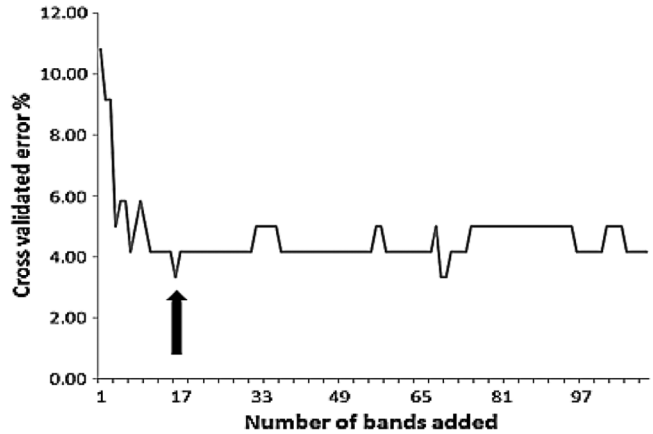

Figure 3 : Results of the backward FS method for selecting the optimal number of bands from the 110 band dataset. The arrow indicates the lowest error of 3%obtained using a linear SVM model with cost parameter of 10.

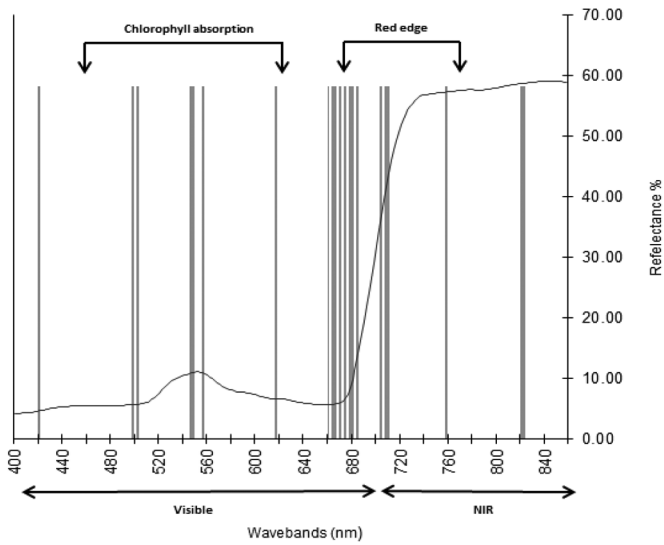

Figure 4 : Distribution of 17 optimal AISA Eagle bands in the VIS and NIR por-tions of the spectrum selected using the SVM-RFE.

accuracy. In fact, the use of more than 17 bands would subsequently result in an ensuing decrease in predictive accuracy up until 69 bands where the accuracy would again improve to 97%.

Ranked variable importance also showed that of the 17 bands selected by the model, wavelengths that had the potential to discriminate bugweed were located in the visible and near-infrared (NIR) regions of the electromagnetic spectrum (Fig. 4). One band (b_5: 419.6 nm) occurred within the blue range (350–450 nm), eight bands (b_{21}: 498.0.1 nm, b_{22}: 502.9 nm, b_{31}: 547.0 nm, b_{45}: 615.6 nm, b_{54}: 659.7 nm, b_{55}: 664.6 nm, b_{56}: 669.5 nm, b_{57}: 674.4 nm) occurred within the chlorophyll absorption regions (450–675 nm) and one band (b_{35}: 566.8 nm) was located within the yellow edge (550–582 nm). Five bands (b_{58}: 679.3 nm, b_{59}: 684.2 nm, b_{63}: 703.8 nm, b_{64}: 708.7 nm, b_{74}: 757.7 nm) were in the red-edge portion of the spectrum (670–753 nm) with the remaining two NIR bands occurring at b_{87}: 821.4 nm and b_{108}: 924.3 respectively.

Table II : Error matrices showing overall and class accuracies using independent test samples for (A) 1 1 0 B and (B) 1 7 B and datasets

(a) Error matrix of bugweed occurrence (Presence or absence)

REFERENCE DATA				
		Bugweed	**Pine**	**Row Total**
CLASSIFIED DATA	**Bugweed**	54	5	59
	Pine	6	55	61
Column Total		60	60	120
Overall Accuracy (%)		90.83		
TSS (%)		82.00		
Sensititvity (%)		90.00		
Specificity (%)		92.00		

(b) Error matrix of bugweed occurrence (Presence or absence)

REFERENCE DATA				
		Bugweed	**Pine**	**Row Total**
CLASSIFIED DATA	**Bugweed**	57	5	62
	Pine	3	55	58
Column Total		60	60	120
Overall Accuracy (%)		93.33		
TSS (%)		87.00		
Sensititvity (%)		95.00		
Specificity (%)		92.00		

The shaded areas in Fig. 4 show the wavelengths of the 17 bands that are significant for the classification of bugweed as determined by the modified SVM-RFE method. Based on the analysis, it is clear that there is a dominance of optimal of bands occurring in the red edge region of the electromagnetic spectrum.

B. Classification Accuracy

Table II shows the results of the linear SVM model using the 110 and 17 band models. The regularization parameters for the 110 and 17 band datasets which yielded the best results were 10 and 100 respectively. The sensitivity of both datasets were high ranging from 90% for the 110 band model and 95% for the 17 band model. This indicates that the proportion of correctly classified reference bugweed trees in relation to all the classified trees in the test data set was very high.

The specificity for both models was also very high at 92% signifying low errors of commission for both the 17 and 110 models and indicating a high probability that a classified pixel in the image was actually represented on the ground. The high values obtained for the TSS statistics (0.82–0.87) in both datasets approach one 1 and are an indication of good model performance showing strong agreement between the actual and predicted values for bugweed, particularly for the 17 band model. Overall accuracy for the 17 band model was slightly higher (93.3%) than the 110 band model (90.8%) indicating an increase in classification accuracy with a smaller subset of bands compared with the original 110 bands.

C. Mapping Bugweed Occurrence

After testing the performance of the SVM classifier utilizing the SVM-RFE reduced band dataset the 17 identified bands were then used to create thematic maps of bugweed occurrence. Visual interpretation of the image classification indicated some important observations. Firstly, the results showed that there was a predominance of bugweed to the north and south of the study area. Towards the north, there were fairly uniform dense bugweed present between certain compartments separated by an open or natural area as well as smaller pockets along roads between certain compartments in the south. The results of the image classification for one compartment located in the south of the study area are presented in Fig. 5. The model was able to detect the distinctive light-green canopies of well-developed bugweed clusters. In some instances the model was even able to isolate individual bugweed trees with canopies equal in size to that of the pixel resolution of the AISA Eagle imagery. What is interesting is that in this compartment the model was able to accurately detect 152 instances of bugweed and of the 152 cases detected, 55 instances were of bugweed having a canopy smaller than 5.7 m^2 (1 pixel). This result is quite significant considering that it has been suggested that weed eradication programs should also target small satellite infestations [70] in view that no early detection system can truly be operational unless it is able to detect small as well as large infestations [19].

IV. Discussion

A. Modeling Bugweed Occurrence Using SVMs

The SVM algorithm applied in this study proved to be a powerful classifier and thus seems promising for hyperspectral image classification within homogenous forestry plantations. The method allowed for the detection, classification and successful mapping of bugweed using the AISA Eagle image data. SVMs were theoretically developed for binary classification scenarios [29] and therefore their application in remote sensing, where a large majority of the land cover classifications involve more than one class (multiclass), are usually limited [63]. However, the binary classification used in the study produced excellent classification results. Other studied have also shown that SVMs can perform better when the number of classifiers, and therefore complexity of the classification problem is reduced [71].

Accuracy assessments show that the SVM algorithm is a robust and accurate method for bugweed image classification using hyperspectral imagery. Overall classification accuracies ranged from 91 to 93% whilst the TSS, an indication of model performance, ranged from 0.82 to 0.87 for all datasets tested theoretically confirming the model's applicability within an operational environment. Previous works using SVMs have also shown successful classification performance for hyperspectral data [24], [29], [31], [58], [71]–[75]. Moreover, the overall accuracies from this study seem to be superior in relation to other weed classification studies applying different classification methodologies. Among these, studies which used Gaussian Maximum Likelihood and linear discriminant analysis [76], Mixture Tuned Match Filtering [16], RandomForest [17] as well as Minimum Noise Fraction, continuum removal and band ratio indices [77] all yielded lower overall classification accuracies than the current study. The potential to use a spatially explicit model for bugweed detection, as well as other nuisance species, is further strengthened if one considers the

Figure 5 : Map showing classified bugweed within compartments after applying the SVM-RFE algorithm to the 17 band AISA Eagle dataset.

relatively high specificity and sensitivity accuracies obtained. The sensitivity and specificity accuracies obtained for the 110 and 17 band datasets are a good indication that the SVM model is not only capable of producing an accurate map of more than one vegetation class, but that the model is also suited for differentiating one particular species with distinctive sets of features, such as bugweed, from other unique species such as pine. However, particular attention was placed on the specificity, which reflects errors of commission, as this illustrates how well the model was able to detect bugweed in the study site. The low specificity consequently confirms the SVMs potential for use as a decision support tool within a vegetation management programme since the extent of bugweed would be adequately detected and mapped. However, caution is advised, as [78] points out that SVM training as well as generalisation performance is highly dependent on the type of kernel function and associated hyper-parameters used for classification. Since the accuracy of the SVM depends on the proper setting of the hyper-parameters, the main challenge for researchers is determining how best to optimize the hyper-parameters for a given application [59].

B. Optimal Waveband Selection for Bugweed Detection

An interesting result from this study was that the SVM-RFE methods selected 17 optimal bands that yielded better classification accuracies than the original 110 band dataset. The results are in contrast to both [58] and [79] who reported improved SVM importance with increased number of bands. The reason for the better performance of the 17 band dataset is that the fewer bands resulted in less noise, enabling the model to limit the use of redundant bands thus improving overall classification accuracy [29]. The results from Fig. 3 also show that the cross validated accuracy obtained at 17 bands were the same as 69 bands (97%). One of the possible reasons for this is that the 69th band occurs at 536 nm (red-green region) which is characterized by high green and red reflectance. When plants are stressed or their leaves appear yellow, as in the case of bugweed, carotenes and xanthophylls are known to be the dominant chemicals responsible for this yellow appearance as both carotenes and xanthophylls absorb blue light and reflect green and red light. [80]. The combination of the green and red light is what then gives the leaves their yellow color. The cross-validated accuracy therefore improved from 95 (66th band) to 97% when the 69th band was included in the iteration as it necessary for detecting bugweed especially if the plant is stressed or is flowering and the leaves are yellow.

More importantly, the study has shown that the 17 band dataset consists of an optimal subset of hyperspectral bands at defined wavelengths within specific regions of the electromagnetic spectrum. However, further in-depth studies would

need to be conducted to determine the potential of for using other hyperspectral or multispectral sensors that have spectral ranges that cover those necessary for bugweed detection. This is one of the reasons why it was necessary to use the entire AISA Eagle dataset covering the visible and NIR regions. There are several operationally viable platforms that could be tested to determine their potential for weed mapping within a commercial forestry plantation and these include the World View, Rapid Eye and QuickBird sensors. Consequently, variable importance also showed that of the 17 bands selected from the model, thirteen bands occurred in the visible region (400–700 nm) and four bands were in located the near-infrared portion (NIR) (700–2500 nm) of the electromagnetic spectrum. [81] points out these regions of the spectrum are defined by (i) vegetation pigment content and by (ii) plant internal structure and the importance of reflectance and shape of individual plant spectral signatures within these regions is in keeping with our understanding of the basis of spectral uniqueness between plant species [19]. Indeed, [82] comment that a large majority of agricultural studies (which include weed management) use spectral measurements in these regions to detect both physiological and biological differences between plant species and other surface features.

These spectral characteristics are of more importance in the red-edge region as this region represents absorption spectra of the visible and reflectance spectra of the NIR portions [76] and subtle differences between species in crown characteristics can show up as large differences in infrared reflectance [83]. The red-edge refers to the point of maximum slope between the red chlorophyll absorption region (680 nm) and the region of high near-infrared reflectance (750 nm) [83]. The red-edge is of significance to researchers as its exact wavelength and strength varies depending on the species considered and as such bands in this region are pivotal to plant species separation and therefore potentially essential for weed identification. Moreover, the spectral reflectance of at least two wavelength bands, usually on either side of the red-edge, enables a variety of vegetation indices to be calculated. Future studies could therefore utilise the normalized differential vegetation index [84], the red ratio vegetation index [85], the green ratio vegetation index [86] or the chlorophyll vegetation index [87] to investigate additional weed spectral and physical characteristics. These indices could be used to determine canopy characteristics and even specific weed properties [82] within a forest compartment. Reducing the dimensionality of the AISA Eagle imagery and isolating a subset of optimal bands may offer an affordable and robust alternative to multispectral systems for both airborne and satellite applications of forestry assessment. Identifying these optimal spectral bands could help forestry managers exploit other optical remote sensing platforms, such as digital multispectral imagers (DMSI), or other commercially accessible hyperspectral sensors, which operate in the desired spectral range for bugweed discrimination.

C. Forestry Management Implications

Mapping any understory invasive species is a challenging exercise. So the occurrence of understory bugweed makes detection with direct optical remote sensing techniques very difficult especially in areas with closed canopies [20]. However, in open canopies that have bugweed growing in the understory, the results from this study have demonstrated that where bugweed dominates the spectral signature, detection is possible. One of the most likely uses for regional bugweed thematic maps from a weed management approach would be to locate and track bugweed infestations and distribution within plantations. Furthermore, the methods described in this study could be used to supplement existing weeding programs or be used as a decision support tool for long term integrated weed control programs [45], [88]. More specifically, the most immediate benefit of applying the methods from this study would be to formulate a framework that prioritizes weeding activities at both the pre-plant and noxious weeding phases. [40] States that pre-plant weeding occurs before tree establishment and if carried out effectively, will not only save costs on future weeding operations but will also promote the sustainable protection of future timber plantations. Conversely, noxious weeding is necessary to mitigate IAPs that have already become established within mature plantations after post-plant weeding and require regular targeted eradication.

The synoptic identification and classification of plantation bugweed could undoubtedly be used to prioritize areas of high infestation on which to focus management and monitoring efforts. This could be done by firstly identifying the bugweed by pinpointing their locality in the plantation, establishing their extent and abundance (single trees or clusters) and then deciding on their potential impact to not only forestry resources but also forestry operations and surrounding ecotones, such as riparian areas [40]. From an ecological stand point, riparian ecosystems (the border of streams and rivers) are extremely vital for fulfilling a variety of ecosystem functions within the plantation yet they are particularly susceptible to weed invasion due to their low-lying position in the landscape and because rivers act as conduits for the dispersal of seeds [89], [90]. These areas act as focal points for further encroachment and potential spreading within the timber plantations and therefore need to be managed just as proactively as the plantation weeding regimes. Whether or not the results of this study prompts a more concerted effort to consider remote sensing technologies as operationally viable for weed management depends on the capacity and will of forestry institutions and resource managers to exploit the availability as well as access to the technology and employ the knowledge for effective use of the tools [91].

D. Challenges to Mapping the Occurrence of Bugweed in Forest Plantations

Results from this study indicate that bugweed within forest stands can be accurately mapped with hyperspectral images acquired at both a high spatial and spectral resolution. However, this comes at the expense of increased computational time and increased classification complexity due to augmented data hyper-dimensionality [76]. Currently, data availability and data cost combined with technical and specialized methodological approaches are the major limitations that persist with regards to operational hyperspectral applications in South Africa [21]. The result is that very few local studies have actually explored the potential of using hyperspectral imagery for classification [92], [94]. These factors have played a crucial role in limiting

the success of developing an operational framework for weed detection in a commercial forestry environment. As [14], points out for remote sensing technologies to be widely accepted by forest companies and the tools to be operationally feasible, methods must allow for the efficient and cost effective mapping of infestations. It should be noted however, the question is not whether one data source is superior to another (e.g. hyperspectral vs multispectral or airborne vs spaceborne) but rather under what conditions a particular sensor can provide the desired information to meet the mapping objective. Evaluating the suitability of remote sensing data for a specific mapping task should include an evaluation of geometric integrity, spatial resolution, spectral resolution, area coverage and image acquisition costs. Furthermore, each of these considerations needs to be evaluated relative to the mapping task at hand [94]. Indeed, [21] was able to show that the generally high classification errors associated with damage discrimination of pine species by *Sirex noctilio* (Eurasian woodwasp) imposes operational limitations on the use of broad band multispectral sensors by forestry companies. For that reason, even though hyperspectral image acquisition may be costly, in certain circumstances, the potential economic benefits gained from having a reliable and repeatable data source to accurately detect non-native weed species are more important than the image and processing costs [16].

V. Conclusion

The primary goal of this study was to demonstrate the utility of SVM methods to analyze high resolution hyperspectral imagery for detecting bugweed, one of the most problematic non-native, invasive species within commercial plantations. Overall, the results of the study showed that the modified SVM-RFE approach is an efficient as well as accurate method for (i) optimal band selection and (ii) detecting the presence of bugweed within mature *P. patula* compartments.

The SVM-RFE approach was able to produce high overall and class accuracies in excess of 90% by using only 17 of the original 272 AISA Eagle spectral bands. A large majority of these bands were situated within the visible and red-edge portion of the electromagnetic spectrum signifying the importance of these regions in detecting the occurrence of bugweed within commercial forestry compartments using hyperspectral imagery. The results from the study have reiterated why SVMs are particularly appealing in classifying remote sensing data. That is, they provide a timely and repeatable product for developing a framework for effective weed management in commercial forestry focusing on weed monitoring, prioritization and eradication. Considering the high overall and class classification accuracies obtained from the study the use of high spatial resolution image data for the classification of nuisance plant species as part of an integrated weed management program should be further pursued by commercial forestry institutions. There are definite management and financial benefits of high resolution weed mapping and monitoring in support of forestry management activities. Yet very few, if any, real-world methods exist to quantify weed abundance at high spatial resolution in regional commercial forest plantations. Lower image acquisition costs combined with some hyperspectral image platforms becoming more commercially accessible should hopefully reinforce hyperspectral sensors as a viable long term management tool for a variety of forestry applications.

References

[1] R. N. Pallet, "Precision forestry for pulpwood re-establishment silviculture," *Southern African Forestry J.*, vol. 203, pp. 33–40, 2005.

[2] State of the Forest Report: A Pilot Report to Test National Criteria and Indicators. Department of Water Affairs and Forestry. Pretoria, South Africa, 2005.

[3] B. A. Moore, Alien Invasive Species: Ipacts on Forests and Forestry "Forest health and biosecurity," FAO Forestry Dept.. Rome, Italy, 2005.

[4] K. Little, J. Kritzinger, and M. Maxfield*, Some Principles of Vegetation Management Explained*. Pietermaritzburg, South Africa: KwaZulu Natal: ICFR, 1997.

[5] K. Hanks, "The Conservation of Agricultural Resources Act Explained—A Legal Solution," Information Brochure, Dept. Water Affairs and Forestry. Pretoria, South Africa, 2009.

[6] E. T. F. Witkowski and R. D. Garner, "Seed production, seed bank dynamics, resprouting and long term response to clearing of the alien invasive *Solunum mauritianum* in temperate to subtropical riparian ecosystem," *South African J. Botany*, vol. 74, pp. 476–484, 2008.

[7] B. Richardson, A. Vanner, J. Ray, N. Davenhill, and G. Coker, "Mechanisms of *Pinus radiata* growth suppression by some common forest weed species," *New Zealand J. Forestry*, vol. 26, pp. 421–437, 1996.

[8] Global invasive species database. Univ. Auckland, New Zealand [Online]. Available: http://www.issg.org/database/

[9] D. M. Richardson and B. W. van Wilgen, "Invasive alien plants in South Africa: How well do we understand the ecological impacts?," *South African J. Science*, vol. 100, pp. 45–52, 2004.

[10] G. L. Anderson, J. H. Everitt, A. J. Richardson, and D. E. Escobar, "Using satellite data to map false broomweed (*Ericameria austrotexana*) infestations on South Texas rangelands," *Weed Science*, vol. 7, pp. 865–871, 1993.

[11] A. Haara and S. Nevalainen, "Detection of defoliated spruces using digital aerial data," *Forest Ecology and Management*, vol. 160, pp. 97–107, 2002.

[12] J. B. Campbell, "Introduction to remote sensing," in *Hyperspectral Remote Sensing*, 3rd ed. New York, NY, USA: Taylor & Francis, 2002, pp. 400–415.

[13] L. W. Lass and T. S. Prather, "Detecting Brazilian pepper (*Schinus terebinthifolius*) with hyperspectral remote sensing technology," *Weed Technology*, vol. 18, pp. 437–442, 2004.

[14] B. Waske, S. Van Der Linden, and J. A. Benediktsson, "Sensitivity of support vector machines to random FS in classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2880–2889, 2010.

[15] X. Miao, P. Gong, S. Swope, R. Pu, R. Carruthers, G. L. Anderson, J. S. Heaton, and C. R. Tracy, "Estimation of yellow star thistle abundance through CASI-2 hyperspectral imagery using linear spectral mixture models," *Remote Sens. Environ.*, vol. 101, pp. 329–341, 2006.

[16] N. F. Glenn, J. T. Mundt, K. T. Weber, T. S. Prather, L. W. Lass, and J. Pettingill, "Hyperspectral data processing for repeat detection of small infestations of leafy spurge," *Remote Sens. Environ.*, vol. 95, pp. 399–412, 2005.

[17] R. L. Lawrence, S. Wood, and R. L. Sheley, "Mapping invasive plants using hyperspectral imagery and Breiman Cutler classification," *Remote Sens. Environ.*, vol. 100, pp. 356–362, 2006.

[18] Y. Hamada, D. A. Stow, L. L. Coulter, J. C. Jafolli, and L. W. Hendricks, "Detecting Tamarisk species (*Tamarix spp.*) in riparian habitats of Southern California using high spatial resolution hyperspectral imagery," *Remote Sens. Environ.*, vol. 109, pp. 237–248, 2007.

[19] E. A. Andrew and S. L. Ustin, "The role of environmental context in mapping invasive plants with hyperspectral image data," *Remote Sens. Environ.*, vol. 112, pp. 4301–4317, 2008.

[20] C. Joshi, J. D. Leeuw, and I. C. van Duren, "Remote sensing and GIS applications for mapping and spatial modelling of invasive species," *Int. Society for Photogrammetry and Remote Sensing*, vol. 35, pp. 669–677, 2004.

[21] R. Ismail, O. Mutanga, and F. Ahmed*, Discriminating Sirex noctilio attack in pine forest plantations in South Africa using high spectral resolution data*. London, U.K.: Taylor and Francis: CRC Press, 2008.

[22] M. Govender, K. Chetty, V. Naiken, and H. Bulcock, "A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation," *Water South Africa*, vol. 34, pp. 147–154, 2007.

[23] R. J. Zomer, A. Trabucco, and S. L. Ustin, "Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing," *J. Environmental Management*, vol. 90, pp. 2170–2177, 2009.

[24] C. Huang, L. S. Davis, and J. R. G. Townsend, "An assessment of support vector machines for land cover classification," *Int. J. Remote Sens.*, vol. 23, no. 4, pp. 725–749, 2002.

[25] T. Kavzoglu and I. Colksen, "A kernel functions analysis for support vector machines for land cover classification," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 11, no. 5, pp. 352–359.

[26] V. N. Vapnik, *The Nature of Statistical Learning*. New York, NY, USA: Springer Verlag, 1995.

[27] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauval, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. 110–122, 2009.

[28] G. Mercier and M. Lennon, "Support vector machines for hyperspectral image classification with spectral-based kernels," in *Proc. IGARSS '03*, Toulouse, France, 2003.

[29] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, 2004.

[30] M. Pal and P. M. Mather, "Support vector machines for classification in remote sensing," *Int. J. Remote Sens.*, vol. 26, no. 5, pp. 1007–1011, 2005.

[31] G. Camps-Vells and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1–12, 2005.

[32] M. Lennon, G. Mercier, and L. Hubert-Moy, "Classification of hyperspectral images with nonlinear filtering and support vector machines," *Proc. IEEE IGARSS '02*, vol. 3, pp. 1670–1672, 2002.

[33] Y. W. Chen and C. J. Lin, "Combining SVMs with various FS strategies," *Strategies*, vol. 324, no. 1, pp. 1–10, 2006.

[34] C. Sanchez-Hernandez, D. S. Boyd, and G. M. Foody, "Mapping specific habitats from remotely sensed imagery: Support vector machine and support vector data descriptionbased classification of coastal saltmarsh habitats," *Ecological Informatics*, vol. 2, pp. 83–88, 2007.

[35] K. Dunne, P. Cunningham, and F. Azuaje, "Solutions to instability problems with sequential wrapper-based approaches to feature selection," Dept. Comput. Sci., Trinity College, Dublin, Ireland, 2002 [Online]. Available: http://www.tara.tcd.ie/bitstream/2262/13144/1/TCD-CS-2002-28.pdf

[36] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, pp. 389–422, 2002.

[37] J. A. N. Van Aardt and M. Norris-Rogers, "Spectral-age interactions in managed, even-aged Eucalyptus plantations: Application of discriminant analysis and classification and regression trees approaches to hyperspectral data," *Int. J. Remote Sens.*, no. 29, pp. 1841–1845, 2008.

[38] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, pp. 273–324, 1997.

[39] Forest Land Types of the Natal Region. Sappi Forests Research, Howick, KwaZulu Natal, South Africa, 1993.

[40] N. Dobyn, Environmental Management Plan: Sappi Forests, Hodgsons, unpublished report, 2009.

[41] H. Jan, M. Zbynek, H. Lucie, V. Kaplan, P. Lukes, and P. Cudlin, "Potentials of the VNIR airborne hyperspectral system AISA Eagle," *GIS Ostrava*, vol. 27, pp. 1–6, 2008.

[42] Galileo Group Inc. [Online]. Available: http://www.galileo-gp.com/aisa_eagle, (Accessed 14 October, 2011), 2007

[43] B. Johnson, R. Joseph, M. Nischan, A. Newbury, J. Kerekes, H. Barclay, B. Willard, and J. J. Zayhowski, "A compact, active hyperspectral imaging system for the detection of concealed targets," *Proc. SPIE*, vol. 3710, pp. 144–153, 1999.

[44] L. A. Klerk, A. Broersen, I. W. Fletcher, R. Van Liere, and R. M. A. Heeren, "Extended data analysis strategies for high resolution imaging MS: New methods to deal with extremely large image hyperspectral datasets," *Int. J. Mass Spectrometry*, vol. 260, pp. 222–236, 2007.

[45] J. M. Goodhall and D. C. Naude, "An ecosystem approach for planning sustainable management of environmental weeds in South Africa," *Agriculture, Ecosystems and Environment*, vol. 68, pp. 109–123, 1998.

[46] ENVI, The Environment for Visualizing Images, ver. 4.7, Research Systems, Inc.. Boulder, USA, 2010.

[47] D. A. Roberts, Y. Yamaguchi, and R. J. P. Lyon, "Calibration of airborne imaging spectrometer data to percent reflectance using field spectral measurements," in *Proc. 19th Int. Symp. Remote Sensing Environment*, Ann Arbor, MI, USA, pp. 21–25.

[48] F. A. Kruse, K. S. Kierein-Young, and J. W. Boardman, "Mineral mapping at Cuprite, Nevada with a 63-channel imaging spectrometer," *Photogramm. Eng. Remote Sens.*, vol. 56, pp. 83–92, 1992.

[49] B. Gao, M. J. Montes, C. O. Davis, and A. F. H. Goetz, "Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean," *Remote Sens. Environ.*, vol. 113, pp. 17–24, 2009.

[50] G. Schaepman-Strub, M. E. Schaepman, T. H. Painter, S. Dangel, and J. V. Martonchik, "Reflectance quantities in optical remote sensing—Definitions and case studies," *Remote Sens. Environ.*, vol. 103, pp. 27–42, 2006.

[51] A. Tashakori and C. Teddie, *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA, USA: Sage, 2003.

[52] L. W. Lass, T. S. Prather, N. F. Glenn, K. T. Weber, J. T. Mundt, and J. Pettingill, "A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (*Centaurea maculosa*) and babysbreath (*Gypsophila paniculata*) with a hyperspectral sensor," *Weed Science*, vol. 53, no. 2, pp. 242–251, 2005.

[53] A. Clay, G. J. Lems, D. E. Clay, F. Forcella, M. M. Ellsbury, and C. G. Carlson, "Sampling weed spatial variability on a field wide scale," *Weed Science*, vol. 47, no. 6, pp. 674–681, 1999.

[54] E. L. Hestir, S. Khanna, M. E. Andrew, M. J. Santos, J. H. Viers, J. A. Greenberg, S. S. Rajapakse, and S. L. Ustin, "Identification of invasive vegetation using hyperspectral remote sensing in the California delta ecosystem," *Remote Sens. Environ.*, vol. 112, pp. 4034–4047, 2008.

[55] ArcGis, Environmental Systems Research Institute (ESRI). ver. 10.1, ArcGis, Redlands, CA, USA, 2012.

[56] R Development Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria, 2008 [Online]. Available: URL http://www.R-project.org.

[57] V. N. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.

[58] M. Pal and P. M. Mather, "Assessment of the effectivness of support vector machines for hyperspectral data," *Future Generation Computer Systems*, vol. 20, pp. 1215–1225, 2004.

[59] V. Cherkasky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 17, pp. 113–126, 2004.

[60] M. D. Sap and M. Noor, "Support vector classification of remote sensing images using improved spectral kernels," *J. Teknologi Maklumat*, vol. 20, no. 1, pp. 14–27, 2008.

[61] C. W. Hsu, C. C. Chang, and C. J. Lin, *A Practical Guide to Support Vector Machines*. Taiwan: Department of Computer Science, National Taiwan University, 2010.

[62] K. Wu and S. Wang, "Choosing the kernel parameters for support vector machines by the inter-cluster distance in the feature space," *Pattern Recognition*, vol. 42, pp. 710–717, 2009.

[63] M. Pal, "Kernel methods in remote sensing: A review," *ISH J. Hydraulic Eng.*, vol. 15, no. 1, 2009.

[64] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognition*, vol. 43, pp. 2367–2379, 2010.

[65] S. B. Serpico and G. Moser, "Extraction of spectral channels from hyperspectral images for classification purposes," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 2, 2007.

[66] I. Guyon and A. Elisseeff, "An introduction to variable and FS," *J. Machine Learning Research*, pp. 1157–1182, 2003.

[67] A. H. Fielding and J. F. Bell, "A review of methods for the assessment of prediction errors in conservation presence/absence models," *Environmental Conservation*, vol. 24, no. 1, pp. 38–49, 1997.

[68] O. Allouche, A. Tsoar, and R. Kadmon, "Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS)," *J. Applied Ecology*, vol. 43, pp. 1223–1232, 2006.

[69] G. Banko, "A review of assessing the accuracy of classifications of remotely sensed data and of methods including remote sensing data in forest inventory," Int. Inst. Applied Systems Analysis. Laxenburg, Austria, 1998, IR-98-081.

[70] M. E. Moody and R. N. Mack, "Controlling the spread of plant invasions: The importance of nascent foci," *J. Applied Ecology*, vol. 25, no. 1009, p. 1021, 1988.

[71] Y. Karimi, S. O. Prasher, R. M. Patel, and S. H. Kim, "Application of support vector machines technology for weed and nitrogen stress detection in corn," *Computers and Electronics in Agriculture*, vol. 51, pp. 99–109, 2006.
[72] M. Chi, R. Feng, and L. Bruzzonne, "Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem," *Advances in Space Research*, vol. 41, pp. 1793–1799.
[73] P. Watanachaturaporn, P. K. Varshney, and M. K. Arora, "Evaluation of factors affecting support vector machines for hyperspectral classification.," in *Proc. ASPRS Annual Conf.*, Denver, CA, USA, May 23–28, 2004.
[74] Y. Bazi and F. Melgani, "Toward an optimal SVM classification system for hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, pp. 3374–3385, 2006.
[75] Q. Guo, M. Kelly, and C. H. Graham, "Support vector machines for predicting distribution of Sudden Oak Death in California," *Ecological Modelling*, vol. 182, pp. 75–90, 2005.
[76] M. Dalponte, L. Bruzzone, L. Vescovo, and D. Gianelle, "The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas," *Remote Sens. Environ.*, vol. 13, pp. 2345–2355, 2009.
[77] E. Underwood, S. Ustin, and D. Dipietro, "Mapping non-native plants using hyperspectral imagery," *Remote Sens. Environ.*, vol. 86, pp. 150–161, 2003.
[78] R. Damasevicius, "Optimization of SVM parameters for promoter recognition in DNA Sequence," in *Int. 20th EURO Mini Conf., Continuous Optimization and Knowledge-Based Technologies*, 2008.
[79] X. Ceamanos, B. Waske, J. A. Benediktsson, J. Chanussot, and J. R. Sviensson*, Ensemble Strategies for Classifying Hyperspectral Remote Sensing Data*. Berlin, Heidelberg: Springer-Verlag, 2009.
[80] L. Kumar, K. Schmidt, S. Dury, and A. Skidmore, "Imaging Spectrometry and Vegetation Science," in *Image Spectrometry, Basic Principles and prospective applications*. Dordrecht, The Netherlands: Springer Netherlands, 2006, ch. 5.
[81] P. Ceccato, N. Gobton, S. Flasse, B. Pinty, and S. Tarantola, "Designing a spectra index to estimate vegetation water content from remote sensing data: Part 1, theoretical approach," *Remote Sens. Environ.*, vol. 82, pp. 188–197, 2002.
[82] I. M. Scotford and P. C. H. Miller, "Applications of spectral reflectance techniques in northern European cereal production: A review," *Biosyst. Eng.*, vol. 90, no. 3, pp. 235–250, 2005.
[83] R. Butler and R. Schlaepfer, "Spruce snag quantification by coupling colour infrared aerial photos and a GIS," *Forest Ecology and Management*, vol. 195, pp. 325–339, 2004.
[84] J. V. Stafford and H. C. Bolam, "Near-ground and aerial radiometry imaging for assessing spatial variability in crop condition," in *Proc. 4th Int. Conf. Precision Agriculture*, Minneapolis, MN, USA, Jul. 1998.
[85] R. H. Biller, "Reduced input of herbicides by use of optoelectronic sensors," *J. Agricultural Engineering Research*, vol. 71, pp. 357–362, 1998.
[86] R. Oberti and J. De Baerdemaeker, "Assessing the nitrogen status of plants by optical measurements," presented at the EuroAgEng 2000 Conf., Warwick, U.K., Jul. 2–7, 2000.
[87] A. A. Gitelson, Y. J. Kaufman, and M. N. Merzlyak, "Use of a green channel in remote sensing of global vegetation from EOS-MODIS," *Remote Sens. Environ.*, vol. 58, pp. 289–292, 1996.
[88] J. M. Goodhall and D. J. Erasmus, "Review of the status and integrated control of the invasive alien weed, *chromolaena odorata*, in South Africa," *Agriculture, Ecosystem and Environment*, vol. 56, pp. 151–164, 1996.
[89] J. T. Du Toit, K. H. Rogers, and H. C. Biggs*, The Kruger Experience: Ecology and Management of Savannah Heterogeneity*. Washington, DC, USA: Island Press, 2003.
[90] D. N. Richardson, P. M. Holmes, K. J. Eslar, S. M. Galatowitsch, J. C. Stromberg, S. P. Kirkman, P. Pysek, and R. J. Hobbs, "Riparian vegetation: Degradation, alien plant invasions, restoration prospects," *Diversity and Distributions*, vol. 13, pp. 126–139, 2007.
[91] A. Elizabeth *et al.*, "Science priorities for reducing the threat of invasive species to sustainable forestry," *Bioscience*, vol. 55, no. 4, pp. 335–348, 2005.
[92] O. Mutanga and A. Skidmore, "Discriminating tropical grass canopies grown under different nitrogen treatments using spectra resampled to HYMAP," *Int. J. Geoinformatics*, vol. 1, no. 2, pp. 21–32, 2005.
[93] O. Mutanga and A. Skidmore, "Integrating imaging spectrometry and neural networks to map grass quality in the Kruger National Park, South Africa," *Remote Sens. Environ.*, vol. 90, no. 1, pp. 104–115, 2004.
[94] M. Madden, "Remote sensing and geographic information system operations for vegetation mapping of invasive exotics," *Weed Technology*, vol. 18, pp. 1457–1463, 2004.

Jonathan Tom Atkinson received the B.Sc. degree in applied environmental science from the University of KwaZulu Natal, South Africa, in 2006, the B.Sc. honors degree in applied environmental science in 2007 and the M.Sc. in environment and society from the University of Pretoria, South Africa, in 2012.

He is currently a GIS Analyst with Sappi Southern Africa Pty Ltd-Forest Division, where he is responsible for addressing a variety of ad hoc as well as sector specific GIS and Remote Sensing queries. His research interests are in the area of strategic environmental and sustainability management specifically related to natural resources.

Riyad Ismail received the Master degree in GIS (*cum laude*) and the Ph.D degree in remote sensing from the University of KwaZulu-Natal, South Africa.

He has over 10 years of experience in implementing spatial technologies (GIS, GPS, and remote sensing) at academic and research institutions. He was recently appointed as a research fellow at the University of KwaZulu-Natal, South Africa, and is currently an analyst for a commercial forestry organization. His research focuses on maximizing the benefit of spatial technologies by using machine learning algorithms.

Mark Robertson received the Ph.D. degree from Rhodes University, Grahamstown, South Africa.

He is an Associate Professor in the Department of Zoology and Entomology at the University of Pretoria, South Africa. His research interests include biological invasions and understanding species distributions.