MDPI

Article

Mapping Invasive Herbaceous Plant Species with Sentinel-2 Satellite Imagery: *Echium plantagineum* in a Mediterranean Shrubland as a Case Study

Patricia Duncan [1,2,3,*] , Erika Podest [4], Karen J. Esler [2,3] , Sjirk Geerts [5] and Candice Lyons [2,6]

1   Department of Civil Engineering and Geomatics, Cape Peninsula University of Technology,
Cape Town 8000, South Africa
2   C.I.B Centre for Invasion Biology, Stellenbosch University, Stellenbosch 7602, South Africa
3   Department of Conservation Ecology and Entomology, Faculty of Agrisciences, Stellenbosch University,
Stellenbosch 7602, South Africa
4   Carbon Cycle and Ecosystems Group, Jet Propulsion Laboratory, California Institute of Technology,
Pasadena, CA 91109, USA
5   Department of Conservation and Marine Sciences, Cape Peninsula University of Technology,
Cape Town 8000, South Africa
6   Biological Sciences Department, University of Cape Town, Cape Town 7701, South Africa
*   Correspondence: duncanp@cput.ac.za

Abstract: Invasive alien plants (IAPs) pose a serious threat to biodiversity, agriculture, health, and economies globally. Accurate mapping of IAPs is crucial for their management, to mitigate their impacts and prevent further spread where possible. Remote sensing has become a valuable tool in detecting IAPs, especially with freely available data such as Sentinel-2 satellite imagery. Yet, remote sensing methods to map herbaceous IAPs, which tend to be more difficult to detect, particularly in shrubland Mediterranean-type ecosystems, are still limited. There is a growing need to detect herbaceous IAPs at a large scale for monitoring and management; however, for countries or organizations with limited budgets, this is often not feasible. To address this, we aimed to develop a classification methodology based on optical satellite data to map herbaceous IAP's using *Echium plantagineum* as a case study in the Fynbos Biome of South Africa. We investigate the use of freely available Sentinel-2 data, use the robust non-parametric classifier Random Forest, and identify the most important variables in the classification, all within the cloud-based platform, Google Earth Engine. Findings reveal the importance of the shortwave infrared and red-edge parts of the spectrum and the importance of including vegetation indices in the classification for discriminating *E. plantagineum.* Here, we demonstrate the potential of Sentinel-2 data, the Random Forest classifier, and Google Earth Engine for mapping herbaceous IAPs in Mediterranean ecosystems.

Keywords: invasive alien plants; vegetation; remote sensing; Sentinel-2; Google Earth Engine; Random Forest; image classification; land cover

1. Introduction

Invasive Alien Plants (IAP) pose a significant threat to many ecosystems [1] and agriculture globally [2]. The growing number of invasive plant species, which include agricultural weeds, compromise ecosystem stability and threaten economic productivity [3–6]. The cost and impact of IAPs is extensive, ranging from reducing grazing land and poisoning livestock [7] to the loss of native species and lowering biodiversity [8–11].

Reliable invasive species distribution maps are important for targeting management of infestations, modelling future invasion risk [12], assisting with policy decisions, strategic allocation of funding, effective implementation of control programmes [13–16] and planning for ecological restoration [17,18]. Traditional mapping of weeds is time-consuming, rapidly outdated, labor-intensive and expensive, and relying heavily on field-based surveys has

associated challenges such as dealing with inaccessible and expansive areas [19]. Remote sensing may offer an alternative approach to the traditional method of on-the-ground weed mapping [20–22]. The application of remote sensing has become increasingly beneficial for mapping and monitoring IAPs due to improved computing capabilities, advanced machine learning algorithms, new sensor technologies and data availability [23–25]. Remote sensing allows for cost-effective, timely, reliable, and repeatable observations at varying landscape scales over time [26]. High spatial and temporal resolution multispectral satellite imagery has contributed to numerous applications of mapping invasive plant species; however, most of these studies have focused on woody species, for example, *Prosopis* in India [27], Bugweed (*Solanum mauritianum*) in South Africa [28], Tamarisk (*Tamarix* spp.) in the USA [29], *Lantana camara* in India [30] as well as in South Africa [31] and *Acacia mearnsii* and *A. dealbata* in South Africa [32]. Despite the many studies mapping invasive alien plant species, detecting herbaceous IAPs still poses a challenge due to the similar spectral signatures of surrounding plant species [33,34]. In addition, herbaceous IAP species are generally smaller than woody invasives, they may be scattered, and they often grow amongst similar vegetation, thus making their detection difficult with multispectral satellite imagery alone [23,25,35].

Even though a large proportion (49%) of remote sensing studies are related to herbaceous invaders, many of these studies include field measurements, airborne (both traditional aircraft and UAV), spaceborne platforms, and incorporate radar, multispectral and hyperspectral data [36]. Multispectral satellite data have been used for the detection of herbaceous IAPs in a range of environments, including *Heracleum mantegazzianum* within plantations, grasslands and pastures in the Czech Republic [37], *Fallopia* spp. in the floodplains of France [38], *Striga hermonthica* in croplands in Kenya [39] and *Parthenium hysterophoru* L. in tropical bush, savannas and forests of South Africa [40]. It is unknown whether these methods can be applied in shrubby Mediterranean ecosystems in which alien herbaceous species might be less conspicuous than in forests, grasslands, or agricultural areas.

South Africa has one of the biggest problems of alien plant invasions globally, with the largest number of invasive plant species occurring in the Fynbos Biome [3,41–43]. Applications of remote sensing to map invasive species within the Fynbos Biome have primarily focused on invasive trees such as those from the genera *Eucalyptus*, *Pinus* and *Acacia* that impact riparian zones and mountain slopes, and pose a threat to water security [44,45]. Within the Fynbos Biome, *E. plantagineum* (Paterson's curse) with its purple flowers is perhaps the most visually obvious invasive annual herbaceous weed; however, the potential impacts it has on the environment, animal health, and the economy have not been quantified. Although *E. plantagineum* is widespread within the Fynbos Biome, the above-ground presence is highly variable in space and over time, which contributes to the challenge of estimating its abundance and extent [46,47]. To date, the species' presence has been mapped using data gathered from roadside surveys [48], which provides a broad-level indication of its distribution at a quarter-degree square (QDS) mapping level [49,50]. The most recent South African Plant Invaders Atlas (SAPIA) information indicates that the distribution of *E. plantagineum* increased from 51 QDS in the year 2000 to 119 QDS in 2016, and it is listed as one of the plant species that has shown the greatest increase in range, as well as reaching "very high local abundances" [51].

Echium plantagineum is a global invasive species of disturbed areas, pastures and cultivated agricultural lands [52]. Originating from the Mediterranean region of Europe, the species has invaded large parts of Australia, New Zealand, North and South America, as well as South Africa, particularly within the Fynbos Biome [51]. In Australia, it is one of the most costly weeds due to its distribution and abundance [7]. The species reduces grain yield, contributes to soil erosion, contaminates grain harvests, and decreases pasture productivity and carrying capacity by competing with more beneficial pasture species [7]. *Echium plantagineum* is estimated to cost the Australian economy AUD 250 million annually [53], impacting the sheep and cattle industry through pasture land degradation, management

costs and contamination of wool by seeds [54]. The species contains pyrrolizidine alkaloids, which are toxic to grazing animals and may result in their death if eaten over prolonged periods [55,56]. *Echium plantagineum* is a prolific seed producer with heavy infestations yielding up to 10,000 seeds per square meter [47,57] which may remain in the soil for up to six years [52,58]. To date, very little research has been conducted on detecting *E. plantagineum* with multispectral satellite imagery. Ullah et al. [59] conducted satellite remote sensing studies using Landsat Thematic Mapper (TM) to detect dense and flowering *E. plantagineum* in Australia; however, since there was no accuracy assessment provided, it is not possible to quantify how successfully the species was detected. In another study, McIntyre [60] determined that *E. plantagineum* could not be distinguished from pastures and crops in high-resolution aerial imagery; however, the species could be more reliably detected and a higher accuracy achieved using EO-1 Hyperion satellite hyperspectral imagery [60,61]. Hyperspectral imagery has been used successfully to detect other herbaceous IAPs such as Musk thistle (*Carduus Nutans*) [62] and Hoary cress (*Cardaria draba*) [63] while using multispectral satellite imagery for detecting herbaceous IAPs such as giant hogweed (*Heracleum mantegazzianum*), knotweed (*Fallopia* sp.) and Striga (*Striga hermonthica*), with mixed results [37–39].

While the most common approach for species detection is based on differences in spectral signatures, phenological approaches, which are based on seasonal events such as flowering, have proven to be effective in detecting some invasive plant species [12,64,65]. Distinct phenological patterns can provide opportunities for remote detection if the invasive species under consideration has a different seasonal or inter-annual growth pattern compared to its co-occurring species [12,35], or floral characteristics that may assist in spectrally discriminating a plant species from its background such as flowering *Acacia salicina* and *Acacia saligna* trees [64], giant hogweed (*Heracleum mantegazzianum*), knotweeds (*Fallopia* spp.) [35] and *E. plantagineum* [59,60]. High temporal resolution satellite sensors such as Sentinel-2 that capture images every few days have the capability of detecting a species in near real time, allowing for opportune interventions and timely species monitoring and management [21,39,66,67]. In terms of supporting long-term and large-scale mapping of IAPs, especially in financially constrained countries, it is necessary to explore freely available multispectral datasets such as Sentinel-2 in conjunction with advanced machine learning algorithms [23].

Other studies in Mediterranean-type ecosystems aimed at detecting IAP species using remote sensing include mostly trees, but also some shrubs and grasses [68–72]. Detection of herbaceous IAPs with remote sensing in a Mediterranean ecosystem using only freely available multispectral satellite imagery has not been performed, probably partly because the success of a remote sensing analysis declines as site complexity increases [73]. The case of *E. plantagineum* within the Fynbos Biome of South Africa provides an ideal study system to address this question. Here, we aim to develop a classification methodology to detect herbaceous IAPs using the cloud-based platform, Google Earth Engine (GEE), a Random Forest classifier, Sentinel-2 data and the species *E. plantagineum* as a case study. Specifically, we aim to determine:

(1) Can Sentinel-2 imagery be used to discriminate the invasive alien plant species, *E. plantagineum*?
(2) What variables are important in detecting *E. plantagineum* in satellite imagery?

2. Materials and Methods

2.1. Study Area

Our study area is situated within three municipalities: the City of Cape Town, Stellenbosch and Drakenstein. These municipal areas are located within the Fynbos Biome which is part of the Cape Floristic Region (CFR) of South Africa. The CFR, with its high native species richness and levels of endemism, is a biodiversity hotspot [74,75]; it is the smallest, but most biologically diverse of all the world's six plant kingdoms, and is recognized as a UNESCO world heritage site [76,77]. The Fynbos Biome is one of the most species-rich

floristic regions in the world [78]; however, it is also the most invaded of all biomes in South Africa [3,41]. The Fynbos Biome is characterized by a Mediterranean-type climate with hot, dry summers and mild wet winters, and has a mean annual precipitation of 480 mm [79]. Fynbos, the dominant vegetation within the Fynbos Biome, is a low growing shrubby vegetation characterized by its fine leaved evergreen shrubs [42,80]. Study sites within the Fynbos Biome with known occurrences of *E. plantagineum* that could easily be accessed and sampled were selected (Figure 1).

Copernicus Sentinel-2 data (2020).
Retrieved from https://scihub.copernicus.eu

0   5   10   15 km
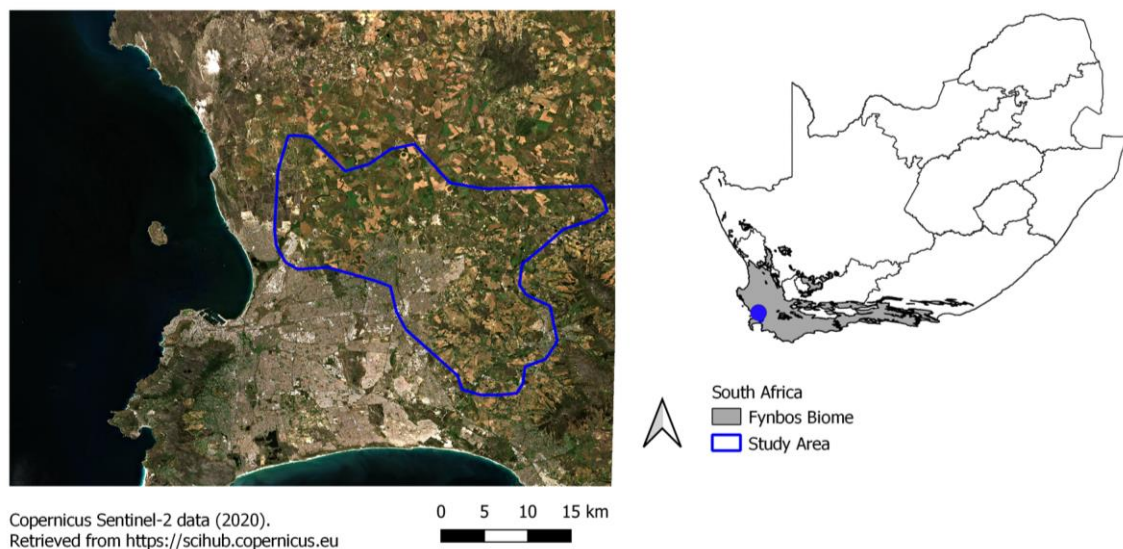
Figure 1. Study area (718 km^2) within the Fynbos Biome of South Africa. Copernicus Sentinel-2 data displayed as an RGB image.

2.2. Study Species

Echium plantagineum is an erect, softly hairy annual, occasionally biennial, herbaceous plant that often occurs in pastures along with other pasture species, weeds, and grasses, but also inhabits road verges, abandoned fields and natural areas in the Fynbos Biome [47], and can grow between 20 and 60 cm tall [81]. The species is an aggressive, drought-tolerant IAP that adapts to many conditions and soil types [58] and readily inhabits fallow land and disturbed areas [7,54]. The main flowering period is from late winter and can last between three and five months [82]. In our study region, the species usually flowers around early or mid-spring (September–October) and lasts for two–three months, but this is variable depending on moisture availability and temperature [58]. During this time, distinctive purple flowers can often be seen in expansive tracts of land where *E. plantagineum* exists. In most instances, however, the species occurs as scattered patches of varying size, and in narrow stretches along roadsides where roadside management (mowing) helps to maintain its presence by encouraging regrowth [57] and removing competition from other species. Seedlings usually emerge and develop as rosettes in late winter, then produce stems which elongate and flower in spring; these plants usually die in the summer months; however, plants growing near areas with a high moisture content (e.g., roadsides, rivers, dams) may continue to flower during summer and die in late summer or early autumn [58].

2.3. Training and Validation Data

The process of image classification requires ground truth data to train and verify model results [83]. Reference data for *E. plantagineum* presence samples were collected using a handheld Garmin 62S Global Positioning System with an accuracy of 3 m. Presence samples were collected during the period of 9–14 October 2020 to capture the species presence during its peak flowering stage, as it is hypothesized that the flowering species is spectrally discernible from co-occurring plant species. Sites of data collection consisted of grazing lands and open fields with *E. plantagineum* present within our study area. The

species commonly occurs with other herbaceous species, graminoids, and bare ground, thus often making their detection in imagery challenging. At the study sites, *E. plantagineum* had more than 50% cover and was the only profusely flowering purple species at the time of data collection. *Echium plantagineum* sample points were collected as individual points where the species dominated and as close to the centre of a plot or field as possible. Since Sentinel-2 has a spatial resolution of 10–20 m, sample points were buffered to 10 m to include surrounding pixels of the class, yet maintaining sample precision by keeping within 20 m. In addition, random sample points were assigned to representative land cover classes present in the Sentinel-2 image; these other land cover classes included bare ground, built-up, agriculture, shrubland and water. Validation points were selected from the labelled samples by randomly splitting the samples into 70% training and 30% validation data [84–86] within GEE (Table 1).

Table 1. Land cover classes and corresponding total number and area of all samples, training samples and number of validation samples.

	Number of Samples		
Land Cover Class	**Training**	**Validation**	**Total**
Bare ground	682	323	1005
Built-up	907	371	1278
Agriculture	725	279	1004
Shrubland	721	317	1038
Water	1006	361	1367
Echium plantagineum	256	119	375
Total number of samples	4297	1170	6067
Area of samples	0.43 km^2 (43 ha)	0.18 km^2 (18 ha)	0.61 km^2 (61 ha)

2.4. Image Acquisition and Processing

The Copernicus Sentinel-2 mission is a constellation of two identical optical satellites, Sentinel-2A and Sentinel-2B, which were launched by the European Space Agency (ESA) in June 2015 and March 2017, respectively [87]. Sentinel-2 satellites provide free and global coverage of multispectral (13 spectral bands) and medium–high spatial resolution (10–60 m) imagery with a temporal resolution of 5 days between the two satellites combined. Furthermore, Sentinel-2 provides novel spectral capabilities with two near infrared (NIR) bands (NIR and NIR narrow), three red-edge bands and two shortwave infrared (SWIR) bands, which are valuable for vegetation detection [87].

Sentinel-2 Level-2A Surface Reflectance imagery for 12 October 2020 was used for the study as this date for imagery coincided with the period when presence samples were collected for *E. plantagineum*. Image bands 1, 9 and 10 were excluded as they are designed for atmospheric features and provide no land cover information. All remaining image bands (Table 2) were then clipped to the extent of the study region.

Sixteen vegetation indices (Table 2) were computed from the Sentinel-2 imagery using the visible, near-infrared (NIR) and red-edge image bands within GEE. These indices were selected based on their ability to separate vegetation from non-vegetation, to account for bare soil, atmospheric influences, and the usefulness of the NIR and red-edge indices for vegetation mapping [20,88]. The most commonly used vegetation index for separating vegetation from non-vegetation and for vegetation vitality and stress is the Normalized Difference Vegetation Index (NDVI) [89]. The Enhanced Vegetation Index (EVI) has proven useful for vegetation classification to simultaneously correct for atmospheric influences and soil background [90]. The Normalized Difference Red Edge Index (NDRE) is similar to NDVI, but uses the ratio of Near-Infrared and the Red Edge, and is known to be effective in detecting healthy vegetation [91]. One of the earliest vegetation indexes for vegetation analysis is the Simple Ratio (SR), which is the ratio formed by spectral ratioing of the NIR band with the red band. The greater the amount of healthy green vegetation present in a pixel, the larger the contrast between NIR and red values, which results in higher NIR-to-Red ratios [92].

Table 2. Sentinel-2 image bands and indices used. Bands + VIs: all variables used (image bands and indices). Number of trees: 160; number of variables: 26. Most important bands + VIs: most important variables used (best image bands and indices). Number of trees: 30; number of variables: 13. Bands: Sentinel-2 only image bands used. Number of trees: 50; number of variables: 10.

Sentinel-2 Image Band or Index	Description of Image Band or Equation of Index	Reference	Bands + VIs	Most Important Bands + VIs	Bands
B2	Blue		x	x	x
B3	Green		x	x	x
B4	Red		x		x
B5	Vegetation Red Edge 1		x	x	x
B6	Vegetation Red Edge 2		x	x	x
B7	Vegetation Red Edge 3		x	x	x
B8	Near Infrared (NIR) wide		x	x	x
B8A	Near Infrared (NIR) narrow		x	x	x
B11	Short Wave Infrared (SWIR) 1		x	x	x
B12	Short Wave Infrared (SWIR) 2		x	x	x
ARVI	(NIR - (2 × Red) + Blue)/(NIR + (2 × Red) + Blue)	[93]	x		
DVI	NIR - Red	[94]	x		
EVI	2.5 × ((NIR - Red)/((NIR) + (6 × Red) - (7.5 × Blue) + 1))	[95]	x		
EVI2	2.5 × ((NIR - Red)/(NIR + 2.4 × Red + 1)	[96]	x		
GDVI	NIR - Green	[94]	x	x	
GNDVI	(NIR - Green)/(NIR + Green)	[97]	x		
IPVI	NIR/(NIR + Red)	[98]	x		
NDGI	(Green - Red)/(Green + Red)	[94]	x		
NDRE	(NIR - Red Edge)/(NIR + Red Edge)	[99]	x	x	
NDVI	(NIR - Red)/(NIR + Red)	[89]	x		
NFRE	(Red Edge - Red)/(Red Edge + Red)	[60]	x	x	
RI	(Red - Green)/(Red + Green)	[100]	x	x	
SAVI	((NIR - Red)/(NIR + Red + L)) × (1 + L). Where L = 0.5	[101]	x		
SR	NIR/Red	[102]	x		
VARI	(Green - Red)/(Green + Red - Blue)	[103]	x		
VDVI	(2 × Green - Red - Blue)/(2 × Green + Red + Blue)	[104]	x		

The Soil Adjusted Vegetation Index (SAVI) corrects for the influence of soil background and is included in this study. In addition, we assessed the inclusion of the Atmospherically Resistant Vegetation Index (ARVI), which is similar to NDVI, but minimizes the effects of atmospheric scattering [90,105]. Furthermore, we included the Green Normalized Difference Vegetation Index (GNDVI), which is like NDVI except that it measures the green spectrum instead of the red spectrum. This index is more sensitive to chlorophyll concentration than NDVI. We also included the Infrared Percentage Vegetation Index (IPVI); this index is the same as NDVI functionally but is computationally faster. Additionally, we included the Normalized Flower Red-Edge (NFRE) as proposed by [60] for flowering *E. plantagineum* and adapted it to Sentinel-2 image bands for testing.

2.5. Image Classification—Random Forest

Based on the success of the Random Forest classifier in detecting other invasive plant species [38,40,64,106,107], we used this classifier in our case study. The Random Forest classifier provides an assessment of the importance of different variables during the classification process; variables or features (e.g., image bands and vegetation indices) with a higher importance value have a greater contribution to classification accuracy. However, variable importance is sensitive to the Random Forest method parameters *number of trees* and *variables per split*. Choosing optimal parameters for the Random Forest classifier can improve stability of variable importance scores and help to discriminate between important and unimportant variables [108]. Hyperparameters are those parameters that are specified by the user to control the training process in the Random Forest algorithm. In this case, the *number of trees* needs to be explicitly defined within GEE, while other parameters, if not specified, use default values. The *number of trees* parameter was incrementally tested within GEE from 10 to 1000 to determine the optimal number of trees which provided the best overall accuracy. The parameter *variables per split* used the square root of the number of variables; the *minimum leaf population*, *bag fraction* and *seed* were set to the default values of 1, 0.5 and 0, respectively, while no limits were set to the maximum number of leaf nodes (*max. nodes*) (Table 3).

Table 3. Description and default values of Random Forest parameters specified within Google Earth Engine [109].

Parameter	Default Value	Description
Number of Trees	No default value. Must be specified	The number of decision trees to create.
Variables Per Split	Square root of the number of variables	The number of variables per split.
Min. Leaf Population	1	Only create nodes whose training set contains at least this number of points.
Bag Fraction	0.5	The fraction of input to bag per tree.
Max. Nodes	No limit	The max. number of leaf nodes in each tree.
Seed	0	The randomization seed.

An advantage of the Random Forest classifier is that the algorithm itself can be used for feature selection [110]. Feature selection methods are used to select a subset of variables from the original variables; the subset variables are those determined to be most valuable for predicting classes in the classification [111,112]. Feature selection provides information on which variables (image bands and vegetation indices) are most relevant for a classification and ranks them in order of importance [110]. By selecting the most optimal variables and removing bands which may be highly correlated, one can avoid a bias in the results. The reduction in data dimension means that processing can be simplified and the performance of the classifier improved [113]. Here, the Random Forest algorithm was used to rank those variables from the set of image bands (10 image bands) and indices (16 indices) (Table 2) which were determined to be most useful in discriminating target classes [110]. The Random Forest algorithm provides a measure of variable importance based on the Mean Decrease in Gini (MDG) value to identify predictor variables. The MDG measures how much a variable reduces the Gini Impurity within a particular class; the larger the MDG value, the purer the variable, and thus the more important the variable [114,115]. The MDG was used to calculate variable importance.

Three classification approaches were evaluated in this study: (a) using Sentinel-2 image bands + VIs, (b) using the most important image bands and VIs, and (c) using Sentinel-2 image bands only. The variables used in the first analysis (Bands and VIs) were all the proposed Sentinel-2 image bands and indices based on their ability to separate vegetation from non-vegetation, to account for bare soil background, atmospheric influences, and their usefulness in vegetation mapping. The second analysis was based on the best variables (top 50%) calculated from the first analysis, while the third analysis was based on using only Sentinel-2 image bands for comparison with the other two analyses (Table 3). Each classification was based on the Random Forest algorithm which was fine-tuned to identify the optimal number of trees used in each classification.

We generated a validation dataset for the accuracy assessment by randomly splitting our samples into 70% for training and 30% for validation to avoid any bias in the results. The accuracy of the classification was assessed using the 30% validation samples within GEE. Classifications were assessed using an error matrix approach, which included the overall accuracy (OA) and Kappa coefficient [116], as well as the user accuracy (UA) and producer accuracy (PA) for each class.

3. Results

We determined that by fine-tuning the hyperparameter to select the optimal number of trees, the overall accuracy could be increased by up to 1.7% (Bands and VIs). When comparing our optimal number of trees to that of a typical moderate size of 100 trees, the accuracy was increased by 0.4% (Figure 2).

The overall classification accuracy for varying number of trees in the Random Forest classification for three different classifications was high overall (Figure 2). The optimized value of 160 trees yielded the best overall classification accuracy at 93% when using all 26 variables (Bands and VIs); fewer trees slightly lowered the overall accuracy (1% decrease in overall accuracy at 20 trees), while with more trees the accuracy would plateau at the expense of computational efficiency. When using only the most important variables (Most

important bands and VIs), the optimal number of trees for the Random Forest classifier was 30, whereas when using only Sentinel-2 image spectral bands (Bands), the optimal number of trees was 50.

Results suggest that all 26 variables (image bands and vegetation indices) were considered important in terms of MDG; however, some variables scored significantly higher than others. The most important variables were the two Sentinel-2 SWIR image bands (B11 and B12 respectively), followed by the red-edge bands (B5 and B6) and the narrow NIR band (B8A). The Green Difference Vegetation Index (GDVI), Red Index (RI), Normalized Flower Red-Edge (NFRE) and the Normalized Difference Red-Edge (NDRE) proved to be the most significant indices evaluated in the classification (Figure 3).

The classification results indicate that using all Bands and VIs outperformed both the analysis using the most important bands and VIs and the analysis using only image bands marginally with an overall accuracy (OA) of 93.11% compared to 92.66% and 92.77%, respectively, while that of the kappa statistic (K) was also highest for the analysis using Bands and VIs (0.92) compared to using only the most important bands + VIs and image bands alone which both achieved marginally lower results (0.91).

The accuracy for *E. plantagineum* ranged between 81.05% and 85.90% for the Producer's Accuracy (PA), while the User's Accuracy (UA) ranged between 56.30% and 64.71% (Table 4). The UA of *E. plantagineum* increased by 8.40% when including vegetation indices in the classification compared to that of using only Sentinel-2 image bands. Similarly, the UA of *E. plantagineum* increased by 6.72% when using image bands and indices (Bands and VIs) compared to using only the top variables (Most important bands and VIs), highlighting the importance of Sentinel-2 image bands and derived vegetation indices for detecting IAPs. Although the classification that relied on only Sentinel-2 image bands achieved a high overall accuracy, Kappa, and producer's accuracy, the inclusion of vegetation indices was important for increasing the user's accuracy. Across all three cases, it was determined that a higher proportion of pixels was misclassified as either agriculture or shrubland compared to other land cover classes, resulting in a lower accuracy for *E. plantagineum* (Table 4). The classification result of the analysis based on Bands and VIs is shown in Figure 4.
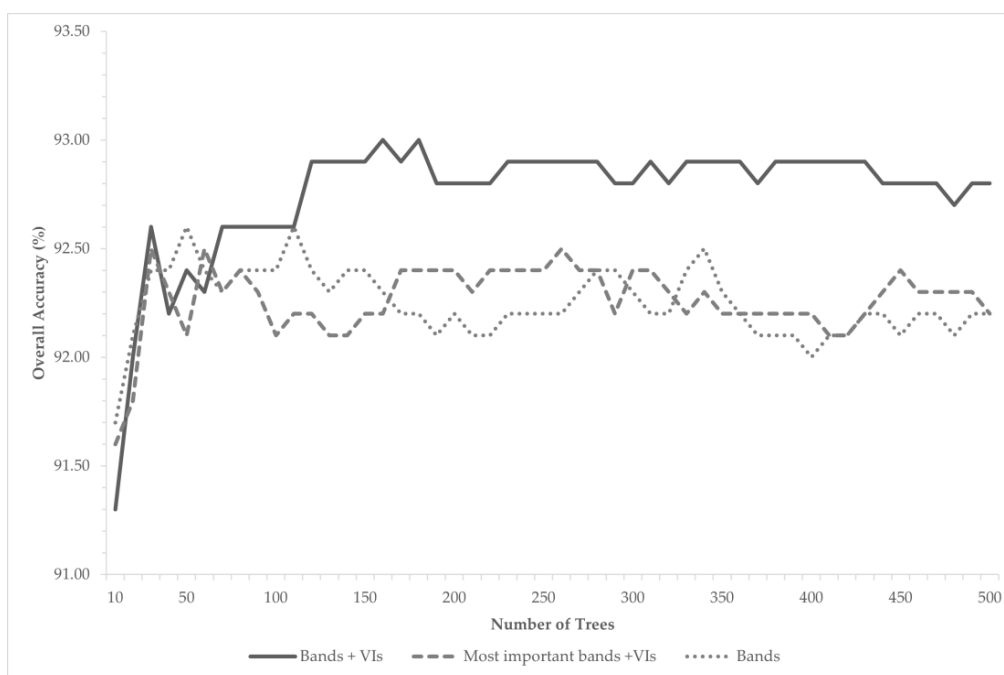
Figure 2. Hyperparameter optimization for number of trees used in the Random Forest model from 10 to 500 trees in increments of 10 for all image bands and indices (Bands + VIs), the important variables (Most important bands + VIs), and only Sentinel-2 image bands (Bands). Refer to Table 2 for description of variables.
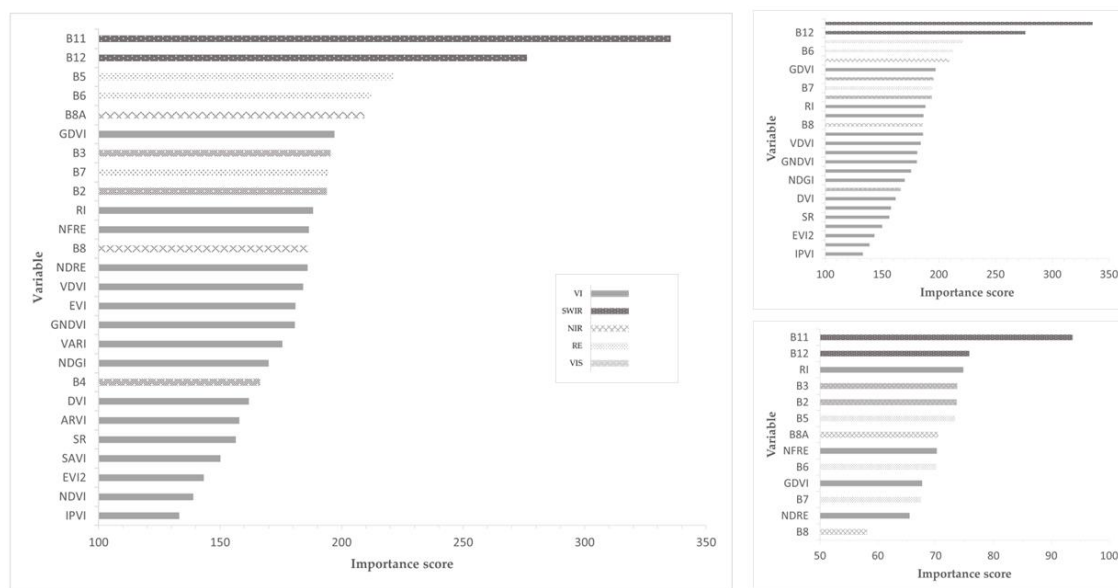
Figure 3. Variable importance as measured by the Random Forest algorithm indicating the usefulness of the different image bands and indices. The higher the value, the higher the predictor importance. **Left**: Variable importance for all proposed variables (Bands + VIs). **Top right**: Variable importance for best predictor variables (Most important bands + VIs). **Bottom right**: Variable importance for Sentinel-2 image bands only (Bands). Refer to Table 2 for description of variables.

Table 4. Confusion matrices for Random Forest classification. Classes are Bare ground (B), Built-up (BU), Agriculture (AG), Shrubland (SH), Water (W), *Echium plantagineum* (EP), Overall accuracy: OA, Kappa statistic: K, User's accuracy: UA, and Producer's accuracy: PA. Bands + VIs: all variables used (image bands and indices). Most important bands and VIs: best image bands and indices; Bands: Sentinel-2 image bands only. Refer to Table 2 for description of variables.

Bands + VIs						
Class	B	BU	AG	SH	W	EP	UA (%)
B	304	19	0	0	0	0	94.12
BU	8	357	5	0	0	1	96.23
AG	0	3	251	16	0	9	89.96
SH	0	3	1	300	5	8	94.64
W	0	0	0	2	359	0	99.45
EP	2	1	20	19	0	77	64.71
PA (%)	96.82	93.21	90.61	89.02	98.63	81.05	
OA (%)			**93.11**				
K			**0.92**				

Most Important Bands + VIs						
Class	B	BU	AG	SH	W	EP	UA (%)
B	299	21	0	3	0	0	92.57
BU	7	358	5	0	0	1	96.50
AG	0	6	250	14	0	9	89.61
SH	0	4	5	299	5	4	94.32
W	0	0	0	2	359	0	99.45
EP	2	2	22	18	0	75	63.03
PA (%)	97.08	91.56	88.65	88.99	98.63	84.27	
OA (%)			**92.66**				
K			**0.91**				

Bands						
Class	B	BU	AG	SH	W	EP	UA (%)
B	306	17	0	0	0	0	94.74
BU	12	353	4	1	0	1	95.15
AG	0	2	258	13	0	6	92.47
SH	0	3	5	300	5	4	94.64
W	0	0	0	3	358	0	99.17
EP	2	1	32	17	0	67	56.30
PA (%)	95.63	93.88	86.29	89.82	98.62	85.90	
OA (%)			**92.77**				
K			**0.91**				
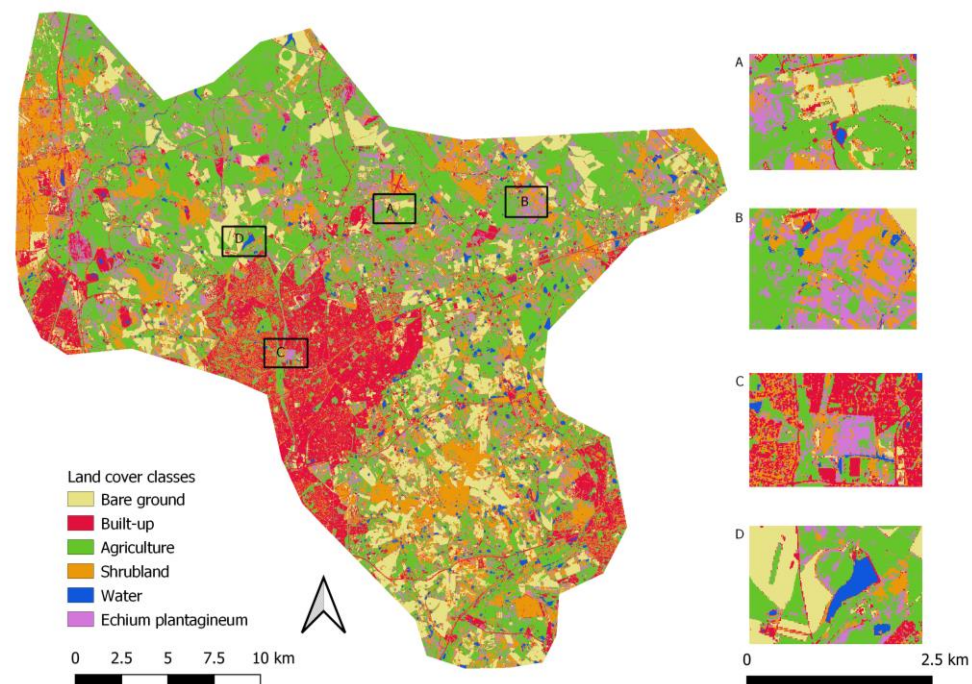
Figure 4. The final classification result (Bands + VIs) for classes Bare ground, Built-up, Agriculture, Shrubland, Water and *Echium plantagineum*. Blocks A, B, C and D are random enlarged parts of the map to show the different land cover classes in more detail.

4. Discussion

The main objective of this study was to test the capabilities of freely available multispectral Sentinel-2 imagery and the cloud-based platform, Google Earth Engine, for detecting herbaceous IAPs in a Mediterranean-type ecosystem. This case study focuses on identifying the species *E. plantagineum* in the Fynbos Biome. The machine learning algorithm, Random Forest with Sentinel-2 image bands and indices within Google Earth Engine were used and results highlight the most important variables used in the classification. Additionally, the Random Forest parameter for the number of trees was fine-tuned and optimized to increase classification accuracy. Even though default values of the Random Forest model usually work reasonably well for image classification [117], we confirmed that fine-tuning the hyperparameter for the number of trees improved the performance of the Random Forest classifier by up to 1.7%. Similarly, hyperparameter optimization increased classification accuracy by 2–3% compared to using default hyperparameters [118].

The most important variables for discrimination of classes were identified using the Random Forest algorithm for feature selection. The most important variables identified were the shortwave infrared, red-edge and near infrared image bands, which is similar to what others have discovered in systems examining invasive trees or shrubs within the South African context [44,45,119]. Similarly, the red-edge and shortwave infrared regions were the most significant for discriminating differences between C3 and C4 grass species using Sentinel-2 imagery [120]. Several other studies have also determined the red-edge band to be critical in distinguishing vegetation classes and plant species [121–123]. This is often ascribed to the fact that this portion of the electromagnetic spectrum is sensitive to plant characteristics such as leaf pigmentation, water content and leaf structure [121,124]. Sentinel-2 has three red-edge bands, all of which were deemed important. Additionally, the vegetation indices that incorporated the red-edge band (B5) were two of the most valuable vegetation indices. Vegetation indices are commonly used as input to classification schemes and have been shown to help differentiate classes [72,125]. Part of this effort was to explore the value in using vegetation indices to classify *E. plantagineum*. The Normalized Flower Red-Edge (NFRE) has been applied to determine a relationship with floral cover of

E. plantagineum originally using hyperspectral imagery [60]; we adapted the index to the closest multispectral image bands of Sentinel-2 data, red-edge (B5) and red (B4) bands, and determined it to be one of the most important vegetation indices to distinguish the species, thus demonstrating the importance of phenological characteristics (such as flowering) for detecting IAP species and confirming that this range of the spectrum is relevant to detecting flowering *E. plantagineum* [61].

We determined that the UA and PA varied with variable importance (Table 4), and that using the most important variables resulted in the PA for *E. plantagineum* increasing by 3.22%, while the UA decreased slightly by 1.68%. Although the producer's accuracy of *E. plantagineum* was 81.05%, showing an omission error of 18.95%, the user's accuracy (64.71%) indicates a commission error of 35.29%, revealing an overestimation when classifying the species. The lower accuracy for the species in all three classifications can be attributed to the difficulty in discriminating *E. plantagineum* from agriculture and shrubland, which may coexist with the species, as well as the fact that the species is commonly interspersed with other vegetation, and that small patches can remain undetected in 10 m spatial resolution imagery. The accuracy may be increased by selecting training for different levels of pureness of *E. plantagineum* to identify the level of accuracy that one can classify *E. plantagineum* from pure homogeneous patches to different levels of mixed patches. In the Australian study to detect *E. plantagineum* using 1 m resolution aerial imagery (red, green, blue and NIR bands), it was determined that the species could not be differentiated from pastures and crops; additionally, the authors reported high errors of commission and omission for the species [60]. The result of our producer's accuracy for *E. plantagineum* was similar to that of the study that achieved 83% using satellite hyperspectral imagery EO-1 Hyperion to detect *E. plantagineum* [60,61]; this is likely due to the difference in spatial resolution of the two sensors. Since EO-1 Hyperion has a spatial resolution of 30 m, there is a higher likelihood of having heterogeneous patches and hence lower classification accuracy; however, the availability of 49 spectral bands may have increased the classification accuracy. Sentinel-2 has a spatial resolution of 10–20 m and thus a lower likelihood of heterogeneous patches, while the low number of bands would have a lower likelihood of detection as opposed to a higher number of bands.

Mapping invasive plant species can be challenging in diverse Mediterranean ecosystems [70,126]; furthermore, the limitations experienced when mapping herbaceous IAPs can be attributed to the varying patch size and density of the species with lower density and smaller patches remaining undetected [35]. In this study, *E. plantagineum* presence samples collected covered a minimum area of 10 m^2 and had at least 50% density coverage; however, the species is commonly interspersed with other herbaceous plants and grasses. Other studies reported that cover densities of IAPs need to be sufficiently high for their detection [28,127,128]. For example, Bugweed (*Solanum mauritianum*) was best detected in WorldView-2 multispectral satellite data when cover density was greater than 76% [28], and spotted knapweed (*Centaurea maculosa*) was detectable using hyperspectral imagery where cover densities were more than 70% and cover was larger than 0.1 ha [128]. Furthermore, the inclusion of regular time series data may improve the ability to the map invasive species [129], since *E. plantagineum* may be in bloom at different times at different sites due to different characteristics of sites (e.g., varying amounts of disturbance or water availability), thus affecting the phenology of the species and hence its detection. Future research may also include assessing the probability of each classified pixel and using this as a measure of its uncertainty and to improve accuracy [115,130]. Alternatively, testing the down scaling of spatial resolution of satellite data using deep learning might be a potential solution (see, for example, [131]).

5. Conclusions

Here, we show that using Sentinel-2 image bands and derived indices in a Random Forest classification holds promise for detecting herbaceous IAP's such as *E. plantagineum*. However, the field samples collected for the species had a minimum density cover of

50%, which may have resulted in lower density patches remaining undetected or being misclassified. For future research, it is recommended that more samples with more than 70% coverage [28,127] of *E. plantagineum* are included within a minimum patch size of 20–30 m^2 for Sentinel-2 imagery. Nevertheless, our results are valuable for initial screening purposes for targeting management of dense infestations and modelling future invasion risk. The results could also be used to supplement the South African Plant Invaders Atlas, which provides a very broad-level indication of the distribution of the species. The classification model may be useful for historical analysis to determine inter-annual drivers of abundance. It is recommended that future research focus on utilizing the classification model and applying it to Sentinel-2 image data from previous years, captured during the same phenological stage (flowering) of *E. plantagineum*, and compare results with those presented here to provide an indication of change in the species over time. This may assist in explaining the relative roles of climate vs. management in determining density or abundance of the species.

References

1. Pyšek, P.; Jarošík, V.; Hulme, P.E.; Pergl, J.; Hejda, M.; Schaffner, U.; Vilà, M. A global assessment of invasive plant impacts on resident species, communities and ecosystems: The interaction of impact measures, invading species' traits and environment. *Glob. Change Biol.* **2012**, *18*, 1725–1737. [CrossRef]
2. Paini, D.R.; Sheppard, A.W.; Cook, D.C.; De Barro, P.J.; Worner, S.P.; Thomas, M.B. Global threat to agriculture from invasive species. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 7575–7579. [CrossRef]
3. Richardson, D.M.; van Wilgen, B.W. Invasive alien plants in South Africa: How well do we understand the ecological impacts? *S. Afr. J. Sci.* **2004**, *100*, 45–52.
4. de Lange, W.J.; van Wilgen, B.W. An economic assessment of the contribution of biological control to the management of invasive alien plants and to the protection of ecosystem services in South Africa. *Biol. Invasions* **2010**, *12*, 4113–4124. [CrossRef]
5. van Wilgen, B.W.; de Lange, W.J. The Costs and Benefits of Biological Control of Invasive Alien Plants in South Africa. *Afr. Entomol.* **2011**, *19*, 504–514. [CrossRef]
6. van Wilgen, B.W.; Fill, J.M.; Baard, J.; Cheney, C.; Forsyth, A.T.; Kraaij, T. Historical costs and projected future scenarios for the management of invasive alien plants in protected areas in the Cape Floristic Region. *Biol. Conserv.* **2016**, *200*, 168–177. [CrossRef]
7. Sheppard, A.W.; Smyth, M. *Echium plantagineum* L.—Paterson's curse. In *Biological Control of Weeds in Australia*; Julien, M.H., McFadyen, R.E., Cullen, J., Eds.; CSIRO Publishing: Clayton, VIC, Australia, 2012; ISBN 9780643099937.
8. Pimentel, D.; McNair, S.; Janecka, J.; Wightman, J.; Simmonds, C.; O'Connell, C.; Wong, E.; Russel, L.; Zern, J.; Aquino, T.; et al. Economic and environmental threats of alien plant, animal, and microbe invasions. *Agric. Ecosyst. Environ.* **2001**, *84*, 1–20. [CrossRef]

9. Monaco, T.J.; Weller, S.C.; Ashton, F.M. *WEED SCIENCE Principles and Practices*; John Wiley & Sons, Inc.: New York, NY, USA, 2002; ISBN 0471370517.

10. Pyšek, P.; Richardson, D.M. Invasive Species, Environmental Change and Management, and Health. *Annu. Rev. Environ. Resour.* **2010**, *35*, 25–55. [CrossRef]

11. Erckie, L.; Adedoja, O.; Geerts, S.; van Wyk, E.; Boatwright, J.S. Impacts of an invasive alien Proteaceae on native plant species richness and vegetation structure. *S. Afr. J. Bot.* **2022**, *144*, 332–338. [CrossRef]

12. Bradley, B.A. Remote detection of invasive plants: A review of spectral, textural and phenological approaches. *Biol. Invasions* **2014**, *16*, 1411–1425. [CrossRef]

13. Van den Berg, E.C.; Kotze, I.; Beukes, H. Detection, Quantification and Monitoring of Prosopis in the Northern Cape Province of South Africa using Remote Sensing and GIS. *S. Afr. J. Geomat.* **2013**, *2*, 68–81.

14. Geerts, S.; Rossenrode, T.; Irlich, U.M.; Visser, V. Emerging Ornamental Plant Invaders in Urban Areas—*Centranthus ruber* in Cape Town, South Africa as a Case Study. *Invasive Plant Sci. Manag.* **2017**, *10*, 322–331. [CrossRef]

15. Afonso, L.; Esler, K.; Gaertner, M.; Geerts, S. The invasive alien *Hypericum canariense* in South Africa: Management, cost, and eradication feasibility. *S. Afr. J. Bot.* **2022**, *146*, 685–694. [CrossRef]

16. Matthys, C.; Jubase, N.; Visser, V.; Geerts, S. Distribution of *Melaleuca rugulosa* (Schlechtendal ex Link) Craven (Myrtaceae) in South Africa: Assessment of invasiveness and feasibility of eradication. *S. Afr. J. Bot.* **2022**, *148*, 228–237. [CrossRef]

17. du Plessis, N.S.; Rebelo, A.J.; Richardson, D.M.; Esler, K.J. Guiding restoration of riparian ecosystems degraded by plant invasions: Insights from a complex social-ecological system in the Global South. *Ambio* **2021**, *51*, 1552–1568. [CrossRef] [PubMed]

18. Holmes, P.M.; Esler, K.J.; Gaertner, M.; Geerts, S.; Hall, S.A.; Nsikani, M.M.; Richardson, D.M.; Ruwanza, S. Biological Invasions and Ecological Restoration in South Africa. In *Biological Invasions in South Africa*; van Wilgen, B.W., Measey, J., Richardson, D.M., Wilson, J.R., Zengeya, T.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; pp. 665–700.

19. Lopez, R.D.; Frohn, R.C. *Remote Sensing for Landscape Ecology: Monitoring, Modeling, and Assessment of Ecosystems*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2017; ISBN 9781315152714.

20. Zwiggelaar, R. A review of spectral properties of plants and their potential use for crop/weed discrimination in row-crops. *Crop. Prot.* **1998**, *17*, 189–206. [CrossRef]

21. Lamb, D.W.; Brown, R.B. Remote-sensing and mapping of weeds in crops. *J. Agric. Eng. Res.* **2001**, *78*, 117–125. [CrossRef]

22. Thorp, K.R.; Tian, L.F. A review on remote sensing of weeds in agriculture. *Precis. Agric.* **2004**, *5*, 477–508. [CrossRef]

23. Royimani, L.; Mutanga, O.; Odindi, J.; Dube, T.; Matongera, T.N. Advancements in satellite remote sensing for mapping and monitoring of alien invasive plant species (AIPs). *Phys. Chem. Earth* **2019**, *1*, 237–245. [CrossRef]

24. Vaz, A.S.; Alcaraz-Segura, D.; Vicente, J.R.; Honrado, J.P. The Many Roles of Remote Sensing in Invasion Science. *Front. Ecol. Evol.* **2019**, *7*, 370. [CrossRef]

25. Huang, C.; Asner, G.P. Applications of remote sensing to alien invasive plant studies. *Sensors* **2009**, *9*, 4869–4889. [CrossRef]

26. Campbell, J.B.; Wyynne, R.H. *Introduction to Remote Sensing*; Guilford Press: New York, NY, USA, 2011; ISBN 9781609181765.

27. Vidhya, R.; Vijayasekaran, D.; Ramakrishnan, S.S. Mapping invasive plant *Prosopis juliflora* in arid land using high resolution remote sensing data and biophysical parameters. *Indian J. Geo-Mar. Sci.* **2017**, *46*, 1135–1144.

28. Peerbhay, K.; Mutanga, O.; Lottering, R.; Bangamwabo, V.; Ismail, R. Detecting bugweed (*Solanum mauritianum*) abundance in plantation forestry using multisource remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2016**, *121*, 167–176. [CrossRef]

29. Evangelista, P.H.; Stohlgren, T.J.; Morisette, J.T.; Kumar, S. Mapping invasive tamarisk (*Tamarix*): A comparison of single-scene and time-series analyses of remotely sensed data. *Remote Sens.* **2009**, *1*, 519–533. [CrossRef]

30. Kimothi, M.M.; Anitha, D.; Vasistha, H.B.; Soni, P.; Chandola, S.K. Remote sensing to map the invasive weed, *Lantana camara* in forests. *Trop. Ecol.* **2010**, *51*, 67–74.

31. Oumar, Z. Assessing the utility of the SPOT 6 sensor in detecting and mapping *Lantana camara* for a community clearing project in KwaZulu-Natal, South Africa. *S. Afr. J. Geomat.* **2016**, *5*, 214–226. [CrossRef]

32. Masemola, C.; Cho, M.A.; Ramoelo, A. Sentinel-2 time series based optimal features and time window for mapping invasive Australian native Acacia species in KwaZulu Natal, South Africa. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *93*, 102207. [CrossRef]

33. Matongera, T.N.; Mutanga, O.; Dube, T.; Sibanda, M. Detection and mapping the spatial distribution of bracken fern weeds using the Landsat 8 OLI new generation sensor. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *57*, 93–103. [CrossRef]

34. Kiala, Z.; Mutanga, O.; Odindi, J.; Masemola, C. Optimal window period for mapping Parthenium weed in South Africa, using high temporal resolution imagery and the ExtraTrees classifier. *Biol. Invasions* **2021**, *23*, 2881–2892. [CrossRef]

35. Müllerová, J.; Brůna, J.; Bartaloš, T.; Dvořák, P.; Vítková, M.; Pyšek, P. Timing Is Important: Unmanned Aircraft vs. Satellite Imagery in Plant Invasion Monitoring. *Front. Plant Sci.* **2017**, *8*, 887. [CrossRef]

36. Vaz, A.S.; Alcaraz-Segura, D.; Campos, J.C.; Vicente, J.R.; Honrado, J.P. Managing plant invasions through the lens of remote sensing: A review of progress and the way forward. *Sci. Total Environ.* **2018**, *642*, 1328–1339. [CrossRef] [PubMed]

37. Müllerová, J.; Pergl, J.; Pyšek, P. Remote sensing as a tool for monitoring plant invasions: Testing the effects of data resolution and image classification approach on the detection of a model plant species *Heracleum mantegazzianum* (giant hogweed). *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *25*, 55–65. [CrossRef]

38. Martin, F.-M.M.; Müllerová, J.; Borgniet, L.; Dommanget, F.; Breton, V.; Evette, A. Using Single- and Multi-Date UAV and Satellite Imagery to Accurately Monitor Invasive Knotweed Species. *Remote Sens.* **2018**, *10*, 1662. [CrossRef]

39. Mudereri, B.T.; Dube, T.; Adel-Rahman, E.M.; Niassy, S.; Kimathi, E.; Khan, Z.; Landmann, T. A comparative analysis of PlanetScope and Sentinel-2 space-borne sensors in mapping Striga weed using Guided Regularised Random Forest classification ensemble. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 701–708. [CrossRef]

40. Royimani, L.; Mutanga, O.; Odindi, J.; Zolo, K.S.; Sibanda, M.; Dube, T. Distribution of *Parthenium hysterophoru L.* with variation in rainfall using multi-year SPOT data and random forest classification. *Remote Sens. Appl. Soc. Environ.* **2019**, *13*, 215–223. [CrossRef]

41. Henderson, L. Invasive, naturalized and casual alien plants in southern Africa: A summary based on the Southern African Plant Invaders Atlas (SAPIA). *Bothalia* **2007**, *37*, 215–248. [CrossRef]

42. Rebelo, A.G.; Boucher, C.; Helme, N.; Mucina, L.; Rutherford, M.C. Fynbos Biome. In *Vegetation Map of South Africa, Lesotho and Swaziland*; Mucina, L., Rutherford, M.C., Eds.; South African National Biodiversity Institute: Pretoria, South Africa, 2006; pp. 52–219.

43. van Wilgen, B.W.; Measey, J.; Richardson, D.M.; Wilson, J.R.; Zengeya, T.A. (Eds.) *Biological Invasions in South Africa*; Springer International Publishing: Cham, Switzerland, 2020; ISBN 978-3-030-32393-6.

44. Holden, P.B.; Rebelo, A.J.; New, M.G. Mapping invasive alien trees in water towers: A combined approach using satellite data fusion, drone technology and expert engagement. *Remote Sens. Appl. Soc. Environ.* **2021**, *21*, 100448. [CrossRef]

45. Mtengwana, B.; Dube, T.; Mkunyana, Y.P.; Mazvimavi, D. Use of multispectral satellite datasets to improve ecological understanding of the distribution of Invasive Alien Plants in a water-limited catchment, South Africa. *Afr. J. Ecol.* **2020**, *58*, 709–718. [CrossRef]

46. Nel, J.L.; Richardson, D.M.; Rouget, M.; Mgidi, T.N.; Mdzeke, N.; Le Maitre, D.C.; van Wilgen, B.W.; Schonegevel, L.; Henderson, L.; Neser, S. A proposed classification of invasive alien plant species in South Africa: Towards prioritizing species and areas for management action. *S. Afr. J. Sci.* **2004**, *100*, 53–64.

47. Sharma, G.P.; Esler, K.J. Phenotypic plasticity among *Echium plantagineum* populations in different habitats of Western Cape, South Africa. *S. Afr. J. Bot.* **2008**, *74*, 746–749. [CrossRef]

48. Henderson, L. Mapping of Invasive Alien Plants: The Contribution of the Southern African Plant Invaders Atlas (SAPIA) to Biological Weed Control. *Afr. Entomol.* **2011**, *19*, 498–503. [CrossRef]

49. Henderson, L. *Alien Weeds and Invasive Plants: A Complete Guide to Declared Weed Invaders of South Africa*; Plant Protection Research Institute: Pretoria, South Africa, 2001.

50. Henderson, L. The Southern African Plant Invaders Atlas (SAPIA) and Its Contribution to Biological Weed Control. 1999, pp. 159–163. Available online: https://www.arc.agric.za/arc-ppri/News%20Articles%20Library/Henderson.pdf (accessed on 1 December 2022).

51. Henderson, L.; Wilson, J.R.U. Changes in the composition and distribution of alien plants in South Africa: An update from the Southern African Plant. *Bothalia* **2017**, *47*, 1–26. [CrossRef]

52. Grigulis, K.; Sheppard, A.W.; Ash, J.E.; Groves, R.H. The comparative demography of the pasture weed *Echium plantagineum* between its native and invaded ranges. *J. Appl. Ecol.* **2001**, *38*, 281–290. [CrossRef]

53. Government of Western Australia. Paterson's Curse: What You Should Know. 2022. Available online: https://www.agric.wa.gov.au/biological-control/patersons-curse-what-you-should-know (accessed on 16 September 2022).

54. Hulting, A.; Krenz, J.; Parker, R. *Paterson's Curse in the Pacific Northwest. A Pacific Northwest Extension*; Oregon State University: Corvallis, OR, USA, 2007.

55. Piggin, C.M. The nutritive value of *Echium plantagineum* L. and *Trifolium subterraneum* L. *Weed Res.* **1977**, *17*, 361–365. [CrossRef]

56. Smyth, M.J.; Sheppard, A.W.; Swirepik, A. The effect of grazing on seed production in *Echium plantagineum*. *Weed Res.* **1997**, *37*, 63–70. [CrossRef]

57. Piggin, C.M. Flowering and seed production of *Echium plantagineum* L. *Weed Res.* **1978**, *18*, 83–87. [CrossRef]

58. Piggin, C.M.; Sheppard, A.W. *Echium plantagineum* L. In *The Biology of Australian Weeds Vol 1*; Groves, R.H., Shepherd, R.C.H., Richardson, R.G., Eds.; R.G. and F.J. Richardson: Melbourne, VIC, Australia, 1995.

59. Ullah, E.; Shepherd, R.C.H.; Baxter, J.T.; Peterson, J.A. Mapping flowering Paterson's curse (*Echium plantagineum*) around Lake Hume, north eastern Victoria, using Landsat TM data. *Plant Prot. Q.* **1989**, *4*, 155–157.

60. McIntyre, D.L. *Application of High Resolution Remote Sensing to Detect and Map the Pasture Weed Paterson's Curse (Echium plantagineum) in Western Australia*; Curtin University: Bentley, WA, Australia, 2015.

61. McIntyre, D.L.; Corner, R.J. Using EO-1 Hyperion satellite hyperspectral imagery to detect the pasture weed Paterson's curse (*Echium plantagineum* L.) in southern Western Australia. In Proceedings of the Twentieth Australasian Weeds Conference, Perth, WA, Australia, 11–15 September 2016; pp. 196–199.

62. Mirik, M.; Ansley, R.J.; Steddom, K.; Jones, D.C.; Rush, C.M.; Michels, G.J.; Elliott, N.C. Remote distinction of a noxious weed (Musk Thistle: *Carduus Nutans*) using airborne hyperspectral imagery and the support vector machine classifier. *Remote Sens.* **2013**, *5*, 612–630. [CrossRef]

63. Mundt, J.T.; Glenn, N.F.; Weber, K.T.; Prather, T.S.; Lass, L.W.; Pettingill, J. Discrimination of hoary cress and determination of its detection limits via hyperspectral image processing and accuracy assessment techniques. *Remote Sens. Environ.* **2005**, *96*, 509–517. [CrossRef]

64. Paz-Kagan, T.; Silver, M.; Panov, N.; Karnieli, A. Multispectral Approach for Identifying Invasive Plant Species Based on Flowering Phenology Characteristics. *Remote Sens.* **2019**, *11*, 953. [CrossRef]

65. Singh, K.K.; Chen, Y.H.; Smart, L.; Gray, J.; Meentemeyer, R.K. Intra-annual phenology for detecting understory plant invasion in urban forests. *ISPRS J. Photogramm. Remote Sens.* **2018**, *142*, 151–161. [CrossRef]

66. Houborg, R.; McCabe, M.F. Daily retrieval of NDVI and LAI at 3 m resolution via the fusion of CubeSat, Landsat, and MODIS data. *Remote Sens.* **2018**, *10*, 890. [CrossRef]

67. Matongera, T.N.; Mutanga, O.; Dube, T.; Lottering, R.T. Detection and mapping of bracken fern weeds using multispectral remotely sensed data: A review of progress and challenges. *Geocarto Int.* **2018**, *33*, 209–224. [CrossRef]

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

69. Lourenço, P.; Teodoro, A.C.; Gonçalves, J.A.; Honrado, J.P.; Cunha, M.; Sillero, N. Assessing the performance of different OBIA software approaches for mapping invasive alien plants along roads with remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *95*, 102263. [CrossRef]

70. Große-Stoltenberg, A.; Hellmann, C.; Thiele, J.; Werner, C.; Oldeland, J. Early detection of GPP-related regime shifts after plant invasion by integrating imaging spectroscopy with airborne LiDAR. *Remote Sens. Environ.* **2018**, *1*, 780–792. [CrossRef]

71. Kattenborn, T.; Lopatin, J.; Förster, M.; Braun, A.C.; Fassnacht, F.E. UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data. *Remote Sens. Environ.* **2019**, *227*, 61–73. [CrossRef]

72. Große-Stoltenberg, A.; Hellmann, C.; Werner, C.; Oldeland, J.; Thiele, J. Evaluation of continuous VNIR-SWIR spectra versus narrowband hyperspectral indices to discriminate the invasive *Acacia longifolia* within a mediterranean dune ecosystem. *Remote Sens.* **2016**, *8*, 334. [CrossRef]

73. Andrew, M.E.; Ustin, S.L. The role of environmental context in mapping invasive plants with hyperspectral image data. *Remote Sens. Environ.* **2008**, *112*, 4301–4317. [CrossRef]

74. Holmes, P.M.; Rebelo, A.G.; Dorse, C.; Wood, J. Can Cape Town's unique biodiversity be saved? Balancing conservation imperatives and development needs. *Ecol. Soc.* **2012**, *17*, 2. [CrossRef]

75. Afonso, L.; Esler, K.J.; Gaertner, M.; Geerts, S. Comparing invasive alien plant community composition between urban, peri-urban and rural areas; the city of Cape Town as a case study. In *Urban Ecology: Emerging Patterns and Social-Ecological Systems*; Verma, P., Singh, P., Singh, R., Rashubanshi, A.S., Eds.; Elsevier: Amsterdam, The Netherlands, 2020; ISBN 978-0-12-820730-7.

76. Davison, A.; Marshak, M. *State of Environment Report*; Communication Department, City of Cape Town: Cape Town, South Africa, 2012.

77. Moran, V.C.; Hoffmann, J.H. Conservation of the fynbos biome in the Cape Floral Region: The role of biological control in the management of invasive alien trees. *BioControl* **2012**, *57*, 139–149. [CrossRef]

78. Goldblatt, P.; Manning, J.C. Plant Diversity of the Cape Region of Southern Africa. *Ann. Mo. Bot. Gard. Press* **2002**, *89*, 281–302. [CrossRef]

79. Mucina, L.; Rutherford, M.C. (Eds.) *The Vegetation of South Africa, Lesotho and Swaziland*; Strelitzia 19, South African National Biodiversity Institute: Pretoria, South Africa, 2006.

80. Manning, J.C.; Paterson-Jones, C. *Field Guide to Fynbos*, 2nd ed.; Struik Nature: Cape Town, South Africa, 2018.

81. Piggin, C.M. The herbaceous species of Echium (Boraginaceae) naturalized in Australia. *Muelleria* **1977**, *3*, 215–244. [CrossRef]

82. IAC. *Biological Control of Echium Species (Including Paterson's Curse/Salvation Jane)*; Industries Assistance Commission Report No. 371; Australian Government Publishing Service: Canberra, ACT, Australia, 1985.

83. Holloway, J.; Mengersen, K. Statistical machine learning methods and remote sensing for sustainable development goals: A review. *Remote Sens.* **2018**, *10*, 1365. [CrossRef]

84. Adam, E.; Mureriwa, N.; Newete, S. Mapping *Prosopis glandulosa* (mesquite) in the semi-arid environment of South Africa using high-resolution WorldView-2 imagery and machine learning classifiers. *J. Arid Environ.* **2017**, *145*, 43–51. [CrossRef]

85. Abdel-Rahman, E.M.; Mutanga, O.; Adam, E.; Ismail, R. Detecting *Sirex noctilio* grey-attacked and lightning-struck pine trees using airborne hyperspectral data, random forest and support vector machines classifiers. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 48–59. [CrossRef]

86. Congalton, R.G. Accuracy assessment and validation of remotely sensed and other spatial information. *Int. J. Wildl. Fire* **2001**, *10*, 321–328. [CrossRef]

87. ESA. Sentinel Online. 2021. Available online: https://sentinel.esa.int/web/sentinel/home (accessed on 31 January 2021).

88. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* **2008**, *1*, 9–23. [CrossRef]

89. Rouse, J.W., Jr.; Hass, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the great plains with ERTS. *Third Earth Resour. Technol. Satell. Symp.* **1973**, *1*, 309–317.

90. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *2017*, 1353691. [CrossRef]

91. Barnes, E.M.; Clarke, T.R.; Richards, S.E.; Colaizzi, P.D.; Haberland, J.; Kostrzewski, M.; Waller, P.; Choi, C.; Riley, E.; Thompson, T.; et al. Coincident detection of crop water status, nitrogen status, and canopy density using ground-based multispectral data. In Proceedings of the Fifth International Conference on Precision Agriculture, American Society of Agronomy, Bloomington, MN, USA, 16–19 July 2000.

92. Chuvieco, E.; Huete, A. *Fundamentals of Satellite Remote Sensing*; CRC Press: Boca Raton, FL, USA, 2009; ISBN 9781420021516.

93. Kaufman, Y.; Tanre, D. Atmospherically resistant vegetation index. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 260–271. [CrossRef]

94. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]

95. Liu, H.Q.; Huete, A. Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 457–465. [CrossRef]

96. Jiang, Z.; Huete, A.R.; Kim, Y.; Didan, K. 2-band enhanced vegetation index without a blue band and its application to AVHRR data. *Remote Sens. Model. Ecosyst. Sustain. IV* **2007**, *6679*, 667905.

97. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]

98. Crippen, R.E. Calculating the vegetation index faster. *Remote Sens. Environ.* **1990**, *34*, 71–73. [CrossRef]

99. Gitelson, A.; Merzlyak, M.N. Quantitative estimation of chlorophyll–A using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B Biol.* **1994**, *22*, 247–252. [CrossRef]

100. Escadafal, R.; Huete, A. Improvement in remote sensing of low vegetation cover in arid regions by correcting vegetation indices for soil noise. *CR Académie Des Sci. Paris* **1991**, *312*, 1385–1391.

101. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]

102. Jordan, C.F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **1969**, *50*, 663–666. [CrossRef]

103. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel Algorithms for Remote Estimation of Vegetation Fraction. *Remote Sens. Environ.* **2002**, *1*, 76–87. [CrossRef]

104. Wang, X.; Wang, M.; Wang, S.; Wu, Y. Extraction of Vegetation Information from Visible Unmanned Aerial Vehicle Images. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 152–159.

105. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A review of vegetation indices. *Remote Sens. Rev.* **1995**, *13*, 95–120. [CrossRef]

106. Ndlovu, H.S.; Sibanda, M.; Odindi, J.; Buthelezi, S.; Mutanga, O. Detecting and mapping the spatial distribution of *Chromoleana odorata* invasions in communal areas of South Africa using Sentinel-2 multispectral remotely sensed data. *Phys. Chem. Earth* **2022**, *126*, 103081. [CrossRef]

107. Makori, D.; Abdel-Rahman, E.M.; Landmann, T.; Mutanga, O.; Odindi, J.; Nguku, E.; Tonnang, H.; Raina, S. Suitability of resampled multispectral datasets for mapping flowering plants in the Kenyan savannah. *PLoS ONE* **2020**, *15*, e0232313. [CrossRef]

108. Genuer, R.; Poggi, J.; Tuleau-malot, C.; Genuer, R.; Poggi, J.; Variable, C.T.; Forests, R.; Genuer, R.; Poggi, J.; Tuleau-malot, C. Variable selection using Random Forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236. [CrossRef]

109. Google. API Reference. 2022. Available online: https://developers.google.com/earth-engine/apidocs/ee-classifier-smilerandomforest (accessed on 26 August 2022).

110. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

111. Liu, H.; Motoda, H. *Feature Extraction, Construction and Selection: A Data Mining Perspective*; Kluwer Academic Publishers: Vancouver, BC, Canada, 1998; ISBN 0792381963.

112. Pal, M.; Foody, G.M. Feature selection for classification of hyperspectral data by SVM. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2297–2307. [CrossRef]

113. Poona, N.K.; Van Niekerk, A.; Nadel, R.L.; Ismail, R. Random Forest (RF) Wrappers for Waveband Selection and Classification of Hyperspectral Data. *Appl. Spectrosc.* **2016**, *70*, 322–333. [CrossRef]

114. Han, H.; Guo, X.; Yu, H. Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. In Proceedings of the 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 26–28 August 2016; Volume 1, pp. 219–224.

115. Belgiu, M.; Drăgu, L.; Drăgut, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]

116. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [CrossRef]

117. Probst, P.; Wright, M.N.; Boulesteix, A.L. Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2019**, *9*, 1–15. [CrossRef]

118. Nygren, R.; Petkov, A. *Evaluation of Hyperparameter Optimization Methods for Random Forest Classifiers*; Kth Royal Institute of Technology School of Electrical Engineering and Computer Science: Stockholm, Sweden, 2021.

119. Rebelo, A.J.; Gokool, S.; Holden, P.B.; New, M.G. Can Sentinel-2 be used to detect invasive alien trees and shrubs in Savanna and Grassland Biomes? *Remote Sens. Appl. Soc. Environ.* **2021**, *23*, 100600. [CrossRef]

120. Shoko, C.; Mutanga, O. Examining the strength of the newly-launched Sentinel 2 MSI sensor in detecting and discriminating subtle differences between C3 and C4 grass species. *ISPRS J. Photogramm. Remote Sens.* **2017**, *129*, 32–40. [CrossRef]

121. Cho, M.A.; Mathieu, R.; Asner, G.P.; Naidoo, L.; van Aardt, J.; Ramoelo, A.; Debba, P.; Wessels, K.; Main, R.; Smit, I.P.J.; et al. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sens. Environ.* **2012**, *125*, 214–226. [CrossRef]

122. Odindi, J.; Mutanga, O.; Rouget, M.; Hlanguza, N. Mapping alien and indigenous vegetation in the KwaZulu-Natal Sandstone Sourveld using remotely sensed data. *Bothalia* **2016**, *46*, 1–9. [CrossRef]

123. Otunga, C.; Odindi, J.; Mutanga, O.; Adjorlolo, C. Evaluating the potential of the red edge channel for C3 (*Festuca* spp.) grass discrimination using Sentinel-2 and Rapid Eye satellite image data. *Geocarto Int.* **2019**, *34*, 1123–1143. [CrossRef]

124. Gitelson, A.A.; Merzlyak, M.N.; Lichtenthaler, H.K. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *J. Plant Physiol.* **1996**, *148*, 501–508. [CrossRef]

125. Rajah, P.; Odindi, J.; Mutanga, O.; Kiala, Z. The utility of Sentinel-2 Vegetation Indices (VIs) and Sentinel-1 Synthetic Aperture Radar (SAR) for invasive alien species detection and mapping. *Nat. Conserv.* **2019**, *35*, 41–61. [CrossRef]
126. César De Sá, N.; Carvalho, S.; Castro, P.; Marchante, E.; Marchante, H. Using Landsat Time Series to Understand How Management and Disturbances Influence the Expansion of an Invasive Tree. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3243–3253. [CrossRef]
127. Ottosen, T.B.; Lommen, S.T.E.; Skjøth, C.A. Remote sensing of cropping practice in Northern Italy using time-series from Sentinel-2. *Comput. Electron. Agric.* **2019**, *157*, 232–238. [CrossRef]
128. Lass, L.W.; Shafii Thill, B.; Prather, T.S. Detecting Spotted Knapweed (*Centaurea maculosa*) with Hyperspectral Remote Sensing. *Weed Technol.* **2002**, *16*, 426–432. [CrossRef]
129. Pastick, N.J.; Dahal, D.; Wylie, B.K.; Parajuli, S.; Boyte, S.P.; Wu, Z. Characterizing land surface phenology and exotic annual grasses in dryland ecosystems using landsat and sentinel-2 data in harmony. *Remote Sens.* **2020**, *12*, 725. [CrossRef]
130. Sage, A.J.; Genschel, U.; Nettleton, D. Tree aggregation for random forest class probability estimation. *Stat. Anal. Data Min.* **2020**, *13*, 134–150. [CrossRef]
131. Chen, M.; Ke, Y.; Bai, J.; Li, P.; Lyu, M.; Gong, Z.; Zhou, D. Monitoring early stage invasion of exotic *Spartina alterniflora* using deep-learning super-resolution techniques based on multisource high-resolution satellite imagery: A case study in the Yellow River Delta, China. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *92*, 102180. [CrossRef]